Title: The seminal odorant binding protein Obp56g is required for mating plug formation and male fertility in *Drosophila melanogaster*

Authors: Nora C. Brown[1], Benjamin Gordon[1,2], Caitlin E. McDonough-Goldstein[3], Snigdha Misra[1,4], Geoffrey D. Findlay[1,5], Andrew G. Clark*[1], Mariana F. Wolfner*[1]

1. Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, United States
2. Present address: Department of Physiology and Biophysics, University of Illinois College of Medicine, Chicago, IL, United States
3. Department of Evolutionary Biology, University of Vienna, Vienna, Austria
4. Present address: University of Petroleum and Energy Studies, Dehradun, UK, India
5. Department of Biology, College of the Holy Cross, Worcester, MA, United States

*Corresponding authors:
Andrew G. Clark
227 Biotechnology Bldg.
526 Campus Road
Ithaca, NY 14853-2703
ac347@cornell.edu

Mariana F. Wolfner
423 Biotechnology Bldg.
526 Campus Road
Ithaca, NY 14853-2703
mfw5@cornell.edu

ORCID IDs:
0000-0001-8567-1273 (NCB)
0000-0002-3856-0500 (BG)
0000-0001-8949-7479 (CEMG)
0000-0001-9435-4464 (SM)
0000-0001-8052-2017 (GDF)
0000-0001-7159-8511 (AGC)
0000-0003-2701-9505 (MFW)

Abstract:

In *Drosophila melanogaster* and other insects, the seminal fluid proteins (SFPs) and male sex pheromones that enter the female with sperm during mating are essential for fertility and induce profound post-mating effects on female physiology and behavior. The SFPs in *D. melanogaster* and other taxa include several members of the large gene family known as odorant binding proteins (Obps). Previous work in *Drosophila* has shown that some *Obp* genes are highly expressed in the antennae and can mediate behavioral responses to odorants, potentially by binding and carrying these molecules to odorant receptors. These observations have led to the hypothesis that the seminal Obps might act as molecular carriers for pheromones or other compounds important for male fertility in the ejaculate, though functional evidence in any species is lacking. Here, we used RNAi and CRISPR/Cas9 generated mutants to test the role of the seven seminal Obps in *D. melanogaster* fertility and the post-mating response (PMR). We found that *Obp56g* is required for male fertility and the induction of the PMR, whereas the other six genes had no effect on fertility when mutated individually. *Obp56g* is expressed in the male's ejaculatory bulb, an important tissue in the reproductive tract that synthesizes components of the mating plug. We found males lacking *Obp56g* fail to form a mating plug in the mated female's reproductive tract, leading to ejaculate loss and reduced sperm storage. We also examined the evolutionary history of these seminal *Obp* genes, as several studies have documented rapid evolution and turnover of SFP genes across taxa. We found extensive lability in gene copy number and evidence of positive selection acting on two genes, *Obp22a* and *Obp51a*. Comparative RNAseq data from the male reproductive tract of multiple *Drosophila* species revealed that *Obp56g* shows high male reproductive tract expression only in species of the *melanogaster* and *obscura* groups, though conserved head expression in all species tested. Together, these functional and expression data suggest that *Obp56g* may have been co-opted for a reproductive function over evolutionary time.

Introduction:

In many taxa, males transfer non-sperm seminal fluid proteins (SFPs) in the ejaculate to females during mating. Odorant binding proteins (Obps) are a common class of SFPs, which have been found in the seminal fluid (or expressed in male reproductive tissues) in a variety of invertebrate species such as mosquitoes (Sirot et al., 2008), honeybees (Baer et al., 2012), flour beetles (Xu et al., 2013), bollworm moths (Sun et al., 2012), tsetse flies (Savini et al., 2021) and *Drosophila* (Begun et al., 2006; Findlay et al., 2008; Karr et al., 2019; Kelleher et al., 2009). Obps have also been described in the seminal fluid of rabbits and the vaginal fluid of hamsters, though vertebrate and insect *Obp* genes are considered non-homologous and have different structures (Mastrogiacomo et al., 2014; Singer et al., 1986; Vieira and Rozas, 2011). Despite their widespread appearance in male seminal fluid across species, the reproductive functions of these Obps are entirely uncharacterized.

In *Drosophila melanogaster*, there are 52 members in the *Obp* gene family, many of which are highly expressed and extremely abundant in olfactory tissues such as antennae and maxillary palps (Rihani et al., 2021; Sun et al., 2018; Vieira and Rozas, 2011). In contrast to odorant receptors, several of which respond to specific odorants *in vivo*, Obps are less well characterized functionally (Ai et al., 2010; Gomez-Diaz et al., 2013; Ha and Smith, 2006; Hallem and Carlson, 2006; Jeong et al., 2013; Sun et al., 2018; Xiao et al., 2019; Xu et al., 2005). Some Obps bind odorants *in vitro*, and mutants of *Obp76a* (*lush*) show abnormal behavioral responses to alcohols and the male sex pheromone cis-vaccenyl acetate (cVA) (Billeter and Levine, 2015; Kim et al., 1998; Xu et al., 2005). These data, combined with the presence of Obps in the aqueous sensillar lymph that surrounds the dendrites of odorant receptor neurons, have led to the model that Obps bind hydrophobic odorants and help transport them across the lymph to their receptors (reviewed in Rihani et al., 2021). However, recent functional data demonstrating robust olfactory responses in the absence of abundant antennal Obps complicate this model and suggest Obps may have roles beyond strictly facilitating chemosensation (Xiao et al., 2019).

Obps are widely divergent at the amino acid level in *Drosophila*, sharing about 20% average pairwise amino acid identity gene family-wide (Hekmat-Scafe et al., 2002; Vieira et al., 2007). However, they share a conserved pattern of 6 cysteines with conserved spacing, which contribute to the formation of disulfide bonds that stabilize the alpha-helical structure (Rihani et al., 2021; Vieira et al., 2007; Vieira and Rozas, 2011). Evolutionarily, divergence in *Obp* gene

copy number in *Drosophila* is consistent with birth-and-death models of gene family evolution, with new members arising via duplication (Rondón et al., 2022; Vieira et al., 2007; Vieira and Rozas, 2011). Genic and expression divergence have been reported for several Obps across *Drosophila*, leading to the hypothesis that turnover in this family may be important for the evolution of substrate preference and niche colonization (Kopp et al., 2008; Matsuo, 2008; Matsuo et al., 2007; Pal et al., 2022; Yasukawa et al., 2010). However, Obps in *Drosophila* and other species have wide expression patterns in larval and adult tissues (including non-chemosensory tissues), suggesting diverse roles for these proteins beyond chemosensation (reviewed in (Rihani et al., 2021)). Indeed, *Obp28a* has been implicated as a target of regulation by the gut microbiota, which stimulates larval hematopoiesis in *Drosophila* and tsetse flies (Benoit et al., 2017).

In *Drosophila*, two olfactory Obps have been implicated in male mating behavior: *Obp76a* (*lush*) and *Obp56h* (Billeter and Levine, 2015; Shorter et al., 2016; Xu et al., 2005). In males, *lush* is required for proper chemosensation of cVA in mated females through the action of *Or67d* in T1 trichoid sensilla (Billeter and Levine, 2015; Kurtovic et al., 2007; Laughlin et al., 2008; Xu et al., 2005). Knockdown of *Obp56h* in males decreases mating latency and alters pheromone profiles, including a strong reduction in the inhibitory sex pheromone 5-tricosene (5-T), indicating *Obp56h* might be involved in sex pheromone production or detection (Shorter et al., 2016).

In addition to the Obps that are transferred in the seminal fluid, intriguingly, several tissues in *D. melanogaster* males produce sex-specific pheromones that are transferred to females during mating. These pheromones include oenocyte-derived 7-tricosene (7-T), ejaculatory bulb-derived cVA and (3R,11Z,19Z)-3-acteoxy-11,19-octacosadien-1-ol (CH503), and accessory gland-derived peptide prohormones (such as Sex Peptide, discussed below) (Brieger and Butterworth, 1970; Everaerts et al., 2010; Guiraudie-Capraz et al., 2007; Scott, 1986; Yew et al., 2009). These molecules have been shown to act individually (in the case of Sex Peptide and CH503) or synergistically in a blend (in the case of cVA and 7-T) to decrease the attractiveness or remating rate of females with other males (reviewed in (Billeter and Wolfner, 2018), (Laturney and Billeter, 2016)). The coincidence of pheromones and Obps being transferred in the seminal fluid during mating have led many to hypothesize that Obps could act as molecular carriers for these molecules in mating, though direct evidence that seminal Obps impact any aspect of female post-mating behavior is lacking.

D. melanogaster SFPs are produced and secreted by the tissues in the male reproductive tract, including the testes, accessory glands (AGs), ejaculatory duct (ED), and ejaculatory bulb (EB) (reviewed in Wigby et al., 2020). Many SFPs are essential for optimal fertility and the induction of the post-mating response (PMR), a collection of behavioral and physiological changes in mated females that include increased egg laying and decreased likelihood of remating (reviewed in (Avila et al., 2011; Wigby et al., 2020)). The induction and maintenance of this response requires the SFPs Sex Peptide (SP) and the long-term response network proteins, which act in a pathway to bind SP to sperm in the female sperm storage organs (Findlay et al., 2014; Ravi Ram and Wolfner, 2009; Singh et al., 2018). Disrupting the presence of sperm in storage, the transfer of SP/network proteins, or the binding and release of SP from sperm leads to a loss of the persistence of the PMR and decreased fertility of the mating pair (Findlay et al., 2014; Kalb et al., 1993; Liu and Kubli, 2003; Misra et al., 2022; Peng et al., 2005; Ravi Ram and Wolfner, 2009; Singh et al., 2018).

A subset of the genes that encode SFPs display interesting evolutionary patterns in many taxa, including elevated sequence divergence consistent with positive selection (or in some cases, relaxed selection), tandem gene duplication, rapid turnover between species, and gene co-option (Ahmed-Braimah et al., 2017; Begun et al., 2006; Begun and Lindfors, 2005; Findlay et al., 2009, 2008; Haerty et al., 2007; McGeary and Findlay, 2020; Mueller et al., 2005; Patlar et al., 2021; Sirot et al., 2014; Swanson et al., 2001; Swanson and Vacquier, 2002). In studies of *Drosophila*, the Obps present in the seminal fluid are composed of both overlapping and distinct sets of proteins between species, mirroring a common feature of SFP evolution: conservation of functional class despite turnover of the individual genes (Findlay et al., 2009, 2008; Karr et al., 2019; Kelleher et al., 2009; Mueller et al., 2004). This pattern is thought to be driven by sexual selection such as sperm competition and male/female intra-sexual conflict, which has been hypothesized to drive molecular arms races between or within the sexes while maintaining functionality of the reproductive system (Avila et al., 2011; Sirot et al., 2015).

Here, we investigate the evolution and reproductive function of seven *D. melanogaster* seminal Obps (Obp8a, Obp22a, Obp51a, Obp56e, Obp56f, Obp56g, and Obp56i) that have been shown to be transferred to females during mating or expressed in SFP-generating tissues (Findlay et al., 2008; Sepil et al., 2019). Using a functional genetic approach, we find that six of the seminal Obps have no or a very marginal effect on the PMR in mated females. However, one Obp,

Obp56g, is required for full male fertility and strong induction of the PMR. We further find that *Obp56g* is expressed in the male ejaculatory bulb, loss of *Obp56g* leads to loss of the mating plug in the female reproductive tract after mating, and this loss leads to a reduction in the number of sperm stored in the mated female. Using comparative RNAseq data across *Drosophila* species, we find *Obp56g* has conserved expression in the head, though expression in the male reproductive tract only in subset of species, suggesting potential co-option of this protein for reproductive function over evolutionary time. Finally, we investigate the molecular evolution of the seminal Obps across a phylogeny of 22 *Drosophila* species. Our results indicate duplication and pseudogenization have played an important role in the evolution of seminal Obps, as well as recurrent positive selection acting on a subset of these genes.

Materials and Methods:

Fly stocks and husbandry:

Flies were reared and mating assays performed on a 12-hour light/dark cycle on standard yeast/glucose media in a 25°C temperature-controlled incubator.

We used the following lines in this study: BL#55079 (*w[*]; TI{w[+mW.hs]=GAL4}Obp56g[1]*) (Jeong et al., 2013); *UAS-CD4-tdGFP* (Han et al., 2011); *LHm pBac{Ubnls-EGFP, ProtB-eGFP}(3)* (a gift from J. Belote and S. Pitnick, Syracuse University) (Manier et al., 2010); Canton-S (CS) ; *w1118* ; BL#25678 (*w[1118]; Df(2R)BSC594/CyO*) (Cook et al., 2012); *w;Gla/CyO* ; *w;;TM3/TM6b* ; BL#3704 (*w[1118]/Dp(1;Y)y[+]; CyO/Bl[1]; TM2/TM6B, Tb[1]*) ; *y1 w1118; attP2{nos-Cas9}/TM6C,Sb Tb)* (Kondo and Ueda, 2013); BL#51324 (*w[1118]; PBac{y[+mDint2] GFP[E.3xP3]=vas-Cas9}VK00027*) ; VDRC#23206 (UAS-Obp56g^{RNAi} from the GD library); BL#49409 (*w[1118]; P{y[+t7.7] w[+mC]=GMR64E07-GAL4}attP2*) (Jenett et al., 2012); *C(1)DX, y[1] w[1] f[1]/FM7c, Kr-GAL4[DC1], UAS-GFP[DC5], sn[+];;;* (a gift from Susan Younger, University of California San Francisco); *Tubulin*-GAL4 (Findlay et al., 2014); BL#35569 (*y[1] w[*] P{y[+t7.7]=nos-phiC31int.NLS}X; PBac{y[+]-attP-9A}VK00027*). We obtained lines of *D. ananassae, D. pseudoobscura, D. mojavensis,* and *D. virilis* from the *Drosophila* Species Stock Center at Cornell University.

To generate males varying in numbers of copies of *Obp56g*, we used a line carrying the *Obp56g*^{1} mutant allele, which is a complete replacement of the *Obp56g* coding sequence with a GAL4 mini-*white* cassette (Jeong et al., 2013). We crossed homozygous *Obp56g*^{1} flies with *Df(2R)BSC594/CyO* to generate trans-heterozygous *Obp56g*^{1} over a deficiency of chromosome

2R, or *Obp56g*^{1} balanced over *CyO* (which have zero and one copy of functional *Obp56g*, respectively). We then crossed *w*^{1118} (the genetic background of the *Obp56g*^{1} null line) with *Df(2R)BSC594/CyO* to obtain *+/Df(2R)* or *+/CyO* males (which have one and two copies of functional *Obp56g*, respectively).

To knock down expression of *Obp56g* in males, we drove a UAS-dsRNA construct against *Obp56g* (VDRC#23206) using the ubiquitous *Tubulin*-GAL4 driver (Lee and Luo, 1999). Control males were the progeny of UAS-*Obp56g*^{RNAi} crossed to *w*^{1118}.

To knock down expression of *Obp56g* in the male ejaculatory duct and bulb, we drove UAS-*Obp56g*^{RNAi} with a *CrebA*-GAL4 enhancer trap driver (Avila et al., 2015; Jenett et al., 2012). Control males were the progeny of *CrebA*-GAL4 crossed to *w*^{1118}.

Construction of gRNA-expressing lines and CRISPR genome editing:

To generate individual *Obp* null alleles, we used a co-CRISPR approach to target each *Obp* gene along with the gene *ebony* as previously described for *Drosophila* (Kane et al., 2017). To this end, we opted for a strategy in which transgenic multiplexed gRNA expressing lines were crossed to germline Cas9 expressing lines (see Figure 2—figure supplement 1 for full crossing scheme).

To generate our gRNA constructs, we used flyCRISPR's Optimal Target Finder tool to design three gRNAs per *Obp* gene (two guides targeting the 5' CDS of the gene, the third guide targeting the 3' end, Table S1) (Gratz et al., 2014). We then integrated these gRNA sequences (and a gRNA targeting *ebony*) into pAC-U63-tgRNA-Rev, a plasmid that expresses multiplexed gRNAs under the control of the U6:3 promoter (Table S2 & S3, supplemental methods) (Kane et al., 2017; Poe et al., 2019). The resulting plasmids were injected into BL#35569 (*y[1] w[*] P{y[+t7.7]=nos-phiC31int.NLS}X; PBac{y[+]-attP-9A}VK00027*) embryos by Rainbow Transgenic Flies, and integrated into the third chromosome attP^{VK27} site via PhiC31-mediated integration.

For the autosomal *Obp* SFP genes, each stable transgenic gRNA line was crossed to *yw;;nos-Cas9attP2* flies in the P0 generation, and the resulting P1 progeny were crossed to *w; CyO/Bl; TM2,e/TM6B,e* as in (Kane et al., 2017). Resulting F1 *ebony/TM6B,e* or *ebony/TM2,e* flies were backcrossed for two generations to *w;Gla/CyO* to isolate mutant *Obp* alleles (and to remove third chromosome *ebony* mutations). The *Obp* mutant lines were then maintained as a

heterozygous stock over *CyO* in a *white*$^{-}$ background (see Figure 2—figure supplement 1 for the detailed crossing scheme). All mutations were validated using PCR and Sanger sequencing with primers that target ~150 bp upstream and downstream of each *Obp* gene (Table S3, Table S4).

For *Obp8a*, which is X-linked, the crossing scheme was the same as above except that we used *w;;vasa-Cas9* to avoid introducing *Obp* mutations on a *yellow*$^{-}$ chromosome (Figure 2—figure supplement 1). Additionally, we used an *FM7c* balancer line instead of *w;Gla/CyO*.

For the mating assays, we used homozygous null *Obp* mutants (*Obp*^{mut}) and their heterozygous *Obp*^{mut}*/CyO* siblings as controls. For *Obp8a* mutants, we used unedited males from sibling lines as controls.

Verifying levels of knockdown:

We used RT-PCR to assess the level of expression of *Obp56g* in our experimental and control knockdown flies. We extracted RNA from whole flies using RNAzol, treated the samples with DNase (Promega), and synthesized cDNA as previously described (Chen et al., 2019), (Sigma-Aldrich). *Obp56g* was then amplified via RT-PCR, using *Rpl32* as a positive control, and dH_2O as a negative control. For *Obp56g* RNAi, we removed the heads of the flies prior to extracting RNA from the rest of the body, which was necessary to increase sensitivity to detect reproductive tract expression, since *Obp56g* is expressed in the head (Galindo and Smith, 2001; Jeong et al., 2013).

Mating assays:

We collected unmated flies under CO_2 anesthesia and aged males and females in separate vials for 3-5 days post-eclosion. We randomly assigned females to a given male genotype and observed single pair copulations, after which we removed the male using an aspirator. The experimenter was then blinded from the genotype of the male for the duration of the experiment. We discarded any mating pair that copulated for an unusually short duration (<10 minutes) as previously described (LaFlamme et al., 2012). Each mating assay was performed two independent times.

Mating latency was measured as the time difference between introducing the male into the vial and the beginning of mating. Mating duration was measured as the time difference between the

end of mating and the beginning of mating. Time data were converted to minutes using the R package chron (version 2.3-58), and statistical differences between male genotypes were tested using Student's T-tests in R (James and Hornik, 2022).

Mating assays (female egg laying, egg hatchability, and female remating rate) were performed as previously described (Findlay et al., 2014). We assessed statistical significance for egg counts using a generalized linear mixed effects model using the lme4 package (version 1.1-30) in R version 4.2.1, where male genotype and day were included as fixed effects, and vial was included as a random effect, as previously described (Bates et al., 2015; Findlay et al., 2014; LaFlamme et al., 2012). Egg laying was modeled using a Poisson distribution, and the fit of the full model was compared against a reduced model where male genotype was dropped, using the R function aov. We accounted for false discovery rate by applying a Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). To assess on which day differences among genotypes were significant, we performed pairwise comparisons on estimated marginal means between days and genotypes using the R package emmeans (version 1.8.1-1) (Lenth et al., 2022). Significance in egg hatchability was assessed the same way, except we used a binomial distribution as previously described (LaFlamme et al., 2012). We assessed statistical significance for differences in female remating rates between two male genotypes using Fisher's exact tests, and tests for equality of proportions when comparing across more than two male genotypes.

To assess mating plug formation and sperm storage, we crossed a *ProtamineB-eGFP* transgene (Manier et al., 2010) into the *Obp56g*^{1} background to visualize sperm directly. We observed single pair matings between CS females and either *Obp56g*^{1}*/Obp56g*^{1}*; ProtB-eGFP* or *Obp56g*^{1}*/CyO; ProtB-eGFP* males. Females were flash frozen in liquid nitrogen immediately after the end of mating. We dissected the lower female reproductive tract (including the bursa, seminal receptacle, and spermathecae) into ice cold PBS, mounted the tissue in a drop of PBS, and added a coverslip. The tissue was imaged on an ECHO-Revolve microscope using a 10X objective with a FITC LED light cube to visualize the autofluorescent mating plug, and each female was scored as having a mating plug present or absent. Statistical significance in mating plug presence vs. absence was assessed using Fisher's exact tests. Sperm counts using these male genotypes were performed similarly, with mated CS females flash frozen either 3 hours or 4 days after the start of mating (ASM). To facilitate sperm counting, the SR was unwound using forceps, and the spermathecal caps were gently crushed under the coverslip to release sperm.

Sperm from both spermathecal caps was counted per individual. Statistical significance in sperm counts was assessed using Student's T-tests in R.

To assess sperm transfer during mating, we flash froze copulating pairs of CS females and either *Obp56g*^{1}*/Obp56g*^{1}*; ProtB-eGFP* or *Obp56g*^{1}*/CyO; ProtB-eGFP* males in liquid nitrogen 12 minutes ASM, a time point when efficient transfer of both sperm and seminal fluid components has finished (Gilchrist and Partridge, 2000; Lung and Wolfner, 2001). Frozen males and females were gently separated at the genitalia, and the female reproductive tract was dissected and scored as described above for the presence/absence of the sperm mass and mating plug.

Expression patterns:

To determine male expression patterns of *Obp56g* in the reproductive tract, we crossed the deletion line of *Obp56g* (BL#55079), which is a promoter-trap GAL4 line, to a *UAS-CD4-tdGFP* line to generate *Obp56g-GAL4 > UAS-CD4-tdGFP* flies (Jeong et al., 2013). Unmated males were aged 3-5 days, and entire reproductive tracts were dissected into ice cold PBS. The tissue was mounted in PBS and a coverslip was added. The tissue was imaged using an ECHO-Revolve microscope as described above, using the FITC light cube to visualize live GFP fluorescence. The ejaculatory bulb is known to autofluorescence due to the seminal protein PEB-me (Lung and Wolfner, 2001), so as a negative control we imaged reproductive tracts from *UAS-CD4-tdGFP* males.

We tested for expression of the other seminal Obps in different parts of the male reproductive tract using semi-quantitative RT-PCR on four dissected tissues: testes (with seminal vesicles), accessory glands, ejaculatory ducts, and ejaculatory bulbs. We dissected each tissue from ~30 3–5-day old Canton-S males directly into RNAzol, and prepared cDNA as described above. As a positive control for each tissue, we amplified *Actin5C*. As a negative control, we prepared RNA samples for each tissue that were not treated with reverse transcriptase. Additionally, we analyzed previously published single nucleus RNAseq data from the Fly Cell Atlas, using scripts from (Raz et al., 2022) to load the loom file, scale, and normalize the expression data from the stringent 10X male reproductive gland sample using Seurat (version 4.2.0), SeuratDisk (version 0.0.0.9020), and ScopeLoomR (version 0.13.0) in R (Hoffman, 2022; Li et al., 2022; Satija et al., 2015).

To examine *Obp* expression patterns across species, we used publicly available RNAseq data from dissected tissues and whole bodies for the following species of *Drosophila: melanogaster, yakuba, ananassae, pseudoobscura, persimilis, willistoni, virilis,* and *mojavensis* (Yang et al., 2018). Gene level read counts were obtained from this study (GSE99574) based on HiSAT2 alignments to the FlyBase 2017_03 annotation. Counts were then normalized within species for genes with at least one read across all samples in DEseq2 with a median ratio method, then log2 normalized with an added count of 1.

To verify the expression patterns seen in the RNAseq dataset, and to determine which tissue of the reproductive tract was responsible for expression, we performed semi-quantitative RT-PCR for *Obp56g* from dissected heads, accessory glands, ejaculatory bulbs, and carcasses from males of *Drosophila* species: *melanogaster, ananassae, pseudoobscura, virilis,* and *mojavensis*. For each species, we reared flies and separated males and females under CO_2 anesthesia and aged the males to sexual maturity (Ahmed-Braimah et al., 2017; Karr et al., 2019; Kelleher et al., 2009; Tsuda et al., 2015). We dissected tissues from ~25 males directly into RNAzol, and prepared cDNA as described above. We designed species-specific primers for *Obp56g* (Table S3) and used *Actin5C* and dH_2O controls.

Western blotting:

To assess the production and transfer of specific seminal proteins, we performed Western blotting on protein extracts from CS females that were mated to either experimental *Obp56g*^{1}*/Df(2R)* or control *Obp56g*^{1}*/CyO* males and flash frozen in liquid nitrogen 35 minutes ASM. For each genotype, we dissected the reproductive tracts from 1 male and 4 mated CS females and performed Western blotting using antibodies against Sex Peptide (SP), CG1656, CG1652, Antares (Antr), CG9997, CG17575, Acp36DE, Ovulin (Acp26Aa), and tubulin as a loading control as previously described (Misra and Wolfner, 2020). Protein extracts were separated on a 12% acrylamide gel, transferred to PVDF membranes, and probed for each seminal protein. Antibodies were used at the following concentrations: Acp26Aa (1:5000), Acp36DE (1:12,000), Antr (1:750), CG9997 (1:750), SP (1:1,000), CG1652 (1:250), CG1656 (1:500), CG17575 (1:500), Tubulin (1:4,000, Sigma-Aldrich T5168) (LaFlamme et al., 2012; Ravi Ram and Wolfner, 2009; Singh et al., 2018).

Evolutionary analysis:

We obtained orthologous coding sequences for each of the seminal Obps from the following 22 *Drosophila* species from NCBI: *melanogaster, simulans, sechellia, erecta, yakuba, ananassae, eugracilis, suzukii, biarmipies, takahashii, elegans, rhopaloa, ficusphila, kikawaii, bipectinata, miranda, pseudoobscura, persimilis, virilis, willistoni, mojavensis,* and *grimshawi*. To do so, we used gene ortholog predictions from the *Drosophila* evolutionary rate covariation ortholog dataset, which was generated using the OrthoFinder2 algorithm (Findlay et al., 2014; Raza et al., 2019). To bolster our ortholog predictions, we performed reciprocal best tBLASTn searches in each of the genomes using the focal *D. melanogaster Obp* gene as the query, retaining only those genes that were reciprocal best hits for study (this filtered ~24% of the predicted orthologs, which were frequently evolutionarily older paralogs from the same genomic cluster). For orthologous gene groups with predicted paralogs, we identified the syntenic region in the target genome by finding orthologs of the flanking genes, assuming conservation of gene order. Additionally, we used RAxML-NG to construct maximum-likelihood phylogenies from the predicted coding sequences to further validate orthology calls for genes with predicted paralogs (Kozlov et al., 2019). Using this syntenic approach, we identified instances where some genes were unannotated by the NCBI Gnomon pipeline. In these situations, we ensured the unannotated genes we retained for our evolutionary analysis had intact open reading frames, splice sites, and lacked premature stop codons. We additionally used InterProScan to ensure these genes had a predicted Obp protein domain (Jones et al., 2014).

We used MUSCLE implemented in MEGA-11 with default settings to align the amino acid sequences, and back-translated the alignment obtain the cDNA alignment (Edgar, 2004; Tamura et al., 2021). We constructed a consensus phylogeny based on a concatenated nucleotide alignment of the *Obp* genes using RAxML-NG, where gaps were used when a particular protein was missing from a species as previously described (Kozlov et al., 2019; McGeary and Findlay, 2020). *Obp51a* was excluded from this concatenated tree due to extensive tandem gene duplication. In RAxML-NG, we used the GTR+Gamma models and performed non-parametric bootstrapping with 1,000 replicates (Kozlov et al., 2019). We used the Transfer Bootstrap Expectation (TBE) as a branch support metric as previously described (Carlisle et al., 2022). We used the top scoring tree topology from RAxML-NG for all analyses run in PAML for genes predicted to be single copy across the *melanogaster* group. For genes with duplications in the *melanogaster* group (*Obp22a* and *Obp51a*), we also constructed gene trees using RAxML-NG, and used those phylogenies in PAML.

For our evolutionary analyses, we used the codeml package in PAML to run branch and sites tests (Edgar, 2004; Kumar et al., 2018; Yang, 2007). For the branch test, we used the consensus phylogeny for all 22 species and compared the likelihood ratio of the "free ratio" model with the M0 model. For the sites tests, we limited species in the analysis to those in the *melanogaster* group to avoid saturation of synonymous sites. For these analyses, we used likelihood ratio tests to compare the M7 with the M8 model. For those genes which showed evidence of positive selection in the M7 vs. M8 comparison, we then performed likelihood ratio tests between models M8 and M8a. For genes in which the M8 model was a significantly better fit, we then used the Bayes empirical Bayes (BEB) predictions to identify specific sites under positive selection. For any genes with significant evidence of positive selection, we detected recombination breakpoints in the *Obp* genes using GARD implemented in DataMonkey, partitioned the genes at the breakpoints and re-ran PAML on each segment separately as previously described (Kosakovsky Pond et al., 2006; McGeary and Findlay, 2020).

Materials availability statement:

All new CRISPR mutants and gRNA lines generated for this study are available upon request.

Results:

Obp56g is required for fecundity and regulates remating rates of mated females

To test the role of the seminal Obps in the long-term post-mating response, we used a co-CRISPR approach to generate individual null alleles in the following genes: *Obp56f, Obp56i, Obp56e, Obp51a, Obp22a,* and *Obp8a* (Table S4). Additionally, we used existing mutant and RNAi lines to perturb *Obp56g* (Jeong et al., 2013). Collectively, we used males of these mutant and RNAi lines to measure the effect of Obp perturbation on egg laying and remating rates of their female mates. Of the seven seminal Obps, only females mated to hemizygous $Obp56g^{1}/Df(2R)$ mutant males laid significantly fewer eggs and were significantly more likely to remate, indicating a loss of the post-mating response (Figures 1B, C & Figure 2 A, B, Figure 1—figure supplement 1). This phenotype was fully recessive, as heterozygous *Obp56g* mutant males ($Obp56g^{1}/CyO$ *or* $Df(2R)/+$) were not significantly different from wildtype (*+/CyO*) males (Figure 1B). We did observe slight changes in egg hatchability, though we note that the fraction of females mated to $Df(2R)/Obp56g^{1}$ males that laid eggs to measure hatchability from is small (Figure 2—figure supplement 2A). None of the other CRISPR mutant lines had a significant effect on egg hatchability, aside from a significant decrease in hatchability in the $Obp8a^{WT}$ line (Figure 2—figure supplement 2B). We observed a small difference in egg numbers and

remating rates between *Obp8a*^{WT} and *Obp8a*$^{\Delta 390}$ lines, but these differences were not consistent across replicates (Figure 2A, B & Figure 1—figure supplement 1C, Table S5)

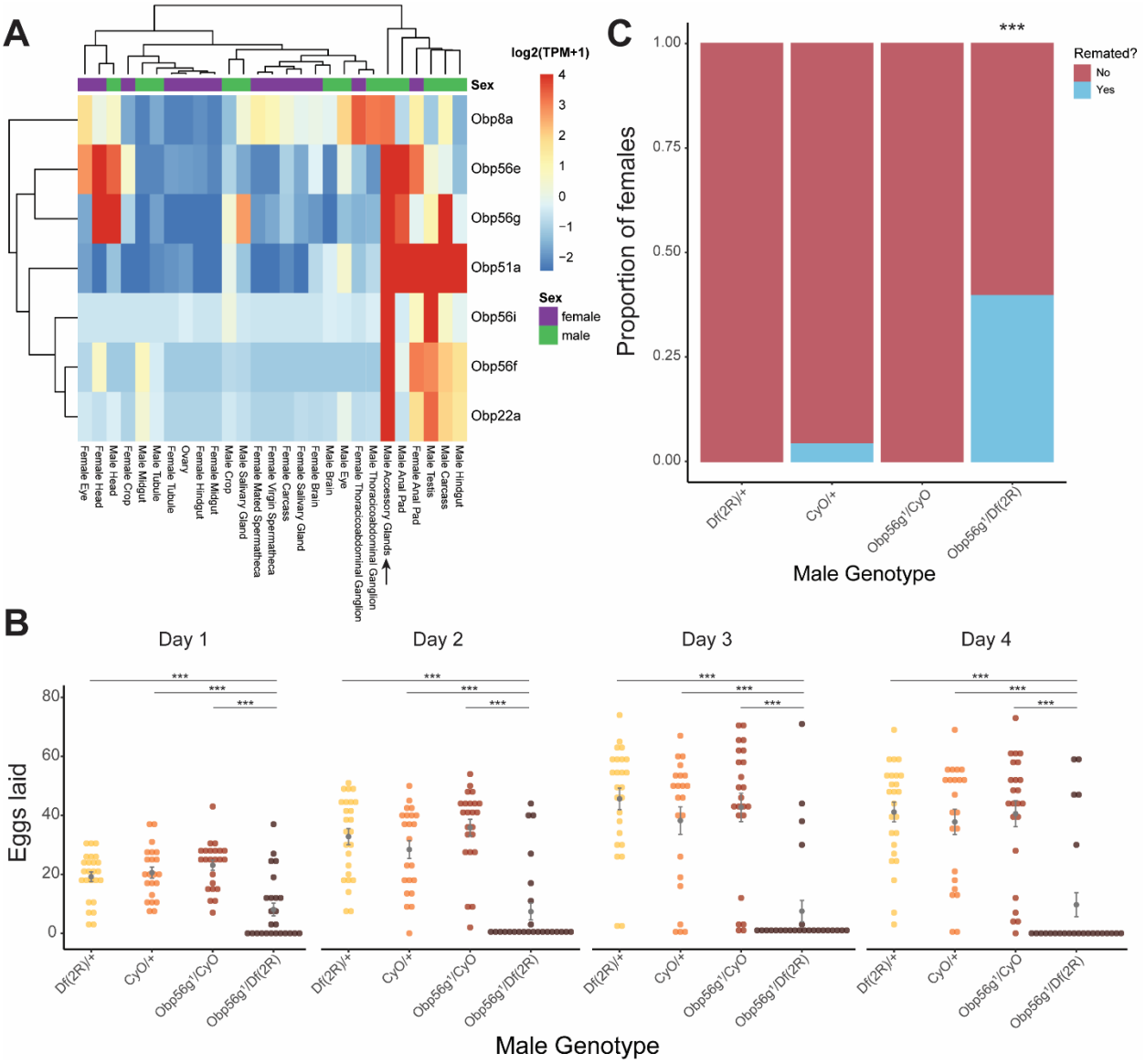


Figure 1: Seminal *Obp* gene expression and fecundity/remating defects in females mated to *Obp56g*^{1} null males. A) Median-centered log2 normalized TPM values for the seven seminal *Obp* genes in adult tissues from FlyAtlas2.0 bulk RNAseq data. Arrow points to male accessory gland sample. B) Egg counts from CS females mated *to Df(2R)/+, CyO/+ Obp56g*^{1}*/CyO*, or *Obp56g*^{1}*/Df(2R)* males from 1-4 days after mating. Significance indicated from pairwise comparisons of male genotypes within days using emmeans on a Poisson linear mixed effects model. Error bars represent mean +/- SEM. C) Proportion of females who did or did not remate with a standard CS male on the fourth day after mating. Significance indicated from tests of equality of proportions. For B and C, n=22-25. Significance levels: *P<0.05, **P<0.01, ***P<0.001, n.s. not significant. One representative biological replicate is shown, but both replicates were significant in the same direction (Figure 1—figure supplement 1A & C).

Figure 1—source data 1: Remating counts and percentages for data shown in Figure 1C.

We tested whether decreased mating duration could account for the decrease in fecundity in females mated to *Obp56g*^{1} mutant males and found no significant difference among the four genotypes tested (Figure 2—figure supplement 3B). Ubiquitous RNAi knockdown of *Obp56g* in males using a *Tubulin*-GAL4 driver recapitulated the phenotype of the hemizygous (*Obp56g*^{1}*/Df(2R)*) mutant, resulting in decreased female egg laying and increased remating rates (Figure 1—figure supplement 2).

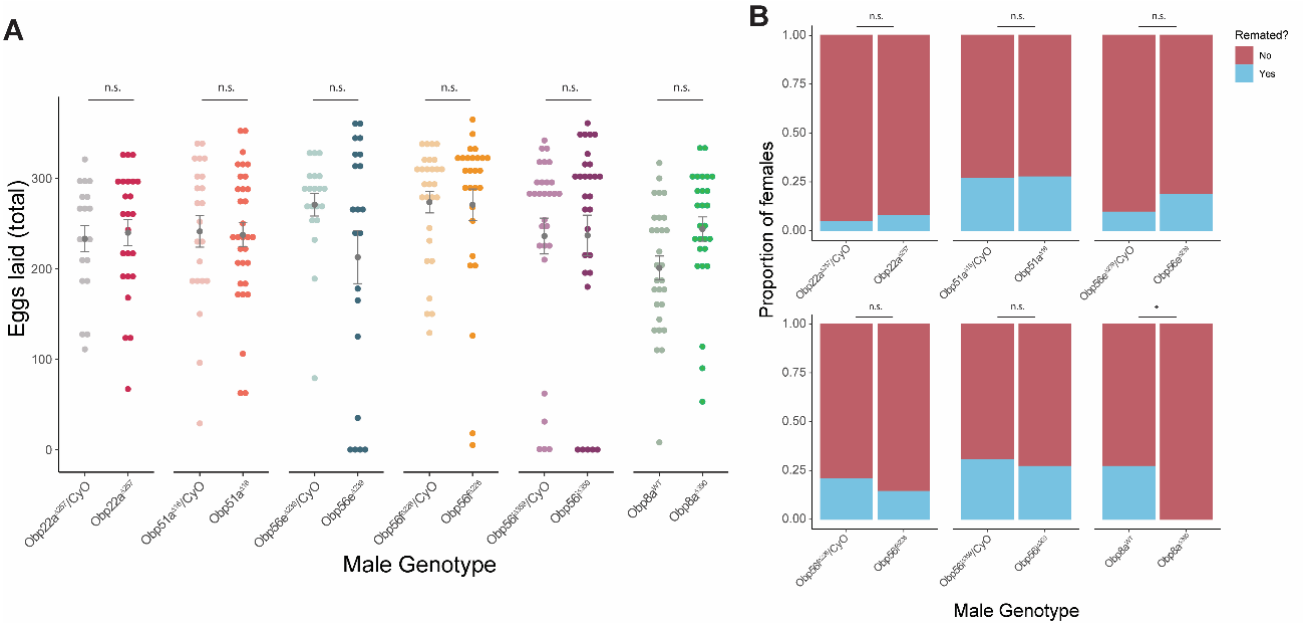


Figure 2: CRISPR/Cas9-generated mutants of *Obp22a, Obp51a, Obp56e, Obp56f, Obp56i*, and *Obp8a* have no or marginal effects on female fecundity and remating rates. A) Egg counts from CS females mated to homozygous null or heterozygous control males (except for *Obp8a*, the control of which is from an unedited sibling line) from 1-4 days after mating. Significance indicated from Poisson linear models with Benjamini-Hochberg corrections for multiple comparisons. Error bars represent mean +/- SEM. B) Proportion of females who did or did not remate with a standard CS male on the fourth day after mating. Significance indicated from Fisher's exact tests with Bejamini-Hochberg correction. Significance levels: *P<0.05, **P<0.01, ***P<0.001, n.s. not significant. For A and B, n=19-32. One representative biological replicate is shown (data from additional replicates can be found in Figure 1—figure supplement 1C, Table S5).

Figure 2—source data 1: Remating counts and percentages for data shown in Figure 2B.

Shorter et al. (2016) reported that male-specific knockdown of *Obp56h*, a paralogous *Obp* gene in the same genomic cluster as *Obp56e, Obp56f, Obp56g,* and *Obp56i*, shortened mating latency times; KD males were faster to mate than control males. RNAseq expression data from

the FlyAtlas2.0 database shows that some of the seminal Obps are co-expressed in other tissues outside of the male reproductive tract, including head tissues (Figure 1A), so we tested whether our mutant lines showed altered mating latency or duration. We did not find any significant differences in either mating latency or duration in any of our mutant lines, aside from a small but statistically significant decrease in mating duration in *Obp8a*^{WT} flies (Figure 2—figure supplement 3).

Obp56g is expressed in the *D. melanogaster* male ejaculatory bulb

While the RNAseq data shown in Figure 1A suggested that *Obp56g* is expressed in the male AG, Findlay et al. (2008) reported that when females are mated to DTA-E males, which are spermless and do not produce main cell accessory gland-derived SFPs (Kalb et al., 1993), transfer of all seminal Obps is lost except for Obp56g. These proteomic data suggest that Obp56g is derived from another (or an additional) tissue within the male reproductive tract. To determine where *Obp56g* is expressed in the male reproductive tract, we crossed the *Obp56g*^{1} mutant line to *UAS-CD4-tdGFP*. We replicated previously published expression patterns for *Obp56g* in the labellum of the proboscis (Figure 3—figure supplement 1), indicating that the promoter-trap GAL4 transgene should recapitulate the true expression patterns of endogenous *Obp56g* (Galindo and Smith, 2001). When we dissected and imaged male reproductive tracts from *Obp56g-GAL4>UAS-CD4-tdGFP* males, we observed strong GFP signal in the ejaculatory bulb epithelium (Figure 3A). The ejaculatory bulb-derived seminal protein PEB-me (also known as *Ebp*) is known to autofluoresce, resulting in autofluorescence of the tissue itself, but the GFP signal we observed in *Obp56g-GAL4>UAS-CD4-tdGFP* males is much stronger than *UAS-CD4-tdGFP* control males (Figure 3B) (Cohen and Wolfner, 2018).

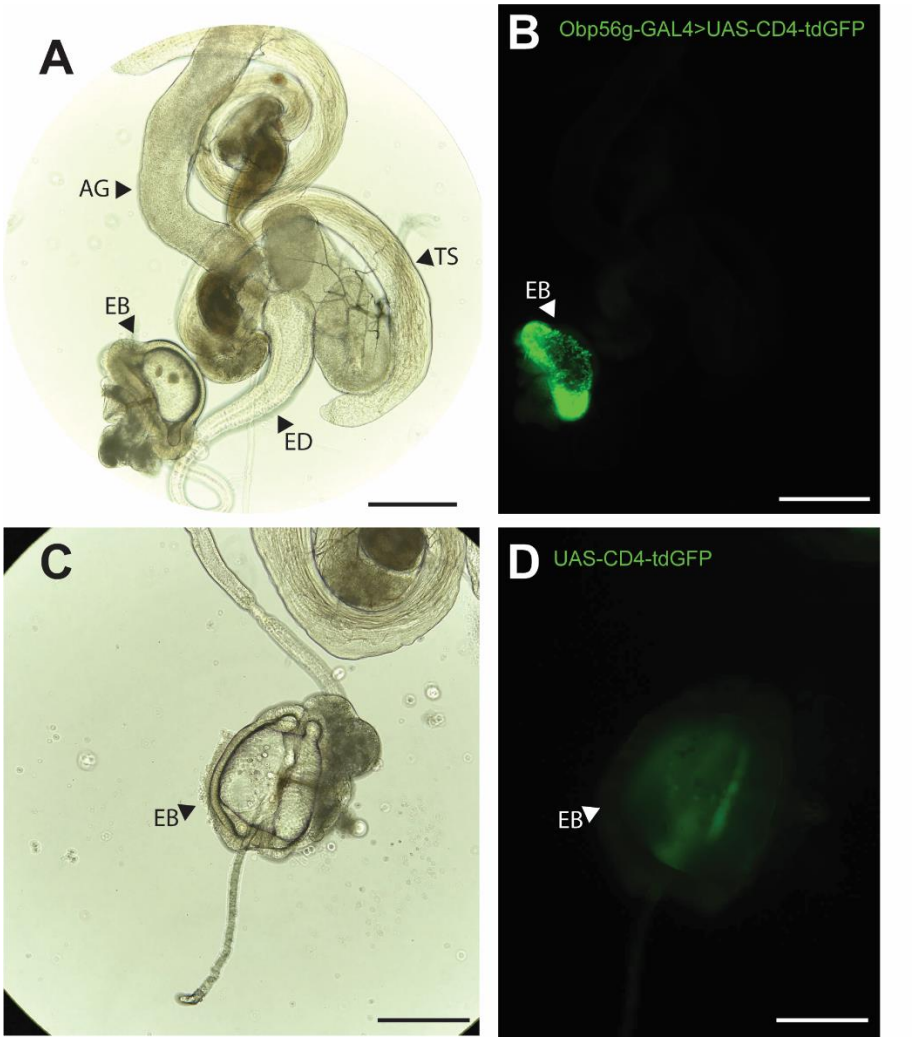


Figure 3: *Obp56g* is expressed in the *Drosophila* male ejaculatory bulb of the reproductive tract. A) Brightfield and B) GFP fluorescent microscopy image of a reproductive tract dissected from a *Obp56g-GAL4>UAS-CD4-tdGFP* male, where the following tissues are labeled: AG, accessory gland. TS, testes. ED, ejaculatory duct. EB, ejaculatory bulb. C) Brightfield and D) GFP fluorescent microscopy images from *UAS-CD4-tdGFP* control males, showing only the EB portion of the tract. Scale bars in A&B=130 um, C&D=70um.

To determine expression patterns for the other seminal Obps, we performed semi-quantitative RT-PCR on dissected testes, accessory gland, ejaculatory duct, and ejaculatory bulb tissues from CS males. Using this approach, we confirmed that *Obp56g* is highly expressed in the ejaculatory bulb, though we also detected expression in the ejaculatory duct and male accessory glands. We observed that the six other *Obp* genes are highly and primarily expressed in the accessory gland and ejaculatory duct (Figure 3—figure supplement 2). We further confirmed these expression patterns in the Fly Cell Atlas scRNAseq data of male reproductive tract tissues (Figure 3—figure supplement 2) (Li et al., 2022).

Obp56g is involved in mating plug formation, ejaculate retention, and sperm storage

Increased egg laying and decreased remating are two phenotypes of the post-mating response that depend on the presence of sperm and SP within the female sperm storage organs (Manning, 1967; Peng et al., 2005). Given that *Obp56g* is expressed in the ejaculatory bulb, and the loss of the post-mating response in *Obp56g* mutant and knockdown males (Figure 1), we wondered whether this loss of fertility could be due to defects in mating plug formation or sperm storage. In order to test this, we crossed a *ProtamineB*-eGFP transgene (Manier et al., 2010), which marks the heads of sperm with GFP, into the $Obp56g^{1}$ mutant line, mated null and control males to females, and directly counted sperm in the female sperm storage organs at 3 hours and 4 days ASM. We also used the autofluorescent nature of PEB-me to score the presence of the mating plug in the female bursa immediately after mating (Lung et al., 2001; Ludwig et al., 1999).

In contrast to $Obp56g^{1}$*/CyO; ProtB-eGFP* control males, which form a fully coagulated mating plug in the female's bursa, we observed homozygous $Obp56g^{1}/Obp56g^{1}$*; ProtB-eGFP* mutant males form much less prominent and non-coagulated mating plugs (Figure 4A & 4B). While the majority of females mated to control males form a mating plug, none of the females mated to $Obp56g^{1}/Obp56g^{1}$*; ProtB-eGFP* males had a fully formed mating plug immediately after the end of mating (Figure 4C). Additionally, at this time point, a subset of females mated to $Obp56g^{1}/Obp56g^{1}$*; ProtB-eGFP* males lacked a sperm mass and had very few or no sperm in their bursa (Figure 4C). To test the possibility that *Obp56g* mutant males have defective sperm transfer, we dissected reproductive tracts from females flash frozen while the flies were still copulating, 12 minutes ASM. In *D. melanogaster*, transfer of mating plug components, SFPs, and sperm begins at 3-5, 3, and 7 minutes, respectively, and is completed by 10 minutes ASM (Gilchrist and Partridge, 2000; Lung and Wolfner, 2001). At this time point, we noted the presence of sperm in the bursa of all females mated to both $Obp56g^{1}/Obp56g^{1}$*; ProtB-eGFP* and $Obp56g^{1}$*/CyO; ProtB-eGFP* males, suggesting the lack of sperm masses immediately after mating is not related to sperm transfer (Figure 4—figure supplement 1). Rather, all females mated to $Obp56g^{1}/Obp56g^{1}$*; ProtB-eGFP* males lacked proper mating plugs at this time point, suggesting loss of the sperm mass is related to issues with ejaculate retention (Figure 4—figure supplement 1). Mutations in the other *Obp* genes had no effect on mating plug formation (Table S6).

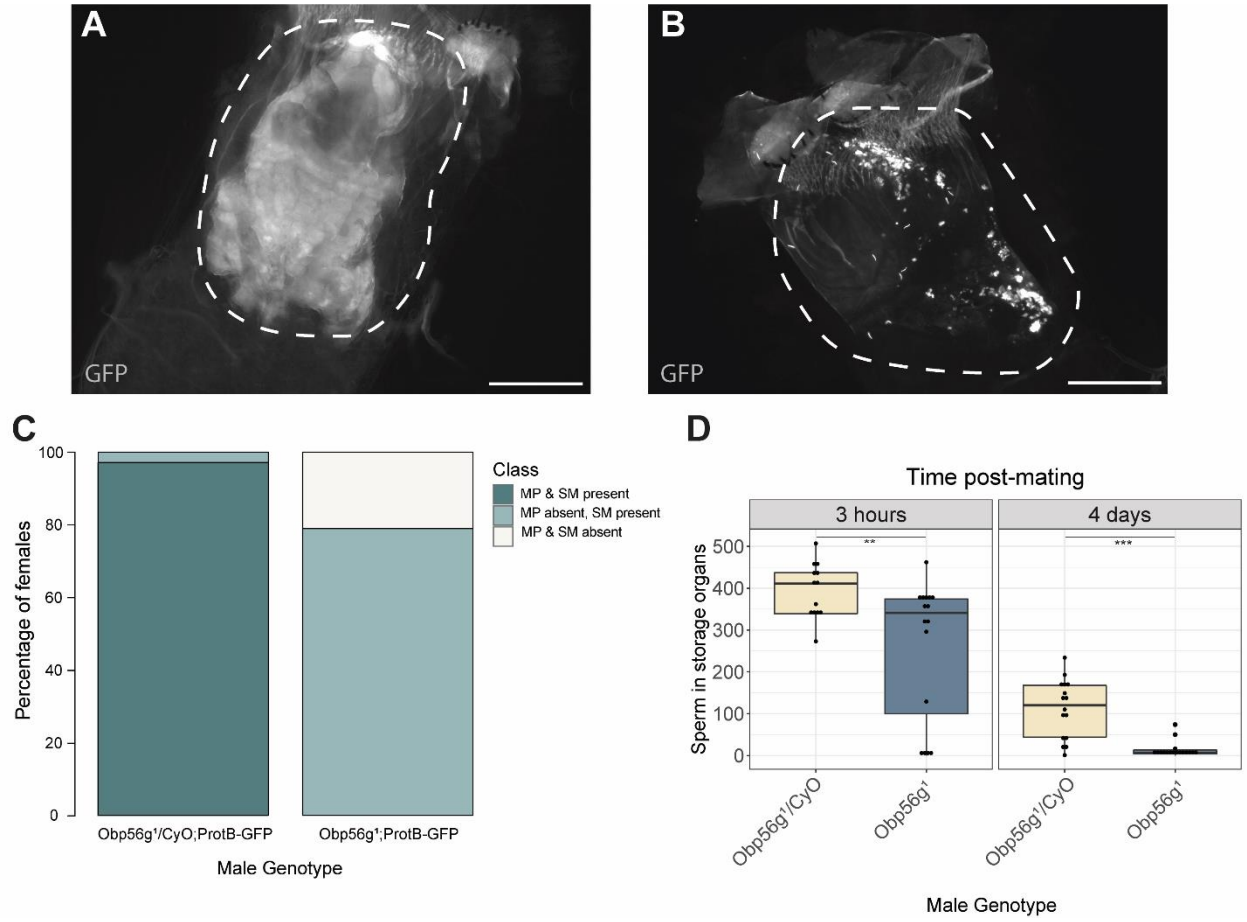


Figure 4: Females mated to *Obp56g*^{1} null males have defects in mating plug formation and sperm storage after mating. A) Fluorescent GFP microscopy image of the bursa of a CS female mated to a *Obp56g*^{1}*/CyO;ProtB-eGFP* control male, with the mating plug surrounded by a dotted white line. Females were frozen in liquid nitrogen immediately after the end of mating. The mating plug is autofluorescent. B) Fluorescent GFP microscopy image of the bursa of a CS female mated to a *Obp56g*^{1}*;ProtB-eGFP* mutant male, where a similar region in the bursa as A) is shown in the dotted white line. C) Proportion of females mated to *Obp56g*^{1}*/CyO;ProtB-eGFP* control or *Obp56g*^{1}*;ProtB-eGFP* mutant males who had mating plugs or sperm masses present or absent immediately after the end of mating (n=35-38). MP, mating plug. SM, sperm mass. D) Box plots of sperm counts in the storage organs of CS females mated to control (*Obp56g*^{1}*/CyO;ProtB-eGFP*) or mutant (*Obp56g*^{1}*;ProtB-eGFP*) males at 3 hours or 4 days post-mating. n=13-17 for each group. Significance indicated from Student's t-tests. Significance levels: *P<0.05, **P<0.01, ***P<0.001, n.s. not significant. Scale bar=130um.

Figure 4—source data 1: Counts and proportions for data shown in Figure 4C.

Previous studies of *D. melanogaster* mating plug proteins Acp36DE and PEB-me reported a reduction in sperm storage when these genes were mutated or knocked down, indicating that integrity of the mating plug is essential for effective sperm storage (Avila et al., 2015; Avila and Wolfner, 2009; Bertram et al., 1996; Neubaum and Wolfner, 1999). At 3 hours and 4 days ASM,

we observed females mated to *Obp56g*^{1}*/Obp56g*^{1}*; ProtB-eGFP* males have significantly fewer sperm in their sperm storage organs than females mated to *Obp56g*^{1}*/CyO; ProtB-eGFP* males, (3 hours mean sperm number *Obp56g*^{1}*/CyO*: 393, mean sperm number *Obp56g*^{1}: 258 $p<0.01$; 4 day mean sperm number *Obp56g*^{1}*/CyO*: 112, mean sperm number *Obp56g*^{1}: 13, $p<0.001$ Figure 4D). These results suggest that the reduction in fecundity we observed in our mating assays is due to issues with sperm retention and subsequent long-term storage in *Obp56g*^{1} mutant males.

We further tested whether male reproductive tract expression of *Obp56g* is required for fertility and mating plug formation by knocking down *Obp56g* using a *CrebA*-GAL4 enhancer-trap driver, which drives expression in the ejaculatory duct and bulb (Avila et al., 2015). We observed that mates of knockdown males showed significantly reduced egg laying and increased remating rates compared to control males, similar to whole body *Obp56g* knockdown and the *Obp56g*^{1} mutant line (Figure 1—figure supplement 3A & C). Additionally, experimental knockdown males had decreased incidence of mating plug formation compared to control males (Figure 1—figure supplement 3B). We also observed instances of ejaculate loss from the bursa of the female after the flies uncoupled, similar to the phenotype previously observed for *PEB-me* knockdown (Figure 1—figure supplement 3D) (Avila et al., 2015). Together, these findings show that ejaculatory duct/bulb expression of *Obp56g* is required for mating plug formation, sperm storage, and the post-mating response.

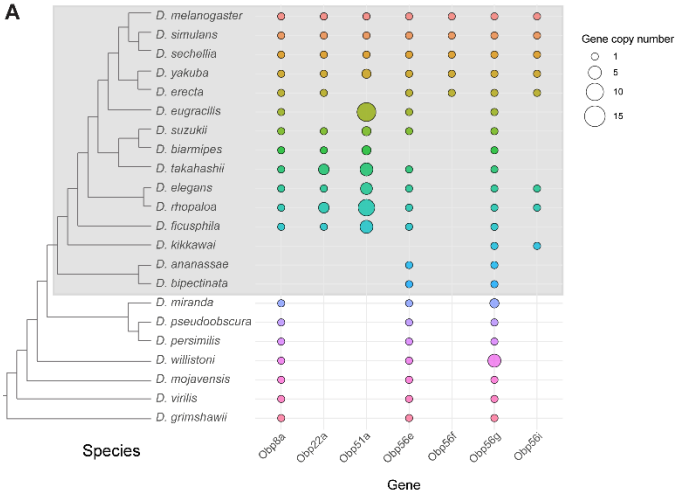
We next tested the possibility that Obp56g may act as a molecular carrier for seminal proteins that promote mating plug formation or the establishment of the post-mating response, such as Sex Peptide. In order to test whether loss of *Obp56g* leads to a loss of particular SFPs in the female reproductive tract after mating, we performed Western blotting on dissected female bursae samples 35 minutes ASM and probed for several SFPs known to be important either for the long term post-mating response or mating plug formation (Avila and Wolfner, 2009; Findlay et al., 2014). We observed no difference in the synthesis of any tested protein in the male reproductive tract between *Obp56g*^{1}*/Df(2R)* and *Obp56g*^{1}*/CyO* males (Figure 4—figure supplement 2, lanes 2 & 3). Rather, we observed a lower signal intensity relative to controls in the bursa of females mated to *Obp56g*^{1}*/Df(2R)* males for CG1652, CG9997, Ovulin, and Acp36DE (and its cleavage products) at 35 minutes ASM, consistent with a defect in ejaculate retention in the mutant condition (Figure 4—figure supplement 2, lanes 4 & 5). In no case did we observe complete loss of any single protein in females mated to *Obp56g*^{1}*/Df(2R)* males,

suggesting that Obp56g likely does not act as the sole or an exclusive carrier for these specific proteins in the seminal fluid.

Seminal Obps have complex evolutionary histories and exhibit evolutionary rate heterogeneity across the *Drosophila* genus

Previous studies have reported elevated rates of divergence and gene turnover of a subset of SFP genes across *Drosophila* (Ahmed-Braimah et al., 2017; Begun et al., 2006; Begun and Lindfors, 2005; Findlay et al., 2008; Mueller et al., 2005; Patlar et al., 2021; Swanson et al., 2001; Wagstaff and Begun, 2005). To examine the evolutionary history of the seminal *Obp* genes, we first identified orthologs of these genes across 22 sequenced species. Combining our orthologous gene predictions with syntenic analysis within each genome allowed us to identify several instances of lineage-specific tandem duplication and loss (Figure 5A, Figure 5—figure supplements 1-6). For example, *Obp8a* and *Obp56e* are single copy and found in most genomes across the genus, with a few predicted losses (Figure 5A, Figure 5—figure supplement 2&4). *Obp56f* and *Obp56i* are also single copy, though restricted to species of the *melanogaster* group (Figure 5A, Figure 5—figure supplement 4&6). *Obp22a* is also only found in *melanogaster* group species and has tandemly duplicated in *D. rhopaloa* and *D. takahashii* (Figure 5—figure supplement 3). *Obp56g* is found in all species across the genus that we examined, and has duplicated several times in the *D. willistoni* lineage to generate four copies (Figure 5A, Figure 5—figure supplement 5). Additionally, in the *obscura* group (*D. miranda, D. pseudoobscura,* and *D. persimilis*), there appears to be an intronless and highly diverged copy of *Obp56g* located immediately adjacent to the conserved gene, possibly the result of a retroduplication. *D. miranda* additionally has a putative Y-linked copy of *Obp56g* which shares 96% amino acid identity with the autosomal copy. *Obp51a*, which is only found in *melanogaster* group species, has the most extreme lability in copy number, ranging from 0 copies to 12 tandem copies in *D. eugracilis* (Figure 5—figure supplement 1). We also found evidence of pseudogenization events in the *Obp22a* and *Obp51a* regions in 5 species, which is consistent with a recent study that found evidence of pseudogenization of *Obp51a* in *repleta* group species (Rondón et al., 2022).

A



B

Gene	M0 ω estimate (genus)	M0 vs. free-ratio		M7 vs. M8		M8 vs. M8a		% of sites in ω > 1	Sites with BEB Pr >0.90
		2*ΔlnL	*p*-value	2*ΔlnL	*p*-value	2*ΔlnL	*p*-value		
Obp8a	0.1830	49.56	0.0245	0.0429	0.978	-	-	-	-
Obp22a	0.5535	46.53	0.00385	34.80	<0.00001	24.52	<0.0001	12%	26E, 40S, 45E, 50R, 129I, 136R, 139G
Obp51a	0.5069	191.66	<0.00001	7.80	0.020	5.280	0.0215	11%	136R
Obp56e	0.1817	93.27	<0.00001	0.0012	0.999	-	-	-	-
Obp56f	0.5079	8.02	0.236	0	0.999	-	-	-	-
Obp56g	0.2355	84.11	0.000514	2.41	0.299	-	-	-	-
Obp56i	0.4589	10.08	0.608	0.641	0.725	-	-	-	-

Figure 5: Dynamic changes in copy number, presence/absence, and evolutionary divergence rates of seminal *Obp* genes across the *Drosophila* genus. A) Inferred copy number of seminal *Obp* genes across Drosophila. Species without a dot represent an inferred loss based on syntenic analysis. Increased size of the dot represents increased gene copy number. Phylogeny on the left from (McGeary and Findlay, 2020). Grey box surrounds species of the *melanogaster* group. B) PAML results for the seminal *Obp* genes from analysis spanning the *Drosophila* genus (M0 ω estimate, M0 vs. free ratio test) or spanning the *melanogaster* group (M7 vs. M8, M8 vs. M8a tests). Bold and red text indicates statistically significant comparisons. Amino acid residues with >0.90 probability of being under positive selection are indicated, with the number/letter indicative of the *D. melanogaster* position within the alignment.

Our syntenic approach also revealed complex evolutionary events for seminal *Obp* genes not found in *D. melanogaster*. *Acp223*, a predicted Obp-like SFP gene with evidence of accessory gland expression in *D. yakuba* and *D. erecta*, resides between *Obp56e* and *Obp56f* (Begun et al., 2006). InterProScan searches of this gene match signal peptide and Obp protein domains, and together with the location in the genome, suggest this gene is an *Obp56* cluster paralog (Begun et al., 2006). Consistent with previous reports of this gene not being present in the *D. melanogaster* genome, we were unable to find hits of this gene in *D. melanogaster* or *D. simulans* genomes using liberal E-value cutoffs in tBLASTn searches, though we found a very diverged noncoding hit in the annotated 3' UTR of *Obp56e* in *D. sechellia* (Begun et al., 2006). Begun et al. (2006) reported finding a partial, noncoding orthologous region in *D. melanogaster*, which we also found in *D. simulans* to be noncoding. We did find orthologs of this gene in other *melanogaster* group species, which showed relatively long branch lengths in phylogenies of all Obp56 cluster genes (Figure 5—figure supplement 7A). In the *Obp51a* cluster, we found previously reported SFPs *Sfp51D* (in *D. simulans*) and *Acp157a* (in *D. yakuba*) ~14 kb upstream of *Obp51a*, which are putative orthologs of each other based on moderate branch support in our phylogenies (Figure 5—figure supplement 7B) (Begun et al., 2006; Findlay et al., 2009).

Consistent with previous results, we were unable to find orthologs of this gene in *D. melanogaster* but found a likely pseudogene in *D. simulans*. Previous work also showed this gene independently duplicated and pseudogenized in *D. yakuba* (Begun et al., 2006). Together, these results illustrate evolutionary lability in presence/absence and copy number of these genes in closely related *Drosophila* species.

Using our high confidence ortholog candidates, we next examined the molecular evolution of these genes across *Drosophila*. Previous reports of *Obp* gene family evolution across *Drosophila* reported heterogenous evolutionary rates for some *Obp* genes across species, but genes without 1:1 orthologs in all 12 *Drosophila* species were excluded from these previous analyses, which included *Obp51a*, *Obp22a*, *Obp56i*, and *Obp8a* (Vieira et al., 2007). We began by using model M0 of PAML to estimate whole-gene ratios of dN/dS (ω) across all species of the phylogeny. Using this approach, we found three *Obp* genes with ω values around ~0.20 (*Obp56g*, *Obp8a*, and *Obp56e*, which are found in species beyond the *melanogaster* group, Figure 5B). Interestingly, the four *Obp* genes restricted to the *melanogaster* group had higher ω values, around ~0.50 (*Obp51a*, *Obp56f*, *Obp56i*, *Obp22a*, Figure 5B) which is much higher than the reported genome-wide average in *D. melanogaster* (Chang and Malik, 2022; Drosophila 12 Genomes Consortium et al., 2007). We then used the “free-ratio” model of PAML to test whether these genes exhibit evolutionary rate heterogeneity across the phylogeny. For all genes except *Obp56f* and *Obp56i*, we found significant evidence of heterogeneity in ω (Figure 5B), indicating these genes have experienced variable selective pressures (and/or variable strengths of selection) across the *Drosophila* genus.

A subset of seminal Obps are evolving under recurrent positive selection

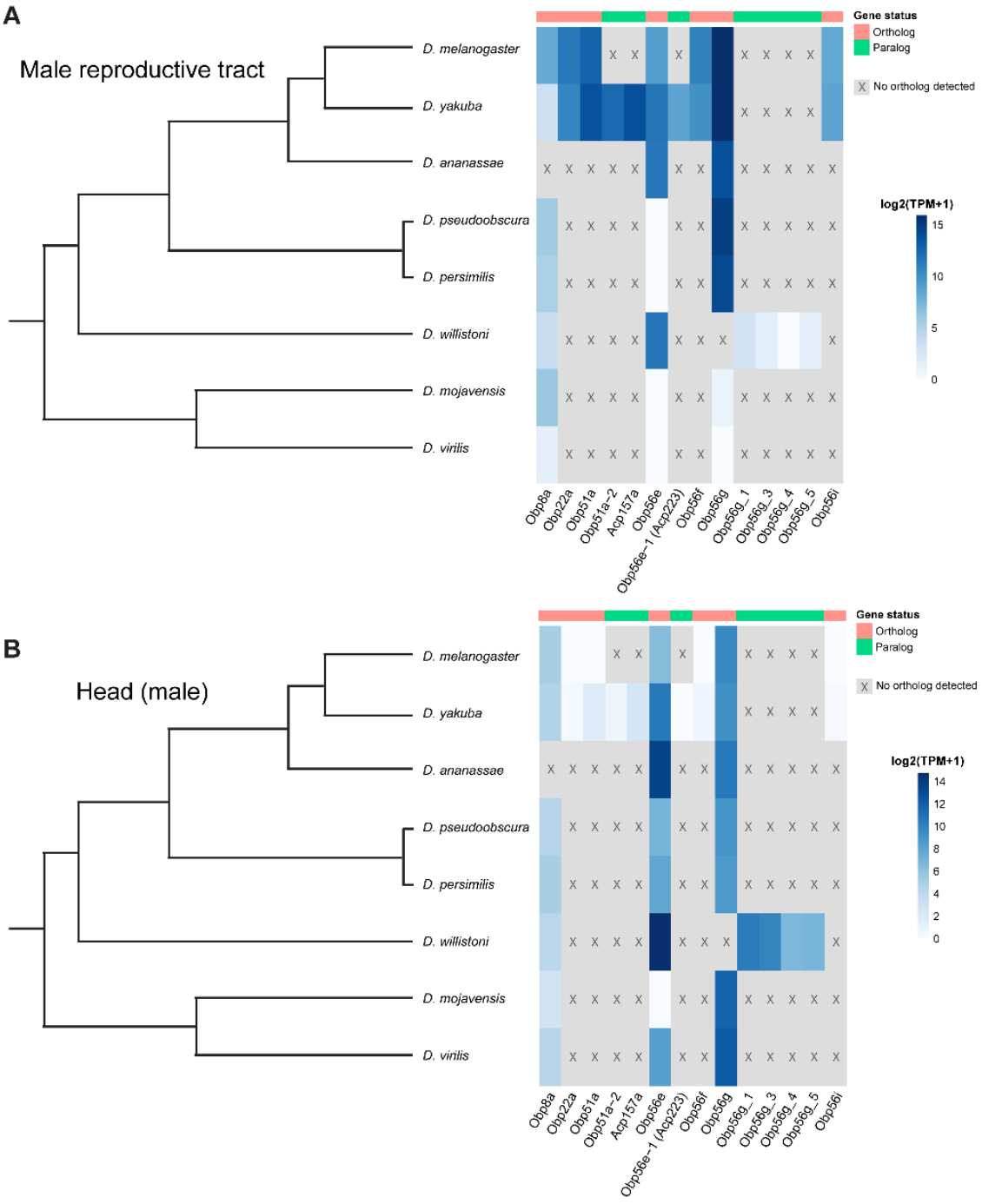
We next tested whether any seminal *Obp* genes show evidence of recurrent positive selection acting on a subset of sites by comparing model M7 and M8 in PAML, limiting our analysis to *melanogaster* group species to avoid synonymous site saturation. Using this approach, we found significant evidence of positive selection for *Obp22a* and *Obp51a*, while the other seminal *Obp* genes are evolving in a manner consistent with purifying selection (Figure 5B). *Obp22a* and *Obp51a* were also significant for the M8/M8a model comparison, implying positive selection rather than neutral divergence accounting for the rapid evolution of sites within these genes. Plotting the ω ratio inferred from the “free-ratio” model onto gene trees for *Obp22a* and *Obp51a* shows multiple branches have ω >1, including those with lineage-specific duplication events (Figure 5—figure supplement 8).

We also used model M8 to infer specific sites under selection for *Obp22a* and *Obp51a* (Figure 5B). We included all detected copies of each gene in our selection analysis, which may have reduced our power to detect specific sites under selection for *Obp51a*, only one of which had posterior probability >0.90. For *Obp22a*, we inferred seven sites under selection (Pr>0.90), which we mapped onto the predicted AlphaFold structure of the protein (Figure 5—figure supplement 9A) (Jumper et al., 2021). We found that these sites are located on the outside-facing region of the protein, away from the hydrophobic binding pocket, which has been found to bind hydrophobic ligands in other Obp proteins such as LUSH (Figure 5—figure supplement 9B) (Laughlin et al., 2008).

Male reproductive tract expression of *Obp56g* is derived in a subset of *Drosophila* species

Individual components of seminal fluid are known to turn over rapidly between species, though the larger biochemical classes these components fall into are conserved between species (Mueller et al., 2005; Swanson et al., 2001; Wigby et al., 2020). Beyond *D. melanogaster*, *Obp56g* has been detected as a seminal protein in *D. simulans*, *D. yakuba*, and *D. pseudoobscura*, but not in more distantly related *Drosophila* species whose seminal fluid proteins have been characterized (*D. mojavensis*, *D. virilis*, and *D. montana*), despite the gene itself being conserved in these species (Ahmed-Braimah et al., 2017; Garlovsky et al., 2020; Kelleher et al., 2009). Considering our findings that *Obp56g* is required for male fertility in *melanogaster*, we were curious to see whether male reproductive tract expression of *D. melanogaster* seminal *Obps* was conserved across the *Drosophila* phylogeny. We therefore leveraged previously published RNAseq data from 8 different *Drosophila* species, focusing specifically on the male head and male reproductive tract samples, which include the accessory glands, ejaculatory ducts, ejaculatory bulbs, and terminal genitalia (Yang et al., 2018). We observed significantly higher expression of *Obp56g* in the male reproductive tract of *D. melanogaster, simulans, yakuba, ananassae, persimilis,* and *pseudoobscura* species, and negligent or zero expression in *D. willistoni, virilis,* and *mojavensis* species (Wilcoxon rank sum test of *melanogaster/obscura* group vs. *repleta* and *virilis* group [excluding *willistoni* which has *Obp56g* duplications], $p<0.001$) , consistent with previous reports that *Obp56g* is a seminal protein in *melanogaster* and *obscura* group species (Figure 6A) (Findlay et al., 2008; Karr et al., 2019). In head tissues, we observed high expression of *Obp56g* in all species (Figure 6B). We confirmed these expression patterns using semi-quantitative RT-PCR on dissected reproductive tract tissues from *melanogaster, ananassae, pseudoobscura, virilis*, and *mojavensis* males,

738 which showed *Obp56g* has conserved reproductive tract expression (in both the accessory
739 gland + ejaculatory duct and ejaculatory bulb tissues) in the *melanogaster* and *obscura* groups,
740 and conserved head expression across all species tested (Figure 6—figure supplement 1).



741

742 **Figure 6:** Seminal *Obp* genes show changes in expression pattern across species from bulk
743 RNAseq data published in Yang et al. (2018). A) log2 normalized TPM expression values
744 (averaged across 4 biological replicates) of seminal *Obp* genes and their associated orthologs
745 and paralogs in male reproductive tissue (including accessory glands, ejaculatory duct,
746 ejaculatory bulb, and terminal genitalia for all species except *D. melanogaster*, which includes

all tissues aside from the genitalia) of different *Drosophila* species. Grey indicates that no ortholog could be detected in that species. B) log2 normalized TPM expression values of seminal *Obp* gene orthologs and paralogs in male head tissue.

Discussion:

Obps have been identified as seminal fluid components in several insect taxa, though their functional importance in reproduction has remained unclear. We found that *Obp56g* is required for mating plug formation, sperm storage, and subsequent male fertility in *D. melanogaster.* Given that the post-mating response depends on sperm, SP, and the long-term response network proteins (Findlay et al., 2014; Manning, 1967; Peng et al., 2005), loss of ejaculate in *Obp56g* mutant males can explain the loss of long-term responses in females that we observed. Recent proteomic evidence has demonstrated that Obp56g is among the most highly abundant SFPs in the mating plug, supporting our inference that it is important for this process (McDonough-Goldstein et al., 2022). We further found *Obp56g* transcripts are primarily derived from the ejaculatory bulb (though transcripts were also detected in the ejaculatory duct and accessory glands), which has previously documented functions in mating plug formation (Avila et al., 2015; Bretman et al., 2010; Lung and Wolfner, 2001). This ejaculatory bulb/duct expression is required for mating plug formation and fertility. We note that *CrebA-GAL4* does not drive expression in the accessory gland (Avila et al., 2015), suggesting that any residual expression in this tissue in these males is not sufficient to induce mating plug formation and the PMR.

We now have functional evidence for a growing list of mating plug and/or EB-derived SFPs, including Acp36DE, PEB-me, EbpII, and Obp56g (Avila et al., 2015; Bretman et al., 2010; Neubaum and Wolfner, 1999). Additionally, approaches such as gas chromatography-mass spectrometry and proteomics have characterized the male- and female-derived compounds and proteins that comprise the mating plug, and experiments dissecting the female tract at different time points after mating have elucidated the timeline of mating plug formation (Avila et al., 2015; Gilchrist and Partridge, 2000; Laturney and Billeter, 2016; Lung and Wolfner, 2001; McDonough-Goldstein et al., 2022). However, we still lack a detailed biochemical understanding of how the mating plug coagulates, as well as the specific mechanistic roles of the proteins highlighted above. For example, does Obp56g bind to and transport a hydrophobic reproductive tract-derived small molecule, as might be expected for an Obp? Does Obp56g concentrate said molecule within the female tract to trigger mating plug formation, or is it merely structural? Or, instead of acting as a structural component, does Obp56g signal to the female tract to secrete

components that aid in mating plug formation? The answers to such questions will provide important insight into a crucial reproductive process in flies and other insect species.

Obp56g has interesting evolutionary characteristics in that the gene itself is conserved widely (and our results show it is under purifying selection in the *melanogaster* group), though its expression pattern in the male reproductive tract is not. Such lineage-specific shifts in expression have been reported for several other reproductive genes in *Drosophila*, including glucose dehydrogenase (*Gld*) in ejaculatory duct tissues of the *melanogaster* group, *jamesbond,* a fatty acid elongase responsible for CH503 production in the ejaculatory bulb, and the Sex Peptide Receptor (*SPR*), which gained expression in the female reproductive tract in the lineage leading to the *melanogaster* group (Cavener, 1985; Ng et al., 2015; Tsuda et al., 2015). Our results also showed that *virilis* and *repleta* group species lack *Obp56g* expression in the male reproductive tract, which is consistent with proteomic and transcriptomic studies that did not detect *Obp56g* as a predicted seminal protein in these species (Ahmed-Braimah et al., 2017; Kelleher et al., 2009). Previous studies have described insemination reactions (*repleta* group) and “dense copulatory plugs” (*virilis* group) in the bursa of females of these species post-mating (Markow and Ankney, 1988; Patterson, 1946). While these structures are very likely composed of ejaculate matter (and female-derived components), whether they are true homologous structures to the *melanogaster* mating plug, which has documented functional roles in promoting sperm storage and in post-mating pheromonal mate guarding, is unclear (Avila et al., 2015; Avila and Wolfner, 2009; Laturney and Billeter, 2016; Neubaum and Wolfner, 1999). A previous study using electron microscopy to analyze post-mating structures in the female bursa in *D. melanogaster* and *D. mojavensis* found the composition, density, and size of these structures to be quite distinct, and characterized them as separate phenomena (termed a “sperm sac” and “true insemination reaction” for *melanogaster* and *mojavensis*, respectively) (Alonso-Pimentel et al., 1994). Interestingly however, several recent studies have shown rapid divergence and anti-aphrodisiac function of pheromonal compounds produced in the ejaculatory bulb or male reproductive tract across *Drosophila* (Chin et al., 2014; Khallaf et al., 2021; Ng et al., 2014). Elucidating the mechanistic function of *Obp56g* will provide interesting insight into whether the rapid turnover of male-specific pheromones is linked to the evolutionary changes in expression we observe for *Obp56g* and the evolutionary turnover in seminal Obps seen across more distant taxa. A further question remains whether *Obp56g* has a conserved function in mating plug formation in the species where the gene is an SFP (and its function in those where it is not), which could help elucidate when and how *Obp56g* acquired its role in reproduction.

Furthermore, whether Obp56g took over a primary role in mating plug formation after it evolved reproductive tract expression, and whether "plugs" or other post-mating structures were fundamentally different prior to this, remains an open question.

Our results also show that when individually knocked out, only *Obp56g* has a strong effect on the PMR and male fertility, while loss of the others has no effect (for *Obp8a*, the mutant had slightly lower remating rates than the control, which is opposite of what is expected for genes involved in PMR phenotypes). These results can be explained in part given our findings that *Obp56g* is the only seminal Obp that is highly expressed in the ejaculatory bulb, which has documented functions in mating plug formation. The other Obps are derived from the accessory gland (*Obp51a, Obp22a, Obp56e, Obp56i, Obp8a*) or the ejaculatory duct (*Obp51a*), which is consistent with previous transcriptomic and proteomic studies of the reproductive tract (Findlay et al., 2008; Li et al., 2022; Majane et al., 2022; Takemori and Yamamoto, 2009). Alternatively, given these genes are in the same gene family, redundancy might mask any individual gene's phenotype, and defects in fertility may only be apparent when these genes are mutated in combination. Indeed, previous studies in *Drosophila* have shown functional redundancy among paralogs of the *Obp50* cluster in male starvation resistance (Johnstun et al., 2021). Evolutionarily, it has been hypothesized that sexual conflict between males and females can drive functional redundancy in the biochemical classes present in seminal fluid through mechanisms of gene duplication, co-option, and gene loss, though this has never been directly functionally tested (Sirot et al., 2015).

Given several previous studies demonstrating elevated divergence of SFP genes in *Drosophila*, we tested whether any of the seminal *Obp* genes are rapidly evolving in the *melanogaster* group. We did not detect positive selection on *Obp56g*, *Obp56e*, *Obp56f*, *Obp56i*, or *Obp8a*, but did detect positive selection acting on *Obp22a* and *Obp51a*. We found that *Obp56g* is highly expressed in head tissues across all the species we tested, raising the possibility that the gene is under pleiotropic constraint for a non-reproductive function, thus limiting its capacity to rapidly diverge (though we did observe a highly diverged paralog of *Obp56g* in the *obscura* clade). Previous studies in *D. melanogaster* have shown *Obp56g* is highly expressed in gustatory sensilla in the labellum in males and females, though functional studies of $Obp56g^{1}$ mutants showed they had normal attractive and aversive behaviors to sucrose and bitter-tasting compounds, respectively (Galindo and Smith, 2001; Jeong et al., 2013). In our assays, $Obp56g^{1}$ mutants did not have significantly altered mating latency or duration times from controls,

indicating it does not play a role in male courtship behavior. Thus, the proboscis-related function of *Obp56g*, and whether it is conserved across species (which would possibly explain our observations of purifying selection acting on the gene), remains unknown. Alternatively, *Obp56g* could possibly be conserved within the *melanogaster* group due to its role in mating plug formation, as it is essential for full male fertility in *D. melanogaster*. Such a hypothesis is consistent with previous findings of conservation among some members of the SP network, whose functions are necessary for successful reproduction in *melanogaster* (McGeary and Findlay, 2020).

Our study also revealed extensive evolutionary lability in copy number of the seminal Obps across species, which appears to be driven by tandem gene duplication, pseudogenization, and gene loss, particularly in the *Obp51a* cluster. Gene duplication has been shown to be a major force in the evolution of female reproductive tract and SFP genes, though the reasons why are less clear (Findlay et al., 2008). There may be selection acting on increased protein abundance, which could be accomplished by gene duplication (Kondrashov et al., 2002). Alternatively, models of sexual conflict propose arms race-style antagonism between males and females, whereby duplication and divergence of reproductive molecules may allow either sex to counter-adapt against the other (Findlay et al., 2008; Kelleher and Markow, 2009; Kelleher and Pennington, 2009; Sirot et al., 2014; Swanson and Vacquier, 2002). Our finding of positive selection acting on *Obp22a* and *Obp51a* suggests the latter may be involved. Studies have also previously demonstrated that relaxed constraint following gene duplication can allow for deleterious or complete loss of function mutations, resulting in gene loss or the formation of pseudogenes, which could explain the patterns of duplication and pseudogenization we observed in the *Obp51a* and *Obp22a* clusters (Birchler and Yang, 2022; Ohno, 1970; Sirot et al., 2015).

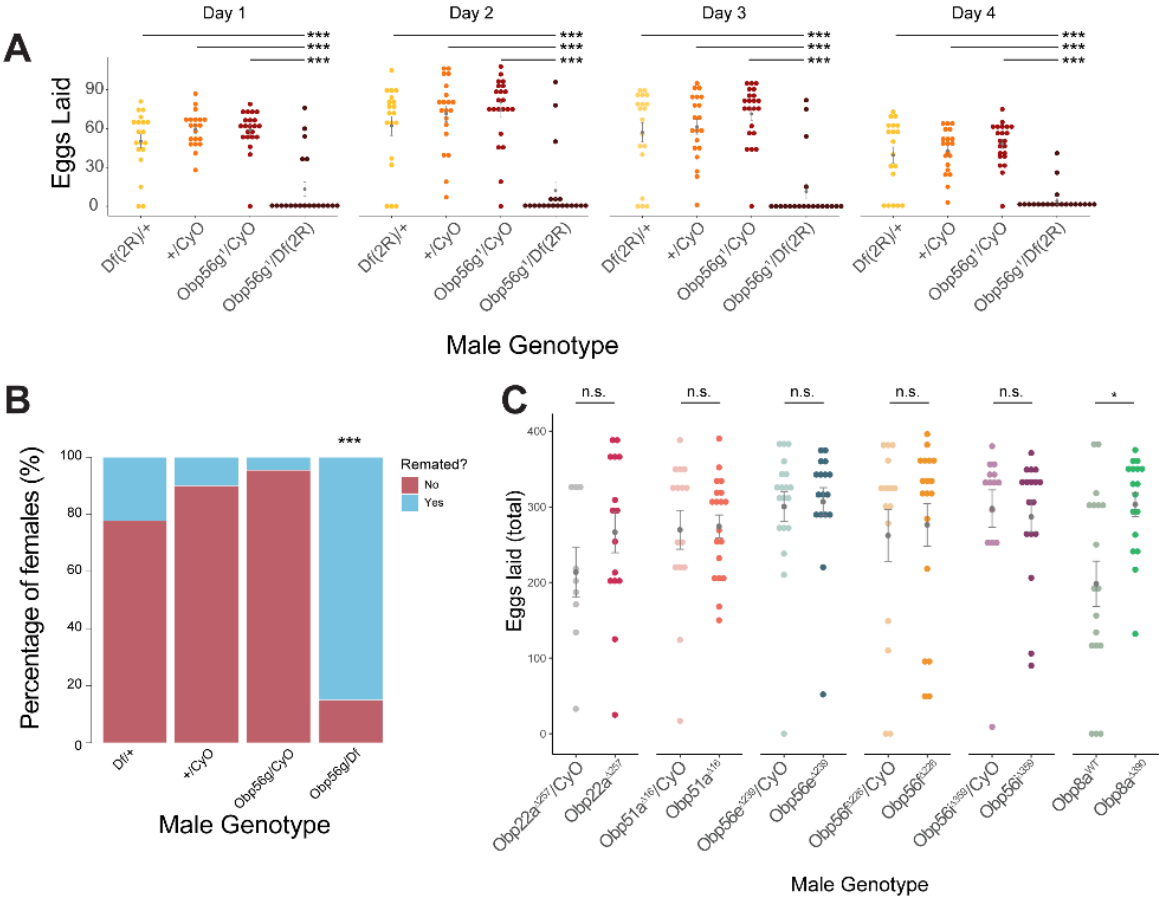
Overall, our study provides new evidence for a novel reproductive role for Obps, highlighting the broad functional diversity for this gene family in *Drosophila*. Additionally, we observed expression shifts, duplication, and divergence in the evolution of these seminal protein genes, highlighting the myriad mechanisms by which reproductive genes can diverge across species. The frequent occurrence of Obps in the seminal fluid across distinct taxa raises the possibility that members of this gene family are repeatedly co-opted into the SFP suite by various means. Functional studies of seminal Obps across these diverged species will provide important

883 comparative data for whether seminal Obps can evolve roles in reproductive processes beyond
884 mating plug formation.
885
886 **Acknowledgements**:
887 We would like to thank Dr. Yasir Ahmed-Braimah for help analyzing FlyAtlas2.0 data, Dr. Jolie
888 Carlisle for help with the evolutionary analysis, Norene Buehner for help with Western blots, and
889 members of the Wolfner and Clark labs for useful comments and advice. We would also like to
890 thank Susan Younger, J. Belote and S. Pitnick, the Vienna *Drosophila* Resource Center, the
891 Bloomington *Drosophila* Stock Center, and the *Drosophila* Species Stock Center for lines. This
892 work was supported by NIH grant R01-HD059060 to AGC and MFW, NIH postdoctoral
893 fellowship F32GM097789 to GDF, and NSF grant 2212972 to GDF.
894
895 **Competing interests**:
896 The authors declare that no competing interests exist.
897
898 **Data availability**:
899 All data generated or analyzed for this study are included in the manuscript, supporting files, or
900 are available on Github. Source data files have been provided for Figure 1C, Figure 2B, Figure
901 4C, Figure 1—figure supplement 1B, figure supplement 2B, Figure 1—figure supplement 3A &
902 B, Figure 3—figure supplement 2, Figure 4—figure supplement 2, and Figure 6—figure
903 supplement 1. All mating data, R code to analyze mating data, RNAseq data across species,
904 and tree files/alignments for use in PAML are available on Github:
905 https://github.com/WolfnerLab/Obps
906
907 The following previously published datasets were used:
908 Delbare SYN, Ahmed-Braimah YH, Wolfner MF, Clark AG. Interactions between the microbiome
909 and mating influence the female's transcriptional profile in Drosophila melanogaster. Sci Rep.
910 2020 Oct 23;10(1):18168. doi: 10.1038/s41598-020-75156-9. PMID: 33097776; PMCID:
911 PMC7584617.
912
913 Yang H, Jaime M, Polihronakis M, Kanegawa K, Markow T, Kaneshiro K, Oliver B. Re-
914 annotation of eight *Drosophila* genomes. Life Sci Alliance. 2018 Dec 24;1(6):e201800156. doi:
915 10.26508/lsa.201800156. PMID: 30599046; PMCID: PMC6305970.
916

Li H, Janssens J, De Waegeneer M, Kolluru SS, Davie K, Gardeux V, Saelens W, David FPA, Brbić M, Spanier K, Leskovec J, McLaughlin CN, Xie Q, Jones RC, Brueckner K, Shim J, Tattikota SG, Schnorrer F, Rust K, Nystul TG, Carvalho-Santos Z, Ribeiro C, Pal S, Mahadevaraju S, Przytycka TM, Allen AM, Goodwin SF, Berry CW, Fuller MT, White-Cooper H, Matunis EL, DiNardo S, Galenza A, O'Brien LE, Dow JAT; FCA Consortium§; Jasper H, Oliver B, Perrimon N, Deplancke B, Quake SR, Luo L, Aerts S, Agarwal D, Ahmed-Braimah Y, Arbeitman M, Ariss MM, Augsburger J, Ayush K, Baker CC, Banisch T, Birker K, Bodmer R, Bolival B, Brantley SE, Brill JA, Brown NC, Buehner NA, Cai XT, Cardoso-Figueiredo R, Casares F, Chang A, Clandinin TR, Crasta S, Desplan C, Detweiler AM, Dhakan DB, Donà E, Engert S, Floc'hlay S, George N, González-Segarra AJ, Groves AK, Gumbin S, Guo Y, Harris DE, Heifetz Y, Holtz SL, Horns F, Hudry B, Hung RJ, Jan YN, Jaszczak JS, Jefferis GSXE, Karkanias J, Karr TL, Katheder NS, Kezos J, Kim AA, Kim SK, Kockel L, Konstantinides N, Kornberg TB, Krause HM, Labott AT, Laturney M, Lehmann R, Leinwand S, Li J, Li JSS, Li K, Li K, Li L, Li T, Litovchenko M, Liu HH, Liu Y, Lu TC, Manning J, Mase A, Matera-Vatnick M, Matias NR, McDonough-Goldstein CE, McGeever A, McLachlan AD, Moreno-Roman P, Neff N, Neville M, Ngo S, Nielsen T, O'Brien CE, Osumi-Sutherland D, Özel MN, Papatheodorou I, Petkovic M, Pilgrim C, Pisco AO, Reisenman C, Sanders EN, Dos Santos G, Scott K, Sherlekar A, Shiu P, Sims D, Sit RV, Slaidina M, Smith HE, Sterne G, Su YH, Sutton D, Tamayo M, Tan M, Tastekin I, Treiber C, Vacek D, Vogler G, Waddell S, Wang W, Wilson RI, Wolfner MF, Wong YE, Xie A, Xu J, Yamamoto S, Yan J, Yao Z, Yoda K, Zhu R, Zinzen RP. Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. Science. 2022 Mar 4;375(6584):eabk2432. doi: 10.1126/science.abk2432. Epub 2022 Mar 4. PMID: 35239393; PMCID: PMC8944923.

951 **Supplemental figures & figure legends**:

952



953

954 **Figure 1—figure supplement 1**: Additional replicate of PMR phenotypes from CS females
955 mated to *Obp56g* and CRISPR mutant males. A) Egg counts from CS females mated *to*
956 *Df(2R)/+, CyO/+ Obp56g*^{1}*/CyO*, or *Obp56g*^{1}*/Df(2R)* males from 1-4 days after mating.
957 Significance indicated from pairwise comparisons of male genotypes within days using
958 emmeans on a Poisson linear mixed effects model. Error bars represent mean +/- SEM. B)
959 Proportion of females who did or did not remate with a standard CS male on the fourth day after
960 mating. Significance indicated from tests of equality of proportions. C) Egg counts from CS
961 females mated to homozygous null or heterozygous control males (except for *Obp8a*, the
962 control of which is from an unedited sibling line) from 1-4 days after mating. Error bars represent
963 mean +/- SEM. Significance indicated from Poisson linear models with Benjamini-Hochberg
964 corrections for multiple comparisons. Significance levels: *P<0.05, **P<0.01, ***P<0.001, n.s.
965 not significant.

966

Figure 1—figure supplement 1—source data 1: Remating counts and percentages for data shown in Figure 1—figure supplement 1B.

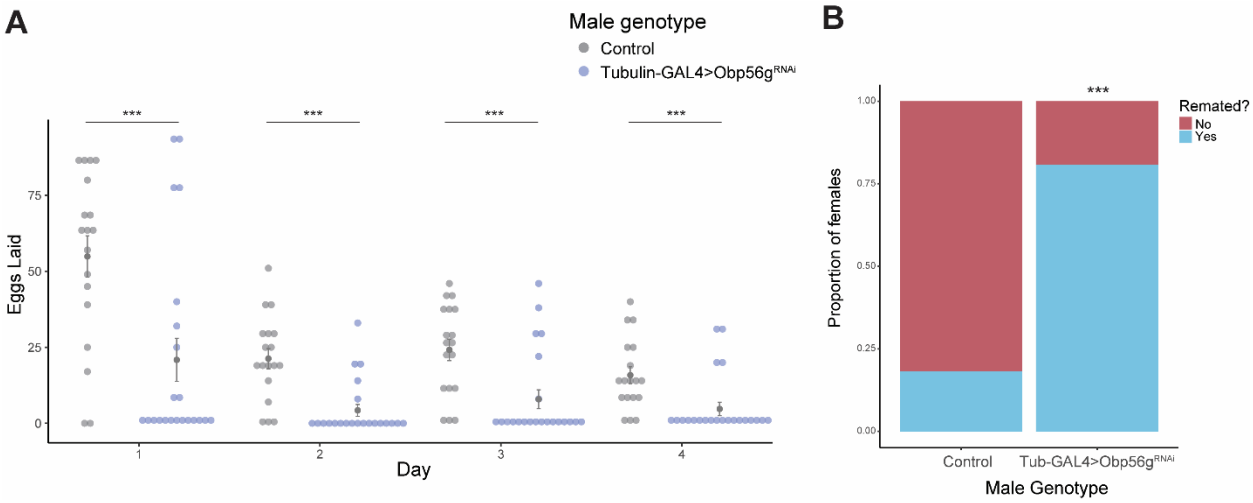


Figure 1—figure supplement 2: Whole body knockdown of *Obp56g* using *Tubulin-GAL4* results in loss of post-mating response phenotypes in females. (A) Counts of eggs from mated CS females over 4 days. Females mated to *Tubulin-GAL4>Obp56g^{RNAi}* males lay significantly fewer eggs than females mated to control males (p<0.001, n=20-24). (B) CS females mated to *Tubulin-GAL4>Obp56g^{RNAi}* males are significantly more likely to remate 4 days post-mating relative to control males (p<0.001, n=26-33). Error bars represent mean +/- SEM. Significance level: * p<0.05, ** p<0.01, *** p<0.001.

Figure 1—figure supplement 2—source data 1: Remating counts and percentages for data shown in Figure 1—figure supplement 2B.

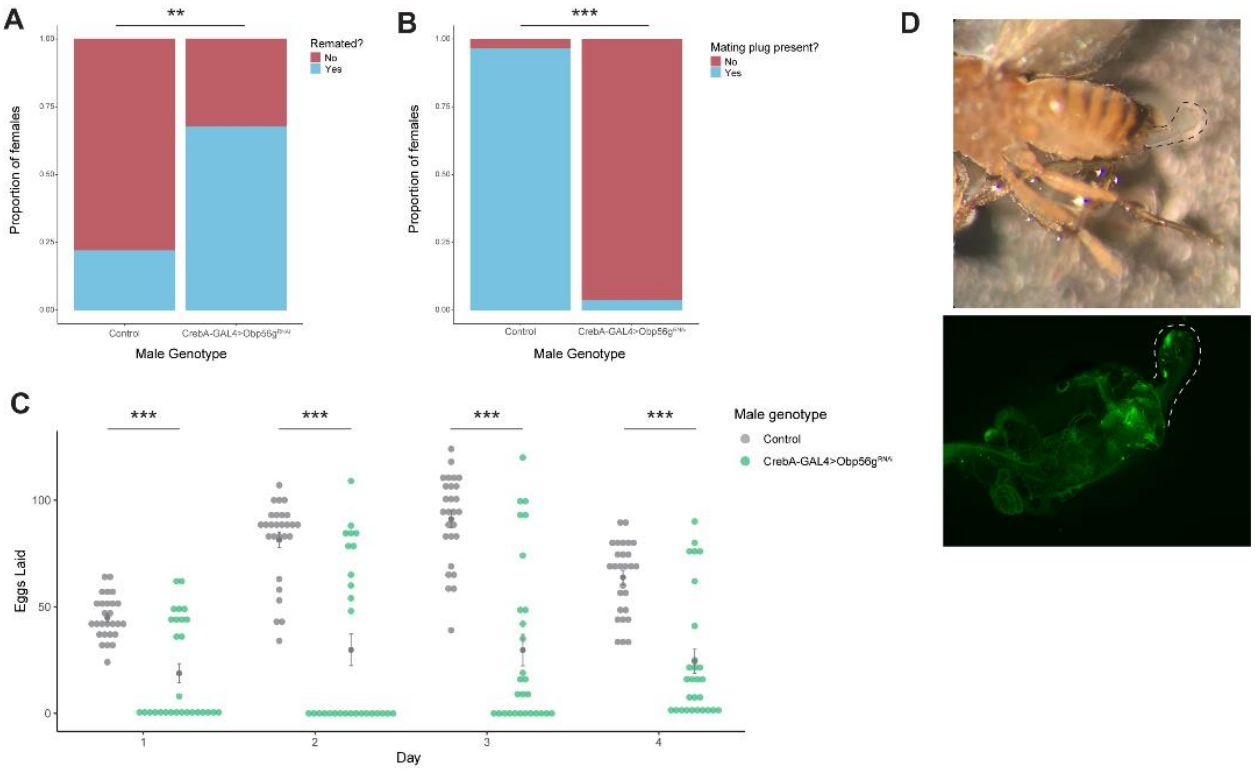


Figure 1—figure supplement 3: Male reproductive tract knockdown of *Obp56g* with *CrebA-GAL4* is required for the post-mating response and mating plug formation. (A) CS females mated to *CrebA-GAL4>Obp56g*^{RNAi} males are significantly more likely to remate 4 days post-mating relative to control males (p=0.001, n=27-28). (B) A significantly reduced proportion of females mated to *CrebA-GAL4>Obp56g*^{RNAi} males have fully formed mating plugs in their bursa immediately after the end of mating relative to females mated to control males (p<0.001, n=27-30). (C) Counts of eggs from mated CS females over 4 days. CS females mated to *CrebA-GAL4>Obp56g*^{RNAi} lay significantly fewer numbers of eggs relative to CS females mated to control males (p<0.001, n=27-28). (D) Ejaculate loss (dotted line) from the bursa observed in females mated to *CrebA-GAL4>Obp56g*^{RNAi} males, with the bursa dissected and imaged for GFP to visualize the autofluorescent speckles that comprise the uncoagulated mating plug. Error bars represent mean +/- SEM. Significance level: * p<0.05, ** p<0.01, *** p<0.001.

Figure 1—figure supplement 3—source data 1: Counts and percentages for data shown in Figure 1—figure supplement 3A & B.

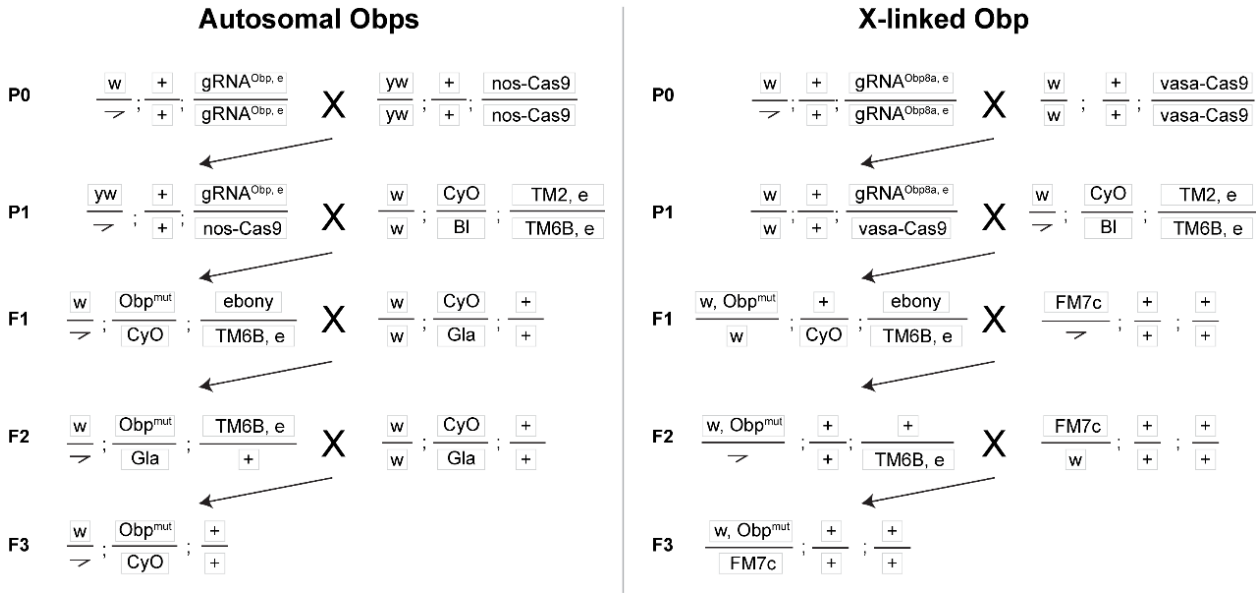


Figure 2—figure supplement 1: Crossing scheme to generate CRISPR mutants in autosomal (*Obp22a, Obp51a, Obp56e, Obp56f, Obp56i*) and X-linked (*Obp8a*) *Obp* genes used in this study, with text boxes representing chromosomes X/Y, 2, and 3 (dot chromosome not shown). The kinked line represents the Y chromosome. *Obp* and *ebony* CRISPR editing takes place in the germline of individuals in the P1 generation. *ebony* editing can happen on either the gRNA or Cas9 chromosomes (written out in the F1 generation as "*ebony*" for simplicity), which are removed from the genetic background before assaying males for reproductive phenotypes.

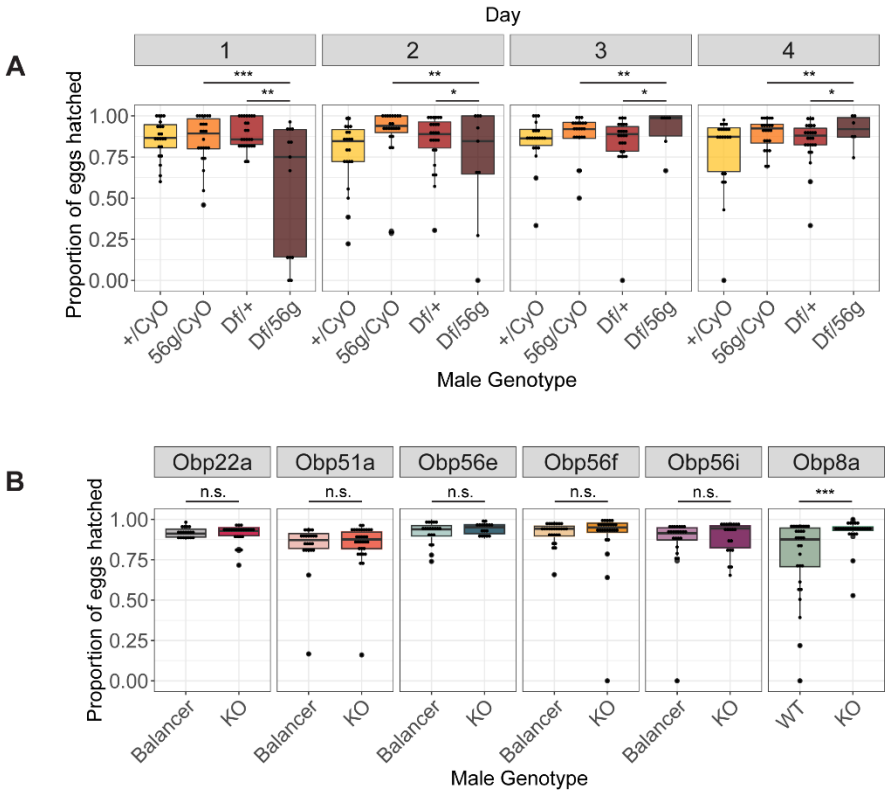


Figure 2—figure supplement 2: Box plots of hatchability estimates from CS females mated to *Obp56g* or CRISPR mutant males. A) Proportion of eggs hatched over 4 days from females mated to *Df(2R)/+, CyO/+ Obp56g*^{1}*/CyO*, or *Obp56g*^{1}*/Df(2R)* males. Significance indicated from pairwise comparisons of male genotypes across days using emmeans on a binomial mixed effects model. B) Proportion of eggs hatched over 4 days from females mated to CRISPR mutant males. Significance indicated from binomial linear models with Benjamini-Hochberg corrections for multiple comparisons. Significance levels: *P<0.05, **P<0.01, ***P<0.001, n.s. not significant.

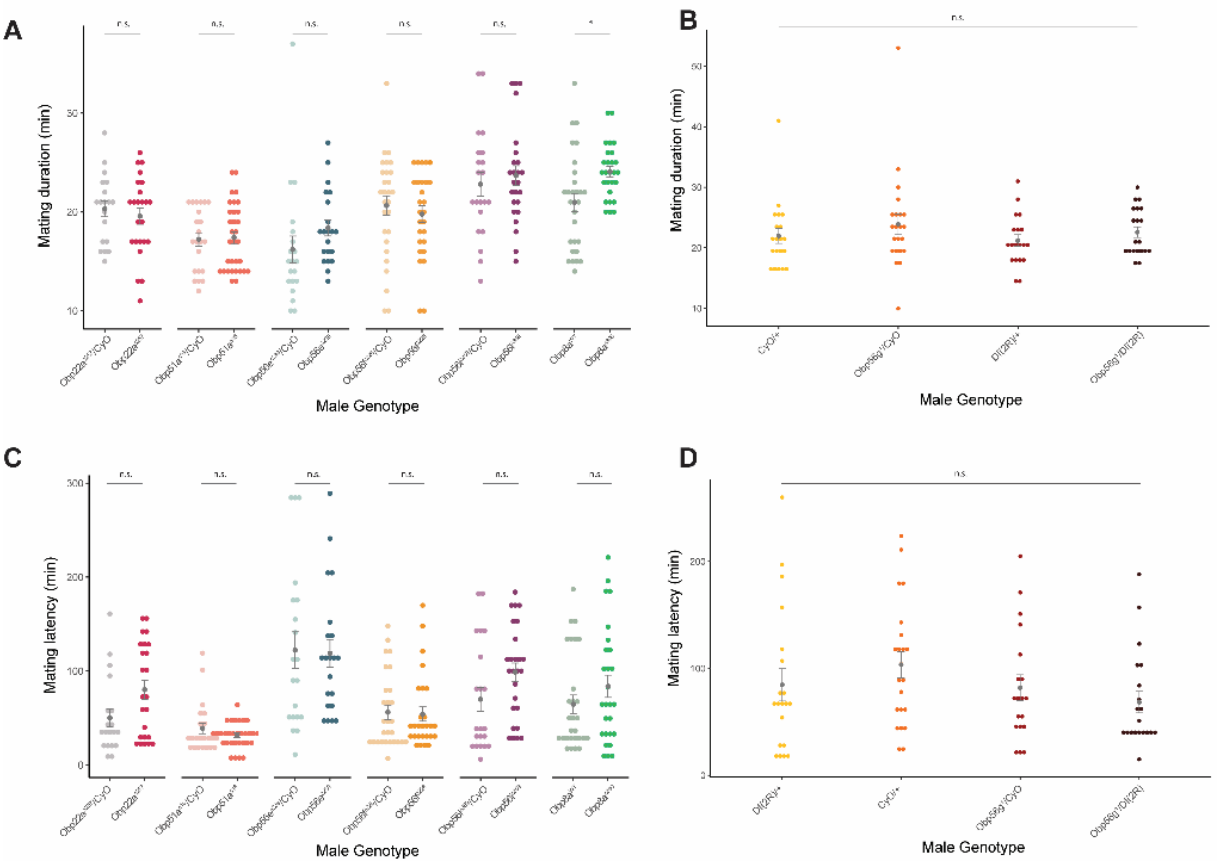


Figure 2—figure supplement 3: Mating latency and duration measurements from *Obp56g*^{1} and CRISPR-generated *Obp* mutants with CS females. (A & B) *Obp8a*$^{\Delta 390}$ mutant flies mate for longer duration than *Obp8a*^{WT} control flies (p<0.05, mean *Obp8a*^{WT} 20.96 minutes, mean *Obp8a*$^{\Delta 390}$ 24.07 minutes), though no other statistically significant differences were observed between mating duration of mutant or control males for other genotypes (p>0.05 for ANOVA [*Obp56g*^{1}] or Benjamini-Hochberg corrected p-values from Student's t-tests [CRISPR mutants]). (C & D) No statistically significant differences observed between mating latency of mutant or control males with CS females (p>0.05 for ANOVA [*Obp56g*^{1}] or Benjamini-Hochberg corrected p-values from Student's t-tests [CRISPR mutants]). n for each genotype ranged from 19 to 32. Error bars represent mean +/- SEM. Significance level: * p<0.05, ** p<0.01, *** p<0.001, n.s. not significant.

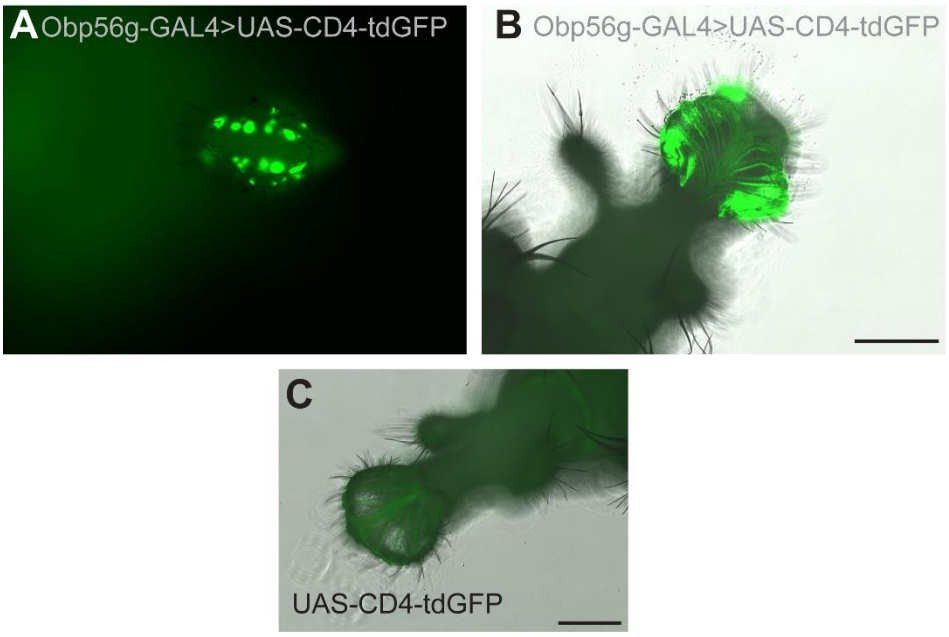


Figure 3—figure supplement 1: Expression of *Obp56g-GAL4* in the gustatory bristles of the labellum. A) GFP expression from *Obp56g-GAL>UAS-CD4-tdGFP* males in the head. This sample is not placed under a coverslip. B) GFP expression in the same genotype, with the proboscis dissected off and gently pressed under a coverslip. C) GFP expression in *UAS-CD4-tdGFP* control male labellum. Scale bar=130um.

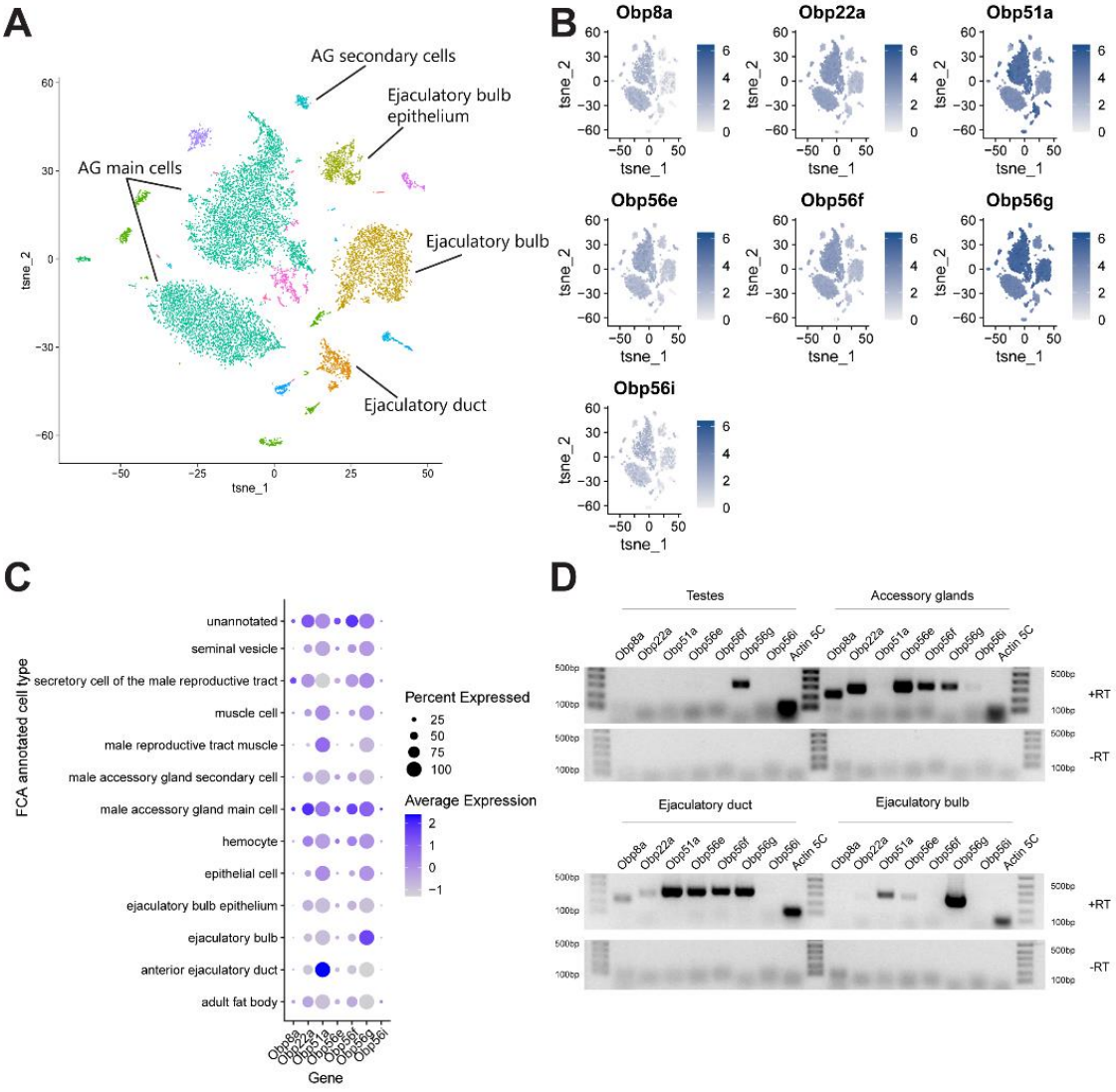


Figure 3—figure supplement 2: *Obp56g* is the most highly expressed seminal Obp in the ejaculatory bulb. A) Seurat tSNE dimensionality reduction plot of single nucleus RNAseq expression data from the male reproductive tract (without testes) and their major cell type annotations according to (Li et al., 2022). B) Feature plots from Seurat showing expression of the seminal *Obp* genes across single nuclei from A. C) Seurat dot plot of scaled average gene expression across annotated cell types for seminal Obps. Dot size indicates the percentage of cells within a cluster that express each *Obp* gene. D) Agarose gel of RT-PCR products of seminal *Obp* genes from microdissected bulk tissues of the *D. melanogaster* male reproductive tract (testes, accessory glands, ejaculatory ducts, and ejaculatory bulbs), with *Actin 5C* used as the positive control for each tissue. Samples treated with reverse transcriptase are above, and those without below (as a negative control), for each tissue type. PCR was performed for 35 cycles.

Figure 3—figure supplement 2—source data 1: Raw and uncropped, labeled gel images for data shown in Figure 3—figure supplement 2D.

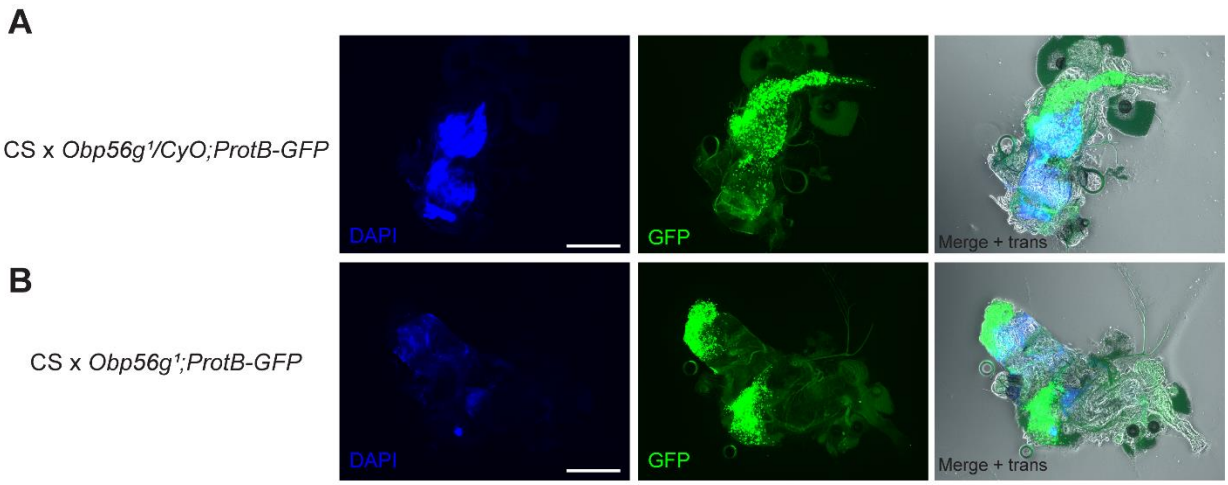


Figure 4—figure supplement 1: *Obp56g*^{1} mutant males do not have gross issues with sperm transfer during mating at the 12-minute ASM time point. A) Representative CS female mated to *Obp56g*^{1}*/CyO;ProtB-eGFP* control males, showing DAPI (mating plug), GFP (sperm heads), and merge + transillumination microscopy images. 9/9 females mated to these males had mating plugs, and 9/9 had sperm masses present in their bursas. B) Representative CS female mated to *Obp56g*^{1}*;ProtB-eGFP* mutant males. 0/10 females had mating plugs, though 10/10 had sperm masses present in their bursas.

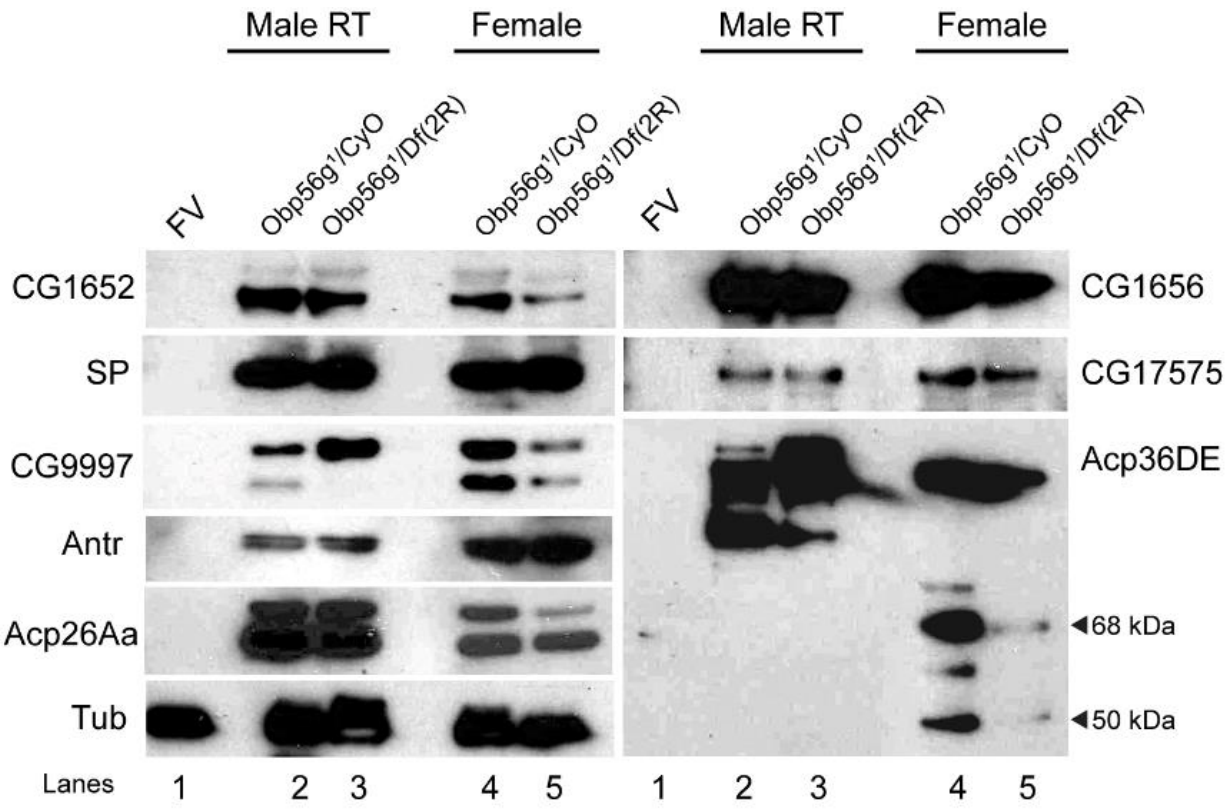


Figure 4—figure supplement 2: Females mated to *Obp56g*^{1} null males have reduced amounts of SFPs in their bursas 35 minutes ASM. Western blots for SFPs in 1) unmated female reproductive tracts, (2-3) male reproductive tracts, or (4-5) mated female reproductive tracts from CS females mated to either *Obp56g*^{1}*/CyO* control or *Obp56g*^{1}*/Df(2R)* males at 35 minutes ASM. All flies are 3-5 days old. Tubulin is shown as a loading control. Cleavage products of Acp36DE (68kDa and 50kDa) are shown with black arrows.

Figure 4—figure supplement 2—source data 1: Raw film images and uncropped, labeled Western blots for data shown in Figure 4—figure supplement 2.

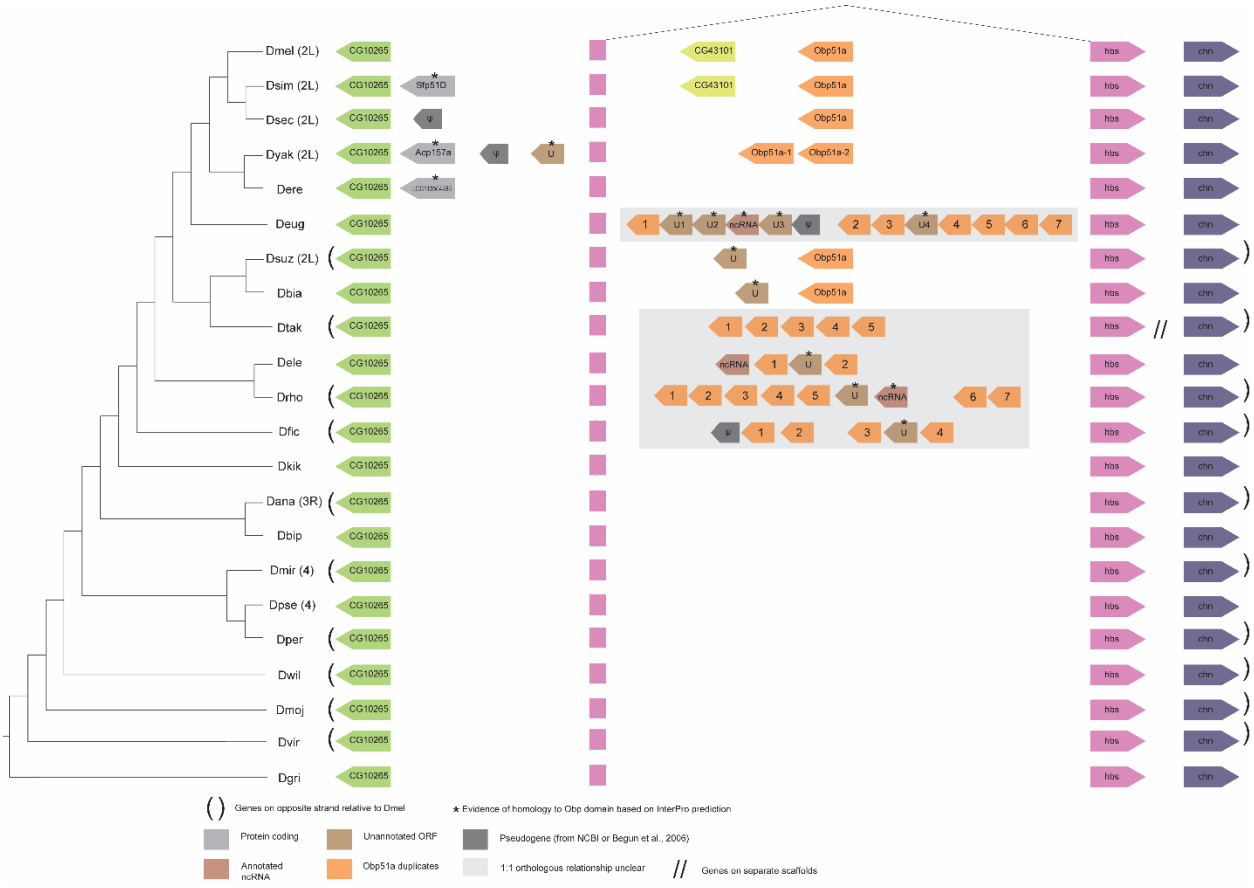


Figure 5—figure supplement 1: Synteny plot for *Obp51a*, phylogeny on the left from (McGeary and Findlay, 2020). Surrounding gene names represent gene names in *D. melanogaster*.

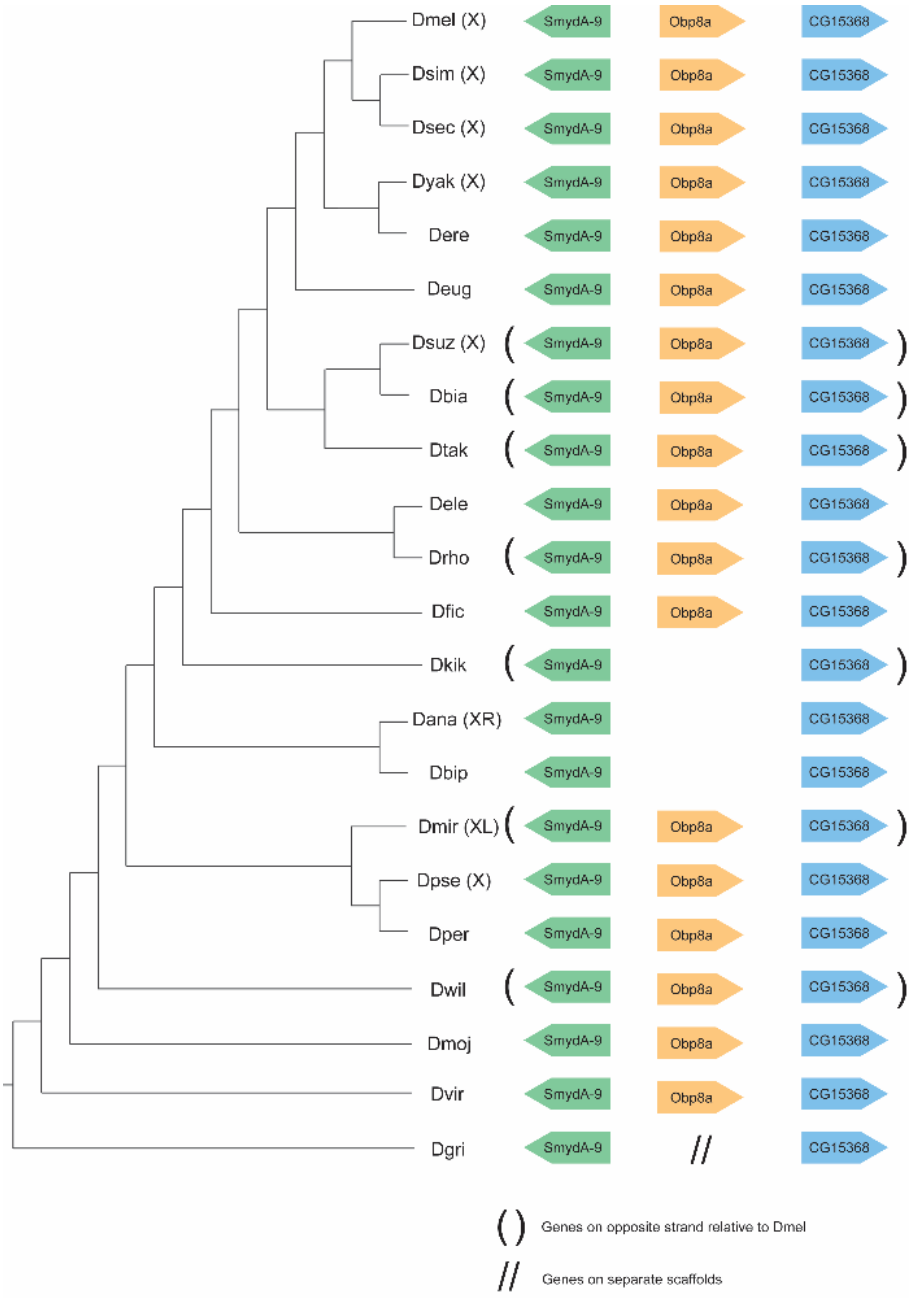


Figure 5—figure supplement 2: Synteny plot for *Obp8a*, phylogeny on the left from (McGeary and Findlay, 2020). Surrounding gene names represent gene names in *D. melanogaster*.

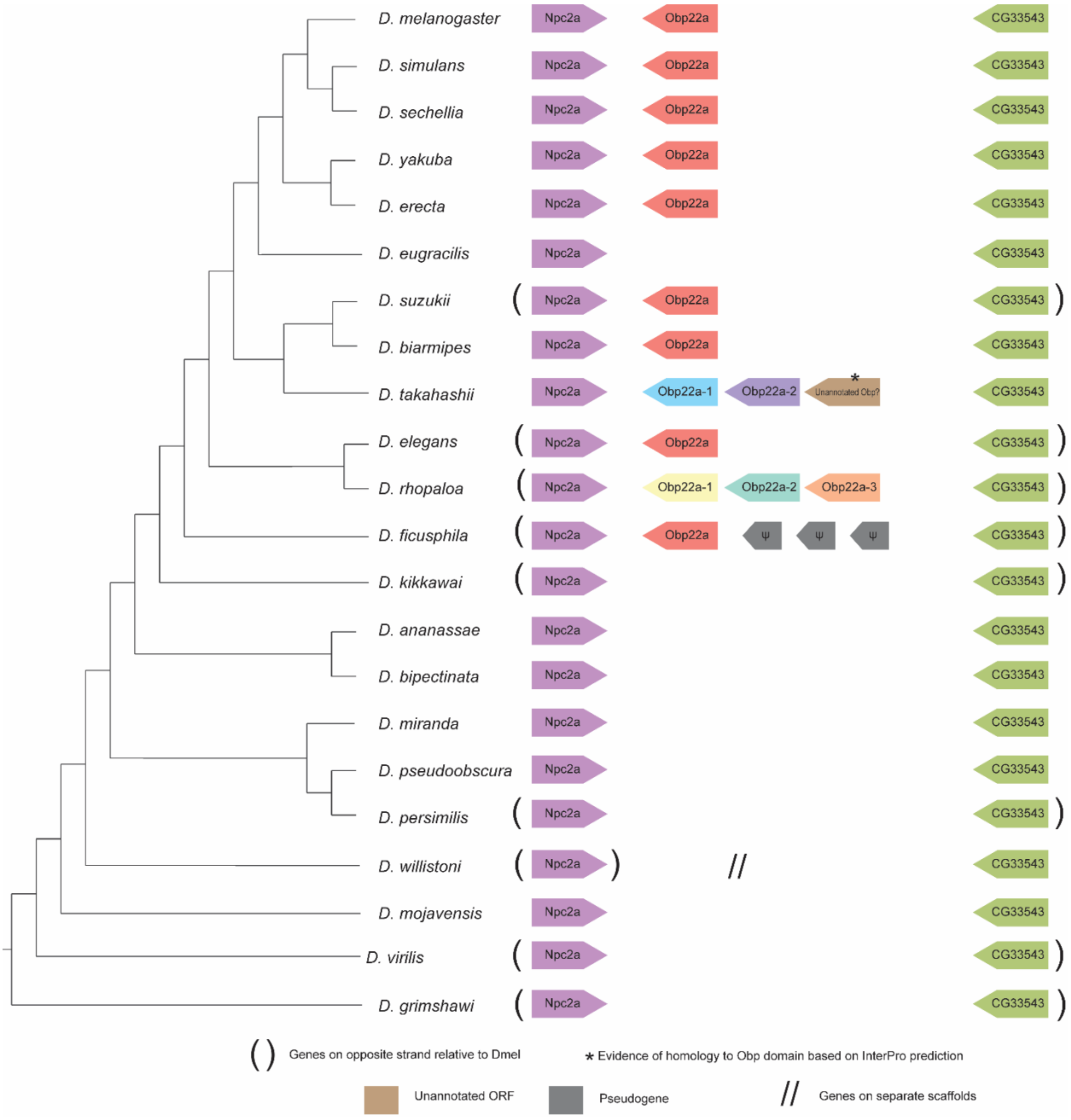


Figure 5—figure supplement 3: Synteny plot for *Obp22a*, phylogeny on the left from (McGeary and Findlay, 2020). Surrounding gene names represent gene names in *D. melanogaster*.

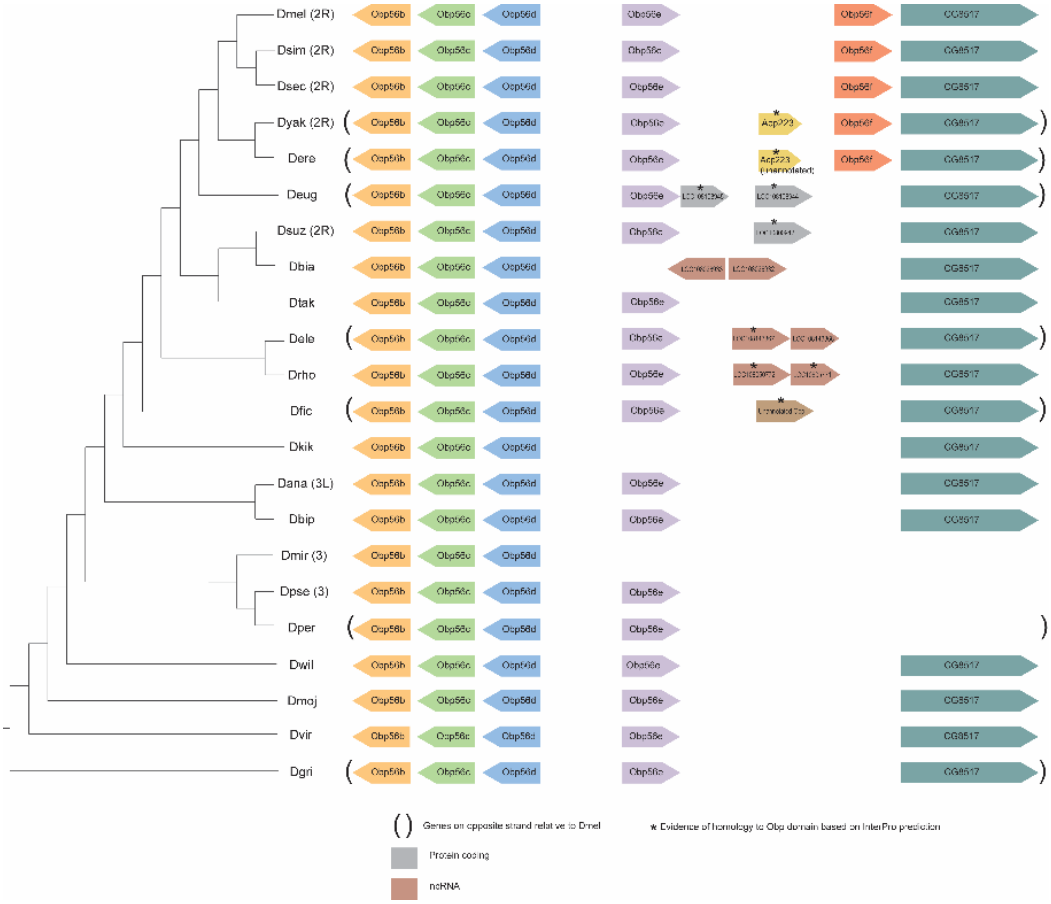


Figure 5—figure supplement 4: Synteny plot for *Obp56e* and *Obp56f*, phylogeny on the left from (McGeary and Findlay, 2020). Surrounding gene names represent gene names in *D. melanogaster*.

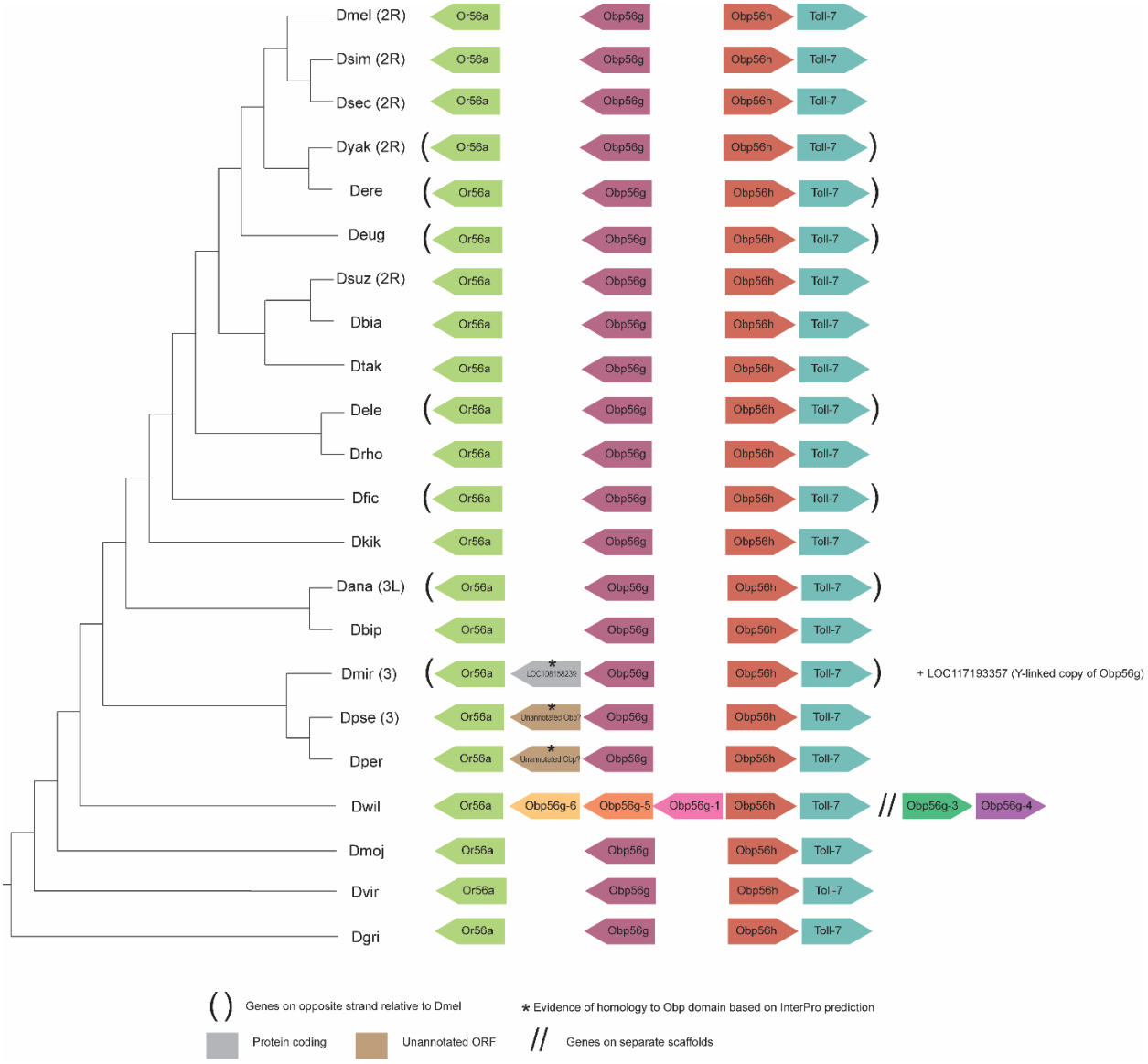


Figure 5—figure supplement 5: Synteny plot for *Obp56g*, phylogeny on the left from (McGeary and Findlay, 2020). Surrounding gene names represent gene names in *D. melanogaster*.

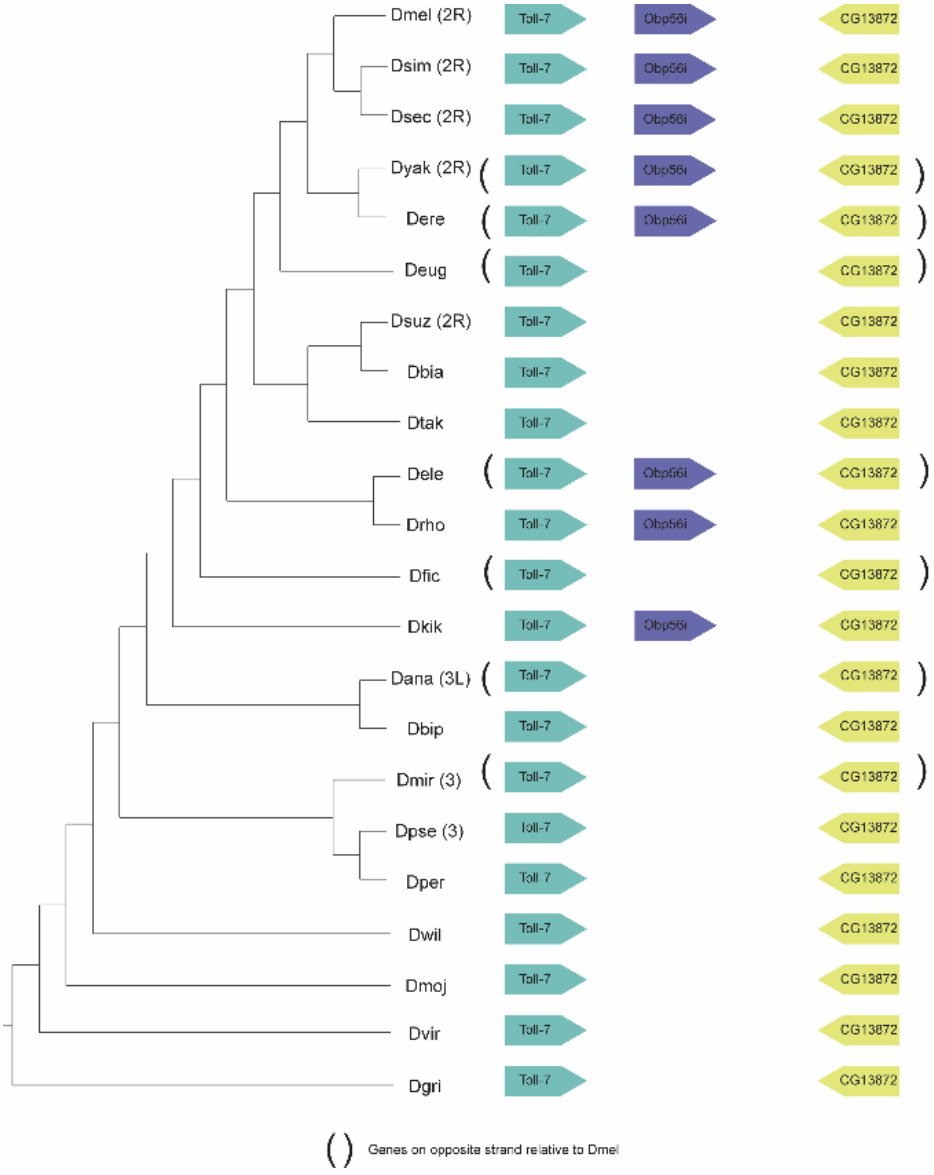


Figure 6—figure supplement 6: Synteny plot for *Obp56i*, phylogeny on the left from (McGeary and Findlay, 2020). Surrounding gene names represent gene names in *D. melanogaster*.

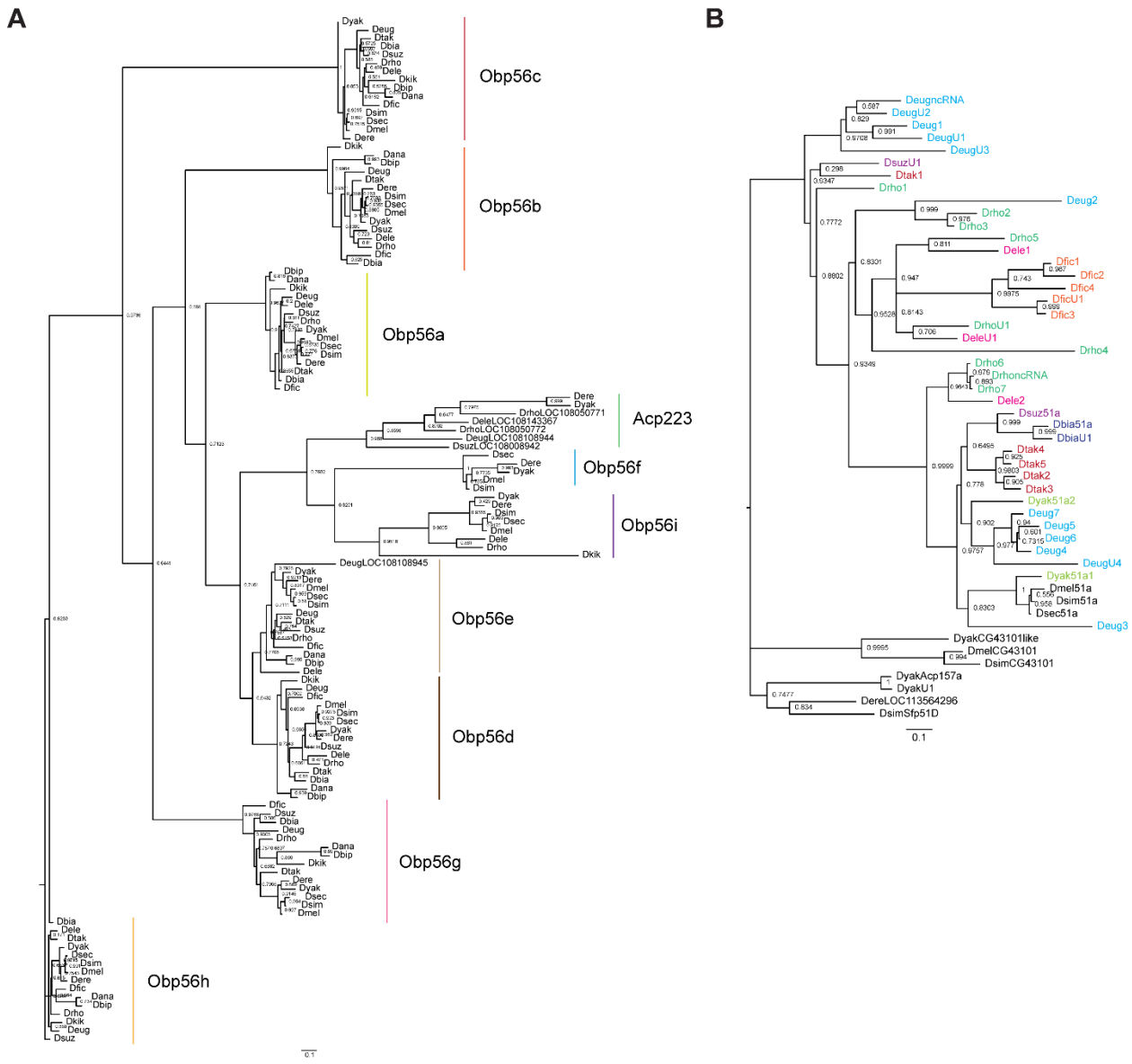


Figure 5—figure supplement 7: RAXML-NG maximum likelihood inferred trees for genes in the A) *Obp56* cluster across *melanogaster* group species, or B) *Obp51a* cluster, where genes are colored as in Figure 5—figure supplement 1. Node values are bootstrap support estimates based on 1,000 replicates. *CG43101* is a gene located next to *Obp51a* in *D. melanogaster*, which has 6 cysteines in a pattern reminiscent of the *Obp* "domain" but is not a predicted *Obp* based on InterProScan searches.

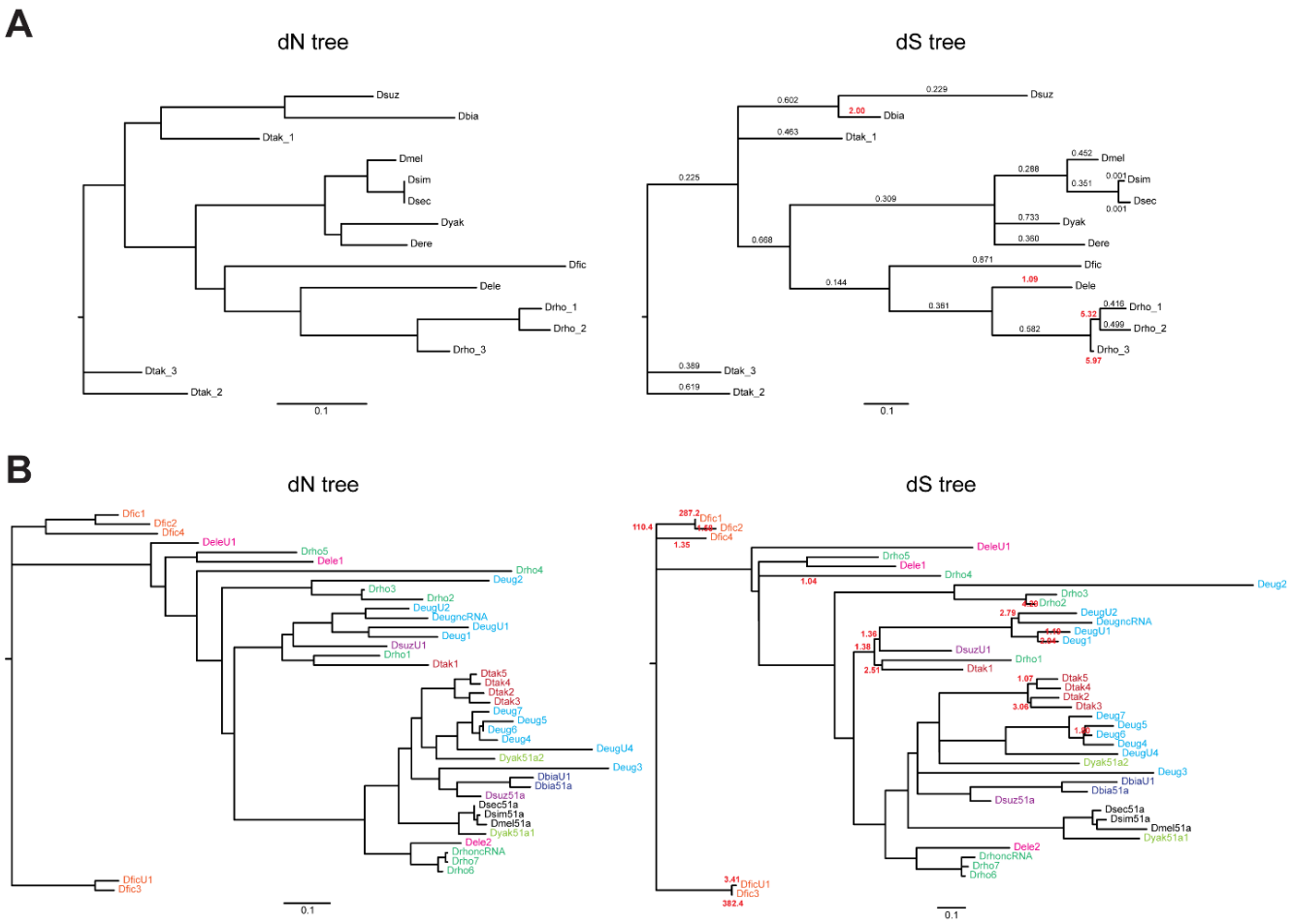


Figure 5—figure supplement 8: A) *Obp22a* and B) *Obp51a* maximum likelihood inferred gene trees, where branch lengths are proportional to either estimates of dN (left) or dS (right) from PAML. Values indicated on the dS tree represent ML-inferred estimates of ω from PAML's free ratio model, where the value is bold and red if ω>1. Values on the *Obp51a* dS tree are only shown if ω >1 for clarity. Genes in B) are color-coded by species if more than one paralog is present in that species' genomes.

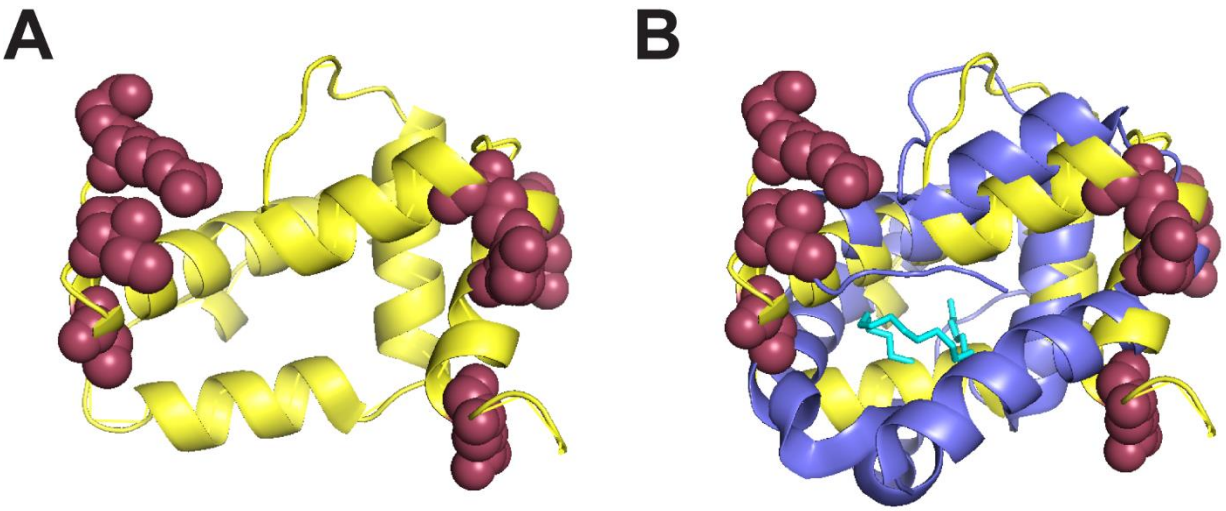


Figure 5—figure supplement 9: Positively selected sites in *Obp22a* cluster on the outward-facing region of the protein. A) Alphafold predicted protein structure (yellow) of Obp22a with the positively selected sites (Pr>0.90 BEB from model M8 of PAML) shown in maroon (Jumper et al., 2021). B) The same structure as A with a superimposed alignment of the crystal structure of Obp76a (LUSH) from (Laughlin et al., 2008), (purple). The cyan molecule represents cVA and the inferred region of the binding pocket.

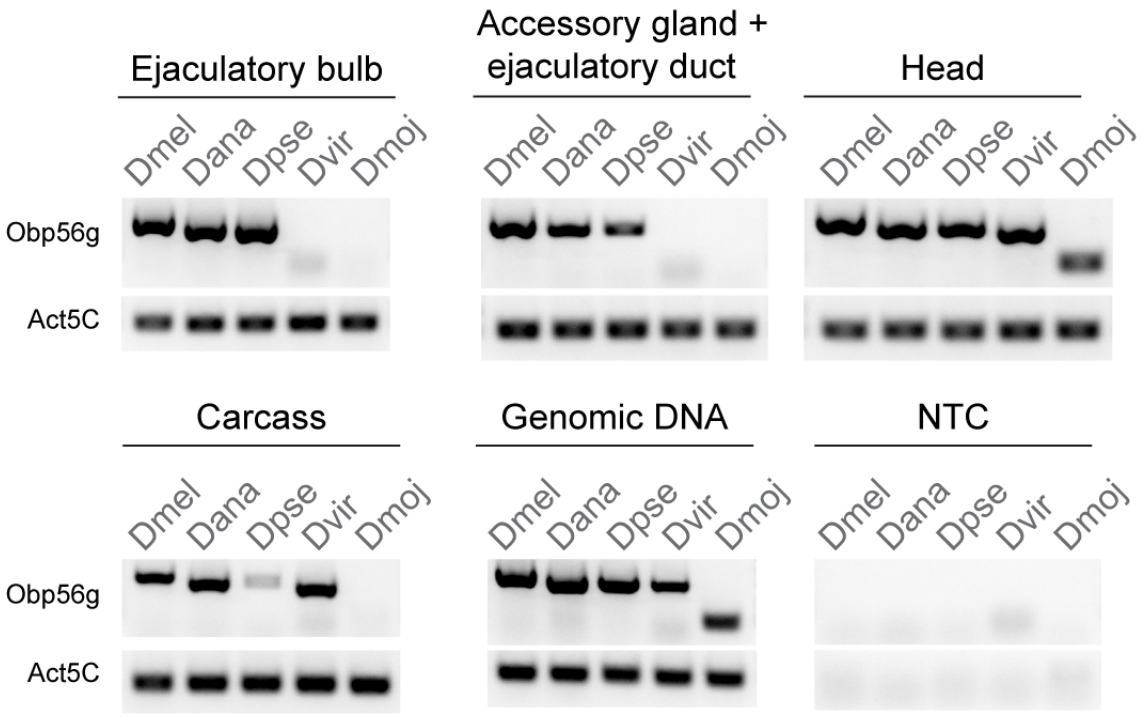


Figure 6—figure supplement 1: Semi-quantitative RT-PCR data from dissected tissues (head, accessory gland + ejaculatory duct, ejaculatory bulb, and carcass) *from D. melanogaster* (Dmel)*, D. ananassae* (Dana)*, D. pseudoobscura* (Dpse)*, D. virilis* (Dvir), and *D. mojavensis* (Dmoj) males after 35 cycles of PCR. NTC = no template control.

Figure 6—figure supplement 1—source data 1: raw and uncropped, labeled gel images for data shown in Figure 6—figure supplement 1.

Supplementary tables:

Gene	**gRNA sequence (5'->3')**
Obp8a	1: GGTGAGGATCGCATGGGCAC 2: GCTGGACAGGATGCAGTTCG 3: ACATGTCCGATGTCATCAAT
Obp22a	1: AATTGTAAGCGAGTGTGCCA 2: GAACAATGTTCATAGGAAGA 3: AAAGTGAGGGGGATAGATAG

Obp51a	1: TGACAGCTAACAACAGAACC 2: GAACGAATGTGCTAAAAAAC 3: TAAATTCTCGTTTCAAGCAC
Obp56e	1: TGAGGCTAAGCAGAGAGCCA 2: CAAGCTATTGCCCTGCGGTC 3: GCCAAGTGTGACTCGACCAA
Obp56f	1: AGCCTGCTTGAAACGGCAGC 2: CACTGCTTACTGGAAGTGAA 3: ATGTTTAGAAGTCTAATGCT
Obp56g	1: GCAAGCCAACATAGACAGTT 2: CGGTGTCACTCCCCAGGATC 3: CGGATCGTTAAGACCCTAAT
Obp56i	1: GGTACAAGCAGGTCCCATTA 2: CGTCATGAGACCGACGACCC 3: CGAAGAACTCGAAATCACAG
ebony (gRNA sequence from Kane et al., 2017)	GCCACAATTGTCGATCGTCA

Table S1: gRNA sequences from flyCRISPR's Optimal Target Finder tool for each *Obp* gene.

Primer	**Primer Sequence (5' -> 3')**	
Primer 1F	TTCCCGGCCGATGCAnnnnnnnnnnnnnnnnnnnnGTTTaAGAGCTAtgctgGAAAcag	n: 20nt gRNA 1
Primer 1R	nnnnnnnnnnnnnnnnnnnnTGCACCAGCCGGGAATC	n: 20nt gRNA 2 (RevComp)
Primer 2F	nnnnnnnnnnnnnnnnnnnnGTTTaAGAGCTAtgctgGAAAcag	n: 20nt gRNA 2

Primer 2R	nnnnnnnnnnnnnnnnnnnnTGCACCAGCCGGGAATC	n: 20nt gRNA 3 (RevComp)
Primer 3F	nnnnnnnnnnnnnnnnnnnnGTTTaAGAGCTAtgctgGAAAcag	n: 20nt gRNA 3
Primer 3R	TTCcagcaTAGCTCTtAAACnnnnnnnnnnnnnnnnnnnnTGCACCAGCCGGGAATC	n: 20nt gRNA 4 (RevComp)

Table S2: Primer sequences for cloning gRNAs from Table S1 into pAC-U63-tgRNA-Rev using pMGC as a PCR template (from Poe et al., 2018).

Gene	**Primer sequence (5' -> 3')**	**Purpose of primer pair**
Obp8a	0F: TCGTAGGTCAGCAGCCCATTAC 0R: TCGCATATGACTTTCAATCCGTGT 1F: CGTGGGAATGATGCGGAGA 1R: CATGGGCAGCATCCTCGAAT	0: Sequencing CRISPR mutants 1: RT-PCR
Obp22a	2F: CCACTTTGTATTGGCAACCGCA 2R: CAGTCCGCCCAACTTTGAGTTT 3F: TGTACTTCTGCTTGGCCTCTC 3R: TTTTGGAAGGATTCTGCACAC	2: Sequencing CRISPR mutants 3: RT-PCR
Obp51a	4F: AGCAATCTCCCTCACGTGATAT 4R: TGCGGCGCTCATGTTTCTTTTA 5F: GGCCTGGTTCTGTTGTTAGC 5R: TCAAGCACTGGAACACCAAG	4: Sequencing CRISPR mutants 5: RT-PCR
Obp56e	6F: ACCTGACAACAAGAAATAACCCGC 6R: CACTAGAGCAAGCGTTCCGTTC 7F: CCCTTGCAGCTCTATCTTTGG 7R: CTTGGTCGAGTCACACTTGG	6: Sequencing CRISPR mutants 7: RT-PCR

Obp56f	8F: GGTAACAGTCCCTGGAAACCGA 8R: GCGCTTTGCCCGGAATAATCTT 9F: TTCATTTTCATCTCTGCTATCTGG 9R: GCCCAATTCACATTTTCCTG	8: Sequencing CRISPR mutants 9: RT-PCR
Obp56g	10F: GTTAGAAACCTTGACAGTGGCA 10R: ATGGGGTAGGCAGTGTATCCCT 11F: AGGGCTACATTCGCATTGAC 11R: ACCTGTCCAAATCCTTTTCG	10: Sequencing CRISPR mutants 11: RT-PCR
Obp56i	12F: ACCTCCATTCGGGTATCTCGAC 12R: GACTGAGTGATGCAAAGCACGT 13F: TGCTGTGCATTATTGTTAGTCG 13R: ACTCGTCATGGGATGTCTCG	12: Sequencing CRISPR mutants 13: RT-PCR
Actin 5C	F: AGCGCGGTTACTCTTTCACCAC R: GTGGCCATCTCCTGCTCAAAGT	RT-PCR control gene
D. ananassae Obp56g	F: TGACTCTGCTGCTTAGCTGC R: GATCCTTGTCCACCTGAGCC	
D. pseudoobscura Obp56g	F: GGAGCCGGAGACATAAGCAA R: GCAGGTTTCCTTTCGCATCC	
D. mojavensis Obp56g	F: AGAAGCCCGAAATGACCCAG R: CTCCAGCTTCACCTCACCAG	
D. virilis Obp56g	F: GCTGCTTCTCGGCTGTCTAA R: CCTTAGCTGGCGCATCCTTA	

Table S3: Primer sequences used in this study.

Gene	**Allele designation**	**Mutant allele description**
Obp8a	*Obp8a$^{\Delta 390}$*	390 bp deletion in exon 2 between gRNA 1 + 3 (95% of non-signal peptide sequence)

Obp22a	*Obp22a*$^{\Delta 257}$	257 bp deletion in exon 2 between gRNA 1 + 3 (86% of non-signal peptide sequence)
Obp51a	*Obp51a*$^{\Delta 16}$	16 bp deletion in middle of signal peptide region of exon 1 within gRNA 1 site (predicted frameshift and early stop codon)
Obp56e	*Obp56e*$^{\Delta 239}$	239 bp deletion in exon 2 between gRNA 1 + 3 (69% of non-signal peptide sequence)
Obp56f	*Obp56f*$^{\Delta 226}$	226 bp deletion in exon 2 between gRNA 2 + 3 (67% of non-signal peptide sequence + 13 bp into 3' UTR)
Obp56g	*Obp56g*$^{\Delta 333}$	333 bp deletion in exon 2 between gRNA 1 + 3 (95% of non-signal peptide sequence + 7 bp into 3' UTR)
Obp56i	*Obp56i*$^{\Delta 359}$	359 bp deletion in exon 2 between gRNA 1 + 3 (98% of non-signal peptide sequence)

Table S4: CRISPR mutant allele summary for each *Obp* gene.

Gene	**4-day receptivity**	**BH-adjusted *p*-value**
Obp8a	KO: 2/17 Control: 10/18	0.069
Obp22a	KO: 1/15 Control: 0/19	1
Obp51a	KO: 3/19 Control: 1/15	1
Obp56e	KO: 0/17 Control:1/19	1
Obp56f	KO: 3/18 Control: 4/15	1
Obp56i	KO: 1/17 Control: 3/14	0.912

Table S5: Four-day receptivity data from an additional replicate using CRISPR mutant males of the genotypes indicated.

Gene	Male genotype	% CS females with mating plugs present after copulation
Obp8a	*Obp8a$^{\Delta 390}$*	100% (n=7)
	Obp8a^{WT}	100% (n=9)
Obp22a	*Obp22a$^{\Delta 257}$*	100% (n=8)
	Obp22a$^{\Delta 257}$/ CyO	100% (n=9)
Obp51a	*Obp51a$^{\Delta 16}$*	100% (n=8)
	Obp51a$^{\Delta 16}$/ CyO	100% (n=8)
Obp56e	*Obp56e$^{\Delta 239}$*	100% (n=8)
	Obp56e$^{\Delta 239}$/ CyO	100% (n=9)
Obp56f	*Obp56f$^{\Delta 226}$*	90% (n=10)
	Obp56f$^{\Delta 226}$/ CyO	100% (n=10)
Obp56g	*Obp56g$^{\Delta 333}$*	0% (n=14)
	Obp56g$^{\Delta 333}$/ CyO	100% (n=11)
Obp56i	*Obp56i$^{\Delta 359}$*	100% (n=10)
	Obp56i$^{\Delta 359}$/ CyO	100% (n=10)

Table S6: Proportion of CS females mated to CRISPR mutant males with morphologically normal mating plugs assessed immediately after the end of mating.

Supplemental methods:

To build our gRNA-expressing vectors, we used pAC-U63-tgRNA-Rev, a plasmid that expresses multiplexed gRNAs separated by rice Gly tRNA sequences, as well as the (F+E)

gRNA scaffold, under the control of the *Drosophila* U6:3 promoter (Poe et al., 2018). We designed Gibson assembly primers containing our gRNA sequences according to (Poe et al., 2018, Table S2). We used these primers to generate PCR products using the pMGC template vector and purified products of the correct size using a gel extraction kit (Poe et al., 2018, Zymo). The empty pAC-U63-tgRNA-Rev plasmid was digested using *Sap*I, and the digested vector and purified PCR products were assembled using the HiFi assembly kit (NEB, NEBuilder). The pAC-U63-tgRNA-Rev and pMGC plasmids were generous gifts from Chun Han at Cornell University.

References:

Ahmed-Braimah YH, Unckless RL, Clark AG. 2017. Evolutionary Dynamics of Male Reproductive Genes in the Drosophila virilis Subgroup. *G3 Bethesda Md* **7**:3145–3155. doi:10.1534/g3.117.1136

Ai M, Min S, Grosjean Y, Leblanc C, Bell R, Benton R, Suh GSB. 2010. Acid sensing by the Drosophila olfactory system. *Nature* **468**:691–695. doi:10.1038/nature09537

Alonso-Pimentel H, Tolbert LP, Heed WB. 1994. Ultrastructural examination of the insemination reaction in Drosophila. *Cell Tissue Res* **275**:467–479. doi:10.1007/BF00318816

Avila FW, Cohen AB, Ameerudeen FS, Duneau D, Suresh S, Mattei AL, Wolfner MF. 2015. Retention of Ejaculate by Drosophila melanogaster Females Requires the Male-Derived Mating Plug Protein PEBme. *Genetics* **200**:1171–1179. doi:10.1534/genetics.115.176669

Avila FW, Sirot LK, LaFlamme BA, Rubinstein CD, Wolfner MF. 2011. Insect Seminal Fluid Proteins: Identification and Function. *Annu Rev Entomol* **56**:21–40. doi:10.1146/annurev-ento-120709-144823

Avila FW, Wolfner MF. 2009. Acp36DE is required for uterine conformational changes in mated Drosophila females. *Proc Natl Acad Sci U S A* **106**:15796–15800. doi:10.1073/pnas.0904029106

Baer B, Zareie R, Paynter E, Poland V, Millar AH. 2012. Seminal fluid proteins differ in abundance between genetic lineages of honeybees. *J Proteomics* **75**:5646–5653. doi:10.1016/j.jprot.2012.08.002

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* **67**. doi:10.18637/jss.v067.i01

1205 Begun DJ, Lindfors HA. 2005. Rapid Evolution of Genomic Acp Complement in the
1206 melanogaster Subgroup of Drosophila. *Mol Biol Evol* **22**:2010–2021.
1207 doi:10.1093/molbev/msi201
1208 Begun DJ, Lindfors HA, Thompson ME, Holloway AK. 2006. Recently Evolved Genes Identified
1209 From Drosophila yakuba and D. erecta Accessory Gland Expressed Sequence Tags.
1210 *Genetics* **172**:1675–1681. doi:10.1534/genetics.105.050336
1211 Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful
1212 Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* **57**:289–300.
1213 Benoit JB, Vigneron A, Broderick NA, Wu Y, Sun JS, Carlson JR, Aksoy S, Weiss BL. 2017.
1214 Symbiont-induced odorant binding proteins mediate insect host hematopoiesis. *eLife*
1215 **6**:e19535. doi:10.7554/eLife.19535
1216 Bertram MJ, Neubaum DM, Wolfner MF. 1996. Localization of the Drosophila male accessory
1217 gland protein Acp36DE in the mated female suggests a role in sperm storage. *Insect*
1218 *Biochem Mol Biol* **26**:971–980. doi:10.1016/S0965-1748(96)00064-1
1219 Billeter J-C, Levine J. 2015. The role of cVA and the Odorant binding protein Lush in social and
1220 sexual behavior in Drosophila melanogaster. *Front Ecol Evol* **3**.
1221 Billeter J-C, Wolfner MF. 2018. Chemical Cues that Guide Female Reproduction in Drosophila
1222 melanogaster. *J Chem Ecol* **44**:750–769. doi:10.1007/s10886-018-0947-z
1223 Birchler JA, Yang H. 2022. The multiple fates of gene duplications: Deletion,
1224 hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance
1225 constraints, and neutral variation. *Plant Cell* **34**:2466–2474. doi:10.1093/plcell/koac076
1226 Bretman A, Lawniczak MKN, Boone J, Chapman T. 2010. A mating plug protein reduces early
1227 female remating in Drosophila melanogaster. *J Insect Physiol* **56**:107–113.
1228 doi:10.1016/j.jinsphys.2009.09.010
1229 Brieger G, Butterworth FM. 1970. Drosophila melanogaster: Identity of Male Lipid in
1230 Reproductive System. *Science* **167**:1262–1262. doi:10.1126/science.167.3922.1262
1231 Carlisle JA, Glenski MA, Swanson WJ. 2022. Recurrent Duplication and Diversification of
1232 Acrosomal Fertilization Proteins in Abalone. *Front Cell Dev Biol* **10**:795273.
1233 doi:10.3389/fcell.2022.795273
1234 Cavener DR. 1985. Coevolution of the glucose dehydrogenase gene and the ejaculatory duct in
1235 the genus Drosophila. *Mol Biol Evol* **2**:141–149.
1236 doi:10.1093/oxfordjournals.molbev.a040344
1237 Chang C-H, Malik HS. 2022. Genetic conflicts between sex chromosomes drive expansion and
1238 loss of sperm nuclear basic protein genes in Drosophila. doi:10.1101/2022.06.08.495379

1239 Chen DS, Delbare SYN, White SL, Sitnik J, Chatterjee M, DoBell E, Weiss O, Clark AG, Wolfner
1240 MF. 2019. Female Genetic Contributions to Sperm Competition in Drosophila
1241 melanogaster. *Genetics* **212**:789–800. doi:10.1534/genetics.119.302284
1242 Chin JS, Ellis SR, Pham HT, Blanksby SJ, Mori K, Koh QL, Etges WJ, Yew JY. 2014. Sex-
1243 specific triacylglycerides are widely conserved in Drosophila and mediate mating
1244 behavior. *eLife* **3**:e01751. doi:10.7554/eLife.01751
1245 Cohen AB, Wolfner MF. 2018. Dynamic changes in ejaculatory bulb size during Drosophila
1246 melanogaster aging and mating. *J Insect Physiol* **107**:152–156.
1247 doi:10.1016/j.jinsphys.2018.04.005
1248 Cook RK, Christensen SJ, Deal JA, Coburn RA, Deal ME, Gresens JM, Kaufman TC, Cook KR.
1249 2012. The generation of chromosomal deletions to provide extensive coverage and
1250 subdivision of the Drosophila melanogaster genome. *Genome Biol* **13**:R21.
1251 doi:10.1186/gb-2012-13-3-r21
1252 Drosophila 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B,
1253 Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN, Pollard DA, Sackton TB,
1254 Larracuente AM, Singh ND, Abad JP, Abt DN, Adryan B, Aguade M, Akashi H, Anderson
1255 WW, Aquadro CF, Ardell DH, Arguello R, Artieri CG, Barbash DA, Barker D, Barsanti P,
1256 Batterham P, Batzoglou S, Begun D, Bhutkar A, Blanco E, Bosak SA, Bradley RK, Brand
1257 AD, Brent MR, Brooks AN, Brown RH, Butlin RK, Caggese C, Calvi BR, Bernardo de
1258 Carvalho A, Caspi A, Castrezana S, Celniker SE, Chang JL, Chapple C, Chatterji S,
1259 Chinwalla A, Civetta A, Clifton SW, Comeron JM, Costello JC, Coyne JA, Daub J, David
1260 RG, Delcher AL, Delehaunty K, Do CB, Ebling H, Edwards K, Eickbush T, Evans JD,
1261 Filipski A, Findeiss S, Freyhult E, Fulton L, Fulton R, Garcia ACL, Gardiner A, Garfield
1262 DA, Garvin BE, Gibson G, Gilbert D, Gnerre S, Godfrey J, Good R, Gotea V, Gravely B,
1263 Greenberg AJ, Griffiths-Jones S, Gross S, Guigo R, Gustafson EA, Haerty W, Hahn
1264 MW, Halligan DL, Halpern AL, Halter GM, Han MV, Heger A, Hillier L, Hinrichs AS,
1265 Holmes I, Hoskins RA, Hubisz MJ, Hultmark D, Huntley MA, Jaffe DB, Jagadeeshan S,
1266 Jeck WR, Johnson J, Jones CD, Jordan WC, Karpen GH, Kataoka E, Keightley PD,
1267 Kheradpour P, Kirkness EF, Koerich LB, Kristiansen K, Kudrna D, Kulathinal RJ, Kumar
1268 S, Kwok R, Lander E, Langley CH, Lapoint R, Lazzaro BP, Lee S-J, Levesque L, Li R,
1269 Lin C-F, Lin MF, Lindblad-Toh K, Llopart A, Long M, Low L, Lozovsky E, Lu J, Luo M,
1270 Machado CA, Makalowski W, Marzo M, Matsuda M, Matzkin L, McAllister B, McBride
1271 CS, McKernan B, McKernan K, Mendez-Lago M, Minx P, Mollenhauer MU, Montooth K,
1272 Mount SM, Mu X, Myers E, Negre B, Newfeld S, Nielsen R, Noor MAF, O'Grady P,

Pachter L, Papaceit M, Parisi MJ, Parisi M, Parts L, Pedersen JS, Pesole G, Phillippy AM, Ponting CP, Pop M, Porcelli D, Powell JR, Prohaska S, Pruitt K, Puig M, Quesneville H, Ram KR, Rand D, Rasmussen MD, Reed LK, Reenan R, Reily A, Remington KA, Rieger TT, Ritchie MG, Robin C, Rogers Y-H, Rohde C, Rozas J, Rubenfield MJ, Ruiz A, Russo S, Salzberg SL, Sanchez-Gracia A, Saranga DJ, Sato H, Schaeffer SW, Schatz MC, Schlenke T, Schwartz R, Segarra C, Singh RS, Sirot L, Sirota M, Sisneros NB, Smith CD, Smith TF, Spieth J, Stage DE, Stark A, Stephan W, Strausberg RL, Strempel S, Sturgill D, Sutton G, Sutton GG, Tao W, Teichmann S, Tobari YN, Tomimura Y, Tsolas JM, Valente VLS, Venter E, Venter JC, Vicario S, Vieira FG, Vilella AJ, Villasante A, Walenz B, Wang J, Wasserman M, Watts T, Wilson D, Wilson RK, Wing RA, Wolfner MF, Wong A, Wong GK-S, Wu C-I, Wu G, Yamamoto D, Yang H-P, Yang S-P, Yorke JA, Yoshida K, Zdobnov E, Zhang P, Zhang Y, Zimin AV, Baldwin J, Abdouelleil A, Abdulkadir J, Abebe A, Abera B, Abreu J, Acer SC, Aftuck L, Alexander A, An P, Anderson E, Anderson S, Arachi H, Azer M, Bachantsang P, Barry A, Bayul T, Berlin A, Bessette D, Bloom T, Blye J, Boguslavskiy L, Bonnet C, Boukhgalter B, Bourzgui I, Brown A, Cahill P, Channer S, Cheshatsang Y, Chuda L, Citroen M, Collymore A, Cooke P, Costello M, D'Aco K, Daza R, De Haan G, DeGray S, DeMaso C, Dhargay N, Dooley K, Dooley E, Doricent M, Dorje P, Dorjee K, Dupes A, Elong R, Falk J, Farina A, Faro S, Ferguson D, Fisher S, Foley CD, Franke A, Friedrich D, Gadbois L, Gearin G, Gearin CR, Giannoukos G, Goode T, Graham J, Grandbois E, Grewal S, Gyaltsen K, Hafez N, Hagos B, Hall J, Henson C, Hollinger A, Honan T, Huard MD, Hughes L, Hurhula B, Husby ME, Kamat A, Kanga B, Kashin S, Khazanovich D, Kisner P, Lance K, Lara M, Lee W, Lennon N, Letendre F, LeVine R, Lipovsky A, Liu X, Liu J, Liu S, Lokyitsang T, Lokyitsang Y, Lubonja R, Lui A, MacDonald P, Magnisalis V, Maru K, Matthews C, McCusker W, McDonough S, Mehta T, Meldrim J, Meneus L, Mihai O, Mihalev A, Mihova T, Mittelman R, Mlenga V, Montmayeur A, Mulrain L, Navidi A, Naylor J, Negash T, Nguyen T, Nguyen N, Nicol R, Norbu C, Norbu N, Novod N, O'Neill B, Osman S, Markiewicz E, Oyono OL, Patti C, Phunkhang P, Pierre F, Priest M, Raghuraman S, Rege F, Reyes R, Rise C, Rogov P, Ross K, Ryan E, Settipalli S, Shea T, Sherpa N, Shi L, Shih D, Sparrow T, Spaulding J, Stalker J, Stange-Thomann N, Stavropoulos S, Stone C, Strader C, Tesfaye S, Thomson T, Thoulutsang Y, Thoulutsang D, Topham K, Topping I, Tsamla T, Vassiliev H, Vo A, Wangchuk T, Wangdi T, Weiand M, Wilkinson J, Wilson A, Yadav S, Young G, Yu Q, Zembek L, Zhong D, Zimmer A, Zwirko Z, Jaffe DB, Alvarez P, Brockman W, Butler J, Chin C,

Gnerre S, Grabherr M, Kleber M, Mauceli E, MacCallum I. 2007. Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450**:203–218. doi:10.1038/nature06341

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792–1797. doi:10.1093/nar/gkh340

Everaerts C, Farine J-P, Cobb M, Ferveur J-F. 2010. Drosophila Cuticular Hydrocarbons Revisited: Mating Status Alters Cuticular Profiles. *PLOS ONE* **5**:e9607. doi:10.1371/journal.pone.0009607

Findlay GD, MacCoss MJ, Swanson WJ. 2009. Proteomic discovery of previously unannotated, rapidly evolving seminal fluid genes in Drosophila. *Genome Res* **19**:886–896. doi:10.1101/gr.089391.108

Findlay GD, Sitnik JL, Wang W, Aquadro CF, Clark NL, Wolfner MF. 2014. Evolutionary Rate Covariation Identifies New Members of a Protein Network Required for Drosophila melanogaster Female Post-Mating Responses. *PLoS Genet* **10**:e1004108. doi:10.1371/journal.pgen.1004108

Findlay GD, Yi X, MacCoss MJ, Swanson WJ. 2008. Proteomics Reveals Novel Drosophila Seminal Fluid Proteins Transferred at Mating. *PLOS Biol* **6**:e178. doi:10.1371/journal.pbio.0060178

Galindo K, Smith DP. 2001. A large family of divergent Drosophila odorant-binding proteins expressed in gustatory and olfactory sensilla. *Genetics* **159**:1059–1072. doi:10.1093/genetics/159.3.1059

Garlovsky MD, Evans C, Rosenow MA, Karr TL, Snook RR. 2020. Seminal fluid protein divergence among populations exhibiting postmating prezygotic reproductive isolation. *Mol Ecol* **29**:4428–4441. doi:10.1111/mec.15636

Gilchrist AS, Partridge L. 2000. Why it is difficult to model sperm displacement in Drosophila melanogaster: the relation between sperm transfer and copulation duration. *Evol Int J Org Evol* **54**:534–542. doi:10.1111/j.0014-3820.2000.tb00056.x

Gomez-Diaz C, Reina JH, Cambillau C, Benton R. 2013. Ligands for Pheromone-Sensing Neurons Are Not Conformationally Activated Odorant Binding Proteins. *PLOS Biol* **11**:e1001546. doi:10.1371/journal.pbio.1001546

Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, O'Connor-Giles KM. 2014. Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila. *Genetics* **196**:961–971. doi:10.1534/genetics.113.160713

1339 Guiraudie-Capraz G, Pho DB, Jallon J-M. 2007. Role of the ejaculatory bulb in biosynthesis of
1340 the male pheromone cis-vaccenyl acetate in Drosophila melanogaster. *Integr Zool* **2**:89–
1341 99. doi:10.1111/j.1749-4877.2007.00047.x
1342 Ha TS, Smith DP. 2006. A Pheromone Receptor Mediates 11-cis-Vaccenyl Acetate-Induced
1343 Responses in Drosophila. *J Neurosci* **26**:8727–8733. doi:10.1523/JNEUROSCI.0876-
1344 06.2006
1345 Haerty W, Jagadeeshan S, Kulathinal RJ, Wong A, Ravi Ram K, Sirot LK, Levesque L, Artieri
1346 CG, Wolfner MF, Civetta A, Singh RS. 2007. Evolution in the Fast Lane: Rapidly
1347 Evolving Sex-Related Genes in Drosophila. *Genetics* **177**:1321–1335.
1348 doi:10.1534/genetics.107.078865
1349 Hallem EA, Carlson JR. 2006. Coding of Odors by a Receptor Repertoire. *Cell* **125**:143–160.
1350 doi:10.1016/j.cell.2006.01.050
1351 Han C, Jan LY, Jan Y-N. 2011. Enhancer-driven membrane markers for analysis of
1352 nonautonomous mechanisms reveal neuron–glia interactions in Drosophila. *Proc Natl*
1353 *Acad Sci U S A* **108**:9673–9678. doi:10.1073/pnas.1106386108
1354 Hekmat-Scafe DS, Scafe CR, McKinney AJ, Tanouye MA. 2002. Genome-Wide Analysis of the
1355 Odorant-Binding Protein Gene Family in Drosophila melanogaster. *Genome Res*
1356 **12**:1357–1369. doi:10.1101/gr.239402
1357 Hoffman P. 2022. SeuratDisk: Interfaces for HDF5-Based Single Cell File Formats.
1358 James D, Hornik K. 2022. chron: Chronological Objects which can Handle Dates and Times.
1359 Jenett A, Rubin GM, Ngo T-TB, Shepherd D, Murphy C, Dionne H, Pfeiffer BD, Cavallaro A, Hall
1360 D, Jeter J, Iyer N, Fetter D, Hausenfluck JH, Peng H, Trautman ET, Svirskas RR, Myers
1361 EW, Iwinski ZR, Aso Y, DePasquale GM, Enos A, Hulamm P, Lam SCB, Li H-H, Laverty
1362 TR, Long F, Qu L, Murphy SD, Rokicki K, Safford T, Shaw K, Simpson JH, Sowell A,
1363 Tae S, Yu Y, Zugates CT. 2012. A GAL4-Driver Line Resource for Drosophila
1364 Neurobiology. *Cell Rep* **2**:991–1001. doi:10.1016/j.celrep.2012.09.011
1365 Jeong YT, Shim J, Oh SR, Yoon HI, Kim CH, Moon SJ, Montell C. 2013. An Odorant Binding
1366 Protein required for suppression of sweet taste by bitter chemicals. *Neuron* **79**:725–737.
1367 doi:10.1016/j.neuron.2013.06.025
1368 Johnstun JA, Shankar V, Mokashi SS, Sunkara LT, Ihearahu UE, Lyman RL, Mackay TFC,
1369 Anholt RRH. 2021. Functional Diversification, Redundancy, and Epistasis among
1370 Paralogs of the Drosophila melanogaster Obp50a-d Gene Cluster. *Mol Biol Evol*
1371 **38**:2030–2044. doi:10.1093/molbev/msab004

1372 Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A,
1373 Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez
1374 R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification.
1375 *Bioinformatics* **30**:1236–1240. doi:10.1093/bioinformatics/btu031
1376 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates
1377 R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A,
1378 Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy
1379 E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D,
1380 Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate
1381 protein structure prediction with AlphaFold. *Nature* **596**:583–589. doi:10.1038/s41586-
1382 021-03819-2
1383 Kalb JM, DiBenedetto AJ, Wolfner MF. 1993. Probing the function of Drosophila melanogaster
1384 accessory glands by directed cell ablation. *Proc Natl Acad Sci U S A* **90**:8093–8097.
1385 Kane NS, Vora M, Varre KJ, Padgett RW. 2017. Efficient Screening of CRISPR/Cas9-Induced
1386 Events in Drosophila Using a Co-CRISPR Strategy. *G3 Bethesda Md* **7**:87–93.
1387 doi:10.1534/g3.116.036723
1388 Karr TL, Southern H, Rosenow MA, Gossmann TI, Snook RR. 2019. The Old and the New:
1389 Discovery Proteomics Identifies Putative Novel Seminal Fluid Proteins in Drosophila. *Mol*
1390 *Cell Proteomics MCP* **18**:S23–S33. doi:10.1074/mcp.RA118.001098
1391 Kelleher ES, Markow TA. 2009. Duplication, Selection and Gene Conversion in a Drosophila
1392 mojavensis Female Reproductive Protein Family. *Genetics* **181**:1451.
1393 doi:10.1534/genetics.108.099044
1394 Kelleher ES, Pennington JE. 2009. Protease Gene Duplication and Proteolytic Activity in
1395 Drosophila Female Reproductive Tracts. *Mol Biol Evol* **26**:2125–2134.
1396 doi:10.1093/molbev/msp121
1397 Kelleher ES, Watts TD, LaFlamme BA, Haynes PA, Markow TA. 2009. Proteomic analysis of
1398 Drosophila mojavensis male accessory glands suggests novel classes of seminal fluid
1399 proteins. *Insect Biochem Mol Biol* **39**:366–371. doi:10.1016/j.ibmb.2009.03.003
1400 Khallaf MA, Cui R, Weißflog J, Erdogmus M, Svatoš A, Dweck HKM, Valenzano DR, Hansson
1401 BS, Knaden M. 2021. Large-scale characterization of sex pheromone communication
1402 systems in Drosophila. *Nat Commun* **12**:4165. doi:10.1038/s41467-021-24395-z
1403 Kim M-S, Repp A, Smith DP. 1998. LUSH Odorant-Binding Protein Mediates Chemosensory
1404 Responses to Alcohols in Drosophila melanogaster. *Genetics* **150**:711–721.
1405 doi:10.1093/genetics/150.2.711

1406 Kondo S, Ueda R. 2013. Highly Improved Gene Targeting by Germline-Specific Cas9
1407 Expression in Drosophila. *Genetics* **195**:715–721. doi:10.1534/genetics.113.156737
1408 Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV. 2002. Selection in the evolution of gene
1409 duplications. *Genome Biol* **3**:research0008.1. doi:10.1186/gb-2002-3-2-research0008
1410 Kopp A, Barmina O, Hamilton AM, Higgins L, McIntyre LM, Jones CD. 2008. Evolution of Gene
1411 Expression in the Drosophila Olfactory System. *Mol Biol Evol* **25**:1081–1092.
1412 doi:10.1093/molbev/msn055
1413 Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW. 2006. Automated
1414 Phylogenetic Detection of Recombination Using a Genetic Algorithm. *Mol Biol Evol*
1415 **23**:1891–1901. doi:10.1093/molbev/msl051
1416 Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and
1417 user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*
1418 **35**:4453–4455. doi:10.1093/bioinformatics/btz305
1419 Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary
1420 Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**:1547–1549.
1421 doi:10.1093/molbev/msy096
1422 Kurtovic A, Widmer A, Dickson BJ. 2007. A single class of olfactory neurons mediates
1423 behavioural responses to a Drosophila sex pheromone. *Nature* **446**:542–546.
1424 doi:10.1038/nature05672
1425 LaFlamme BA, Ravi Ram KR, Wolfner MF. 2012. The Drosophila melanogaster Seminal Fluid
1426 Protease “Seminase” Regulates Proteolytic and Post-Mating Reproductive Processes.
1427 *PLOS Genet* **8**:e1002435. doi:10.1371/journal.pgen.1002435
1428 Laturney M, Billeter J-C. 2016. Drosophila melanogaster females restore their attractiveness
1429 after mating by removing male anti-aphrodisiac pheromones. *Nat Commun* **7**:12322.
1430 doi:10.1038/ncomms12322
1431 Laughlin JD, Ha TS, Jones DNM, Smith DP. 2008. Activation of Pheromone-Sensitive Neurons
1432 Is Mediated by Conformational Activation of Pheromone-Binding Protein. *Cell* **133**:1255–
1433 1265. doi:10.1016/j.cell.2008.04.046
1434 Lee T, Luo L. 1999. Mosaic Analysis with a Repressible Cell Marker for Studies of Gene
1435 Function in Neuronal Morphogenesis. *Neuron* **22**:451–461. doi:10.1016/S0896-
1436 6273(00)80701-1
1437 Lenth RV, Buerkner P, Giné-Vázquez I, Herve M, Jung M, Love J, Miguez F, Riebl H, Singmann
1438 H. 2022. emmeans: Estimated Marginal Means, aka Least-Squares Means.

1439 Li H, Janssens J, De Waegeneer M, Kolluru SS, Davie K, Gardeux V, Saelens W, David FPA,
1440 Brbić M, Spanier K, Leskovec J, McLaughlin CN, Xie Q, Jones RC, Brueckner K, Shim J,
1441 Tattikota SG, Schnorrer F, Rust K, Nystul TG, Carvalho-Santos Z, Ribeiro C, Pal S,
1442 Mahadevaraju S, Przytycka TM, Allen AM, Goodwin SF, Berry CW, Fuller MT, White-
1443 Cooper H, Matunis EL, DiNardo S, Galenza A, O'Brien LE, Dow JAT, FCA Consortium§,
1444 Jasper H, Oliver B, Perrimon N, Deplancke B, Quake SR, Luo L, Aerts S, Agarwal D,
1445 Ahmed-Braimah Y, Arbeitman M, Ariss MM, Augsburger J, Ayush K, Baker CC, Banisch
1446 T, Birker K, Bodmer R, Bolival B, Brantley SE, Brill JA, Brown NC, Buehner NA, Cai XT,
1447 Cardoso-Figueiredo R, Casares F, Chang A, Clandinin TR, Crasta S, Desplan C,
1448 Detweiler AM, Dhakan DB, Donà E, Engert S, Floc'hlay S, George N, González-Segarra
1449 AJ, Groves AK, Gumbin S, Guo Y, Harris DE, Heifetz Y, Holtz SL, Horns F, Hudry B,
1450 Hung R-J, Jan YN, Jaszczak JS, Jefferis GSXE, Karkanias J, Karr TL, Katheder NS,
1451 Kezos J, Kim AA, Kim SK, Kockel L, Konstantinides N, Kornberg TB, Krause HM, Labott
1452 AT, Laturney M, Lehmann R, Leinwand S, Li J, Li JSS, Li Kai, Li Ke, Li L, Li T,
1453 Litovchenko M, Liu H-H, Liu Y, Lu T-C, Manning J, Mase A, Matera-Vatnick M, Matias
1454 NR, McDonough-Goldstein CE, McGeever A, McLachlan AD, Moreno-Roman P, Neff N,
1455 Neville M, Ngo S, Nielsen T, O'Brien CE, Osumi-Sutherland D, Özel MN, Papatheodorou
1456 I, Petkovic M, Pilgrim C, Pisco AO, Reisenman C, Sanders EN, Dos Santos G, Scott K,
1457 Sherlekar A, Shiu P, Sims D, Sit RV, Slaidina M, Smith HE, Sterne G, Su Y-H, Sutton D,
1458 Tamayo M, Tan M, Tastekin I, Treiber C, Vacek D, Vogler G, Waddell S, Wang W,
1459 Wilson RI, Wolfner MF, Wong Y-CE, Xie A, Xu J, Yamamoto S, Yan J, Yao Z, Yoda K,
1460 Zhu R, Zinzen RP. 2022. Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult
1461 fruit fly. *Science* **375**:eabk2432. doi:10.1126/science.abk2432
1462 Liu H, Kubli E. 2003. Sex-peptide is the molecular basis of the sperm effect in Drosophila
1463 melanogaster. *Proc Natl Acad Sci U S A* **100**:9929–9933. doi:10.1073/pnas.1631700100
1464 Lung O, Wolfner MF. 2001. Identification and characterization of the major Drosophila
1465 melanogaster mating plug protein. *Insect Biochem Mol Biol* **31**:543–551.
1466 doi:10.1016/s0965-1748(00)00154-5
1467 Majane AC, Cridland JM, Begun DJ. 2022. Single-nucleus transcriptomes reveal evolutionary
1468 and functional properties of cell types in the Drosophila accessory gland. *Genetics*
1469 **220**:iyab213. doi:10.1093/genetics/iyab213
1470 Manier MK, Belote JM, Berben KS, Novikov D, Stuart WT, Pitnick S. 2010. Resolving
1471 mechanisms of competitive fertilization success in Drosophila melanogaster. *Science*
1472 **328**:354–357. doi:10.1126/science.1187096

1473 Manning A. 1967. The control of sexual receptivity in female Drosophila. *Anim Behav* **15**:239–
1474 250. doi:10.1016/0003-3472(67)90006-1
1475 Markow TA, Ankney PF. 1988. Insemination Reaction in Drosophila: Found in Species Whose
1476 Males Contribute Material to Oocytes Before Fertilization. *Evolution* **42**:1097–1101.
1477 doi:10.2307/2408926
1478 Mastrogiacomo R, D′Ambrosio C, Niccolini A, Serra A, Gazzano A, Scaloni A, Pelosi P. 2014.
1479 An Odorant-Binding Protein Is Abundantly Expressed in the Nose and in the Seminal
1480 Fluid of the Rabbit. *PLOS ONE* **9**:e111932. doi:10.1371/journal.pone.0111932
1481 Matsuo T. 2008. Rapid Evolution of Two Odorant-Binding Protein Genes, Obp57d and Obp57e,
1482 in the Drosophila melanogaster Species Group. *Genetics* **178**:1061–1072.
1483 doi:10.1534/genetics.107.079046
1484 Matsuo T, Sugaya S, Yasukawa J, Aigaki T, Fuyama Y. 2007. Odorant-binding proteins
1485 OBP57d and OBP57e affect taste perception and host-plant preference in Drosophila
1486 sechellia. *PLoS Biol* **5**:e118. doi:10.1371/journal.pbio.0050118
1487 McDonough-Goldstein CE, Pitnick S, Dorus S. 2022. Drosophila female reproductive glands
1488 contribute to mating plug composition and the timing of sperm ejection. *Proc R Soc B*
1489 *Biol Sci* **289**:20212213. doi:10.1098/rspb.2021.2213
1490 McGeary MK, Findlay GD. 2020. Molecular evolution of the sex peptide network in Drosophila. *J*
1491 *Evol Biol* **33**:629–641. doi:10.1111/jeb.13597
1492 Misra S, Buehner NA, Singh A, Wolfner MF. 2022. Female factors modulate Sex Peptide's
1493 association with sperm in Drosophila melanogaster. *BMC Biol* **20**:279.
1494 doi:10.1186/s12915-022-01465-2
1495 Misra S, Wolfner MF. 2020. Drosophila seminal sex peptide associates with rival as well as own
1496 sperm, providing SP function in polyandrous females. *eLife* **9**:e58322.
1497 doi:10.7554/eLife.58322
1498 Mueller JL, Ram KR, McGraw LA, Bloch Qazi MC, Siggia ED, Clark AG, Aquadro CF, Wolfner
1499 MF. 2005. Cross-Species Comparison of Drosophila Male Accessory Gland Protein
1500 Genes. *Genetics* **171**:131–143. doi:10.1534/genetics.105.043844
1501 Mueller JL, Ripoll DR, Aquadro CF, Wolfner MF. 2004. Comparative structural modeling and
1502 inference of conserved protein classes in Drosophila seminal fluid. *Proc Natl Acad Sci*
1503 **101**:13542–13547. doi:10.1073/pnas.0405579101
1504 Neubaum DM, Wolfner MF. 1999. Mated Drosophila melanogaster females require a seminal
1505 fluid protein, Acp36DE, to store sperm efficiently. *Genetics* **153**:845–857.

Ng SH, Shankar S, Shikichi Y, Akasaka K, Mori K, Yew JY. 2014. Pheromone evolution and sexual behavior in Drosophila are shaped by male sensory exploitation of other males. *Proc Natl Acad Sci U S A* **111**:3056–3061. doi:10.1073/pnas.1313615111

Ng WC, Chin JSR, Tan KJ, Yew JY. 2015. The fatty acid elongase Bond is essential for Drosophila sex pheromone synthesis and male fertility. *Nat Commun* **6**:8263. doi:10.1038/ncomms9263

Ohno S. 1970. Evolution by Gene Duplication.

Pal S, Oliver B, Przytycka TM. 2022. Stochastic Modeling of Gene Expression Evolution Uncovers Tissue- and Sex-Specific Properties of Expression Evolution in the Drosophila Genus. *J Comput Biol J Comput Mol Cell Biol*. doi:10.1089/cmb.2022.0121

Patlar B, Jayaswal V, Ranz JM, Civetta A. 2021. Nonadaptive molecular evolution of seminal fluid proteins in Drosophila. *Evol Int J Org Evol* **75**:2102–2113. doi:10.1111/evo.14297

Patterson JT. 1946. A New Type of Isolating Mechanism in Drosophila. *Proc Natl Acad Sci U S A* **32**:202–208. doi:10.1073/pnas.32.7.202

Peng J, Chen S, Büsser S, Liu H, Honegger T, Kubli E. 2005. Gradual Release of Sperm Bound Sex-Peptide Controls Female Postmating Behavior in Drosophila. *Curr Biol* **15**:207–213. doi:10.1016/j.cub.2005.01.034

Poe AR, Wang B, Sapar ML, Ji H, Li K, Onabajo T, Fazliyeva R, Gibbs M, Qiu Y, Hu Y, Han C. 2019. Robust CRISPR/Cas9-Mediated Tissue-Specific Mutagenesis Reveals Gene Redundancy and Perdurance in Drosophila. *Genetics* **211**:459–472. doi:10.1534/genetics.118.301736

Ravi Ram KR, Wolfner MF. 2009. A network of interactions among seminal proteins underlies the long-term postmating response in Drosophila. *Proc Natl Acad Sci* **106**:15384–15389. doi:10.1073/pnas.0902923106

Raz AA, Vida GS, Stern SR, Mahadevaraju S, Fingerhut JM, Viveiros JM, Pal S, Grey JR, Grace MR, Berry CW, Li H, Janssens J, Saelens W, Shao Z, Hun C, Yamashita YM, Przytycka TM, Oliver B, Brill JA, Krause HM, Matunis EL, White-Cooper H, DiNardo S, Fuller MT. 2022. Emergent dynamics of adult stem cell lineages from single nucleus and single cell RNA-Seq of Drosophila testes. *bioRxiv* 2022.07.26.501581. doi:10.1101/2022.07.26.501581

Raza Q, Choi JY, Li Y, O'Dowd RM, Watkins SC, Chikina M, Hong Y, Clark NL, Kwiatkowski AV. 2019. Evolutionary rate covariation analysis of E-cadherin identifies Raskol as a regulator of cell adhesion and actin dynamics in Drosophila. *PLoS Genet* **15**:e1007720. doi:10.1371/journal.pgen.1007720

1540 Rihani K, Ferveur J-F, Briand L. 2021. The 40-Year Mystery of Insect Odorant-Binding Proteins.
1541 *Biomolecules* **11**:509. doi:10.3390/biom11040509
1542 Rondón JJ, Moreyra NN, Pisarenco VA, Rozas J, Hurtado J, Hasson E. 2022. Evolution of the
1543 odorant-binding protein gene family in Drosophila. *Front Ecol Evol* **10**.
1544 Satija R, Farrell JA, Gennert D, Schier AF, Regev A. 2015. Spatial reconstruction of single-cell
1545 gene expression data. *Nat Biotechnol* **33**:495–502. doi:10.1038/nbt.3192
1546 Savini G, Scolari F, Ometto L, Rota-Stabelli O, Carraretto D, Gomulski LM, Gasperi G, Abd-Alla
1547 AMM, Aksoy S, Attardo GM, Malacrida AR. 2021. Viviparity and habitat restrictions may
1548 influence the evolution of male reproductive genes in tsetse fly (Glossina) species. *BMC*
1549 *Biol* **19**:211. doi:10.1186/s12915-021-01148-4
1550 Scott D. 1986. Sexual mimicry regulates the attractiveness of mated Drosophila melanogaster
1551 females. *Proc Natl Acad Sci U S A* **83**:8429–8433. doi:10.1073/pnas.83.21.8429
1552 Sepil I, Br H, R D, Ml T, Pd C, R K, R F, Bm K, S W. 2019. Quantitative Proteomics
1553 Identification of Seminal Fluid Proteins in Male Drosophila melanogaster. *Mol Cell*
1554 *Proteomics MCP* **18**. doi:10.1074/mcp.RA118.000831
1555 Shorter JR, Dembeck LM, Everett LJ, Morozova TV, Arya GH, Turlapati L, St. Armour GE,
1556 Schal C, Mackay TFC, Anholt RRH. 2016. Obp56h Modulates Mating Behavior in
1557 Drosophila melanogaster. *G3 GenesGenomesGenetics* **6**:3335–3342.
1558 doi:10.1534/g3.116.034595
1559 Singer AG, Macrides F, Clancy AN, Agosta WC. 1986. Purification and analysis of a
1560 proteinaceous aphrodisiac pheromone from hamster vaginal discharge. *J Biol Chem*
1561 **261**:13323–13326.
1562 Singh A, Buehner NA, Lin H, Baranowski KJ, Findlay GD, Wolfner MF. 2018. Long-term
1563 interaction between Drosophila sperm and sex peptide is mediated by other seminal
1564 proteins that bind only transiently to sperm. *Insect Biochem Mol Biol* **102**:43–51.
1565 doi:10.1016/j.ibmb.2018.09.004
1566 Sirot LK, Findlay GD, Sitnik JL, Frasheri D, Avila FW, Wolfner MF. 2014. Molecular
1567 Characterization and Evolution of a Gene Family Encoding Both Female- and Male-
1568 Specific Reproductive Proteins in Drosophila. *Mol Biol Evol* **31**:1554–1567.
1569 doi:10.1093/molbev/msu114
1570 Sirot LK, Poulson RL, McKenna MC, Girnary H, Wolfner MF, Harrington LC. 2008. Identity and
1571 transfer of male reproductive gland proteins of the dengue vector mosquito, Aedes
1572 aegypti: potential tools for control of female feeding and reproduction. *Insect Biochem*
1573 *Mol Biol* **38**:176. doi:10.1016/j.ibmb.2007.10.007

1574 Sirot LK, Wong A, Chapman T, Wolfner MF. 2015. Sexual Conflict and Seminal Fluid Proteins:
1575 A Dynamic Landscape of Sexual Interactions. *Cold Spring Harb Perspect Biol*
1576 **7**:a017533. doi:10.1101/cshperspect.a017533
1577 Sun JS, Xiao S, Carlson JR. 2018. The diverse small proteins called odorant-binding proteins.
1578 *Open Biol* **8**:180208. doi:10.1098/rsob.180208
1579 Sun Y-L, Huang L-Q, Pelosi P, Wang C-Z. 2012. Expression in Antennae and Reproductive
1580 Organs Suggests a Dual Role of an Odorant-Binding Protein in Two Sibling Helicoverpa
1581 Species. *PLOS ONE* **7**:e30040. doi:10.1371/journal.pone.0030040
1582 Swanson WJ, Clark AG, Waldrip-Dail HM, Wolfner MF, Aquadro CF. 2001. Evolutionary EST
1583 analysis identifies rapidly evolving male reproductive proteins in Drosophila. *Proc Natl*
1584 *Acad Sci U S A* **98**:7375–7379. doi:10.1073/pnas.131568198
1585 Swanson WJ, Vacquier VD. 2002. The rapid evolution of reproductive proteins. *Nat Rev Genet*
1586 **3**:137–144. doi:10.1038/nrg733
1587 Takemori N, Yamamoto M-T. 2009. Proteome mapping of the Drosophila melanogaster male
1588 reproductive system. *Proteomics* **9**:2484–2493. doi:10.1002/pmic.200800795
1589 Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis
1590 Version 11. *Mol Biol Evol* **38**:3022–3027. doi:10.1093/molbev/msab120
1591 Tsuda M, Peyre J-B, Asano T, Aigaki T. 2015. Visualizing Molecular Functions and Cross-
1592 Species Activity of Sex-Peptide in Drosophila. *Genetics* **200**:1161–1169.
1593 doi:10.1534/genetics.115.177550
1594 Vieira FG, Rozas J. 2011. Comparative Genomics of the Odorant-Binding and Chemosensory
1595 Protein Gene Families across the Arthropoda: Origin and Evolutionary History of the
1596 Chemosensory System. *Genome Biol Evol* **3**:476–490. doi:10.1093/gbe/evr033
1597 Vieira FG, Sánchez-Gracia A, Rozas J. 2007. Comparative genomic analysis of the odorant-
1598 binding protein family in 12 Drosophila genomes: purifying selection and birth-and-death
1599 evolution. *Genome Biol* **8**:R235. doi:10.1186/gb-2007-8-11-r235
1600 Wagstaff BJ, Begun DJ. 2005. Comparative Genomics of Accessory Gland Protein Genes in
1601 Drosophila melanogaster and D. pseudoobscura. *Mol Biol Evol* **22**:818–832.
1602 doi:10.1093/molbev/msi067
1603 Wigby S, Brown NC, Allen SE, Misra S, Sitnik JL, Sepil I, Clark AG, Wolfner MF. 2020. The
1604 Drosophila seminal proteome and its role in postcopulatory sexual selection. *Philos*
1605 *Trans R Soc B Biol Sci* **375**:20200072. doi:10.1098/rstb.2020.0072
1606 Xiao S, Sun JS, Carlson JR. 2019. Robust olfactory responses in the absence of odorant
1607 binding proteins. *eLife* **8**:e51040. doi:10.7554/eLife.51040

Xu J, Baulding J, Palli SR. 2013. Proteomics of Tribolium castaneum seminal fluid proteins: identification of an angiotensin-converting enzyme as a key player in regulation of reproduction. *J Proteomics* **78**:83–93. doi:10.1016/j.jprot.2012.11.011

Xu P, Atkinson R, Jones DNM, Smith DP. 2005. Drosophila OBP LUSH Is Required for Activity of Pheromone-Sensitive Neurons. *Neuron* **45**:193–200. doi:10.1016/j.neuron.2004.12.031

Yang H, Jaime M, Polihronakis M, Kanegawa K, Markow T, Kaneshiro K, Oliver B. 2018. Re-annotation of eight Drosophila genomes. *Life Sci Alliance* **1**. doi:10.26508/lsa.201800156

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol* **24**:1586–1591. doi:10.1093/molbev/msm088

Yasukawa J, Tomioka S, Aigaki T, Matsuo T. 2010. Evolution of expression patterns of two odorant-binding protein genes, Obp57d and Obp57e, in Drosophila. *Gene* **467**:25–34. doi:10.1016/j.gene.2010.07.006

Yew JY, Dreisewerd K, Luftmann H, Müthing J, Pohlentz G, Kravitz EA. 2009. A new male sex pheromone and novel cuticular cues for chemical communication in Drosophila. *Curr Biol CB* **19**:1245–1254. doi:10.1016/j.cub.2009.06.037