**fs(1)A1304** is a 5′ UTR deletion of the essential gene *small ovary* in *Drosophila*

Myles Hammond¹, Jillian G. Gomez¹, Brian Oliver², Steve Kucera¹ and Leif Benner²,³§
¹Department of Biology, The University of Tampa, Tampa, FL, 33606
²Section 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
³Department of Biology, Johns Hopkins University, Baltimore, MD, 21218
§To whom correspondence should be addressed: leif.benner@gmail.com

**Abstract**

X-linked female sterile screens in *Drosophila* have led to a tremendous increase in our understanding of the genetic control of oogenesis (Gans et al. 1975; Mohler 1977; Komitopoulou et al. 1983). However, many of the loci in these screens have not been mapped to a single gene and therefore remain a rich resource for further elucidating the genetic control of female fertility. *fs(1)A1304* is one such allele that is germine dependent and results in a degenerative ovary phenotype (Gans et al. 1975; Khipple and King 1976; Mulligan 1981; Wieschaus et al. 1981; Mulligan and Rasch 1985; Lamnissou and Gelti-Douka 1985). We were interested in determining the mutation that leads to sterility in *fs(1)A1304* females. Previous recombination mapping had placed *fs(1)A1304* at 19±2 cM on the X chromosome (Gans, Audit, and Masson 1975; Khipple and King 1976). We confirmed the previous mapping interval by meiotically mapping *fs(1)A1304*.

**Figure 1:** A) Complementation mapping of *fs(1)A1304*. Boxes represent either the portion of the chromosome deleted or duplicated. For deficiencies, green indicates complementing deletions and black indicates non-complementing deletions. For duplications, green indicates rescuing fragments while black indicates non-rescuing fragments. Numbers indicate genomic coordinates in bases along the X chromosome. B) Genomic PCR of wildtype (WT) and *sovA1304-1* flies. Primers were designed to amplify genomic DNA encoding the 5′ UTR region of the *sov-RA* transcript. C) Cartoon of the *sov* locus. Dark blue represents the *sov* gene region with the left arrow representing the *sov-RA* transcriptional start site and right arrow representing the *sov-RB/RC* transcriptional start site. Green represents the *CR43496* gene region with the arrow representing the transcriptional start site. Red box represents the deleted segment in *sovA1304-1* flies. Small rectangles represent untranslated regions while large boxes represent translated regions. Numbers indicate genomic coordinates in bases along the X chromosome.

**Description**

X-linked female sterile screens in *Drosophila* have led to a tremendous increase in our understanding of the genetic control of oogenesis (Gans et al. 1975; Mohler 1977; Komitopoulou et al. 1983). However, many of the loci in these screens have not been mapped to a single gene and therefore remain a rich resource for further elucidating the genetic control of female fertility. *fs(1)A1304* is one such allele that is germine dependent and results in a degenerative ovary phenotype (Gans et al. 1975; Khipple and King 1976; Mulligan 1981; Wieschaus et al. 1981; Mulligan and Rasch 1985; Lamnissou and Gelti-Douka 1985). We were interested in determining the mutation that leads to sterility in *fs(1)A1304* females. Previous recombination mapping had placed *fs(1)A1304* at 19±2 cM on the X chromosome (Gans, Audit, and Masson 1975; Khipple and King 1976). We confirmed the previous mapping interval by meiotically mapping *fs(1)A1304*.
to the right of crossveinless (12 cM) and to the left of singed (22 cM). We began complementation tests for female sterility with known deficiencies tiling the crossveinless and singed region and placed the lesion within a roughly 235 kb region (Figure 1A, non-complementing Df(1)BSC276, BSC285, BSC286, BSC297, BSC351, BSC355, and sov) (Parks et al. 2004; Cook et al. 2012). Two duplications within this narrow region rescued fs(1)A1304\(^1\) sterility and thus further narrowed down the possible location of the causal mutation (Figure 1A, Dp(1;3)DC486 and Dp(1;3)DC026) (Venken et al. 2010). The mapping results were somewhat ambiguous within this narrow region (discussed below). However, the smallest non-complementing deficiency, Df(1)sov, contains only the protein coding gene small ovary (sov) and non-coding RNA gene CR43496. We therefore decided to complementation test fs(1)A1304\(^1\) with known alleles of sov. Flies homozygous for hypomorphic alleles of sov show a similar female sterility phenotype to flies bearing fs(1)A1304\(^1\) while amorphic sov alleles are embryonic lethal (Wayne et al. 1995; Jankovics et al. 2018; Benner et al. 2019). We found that amorphic alleles sov^E442 and sov^ML150 failed to complement fs(1)A1304\(^1\) female sterility while the hypomorphic sov^2 complemented fs(1)A1304\(^2\) sterility. Collectively this indicates that fs(1)A1304\(^1\) is a sov allele (sov^A1304-1).

To determine the molecular lesion, we performed paired-end DNA sequencing on sov^A1304-1 females. The sov locus contains three annotated transcripts; sov-RA has an annotated upstream transcriptional start site while sov-RB/RC are annotated to use a downstream transcriptional start site (Thurmond et al. 2019). Our sequencing data suggested that sov^A1304-1 flies contained a deletion within the sov gene region that would delete a majority of the sov-RA 5′ UTR. Genomic PCR of this potential deletion confirmed the presence of a deletion in sov^A1304-1 flies (Figure 1B). Sanger sequencing of the sov^A1304-1 genomic PCR product showed that there was a 324 nucleotide deletion (chrX:6,756,385-6,756,709) and a 10 nucleotide insertion (TCACCTTCCG) in the sov-RA 5′ UTR and would therefore remove most of the annotated 5′ UTR and donor splice site (Figure 1C).

We are unsure why a duplication (Dp(1;3)DC026) and a deficiency (Df(1)BSC535) to the left of the sov region rescued and failed to complement sov^A1304-1, respectively. We also found that the small duplication of just sov and CR43496 (Dp(1;3)sov^{CH322:191E24}) failed to rescue. We were not able to find any deleterious mutations or structural variants in our sequencing data to the left of sov that might indicate the presence of a second-site suppressor or long-range genomic interactions with the sov locus that are necessary for its proper expression. It is interesting that sov^A1304-1 had not been previously mapped to sov since the Mohler and Gans X-linked female sterile collections had been previously complementation tested inter se (Perrimon et al. 1986). We found that one of the original Mohler alleles, sov^2, complemented sov^A1304-1 sterility and is thus possible that the other two Mohler alleles, sov^1 and sov^3, behaved similarly, providing an explanation as to why sov^A1304-1 was not previously recognized as belonging to the sov locus. It would be interesting to determine if the 5′ UTR deletion of the sov-RA transcript found in sov^A1304-1 flies affects sov activity in other tissues of the body other than the ovary. There is no indication that sov-RA, or sov-RB/RC, is differentially expressed in the ovary or other adult tissues (Benner et al. 2019). Pole cell transplantation studies of sov^A1304-1 indicated that defects are germline dependent (Wieschaus et al. 1981; Lamnissou and Gelti-Douka 1985), however, sov is an essential gene that has been shown to dominantly suppress position-effect variegation in tissues such as the eye (Jankovics et al. 2018; Benner et al. 2019). It is possible that the deletion solely affects sov-RA and that the Drosophila ovary is more sensitive to loss of sov-RA, or sov transcripts in general, in comparison to other tissues since sov^A1304-1 females are viable but sterile. However, we have not directly measured the deletions effects on sov-RB/RC transcript levels, which might also be perturbed. The nature of the sov^A1304-1 deletion therefore provides a unique mechanism to further elucidate the function of Sov at potentially both the transcript and regulatory level in Drosophila.

**Methods**

Request a detailed protocol

Flies were cultured on ‘Fly Food A’ (LabExpress, Ann Arbor, MI) under standard laboratory conditions at 25°C. Genomic DNA was extracted from 30 homozygous fs(1)A1304\(^1\) flies with a Qiagen DNeasy Blood and Tissue Kit (Hilden, Germany) according to the manufacturers insect protocol. DNA-sequencing libraries were made with Illumina Nextera DNA Library Prep Kit (San Diego, CA). 50 nucleotide paired-end sequencing was performed (Illumina HiSeq 2500, CASAVA base calling). Sequencing reads were mapped with Hisat2 to the FlyBase r6.25 genome and are available at the SRA (SRP238927) (Kim et al. 2015; Thurmond et al. 2019). Variant calling was completed with mpileup and bcftools from SAMtools within the X chromosome region 6625450-6660753 (Li et al. 2009; Li 2011) followed with variant annotation software snpEFF (Cingolani et al. 2012). For structural variant calling, we used BreakDancer software (Chen et al. 2009). Sanger sequencing was completed by Genewiz (Plainfield, NJ).

**Reagents**

Deficiencies and duplications in order as they appear in Figure 1 (top to bottom).
Deficiencies:

\[ D_f(1)\text{ED6802} = \text{BDSC 8949} \text{ (or FBst0008949)} \]
\[ D_f(1)\text{BSC654} = \text{BDSC 26506} \text{ (or FBst0026506)} \]
\[ D_f(1)\text{dx81} = \text{BDSC 5281} \text{ (or FBst0005281)} \]
\[ D_f(1)\text{ED418} = \text{BDSC 8032} \text{ (or FBst008032)} \]
\[ D_f(1)\text{ED6829} = \text{BDSC 8947} \text{ (or FBst008947)} \]
\[ D_f(1)\text{Exel6238} = \text{BDSC 7712} \text{ (or FBst007712)} \]
\[ D_f(1)\text{BSC640} = \text{BDSC 25730} \text{ (or FBst0025730)} \]
\[ D_f(1)\text{Exel6239} = \text{BDSC 7713} \text{ (or FBst007713)} \]
\[ D_f(1)\text{Exel6240} = \text{BDSC 7714} \text{ (or FBst007714)} \]
\[ D_f(1)\text{e02477-d06059} = \text{BDSC 39617} \text{ (or FBst0039617)} \]
\[ D_f(1)\text{BSC535} = \text{BDSC 25063} \text{ (or FBst0025063)} \]
\[ D_f(1)\text{BSC285} = \text{BDSC 23670} \text{ (or FBst0023670)} \]
\[ D_f(1)\text{BSC351} = \text{BDSC 24375} \text{ (or FBst0024375)} \]
\[ D_f(1)\text{BSC297} = \text{BDSC 23681} \text{ (or FBst0023681)} \]
\[ D_f(1)\text{BSC286} = \text{BDSC 23671} \text{ (or FBst0023671)} \]
\[ D_f(1)\text{BSC276} = \text{BDSC 23661} \text{ (or FBst0023661)} \]
\[ D_f(1)\text{sov} = \text{Benner et al., 2019} \]
\[ D_f(1)\text{ED6878} = \text{BDSC 9625} \text{ (or FBst009625)} \]
\[ D_f(1)\text{BSC882} = \text{BDSC 30587} \text{ (or FBst0030587)} \]
\[ D_f(1)\text{BSC867} = \text{BDSC 29990} \text{ (or FBst0029990)} \]
\[ D_f(1)\text{Sxl-bt} = \text{BDSC 3196} \text{ (or FBst003196)} \]
\[ D_f(1)\text{Sxl}^{D7B0} = \text{BDSC 58489} \text{ (or FBst0058489)} \]

Duplications:

\[ D_p(1;3)\text{DC158} = \text{BDSC 30296} \text{ (or FBst0030296)} \]
\[ D_p(1;3)\text{DC159} = \text{BDSC 32268} \text{ (or FBst0032268)} \]
\[ D_p(1;3)\text{DC160} = \text{BDSC 30297} \text{ (or FBst0030297)} \]
\[ D_p(1;3)\text{DC026} = \text{BDSC 30226} \text{ (or FBst0030226)} \]
\[ D_p(1;3)\text{DC486} = \text{BDSC 32306} \text{ (or FBst0032306)} \]
\[ D_p(1;3)\text{sov}^{CH322-19\text{E24}} = \text{Venken et al., 2010} \text{ (or FBal0243261)} \]
\[ D_p(1;3)\text{DC163} = \text{BDSC 32269} \text{ (or FBst0032269)} \]
\[ D_p(1;3)\text{DC164} = \text{BDSC 32270} \text{ (or FBst0032270)} \]

Alleles:

\[ fs(1)\text{A1304}^1 \text{ (sov}^{A1304-1}) = \text{BDSC 4314} \text{ (or FBst0004314)} \]
\[ sov^2 = \text{BDSC 4611} \text{ (or FBst0004611)} \]
\[ sov^{EA42} \text{ (synonymous with l(1)6Dc}^3) = \text{FBal0007068} \]
\[ sov^{ML150} = \text{BDSC 4591} \text{ (or FBst0004591)} \]

Primer \[ fs(1)\text{A1304}^1 \text{ Forward} = TGACCATGTTGCATCTAAGCCA \]
Primer \[ fs(1)\text{A1304}^1 \text{ Reverse} = AGTAGAGCTCGCAATACGCC \]

Acknowledgments: Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) were used in this study. Sequencing was performed by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Genomics Core, under the direction of Harold Smith. Genetic and genomic information was obtained from FlyBase (U41
This work utilized the computational resources of the NIH High-Performance Computing Biowulf cluster (http://hpc.nih.gov).

References


**Funding:** This research was supported in part by the Intramural Research Program of the NIH, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) to BO and by an Undergraduate Research and Inquiry grant from The University of Tampa to MH.

**Author Contributions:** Myles Hammond: Data curation, Formal analysis, Writing - original draft, Writing - review and editing, Investigation. Jillian G. Gomez: Data curation, Formal analysis, Writing - review and editing, Investigation. Brian Oliver: Writing - review and editing, Funding acquisition, Conceptualization, Project administration, Supervision. Steve Kucera: Conceptualization, Writing - review and editing, Funding acquisition, Project administration, Supervision. Leif Benner: Conceptualization, Data curation, Formal analysis, Investigation, Writing - review and editing, Supervision.

**Reviewed By:** Anonymous and Steven Marygold

**History:** Received March 19, 2020 Revision received April 10, 2020 Accepted May 5, 2020 Published May 6, 2020

**Copyright:** © 2020 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Citation:** Hammond, M; Gomez, JG; Oliver, B; Kucera, S; Benner, L (2020). fs(1)A13041 is a 5′ UTR deletion of the essential gene small ovary in *Drosophila*. microPublication Biology. https://doi.org/10.17912/micropub.biology.000246