1 **TITLE**
2 Regional specialization, polyploidy, and seminal fluid transcripts in the Drosophila female
3 reproductive tract
4
5 **AUTHORS**
6 Rachel C. Thayer[1], Elizabeth S. Polston[1], Jixiang Xu[1], and David J. Begun[1]
7 1 Department of Evolution and Ecology, University of California, Davis, CA, USA
8 Corresponding email: thayerr@berkeley.edu
9
10 **KEYWORDS**
11 spermatheca, seminal receptacle, gene expression, marker gene, seminal fluid protein, uterus,
12 oviduct, parovaria, accessory gland, secretion, endoreplication
13
14 **ABSTRACT**
15 Internal fertilization requires the choreographed interaction of female cells and molecules with
16 seminal fluid and sperm. In many animals, including insects, the female reproductive tract is
17 physically subdivided into sections that carry out specialized functions. For example, females of
18 many species have specialized organs for sperm storage. *Drosophila melanogaster* is a premier
19 model system for investigating many aspects of animal reproduction. Nevertheless, in contrast
20 to males, much of the basic biology of the *D. melanogaster* female reproductive tract remains
21 poorly understood or completely unknown. Here we use single-cell RNA-seq data and in situ
22 hybridization to reveal a rich and previously unknown female reproductive tract cell diversity,
23 including widespread variation in ploidy levels. We find that many so-called seminal fluid protein
24 genes appear to be transcribed in specialized cells of the female reproductive tract, motivating a
25 re-evaluation of the functional and evolutionary biology of this major class of proteins.
26
27
28 **INTRODUCTION**
29
30 The *Drosophila* female reproductive tract is morphologically and functionally complex. Below the
31 ovaries, the lower reproductive tract (FRT) is composed of 5 morphologically distinct organs
32 (Fig. 1A-B) with various critical functions. The uterus (UT) is a muscular organ that changes
33 conformation to transit eggs and allow sperm to access the storage organs (Mattei et al. 2015;
34 Adams and Wolfner 2007). To maneuver sperm and eggs, the uterine wall has functionally
35 distinctive regions. These include the oviduct valve flap (OVF), which is closed prior to mating,
36 and which can block access to the oviduct and sperm storage organs, and the specialized
37 vaginal intima (SVI), a region in the lower uterine wall. The uterus is also the site of fertilization
38 and mating plug formation, and is the first tissue contacted by male products during copulation.
39 The oviduct (OV) transfers eggs into the uterus in response to octopaminergic neural signaling
40 (Rubinstein and Wolfner 2013). Sperm from separate males compete for occupancy of the
41 seminal receptacle (SR), which stores the sperm set primarily used for fertilization (Manier et al.
42 2010). The seminal receptacle is a long, coiled, tubular organ and has two morphologically
43 distinct regions: a narrow proximal section and a contrastingly thicker distal section, which also
44 differ in the arrangement of microvilli on the inner surface, the density of secretory vacuoles,

45   and the postures adopted by occupying sperm within each section (Y. Heifetz and Rivlin 2010;
46   Mayhew and Merritt 2013). The paired spermathecae (ST) are sclerotized, glandular capsules
47   that also store sperm. The spermathecae are capped by secretory cells that are required for
48   fertilization and ovulation (Schnakenberg, Matias, and Siegal 2011; Sun and Spradling 2013),
49   as are the secretions of the bilaterally paired parovaria (PV, (Sun and Spradling 2012), also
50   known as the female accessory glands. A reproductive-associated fat body (FB) is found
51   agglomerated over the 4 glands (Fig. 2C).
52
53   In addition to its core functions managing sperm and eggs, the FRT is the site of high-stakes
54   interactions with male products, specifically seminal fluid proteins (SFPs). SFPs are proteins
55   that males transfer to females during copulation; there are >290 in *D. melanogaster* (Sepil et al.
56   2019). Upon recognizing the receipt of male products, the FRT becomes the initiation site for a
57   cascade of systemic post-mating responses: increased activity and reduced sleep (Isaac et al.
58   2010), altered neural gene expression (Yael Heifetz et al. 2014), reduced receptivity to remating
59   (Chen et al. 1988), increased aggression (Bath et al. 2017), improved learning (Scheunemann
60   et al. 2019), reduced immune capacity despite systemic upregulation of immunity genes
61   (Fedorka et al. 2007), and even reduced lifespan (Chapman et al. 1995). Taken together, it is
62   clear that events in the lower FRT, including signal transduction in response to mating, can have
63   profound, systemic consequences for female biology.
64
65   Despite model organism status, only six gene expression surveys of the female *D.*
66   *melanogaster* somatic reproductive tract are reported (Prokupek et al. 2009; Leader et al. 2018;
67   Mack et al. 2006; McDonough-Goldstein, Borziak, et al. 2021; Allen and Spradling 2008;
68   Robinson et al. 2013). Of these, four used old methods that are incomplete or biased and five of
69   these studies either aggregated all somatic reproductive tissues or included only the
70   spermatheca. Only one unbiased transcriptomic analysis of each lower reproductive tract tissue
71   exists (McDonough-Goldstein, Borziak, et al. 2021). Recent single-cell transcriptome
72   investigation of D. melanogaster by the Fly Cell Atlas consortium excluded the somatic female
73   reproductive tract (Li et al. 2022), leaving gene expression at cellular resolution wholly
74   unexplored in these tissues. This state of affairs stands in contrast to the level of research
75   investment in male somatic reproductive tissue–the accessory glands–which have been
76   characterized in every generation of the Fly Atlas project (Chintapalli, Wang, and Dow 2007;
77   Leader et al. 2018) and in many independent transcriptomic investigations. Given the similarities
78   between fly and mammal reproductive tracts, such as their highly secretory nature and shared
79   enrichments for functional classes of expressed products (e.g. serine proteases, (Muytjens, Yu,
80   and Diamandis 2018; Lawniczak and Begun 2007; Kelleher and Pennington 2009), the lack of
81   FRT data contributes to the general research shortfall in female subjects in the biomedical
82   literature (Beery and Zucker 2011; Woitowich, Beery, and Woodruff 2020; Orr et al. 2020). In
83   direct consequence, marker genes for key female organs and cell types are generally unknown
84   and functional tools such as GAL4 drivers are lacking, leaving outstanding unanswered
85   questions about female reproduction. To address this gap in the literature, we characterized
86   gene expression and cell type diversity in the *Drosophila* female reproductive tract, using single-
87   nucleus RNA-sequencing and fluorescent *in situ* hybridization chain reaction.
88

METHODS

Fly strains and husbandry: We sampled unmated, 2-3 day old female flies from 30 isofemale lines that were established from gravid females collected from Fairfield, Maine (September, 2011) and Panama City, Panama (January 2012) (Zhao et al. 2015). Flies were housed on standard yeast-cornmeal-agar food at 25° C on a 12-hour light:dark cycle.

Nuclei isolation and sequencing: Whole lower reproductive tracts were dissected in chilled Schneider's insect medium. Five reproductive tracts per each of 15 isofemale lines (i.e. 15 genotypes; 75 animals) from Maine were pooled and processed together. The same sampling strategy was followed separately with 15 genotypes from Panama, resulting in two sequencing libraries that together included 150 flies from 30 genotypes. Nuclei were isolated using a modified protocol from (Martelotto n.d.; Majane, Cridland, and Begun 2022). Briefly, tissues were dounced in lysis buffer, passed through a sterile filter, and centrifuged. The clarified preparation was stained with 10 µg/ml of DAPI prior to FACS analysis and selection by flow cytometry technical personnel using the Beckman Coulter "Astrios" cell sorter. The Astrios was pre-configured for rapid, chilled processing of the nuclei preparations using the 70 µm nozzle at 60 psi fluid pressure. The total cellular preparation was assessed for laser light scatter (forward vs. side angle scatter) and gated based on laser scatter to exclude debris and aggregates. Putative single nuclei were assessed for DAPI intensity following illumination with a 405 nm laser and detection in a 450/50 restricted photodetector (Supp figures x). The majority of the sample (59%) exhibited little or no fluorescence (DAPI negative, median fluorescent intensity [MFI] = 39) or low fluorescence (DAPI low, 33.7%, MFI = 571). The DAPI bright positive nuclei (fluorescent intensity > 5000), which comprised ~4.8% of the total cellular preparation, were collected for subsequent genomic analysis. A total of at least 100,000 DAPI+ nuclei were collected per sample. Barcoded 3' single nucleus libraries were prepared with the Chromium Next GEM Single Cell 3' kit v3.1 (10X Genomics, Pleasanton, California) according to the manufacturer recommendations. Library quality was assessed with a Bioanalyzer 2100 (Agilent, Santa Clara, CA). The libraries were sequenced on a NovaSeq 6000 sequencer (Illumina, San Diego, CA) with paired-end 150 bp reads. The sequencing generated approximately 12,671 / 13,554 reads per cell and 128 / 220 million reads per library for Maine and Panama respectively. The Maine and Panama libraries recovered 10,073 and 16,207 nuclei, respectively.

RNAseq data preparation: Alignment, barcode removal, and UMI counting were done in CellRanger (6.1.2) with the "count" command. Reads were aligned to version 6.41 of the *D. melanogaster* genome (FlyBase, downloaded August 9, 2021), with an index built using CellRanger "mkref." To the FlyBase *D. melanogaster* 6.41 GTF file, we added *de novo* gene annotations from (Cridland et al. 2022) and de novo and newly assembled gene annotation models from (Lombardo et al. 2023). Because the cellranger pipeline does not accept annotations with unknown strandedness, we edited the annotations for all single exon de novo genes to be on the '+' strand. We then used CellRanger "mkgtf" to remove non-polyA transcripts that overlapped with protein-coding gene models.

RNAseq data reduction: We used R 4.1.2 to curate and analyze the dataset. SoupX v1.5.2 (Young and Behjati 2020) was used with default parameters to remove cell-free mRNA contamination. Using Seurat v4.3.0 (Satija et al. 2015), we applied filters to remove genes detected in fewer than 3 nuclei, nuclei with >30,000 RNAs, nuclei with >275 genes, and nuclei

133   with >2% mitochondrial gene expression. Because we recovered more nuclei than the 10,000
134   maximum that 10x Genomics recommends per sequencing lane, crowding and doublets were
135   likely–an especially devilish problem given that little information that could be used to guide
136   clustering was previously available (e.g. how many cell types are present; validated marker
137   genes). To remove doublets, following preliminary clustering in Seurat, we used doubletfinder
138   (McGinnis, Murrow, and Gartner 2019), a high-performing tool (Xi and Li 2021), according to
139   default settings and assuming a 7-8% percent doublet rate in the ME and PAN libraries,
140   respectively, per the 10x Genomics user guide. However, doubletfinder heavily penalized
141   polyploid cell types, to such an extent that accepting all of doubletfinder's designations would
142   result in complete loss of the PSC and SSC clusters. Therefore, we did not remove putative
143   doublets from the two clusters corresponding to the polyploid spermathecal and parovaria
144   secretory cells. The polyploid fat cell cluster was also heavily penalized (about 18% of cells
145   flagged by doubletfinder). Though this aggressive pruning may also include false positive
146   doublet identifications, it was not so severe that the cluster, nor its signatures of polyploidy
147   including high average RNA counts, were lost. To retain a conservative set of high-confidence
148   singlets, we removed all putative doublets from the fat body cluster.
149   RNAseq clustering and data analysis: Prior to clustering, singlets from the ME and PAN libraries
150   were combined using "merge" in Seurat. Data were normalized and scaled using the following
151   functions in Seurat: "NormalizeData", "FindVariableFeatures" set to 12,300 features, and
152   "ScaleData" with the argument to regress out variation in the percentage of mitochondrial reads.
153   Nuclei were clustered using "FindNeighbors," "FindClusters", and "FindUMAP" with principal
154   components 1 to 75. We used "clustree" (Zappia and Oshlack 2018) to evaluate cluster stability
155   across resolution settings from 0.5 to 4 before proceeding with a resolution of 1.5. Cluster
156   similarity was evaluated using the Seurat function "BuildClusterTree" and principal components
157   1 to 75. Genes were called as 'expressed' using two different thresholds. For analyses focused
158   on determining the typical attributes of a cluster (GO enrichment analysis), we called a gene
159   'expressed' if it had >0 logCPM in >50% of cells per cluster. This strict threshold conservatively
160   includes genes whose expression we can be confident is typical for the cluster. However we
161   observed that, perhaps owing to the gene dropout challenges inherent to single cell sequencing,
162   this threshold was too strict to detect some externally validated expressed genes. Antibody
163   staining (Rezával et al. 2012) indicates that all FRT neurons express *ppk*, while *fru* and *dsx*
164   expression can be used to identify subsets of FRT neurons. In our dataset, the neuron cluster
165   robustly expresses both *fru* and *dsx*, but *ppk* is only detectable in 2/34 cells (just under 6% of
166   cells). Therefore, for analyses focused on gene discovery (e.g. to report the total set of
167   expressed genes / SFPs), we applied an empirically-calibrated threshold wherein a gene is
168   considered 'expressed' if it is detected at least as strongly as was *ppk* in the neuron cluster;
169   specifically, >1 logCPM in >6% of cells per cluster.
170   HCR in situ hybridization: Lower reproductive tracts were dissected from flies with matched
171   genotypes, ages, and mating status to the sequenced specimens. Tissues were fixed with 3.2%
172   paraformaldehyde for 40 minutes, methanol dehydrated, and stored at -20°C until the
173   hybridization reaction. Prospective gene targets were identified using the Seurat
174   "FindAllMarkers" and "FindMarkers" functions, and prioritized using the Seurat visualization
175   tools "VlnPlot" and "DotPlot" (Supplement) along with practical considerations–transcript length
176   and an absence of overlapping gene annotations. Given the large number of putative cell types

177   to be annotated in a thick, whole-mount tissue with melanized and chitinous regions prone to
178   autofluorescence, we used a high-throughput approach with multiplexed probes and a split-
179   initiator scheme that minimizes background fluorescence (Choi et al. 2018). Probes were
180   designed using previously published software (Kuehn et al. 2022) implemented with
181   (https://github.com/rwnull/HCRProbeMakerCL) with the following settings: 0 basepair 5' offset, 5
182   maximum A/T homopolymer length, 4 maximum C/G homopolymer length, between 35-55% CG
183   bases in probe, BLAST against v 6.41 of the D. melanogaster reference genome, and a
184   maximum of 40 probe pairs. For short genes that otherwise would not accommodate enough
185   total probe pairs (*CG6055, CG7443, CG17108, CG17239,* and *red*) we relaxed the limits to 31-
186   61% CG probe content. We used transcript sequences from the Drosophila genome v. 6.41 and
187   substituted in SNP alleles that occur at greater than 50% frequencies in the east coast source
188   populations per (Svetec et al. 2016; Reinhardt et al. 2014). Probes were preferentially designed
189   against only the CDS unless CDS length was too short. Probes were ordered from Integrated
190   DNA Technologies, Inc., (Coralville, Iowa, USA) in three-gene oligo pools, where the probe
191   sequences for three genes were designed to each use different initiation sequences and were
192   ordered pooled in a single tube. Probe sequences are listed in (Supp Table 1). Successful
193   probe sets had from 8-40 probe pairs per gene. We used three amplifier hairpin designs from
194   Molecular Instruments Inc., (Los Angeles, CA, USA), specifically the B1 initiation sequence
195   conjugated to AlexaFluor 647, B2= AlexaFluor488, and B4 =AlexaFluor 546. Amplifier hairpin
196   and initiation sequences were developed by (Choi, Beck, and Pierce 2014). Hybridization
197   reactions followed the (Bruce et al. 2021) protocol, using 3x probe concentrations, no sonication
198   step, a 200 uL reaction volume for hairpin amplification, and 20+ hour incubation times for the
199   hybridization and amplification steps. For small genes with ten or fewer probe pairs (*CG7443,*
200   *CG43112, EbpIII*), we doubled the concentration of the probe within the oligo pool during probe
201   design, resulting in an effective 6x probe concentration. In other words, we optimized the
202   protocol for bright signal in a high-level whole-tissue image, to be able to detect rare or disperse
203   cell types for which we had no a priori information regarding which region to focus on while
204   imaging. Negative controls received the same treatment, including the application of hairpins,
205   except that no probes were applied. We also used a second negative control strategy, using
206   previously validated probe sets against two genes, *wg* and *snail*, which were not expressed in
207   our single nuclei dataset (Supplement, 'Pool B' section).
208   Confocal microscopy and image processing: Samples were mounted in 50% glycerol / PBS and
209   imaged on a Zeiss Airyscan 980 in confocal mode, using the 355, 488, 561, and 639 nm lasers.
210   Consistent with our objective to locate cell types, rather than to precisely quantify signal, we
211   optimized laser power and gain settings to detect signal for each sample, allowing variation in
212   excitation between gene sets and replicates. Consequently, relative brightness may be
213   compared within but not among images in our dataset. We imaged negative controls with both
214   maximal and typical (mean and median) settings for laser power and gain (see full presentation
215   of negative controls in Supplement). Images were processed in ImageJ v2.3.0 as follows: for
216   each laser channel, Z-series were projected using the maximum intensity setting, and the
217   'Maximum' slider in the 'Brightness & Contrast' settings was adjusted to set the brightest pixels
218   at saturation, while viewing the image with the 'HiLo' lookup table. As with gain and laser power,
219   brightness was optimized separately for each sample and color channel, with maximal and
220   typical adjustments applied to the respective negative controls. All brightness adjustments were

221 uniformly applied to the entire image. Color channels were imported as layers to Adobe
222 Photoshop 2015.5.2 and merged using the 'screen' overlay setting.
223
224
225 **RESULTS**
226
227 **Nuclei clustering and annotation identify novel cell types**
228 To characterize gene expression among cell types in the Drosophila female somatic
229 reproductive tract (FRT, Fig. 1B), we generated two single-nucleus sequencing libraries from
230 unmated females. To ensure that our results, including putative cell type designations with their
231 marker genes and transcriptomic profiles, would be robust to potential genotypic variation, we
232 sampled from 30 wild-derived isofemale line genotypes from the North American east coast,
233 evenly represented in these libraries (Methods). Sequencing generated approximately 128/220
234 million reads per library, with an average of 12,671/13,554 reads per nucleus among the 10,073
235 /16,207 recovered nuclei per library. To discover female reproductive tract cell types we
236 clustered nuclei by gene expression principal components of variation, resulting in 37 putative
237 cell types (Fig. 1C-D).
238
239 We then set out to first determine which of our inferred clusters were likely to represent
240 previously described cell types in the species. Using the known marker *Send 1* for spermathecal
241 secretory cells (SSC, Fig. 2C (Schnakenberg, Matias, and Siegal 2011) and known markers *fru*,
242 *dsx*, and *ppk* for neurons (Rezával et al. 2012) we assigned clusters to these cell identities. We
243 then investigated which of our clusters might correspond to cell types previously identified in Fly
244 Cell Atlas (Li et al. 2022). To do so, we extracted marker genes for each cluster using Seurat's
245 "findallmarkers" function and then investigated their expression in Fly Cell Atlas, followed by
246 reciprocally checking top marker genes of potentially matching Fly Cell Atlas clusters in our
247 dataset. The following matched sets of marker genes and cell types in Fly Cell Atlas allowed us
248 to annotate multiple cell-type clusters in our dataset: oenocytes, *FASN2, FASN3, LpR1*;
249 hemocytes, *Hml, Ppn, Nimc1*; muscle, *bt, sls, up, Octalpha2R, CG44422, CG45076*; fat body,
250 *CG13315, CG4716, Ubx, apolpp*; sensory neuron annotation, *para, Rdl,* and *futsch*. Another
251 cluster, "?-1" (Fig. 1C), shared some marker genes (*CG8012, CG5162, wat*) with adult tracheal
252 cells, but its top two marker genes (*Antp*–see Supplement, and *CG15353*) were instead only
253 expressed in an unannotated cell type in the Fly Cell Atlas tracheal dissection dataset. Within
254 Fly Cell Atlas, this latter, unannotated cell type was distinguished from adult tracheal cells by the
255 expression of *wgn, pk,* and *gol,* which are unexpressed in Fly Cell Atlas tracheal cells and in our
256 cluster. While some type of tracheal identity is most likely for this cluster, we have taken the
257 conservative approach of leaving it unannotated in our figures. Finally, we determined whether
258 for the unidentified clusters, our top 4 marker genes corresponded to gene expression inferred
259 from individual organ dissection of the FRT (McDonough-Goldstein, Borziak, et al. 2021). These
260 comparisons allowed us to identify 3 uterine clusters, 1 seminal receptacle cluster, 1 parovaria
261 cluster, and 3 oviduct clusters. Additionally, cross-checking confirmed that our top spermathecal
262 secretory cell marker genes were most highly expressed in the ST bulk tissue transcriptome.
263 For the remaining clusters, our marker genes did not all exhibit their highest expression in bulk
264 organ-level FRT transcriptomes (McDonough-Goldstein, Borziak, et al. 2021). Thus, in total,

265 based on prior literature we identified 5 cell types, with suggestive or tissue-level annotations for
266 an additional 14 clusters (including "?-1"). The remaining 18 clusters are unlikely to correspond
267 to cell types with known molecular identities in this model species.
268
269 To further investigate the unknown clusters, we performed fluorescent *in situ* hybridization chain
270 reaction using subsets of cluster marker genes (Fig. 2, Supp). By targeting combinations of
271 marker genes, we annotated 13 additional clusters (Fig. 1 C-D) and discovered extensive novel
272 cell type heterogeneity within tissues and striking spatial regionalization. For example, previous
273 research described two morphologically distinct regions in the seminal receptacle: a narrow
274 proximal section and a contrastingly thicker distal section, which also differ in the arrangement
275 of microvilli on the inner surface, the density of secretory vacuoles, and the postures adopted by
276 occupying sperm within each section (Y. Heifetz and Rivlin 2010). We find molecular markers
277 that are specific to the seminal receptacle and distinguish these two sections: the proximal SR is
278 *beat-VII* positive, while the distal section is marked by *CG6055* (Fig. 2A-B). We also find
279 molecular markers and cell types that map onto spatially discrete and thus, likely functionally
280 specialized regions in the uterus, including a previously undescribed region in the uterine wall
281 encircling the entrance to the spermathecal ducts, marked by *cyp313a3* and *rk* (UT-into-ST, Fig.
282 2D-E and Supplement, pool H section). *CG9701* marks the cells of the oviduct valve flap
283 (Adams and Wolfner 2007), a portion of the inner, anterior uterine wall that changes
284 conformation to allow access to the oviduct and storage organs (UT-a3, Fig. 2E). *obst-E* marks
285 the UT-OVF + UT-into-ST clusters and adds a third expression domain that corresponds to UT-
286 a4 cells, which is located along the portion of the anterior uterine wall that becomes the papillate
287 elevation region after mating (Supplement, pool H section, (Adams and Wolfner 2007).
288
289 Next, we sought to use our cell-level data to significantly augment organ-level expression data
290 from the FRT (McDonough-Goldstein, Borziak, et al. 2021) (Leader et al. 2018)) and distinguish
291 between transcripts produced in reproductive cells from those produced by connective tissues
292 and other unintentional 'by-catch' in bulk tissue transcriptome sequencing. For example,
293 (McDonough-Goldstein, Borziak, et al. 2021) reported a significant enrichment of expressed
294 receptor genes unique to the seminal receptacle, with 10 SR receptor genes detected – an
295 exciting prospect, since females are thought to respond to the receipt of seminal fluid proteins,
296 yet only one SFP receptor has been identified (*Sex peptide receptor*, (Yapici et al. 2008)). In our
297 data, five of these putative SR receptor genes are unexpressed in any SR cluster, and 4 are
298 instead expressed by hemocytes. Moreover, our top hemocyte marker genes are most strongly
299 expressed in SR transcriptomes from the (McDonough-Goldstein, Borziak, et al. 2021) bulk-
300 tissue analysis. Together, these observations suggest that hemocytes may be
301 disproportionately captured alongside the SR in bulk tissue dissections, and that the observed
302 enrichment for receptor genes in the bulk-tissue SR transcriptome may be artefactual. To
303 identify a high-confidence set of female reproductive tract genes we report ~5960 genes
304 expressed in our 22 validated reproductive (e.g. non-muscle, -hemocyte, -oenocyte, -glia)
305 clusters (Supplement Table 2). Additionally, our findings greatly expand the number of female
306 reproductive tract marker genes (Supplement Table 3), which can facilitate future resource
307 development and functional analyses. This list includes a much needed, perfectly specific
308 marker gene for parovaria secretory cells (*CG42780*, Fig. 2C,H), as well as marker genes for

309   spermathecal epithelial cells, the glandular ducts, and a variety of morphologically specialized
310   regions throughout the seminal receptacle and uterus. Indeed, many FRT marker genes
311   identified in our analysis are not currently associated with any functional annotation,
312   demonstrating the great potential value of these data.
313
314

315   **Polyploidy is pervasive**
316   While the spermathecal secretory cells have previously been recognized as polyploid (Mayhew
317   and Merritt 2013; Almeida Machado Costa et al. 2022), we find that polyploidy is common in the
318   reproductive tract, occurring in at least 5 of the 6 major reproductive tissues, and spanning four
319   different ploidy levels. In aggregate, our data suggest that more than half of the FRT cells are
320   polyploid. Four lines of evidence support this conclusion. First, we noticed disparate nuclei sizes
321   upon visual inspection of dissociated nuclei; this was also noticeable in confocal micrographs
322   stained with DAPI. Second, quantitative evidence of polyploidy was observed while FACS
323   sorting the nuclear preparation prior to sequencing, evidenced by discrete strata of DAPI
324   fluorescent intensity (Fig. 3A). DAPI bright positive nuclei exhibited at least four levels of
325   increasing intensity, corresponding to 4 quantized levels of DNA abundance per nucleus, with
326   ~29% of the nuclei exhibiting a median fluorescent intensity (MFI) of 7,061, 20% with MFI of
327   11,548, 25% with MFI of 19,754 and a lesser population of 4% of total nuclei at an MFI of
328   40,112 intensity units (Supplement). Notably, median fluorescence intensity approximately
329   doubles with each step. Further evidence comes from highly variable total RNA counts per cell
330   type (Fig. 3B), where SSC, PSC, and fat cells in particular stand out as having very high total
331   RNA counts per cell; certain oviduct and uterine cells are also elevated. Finally, based on
332   expressed genes per cell type, 14 clusters had significant Gene Ontology term enrichments for
333   'polytene chromosome', 'polytene puff', and/or 'polytene band' (Fig. 3B), a chromosomal
334   morphology associated with polyploidy in Drosophila. To further investigate this we checked the
335   expression of genes that regulate or initiate endoreplication (Fig. 3C, (Almeida Machado Costa
336   et al. 2022). These genes are associated with in-progress endoreplication and are not
337   necessarily expected to mark cells that previously completed a limited round of endoreplication.
338   When endoreplication is underway, overexpression of *Myc* can stimulate additional rounds of
339   endocycling. Several signaling pathways can regulate endocycle in specific contexts (Almeida
340   Machado Costa et al. 2022). Patterns of gene expression in our data suggest that the JNK
341   pathway (terminal kinase *Bsk*), the Hippo pathway (terminal member *Yki*), and the MAPK
342   pathway (activated by *KDM5*) may influence endocycling in the FRT. One recognized function of
343   polyploidy is to achieve a higher transcriptional output, especially in secretory cells; in this
344   context, the extensive polyploidization of the FRT is consistent with the volume of secretions
345   that it produces.
346
347

348   **Female reproductive cell types specialize in producing seminal fluid proteins**
349   While seminal fluid protein gene (SFP) functions are typically interpreted with respect to their
350   production and transfer by males and their influence on females, we find many SFP genes
351   expressed in unmated female reproductive tracts. Importantly, certain FRT cell types are
352   enriched for SFP transcription (Fig. 4C). For example, in cluster "SR-dist2", a cell type in the

353   distal seminal receptacle, 30.4% of mRNAs in the median nucleus are from SFP genes, and
354   SFP transcripts comprise up to 65% of mRNAs in some nuclei. This result is driven by
355   transcripts of 11 very highly expressed SFP genes, 22 with moderate expression, and an
356   additional 15 detected (total 48 expressed SFP genes). Cluster "?-5" also appears specialized
357   for SFP transcription, with a median of 11% of mRNAs aligning to SFP genes; this cluster
358   expresses the greatest number of SFP genes relative to other FRT clusters (60 genes; 20% of
359   the 292 *D. melanogaster* SFPs). Other clusters for which >5% of mRNAs align to SFPs in the
360   median cell are the spermathecal secretory cells (9%), the ST epithelial cells (5%) and the
361   remaining SR clusters (SR-prox 1-3 and -dist1, 5-8%). In total, 78 of the 292 *D. melanogaster*
362   SFPs (>26% of SFPs) are expressed in one or more FRT cell types. SFPs are enriched in
363   different cell types than other biologically relevant gene sets (Fig. 4), and their expression is not
364   a simple correlate of high overall levels of transcription (Fig. 3B), mating responsivity (Fig. 4 A-
365   B), nor secretion (Fig. 4 D).
366
367   To further investigate the prevalence of FRT-expressed SFPs, we looked at previously
368   published FRT transcriptomes (McDonough-Goldstein, Borziak, et al. 2021; Cridland and Begun
369   2023). These two studies complement our work by supplying gene expression profiles for the
370   mated-state FRT (McDonough-Goldstein, Borziak, et al. 2021) and by the use of contrasting
371   female and male genotypes to distinguish male-contributed RNAs from endogenous female
372   RNAs in the FRT after mating (Cridland and Begun 2023). Taking evidence from these studies
373   together with our results, at least 127 of the 292 known SFP genes–43% of SFPs–are
374   expressed by the female reproductive tract. Intriguingly, this set of female-expressed SFPs
375   includes several individually notable SFPs, specifically, SFPs involved in sperm storage (*est-6,*
376   *Acp62F, Pde1c*); in binding Sex Peptide to sperm and, ergo, the maintenance of the long-term
377   post-mating response (*lectin-46Ca, lectin-46Cb*); in influencing female remating receptivity
378   (*CG10433*, and nonsignificant trend *CG32833*); and *Dup99B,* which is a predicted duplicate of
379   *Sex Peptide* that can weakly elicit aspects of the systemic post-mating response (Saudan et al.
380   2002). If we further include suggestive evidence of gene expression (i.e. detection in female
381   carcass (McDonough-Goldstein, Borziak, et al. 2021); detection in mated female FRTs in study
382   designs that cannot exclude the possibility the RNAs were transferred by males), there may be
383   >160 SFPs produced by females, i.e., the majority of *melanogaster* SFPs. The functions of most
384   SFPs are unknown; these data suggest that investigating their function in females is necessary
385   to attain a realistic view of their biological roles.
386
387
388
389   **DISCUSSION**
390
391   The work presented here has several major implications for our understanding of female insect
392   reproduction. We provide transcriptomes, functional gene set enrichments, and marker genes
393   for 22 reproductive cell types, all but one of which (SSC) was previously only morphologically
394   defined (ST cap epithelial cells, PSC, and morphologically described cell types within SR and
395   UT) or entirely unknown (cell type surrounding the entry to the ST, "UT-into-ST"). Many of these

396 new marker genes are unnamed genes for which no functional information was previously
397 available, so their status as marker genes in the FRT adds substantial biological insight.
398
399 SFP genes have canonically been primarily thought of as coding for male-expressed proteins.
400 Indeed, recent analyses have established criteria for SFP identification that include evidence of
401 transfer from males to females during copulation (Wigby et al. 2020). Whereas several SFPs
402 play a role in initiating female post-mating responses (Liu et al. 2014; Isaac et al. 2010; Chen et
403 al. 1988; Chapman et al. 1995; Pilpel et al. 2008; Avila and Wolfner 2009; Rubinstein and
404 Wolfner 2013), the past 30 years of literature have often explored the rhetorical framing that
405 males use SFPs to control, manipulate, or harm females (Chapman et al. 1995; Rice 1996;
406 Wigby and Chapman 2004; Hollis et al. 2019; Gioti et al. 2012). From this perspective, SFPs are
407 an important substrate of interlocus sexual conflict against which females must continually
408 evolve defensive countermeasures, which has been hypothesized to explain observations that
409 SFPs evolve rapidly (Haerty et al. 2007; Patlar et al. 2021). Our finding that females transcribe
410 >40% of all SFP genes, and that SFPs are enriched in specialized female cell types, suggests
411 this simple view is incorrect. For perspective, we can tally the SFPs asserted to influence post-
412 mating female biology (Wigby et al. 2020). We liberally include non-significant trending results
413 (*CG32833*, (Sirot et al. 2014)); suggestive evidence from an association study (*msopa*,
414 (Fiumera, Dumont, and Clark 2007)) that was not supported by gene knockdown (Patlar and
415 Civetta 2022); conclusions based on whole-animal manipulations for which results could
416 conceivably be mediated by male expression in non-reproductive tissues (*CG10433*, (Liu et al.
417 2014)); genes whose documented effect opposes male interests (i.e. accelerating female
418 remating receptivity, *est-6 (Gilbert, Richmond, and Sheehan 1981)*; and genes that we now
419 recognize are also expressed by the FRT (*CG10433, est-6, Acp62F, Dup99B, Pde1c, lectin-*
420 *46Ca, lectin-46Cb, CG32833*), to reach a final tally of only 24 genes (8% of SFPs). Importantly,
421 these constitute a far smaller sample of SFPs than the female-expressed subset. Thus, the
422 premise that the majority of SFPs have strictly male functions directed at manipulating the
423 female post-mating response in favor of male interests, should be re-evaluated (e.g. (Hopkins
424 and Perry 2022). As for how typical it is that SFPs function in facilitating post-mating responses,
425 >26% are found expressed in females that have never mated, but it remains possible that the
426 proteins are secreted, trafficked to their proper destination, or activated only following mating,
427 consistent with (Sanchez-Lopez et al. 2022). Investigating SFP function in females is necessary
428 to attain a realistic view of their biological roles, and expression patterns in females furnish
429 helpful clues–it is striking that SFPs are enriched in both sperm storage organs, consistent with
430 possible roles in moving, activating, or nourishing sperm.
431
432 While it is well appreciated that lower reproductive tract functions include interacting with male
433 products, the FRT's potential interactions with eggs, beyond physically moving them from
434 ovulation through oviposition, are poorly understood. In many animals, eggs are still under
435 construction post-ovulation, with major contributions to egg development occurring in the
436 somatic reproductive tract. For example, in birds, biomineralization of the eggshell is
437 accomplished by the uterus (Gautron et al. 2021). In many insects, the parovaria (i.e. female
438 accessory glands) produce additional substances applied to the egg's exterior, such as
439 adhesives or silk to anchor the egg after oviposition; or, particularly among aquatic insects,

440 gelatinous substances that hold a clutch of eggs together (Lancaster and Downes 2013). These
441 additional coatings can also enhance desiccation resistance and deter predation or parasitism
442 by their sticky or tough textures. In *Drosophila*, the eggshell is produced pre-ovulation by follicle
443 cells, including the deposition of three major layers and enzymatic treatments that covalently
444 crosslink the eggshell proteins. Prior work has identified proteins that constitute the eggshell via
445 mass spectrometry and gel electrophoresis (Fakhouri et al. 2006; Waring and Mahowald 1979).
446 However, owing to technical limitations on solubilizing or separating proteins after cross-linking,
447 these studies focused on pre-ovulation eggs, meaning that if subsequent contributions or
448 modifications occur post-ovulation in the lower reproductive tract, those effects would go
449 undetected. The FRT expresses more than 1000 genes with secretion signal peptides, resulting
450 in a secretion-rich environment in its lumen, including many enzymes, antifungal proteins, and
451 proteins of unknown function (McDonough-Goldstein, Whittington, et al. 2021; McDonough-
452 Goldstein, Pitnick, and Dorus 2022). It would be surprising if the egg were wholly unaffected by
453 being washed through this protein milieu. Certain FRT cell types that may be especially relevant
454 to post-ovulation egg development, specifically the parovaria secretory cells and the two distinct
455 cell types in the posterior uterus, show enriched transcription of non-SFP secreted genes (Fig.
456 4D). One can easily imagine the posterior uterine cells could apply a final protective coating or
457 lubricant during oviposition. Interactions between FRT secretions and the egg are a promising
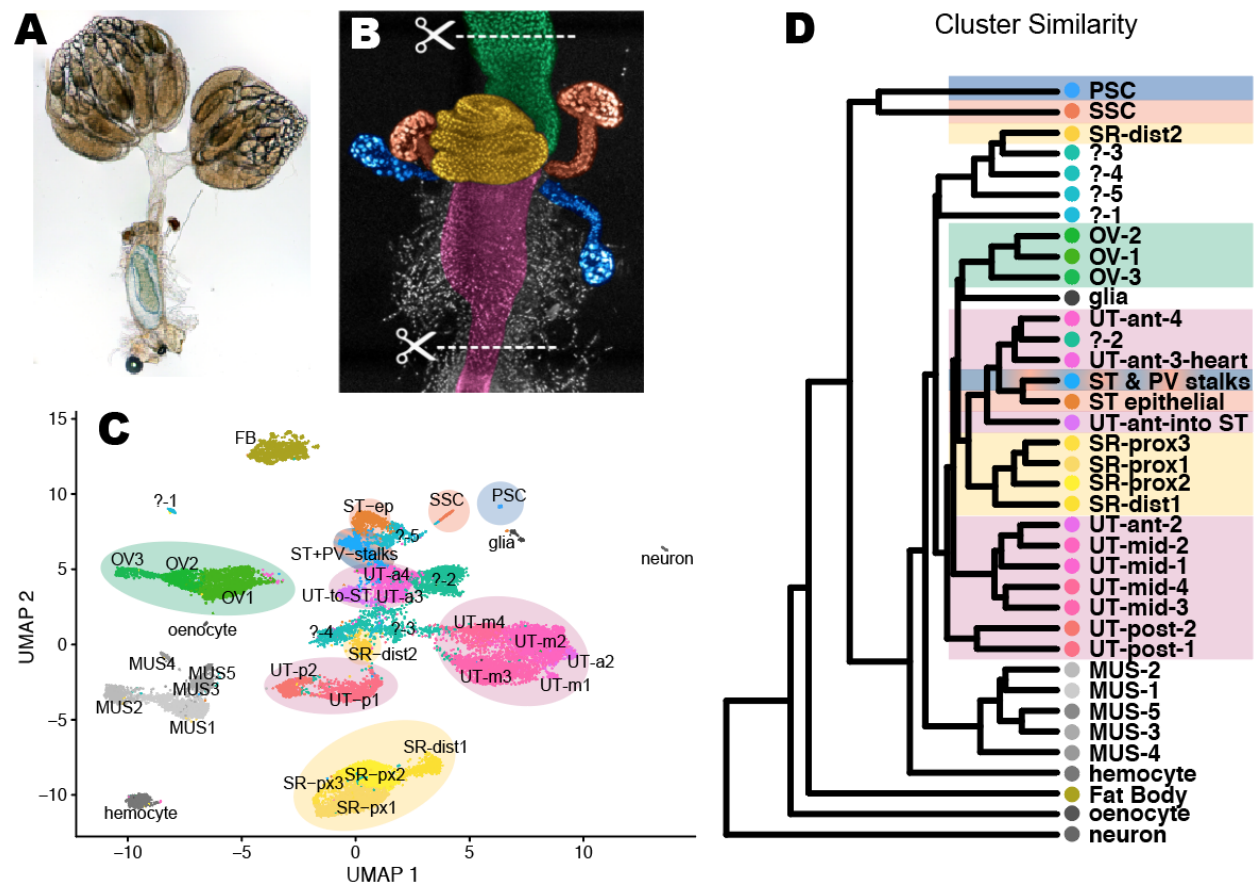458 direction for future work.
459
460

461 **ACKNOWLEDGEMENTS**
473

474 **DECLARATION OF INTERESTS**
475 The authors declare no competing interests.
476

477 **FIGURES**
478
479
480

Figure 1: Organs and cell types in the somatic female reproductive tract. *A)* The complete *Drosophila* female reproductive tract, including the ovaries and a descended egg held in the uterus. *B)* Micrograph of the sequenced region of the lower FRT, with a DAPI stain and overlay colors showing organs: green = oviduct (OV), yellow = seminal receptacle (SR), orange = spermathecae (ST), blue = parovaria (PV, also called female accessory glands), and magenta = uterus (UT). *C)* Annotated UMAP plot of all sequenced nuclei. 'SSC' = spermathecal secretory cells. 'PSC' = parovaria secretory cells. 'MUS' = muscle. *D)* Cluster similarity, based on gene expression principal components analysis.

491
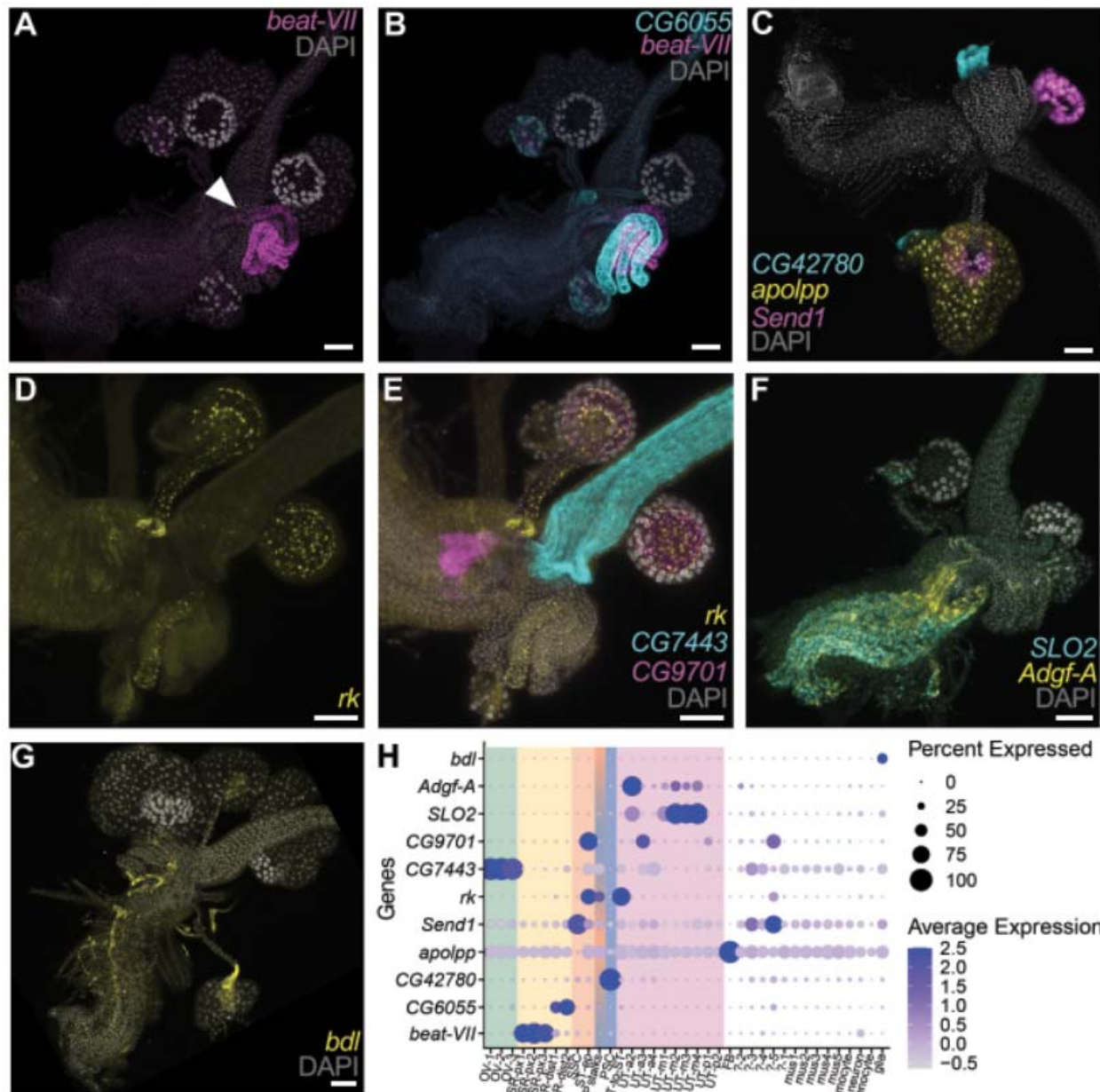
Figure 2: Cluster annotation with *in situ* hybridization chain reactions. *A-G)* Lower reproductive tracts from non-mated females processed for *in situ* hybridization chain reaction, with a nuclear stain (DAPI, grey). *A-B)* The proximal seminal receptacle is *beat-VII* positive (magenta), while the distal SR is *CG6055* positive (cyan). White arrowhead indicates the point at which the seminal receptacle joins the uterus. *C) CG42780* (cyan) is a marker gene for parovaria secretory cells. *Send1* expression (magenta) confirms the spermathecal secretory cell annotation. *Apolipophorin* (yellow) is a marker gene for the reproductive-associated fat body, which was dissected away from one lateral side of the tissue. *D-E)* A ring of cells in the uterine wall that encircle the entry point to the spermatheca and the stalks of the spermathecae and parovaria are all *rk* positive (yellow). *CG9701* expression (magenta) marks the oviduct valve flap. *In situ* probes against *rk* and *CG9701* transcripts co-localize in the spermathecal epithelial cells. *CG7443* (cyan) is an oviduct marker gene. *F) SLO2* (cyan) and *Adgf-A* (yellow) are co-

504   expressed in much of the uterus, with relatively stronger *Adgf-2* expression in a subset of the
505   anterior uterus versus higher *SLO2* expression through the middle uterus. *G) borderless*
506   (yellow) is a marker gene for glia. *H)* Expression across clusters for the marker genes targeted
507   with *in situ* hybridization chain reaction. Scale bar = 50 micrometers.
508
509
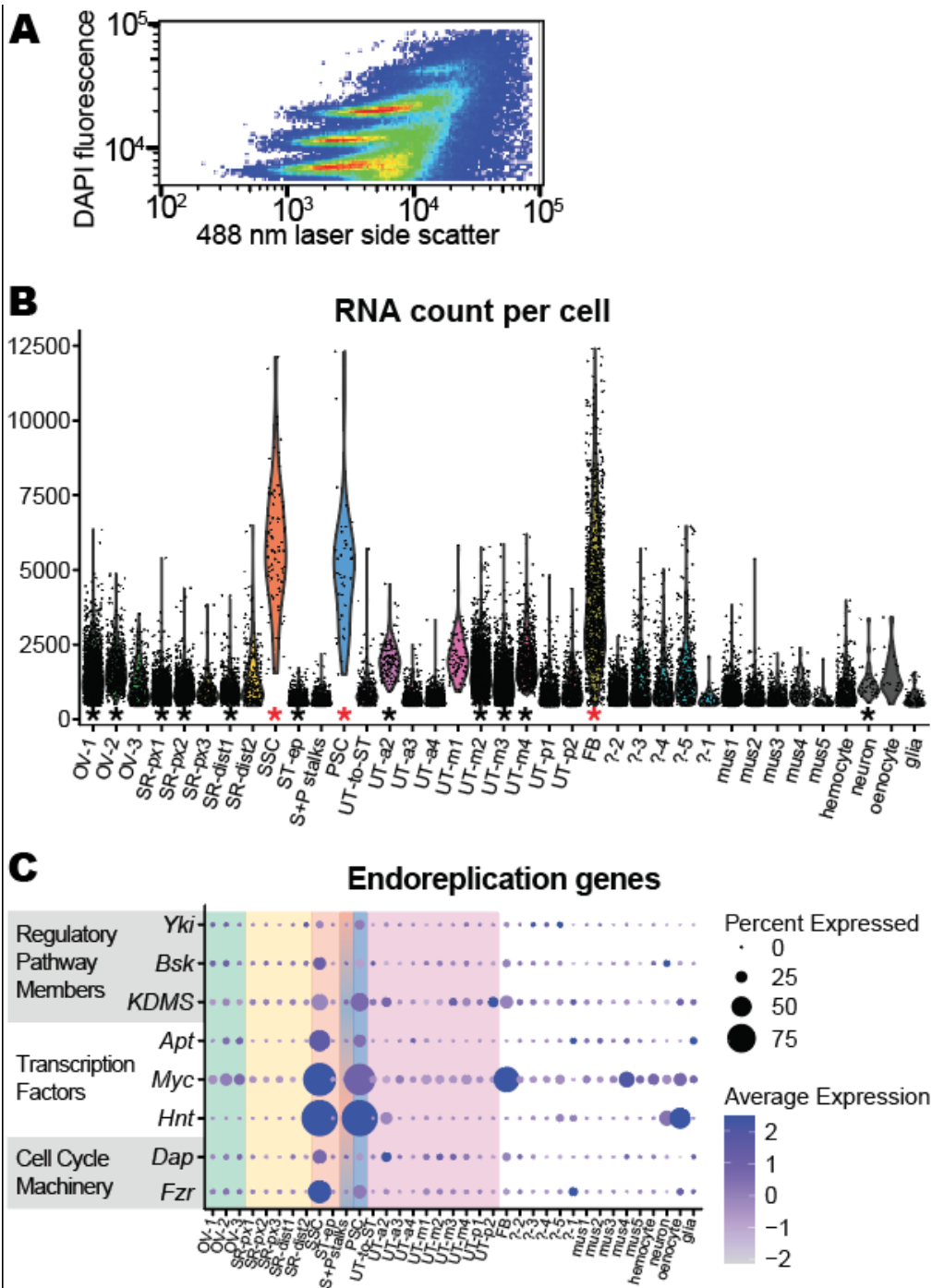
510
Figure 3: Polyploidy is common in the FRT. *A)* Nuclei segregate into strata during FACs sorting
512   by side scatter (size) and DAPI fluorescent intensity, with nuclei frequency shown by heat map

513   colors. *B)* RNA count per cell varies dramatically among clusters. Asterisks above the cluster ID

514   along the x-axis indicate significant Gene Ontology enrichments for "polytene chromosomes," a

515   chromosomal morphology associated with polyploidy in Drosophila (red = adjusted $p < 10^{-4}$,

516   black = adjusted $p < 0.01$). *C)* Expression of genes associated with the endoreplication process

517   (Almeida Machado Costa et al. 2022).

518
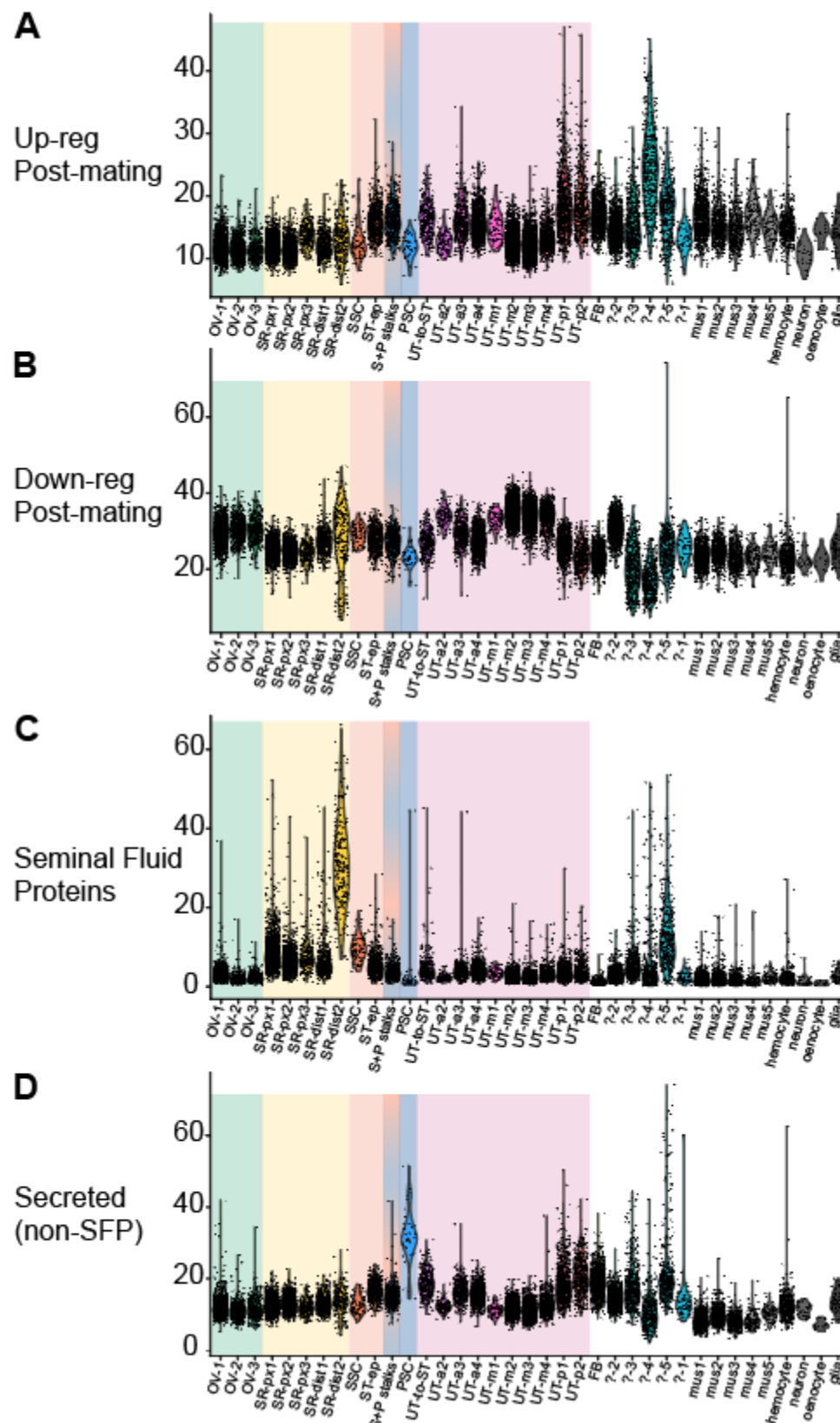
Figure 4: The expression of biologically relevant gene sets varies among cell types in the FRT.
A-D) Each point is one cell; each violin is the distribution of cell values for one cell type cluster.

522   The y-axis gives the percent of RNA read counts per cell that align to genes belonging to the
523   relevant functional category. Background colors correspond to organ-level identity of the
524   clusters, as given in Fig. 1B. *A)* Expression of genes that are up-regulated in response to mating
525   in (McDonough-Goldstein, Borziak, et al. 2021). *B)* Expression of genes that are down-regulated
526   in response to mating. *C)* Expression of genes which produce seminal fluid proteins (SFPs, as
527   identified in (Wigby et al. 2020). The seminal receptacle, in particular, vigorously transcribes
528   seminal fluid protein genes. *D)* Transcription of genes that have a secretion signal peptide
529   detected by the SignalP algorithm (Almagro Armenteros et al. 2019), excluding seminal fluid
530   genes.

531
532
533
534   **REFERENCES**

535   Adams, Erika M., and Mariana F. Wolfner. 2007. "Seminal Proteins but Not Sperm Induce
536   Morphological Changes in the Drosophila Melanogaster Female Reproductive Tract during
537   Sperm Storage." *Journal of Insect Physiology* 53 (4): 319–31.
538   Allen, Anna K., and Allan C. Spradling. 2008. "The Sf1-Related Nuclear Hormone Receptor
539   Hr39 Regulates Drosophila Female Reproductive Tract Development and Function."
540   *Development* 135 (2): 311–21.
541   Almagro Armenteros, José Juan, Konstantinos D. Tsirigos, Casper Kaae Sønderby, Thomas
542   Nordahl Petersen, Ole Winther, Søren Brunak, Gunnar von Heijne, and Henrik Nielsen.
543   2019. "SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks."
544   *Nature Biotechnology* 37 (4): 420–23.
545   Almeida Machado Costa, Caique, Xian-Feng Wang, Calder Ellsworth, and Wu-Min Deng. 2022.
546   "Polyploidy in Development and Tumor Models in Drosophila." *Seminars in Cancer Biology*
547   81 (June): 106–18.
548   Avila, Frank W., and Mariana F. Wolfner. 2009. "Acp36DE Is Required for Uterine
549   Conformational Changes in Mated Drosophila Females." *Proceedings of the National*
550   *Academy of Sciences of the United States of America* 106 (37): 15796–800.
551   Bath, Eleanor, Samuel Bowden, Carla Peters, Anjali Reddy, Joseph A. Tobias, Evan Easton-
552   Calabria, Nathalie Seddon, Stephen F. Goodwin, and Stuart Wigby. 2017. "Sperm and Sex
553   Peptide Stimulate Aggression in Female Drosophila." *Nature Ecology & Evolution* 1 (6):
554   0154.
555   Beery, Annaliese K., and Irving Zucker. 2011. "Sex Bias in Neuroscience and Biomedical
556   Research." *Neuroscience and Biobehavioral Reviews* 35 (3): 565–72.
557   Bruce, H. S., G. Jerz, S. Kelly, J. McCarthy, and A. Pomerantz. 2021. "Hybridization Chain
558   Reaction (HCR) in Situ Protocol." https://www.researchgate.net/profile/Heather-Bruce-
559   5/publication/353582040_Hybridization_Chain_Reaction_HCR_In_Situ_Protocol/links/6137
560   c3db9520966a6b039e53/Hybridization-Chain-Reaction-HCR-In-Situ-Protocol.pdf.
561   Chapman, T., L. F. Liddle, J. M. Kalb, M. F. Wolfner, and L. Partridge. 1995. "Cost of Mating in
562   Drosophila Melanogaster Females Is Mediated by Male Accessory Gland Products." *Nature*
563   373 (6511): 241–44.
564   Chen, P. S., E. Stumm-Zollinger, T. Aigaki, J. Balmer, M. Bienz, and P. Böhlen. 1988. "A Male
565   Accessory Gland Peptide That Regulates Reproductive Behavior of Female D.
566   Melanogaster." *Cell* 54 (3): 291–98.
567   Chintapalli, Venkateswara R., Jing Wang, and Julian A. T. Dow. 2007. "Using FlyAtlas to
568   Identify Better Drosophila Melanogaster Models of Human Disease." *Nature Genetics* 39
569   (6): 715–20.

Choi, Harry M. T., Victor A. Beck, and Niles A. Pierce. 2014. "Next-Generation in Situ Hybridization Chain Reaction: Higher Gain, Lower Cost, Greater Durability." *ACS Nano* 8 (5): 4284–94.

Choi, Harry M. T., Maayan Schwarzkopf, Mark E. Fornace, Aneesh Acharya, Georgios Artavanis, Johannes Stegmaier, Alexandre Cunha, and Niles A. Pierce. 2018. "Third-Generation in Situ Hybridization Chain Reaction: Multiplexed, Quantitative, Sensitive, Versatile, Robust." *Development* 145 (12). https://doi.org/10.1242/dev.165753.

Cridland, Julie M., and David J. Begun. 2023. "Male-Derived Transcripts Isolated from the Mated Female Reproductive Tract in Drosophila Melanogaster." *G3* , September. https://doi.org/10.1093/g3journal/jkad202.

Cridland, Julie M., Alex C. Majane, Li Zhao, and David J. Begun. 2022. "Population Biology of Accessory Gland-Expressed de Novo Genes in Drosophila Melanogaster." *Genetics* 220 (1). https://doi.org/10.1093/genetics/iyab207.

Fakhouri, Mazen, Maggie Elalayli, Daniel Sherling, Jacklyn D. Hall, Eric Miller, Xutong Sun, Lance Wells, and Ellen K. LeMosy. 2006. "Minor Proteins and Enzymes of the Drosophila Eggshell Matrix." *Developmental Biology* 293 (1): 127–41.

Fedorka, Kenneth M., Jodell E. Linder, Wade Winterhalter, and Daniel Promislow. 2007. "Post-Mating Disparity between Potential and Realized Immune Response in Drosophila Melanogaster." *Proceedings. Biological Sciences / The Royal Society* 274 (1614): 1211–17.

Fiumera, Anthony C., Bethany L. Dumont, and Andrew G. Clark. 2007. "Associations between Sperm Competition and Natural Variation in Male Reproductive Genes on the Third Chromosome of Drosophila Melanogaster." *Genetics* 176 (2): 1245–60.

Gautron, J., L. Stapane, N. Le Roy, Y. Nys, A. B. Rodriguez-Navarro, and M. T. Hincke. 2021. "Correction to: Avian Eggshell Biomineralization: An Update on Its Structure, Mineralogy and Protein Tool Kit." *BMC Molecular and Cell Biology* 22 (1): 14.

Gilbert, Donald G., Rollin C. Richmond, and Kathy B. Sheehan. 1981. "Studies of Esterase 6 inDrosophila Melanogaster. VII. Remating Times of Females Inseminated by Males Having Active or Null Alleles." *Behavior Genetics* 11 (3): 195–208.

Gioti, A., S. Wigby, B. Wertheim, E. Schuster, P. Martinez, C. J. Pennington, L. Partridge, and T. Chapman. 2012. "Sex Peptide of Drosophila Melanogaster Males Is a Global Regulator of Reproductive Processes in Females." *Proceedings. Biological Sciences / The Royal Society* 279 (1746): 4423–32.

Haerty, Wilfried, Santosh Jagadeeshan, Rob J. Kulathinal, Alex Wong, Kristipati Ravi Ram, Laura K. Sirot, Lisa Levesque, et al. 2007. "Evolution in the Fast Lane: Rapidly Evolving Sex-Related Genes in Drosophila." *Genetics* 177 (3): 1321–35.

Heifetz, Yael, Moshe Lindner, Yuval Garini, and Mariana F. Wolfner. 2014. "Mating Regulates Neuromodulator Ensembles at Nerve Termini Innervating the Drosophila Reproductive Tract." *Current Biology: CB* 24 (7): 731–37.

Heifetz, Y., and P. K. Rivlin. 2010. "Beyond the Mouse Model: Using Drosophila as a Model for Sperm Interaction with the Female Reproductive Tract." *Theriogenology* 73 (6): 723–39.

Hollis, Brian, Mareike Koppik, Kristina U. Wensing, Hanna Ruhmann, Eléonore Genzoni, Berra Erkosar, Tadeusz J. Kawecki, Claudia Fricke, and Laurent Keller. 2019. "Sexual Conflict Drives Male Manipulation of Female Postmating Responses in Drosophila Melanogaster." *Proceedings of the National Academy of Sciences of the United States of America* 116 (17): 8437–44.

Hopkins, Ben R., and Jennifer C. Perry. 2022. "The Evolution of Sex Peptide: Sexual Conflict, Cooperation, and Coevolution." *Biological Reviews of the Cambridge Philosophical Society* 97 (4): 1426–48.

Isaac, R. Elwyn, Chenxi Li, Amy E. Leedale, and Alan D. Shirras. 2010. "Drosophila Male Sex Peptide Inhibits Siesta Sleep and Promotes Locomotor Activity in the Post-Mated Female." *Proceedings. Biological Sciences / The Royal Society* 277 (1678): 65–70.

621   Kelleher, Erin S., and James E. Pennington. 2009. "Protease Gene Duplication and Proteolytic
622   Activity in Drosophila Female Reproductive Tracts." *Molecular Biology and Evolution* 26 (9):
623   2125–34.
624   Kuehn, Emily, David S. Clausen, Ryan W. Null, Bria M. Metzger, Amy D. Willis, and B. Duygu
625   Özpolat. 2022. "Segment Number Threshold Determines Juvenile Onset of Germline
626   Cluster Expansion in Platynereis Dumerilii." *Journal of Experimental Zoology. Part B,*
627   *Molecular and Developmental Evolution* 338 (4): 225–40.
628   Lancaster, Jill, and Barbara J. Downes. 2013. *Aquatic Entomology*. OUP Oxford.
629   Lawniczak, Mara K. N., and David J. Begun. 2007. "Molecular Population Genetics of Female-
630   Expressed Mating-Induced Serine Proteases in Drosophila Melanogaster." *Molecular*
631   *Biology and Evolution* 24 (9): 1944–51.
632   Leader, David P., Sue A. Krause, Aniruddha Pandit, Shireen A. Davies, and Julian A. T. Dow.
633   2018. "FlyAtlas 2: A New Version of the Drosophila Melanogaster Expression Atlas with
634   RNA-Seq, miRNA-Seq and Sex-Specific Data." *Nucleic Acids Research* 46 (D1): D809–15.
635   Li, Hongjie, Jasper Janssens, Maxime De Waegeneer, Sai Saroja Kolluru, Kristofer Davie,
636   Vincent Gardeux, Wouter Saelens, et al. 2022. "Fly Cell Atlas: A Single-Nucleus
637   Transcriptomic Atlas of the Adult Fruit Fly." *Science* 375 (6584): eabk2432.
638   Liu, Chen, Jia-Lin Wang, Ya Zheng, En-Juan Xiong, Jing-Jing Li, Lin-Ling Yuan, Xiao-Qiang Yu,
639   and Yu-Feng Wang. 2014. "Wolbachia-Induced Paternal Defect in Drosophila Is Likely by
640   Interaction with the Juvenile Hormone Pathway." *Insect Biochemistry and Molecular Biology*
641   49 (June): 49–58.
642   Lombardo, Kaelina D., Hayley K. Sheehy, Julie M. Cridland, and David J. Begun. 2023.
643   "Identifying Candidate de Novo Genes Expressed in the Somatic Female Reproductive
644   Tract of Drosophila Melanogaster." *G3* 13 (8). https://doi.org/10.1093/g3journal/jkad122.
645   Mack, Paul D., Anat Kapelnikov, Yael Heifetz, and Michael Bender. 2006. "Mating-Responsive
646   Genes in Reproductive Tissues of Female Drosophila Melanogaster." *Proceedings of the*
647   *National Academy of Sciences of the United States of America* 103 (27): 10358–63.
648   Majane, Alex C., Julie M. Cridland, and David J. Begun. 2022. "Single-Nucleus Transcriptomes
649   Reveal Evolutionary and Functional Properties of Cell Types in the Drosophila Accessory
650   Gland." *Genetics* 220 (2). https://doi.org/10.1093/genetics/iyab213.
651   Manier, Mollie K., John M. Belote, Kirstin S. Berben, David Novikov, Will T. Stuart, and Scott
652   Pitnick. 2010. "Resolving Mechanisms of Competitive Fertilization Success in Drosophila
653   Melanogaster." *Science* 328 (5976): 354–57.
654   Martelotto, Luciano. n.d. "'Frankenstein'protocol for Nuclei Isolation from Fresh and Frozen
655   Tissue for snRNAseq." Accessed September 15, 2023.
656   https://scholar.archive.org/work/x2xcq7kxwzgz3dj4afz4g5x3s4/access/wayback/https://ww
657   w.protocols.io/view/frankenstein-protocol-for-nuclei-isolation-from-f-3fkgjkw.pdf.
658   Mattei, Alexandra L., Mark L. Riccio, Frank W. Avila, and Mariana F. Wolfner. 2015. "Integrated
659   3D View of Postmating Responses by the Drosophila Melanogaster Female Reproductive
660   Tract, Obtained by Micro-Computed Tomography Scanning." *Proceedings of the National*
661   *Academy of Sciences of the United States of America* 112 (27): 8475–80.
662   Mayhew, Mark Leonard, and David John Merritt. 2013. "The Morphogenesis of Spermathecae
663   and Spermathecal Glands in Drosophila Melanogaster." *Arthropod Structure &*
664   *Development* 42 (5): 385–93.
665   McDonough-Goldstein, Caitlin E., Kirill Borziak, Scott Pitnick, and Steve Dorus. 2021.
666   "Drosophila Female Reproductive Tract Gene Expression Reveals Coordinated Mating
667   Responses and Rapidly Evolving Tissue-Specific Genes." *G3* 11 (3).
668   https://doi.org/10.1093/g3journal/jkab020.
669   McDonough-Goldstein, Caitlin E., Scott Pitnick, and Steve Dorus. 2022. "Drosophila Female
670   Reproductive Glands Contribute to Mating Plug Composition and the Timing of Sperm
671   Ejection." *Proceedings. Biological Sciences / The Royal Society* 289 (1968): 20212213.

672 McDonough-Goldstein, Caitlin E., Emma Whittington, Erin L. McCullough, Sharleen M. Buel,
673   Scott Erdman, Scott Pitnick, and Steve Dorus. 2021. "Pronounced Postmating Response in
674   the Drosophila Female Reproductive Tract Fluid Proteome." *Molecular & Cellular*
675   *Proteomics: MCP* 20 (September): 100156.
676 McGinnis, Christopher S., Lyndsay M. Murrow, and Zev J. Gartner. 2019. "DoubletFinder:
677   Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors."
678   *Cell Systems* 8 (4): 329–37.e4.
679 Muytjens, Carla M. J., Yijing Yu, and Eleftherios P. Diamandis. 2018. "Functional Proteomic
680   Profiling Reveals KLK13 and TMPRSS11D as Active Proteases in the Lower Female
681   Reproductive Tract." *F1000Research* 7 (October): 1666.
682 Orr, Teri J., Mercedes Burns, Kristen Hawkes, Kay E. Holekamp, Kristin A. Hook, Chloe C.
683   Josefson, Abigail A. Kimmitt, et al. 2020. "It Takes Two to Tango: Including a Female
684   Perspective in Reproductive Biology." *Integrative and Comparative Biology* 60 (3): 796–
685   813.
686 Patlar, Bahar, and Alberto Civetta. 2022. "Seminal Fluid Gene Expression and Reproductive
687   Fitness in Drosophila Melanogaster." *BMC Ecology and Evolution* 22 (1): 20.
688 Patlar, Bahar, Vivek Jayaswal, José M. Ranz, and Alberto Civetta. 2021. "Nonadaptive
689   Molecular Evolution of Seminal Fluid Proteins in Drosophila." *Evolution; International*
690   *Journal of Organic Evolution* 75 (8): 2102–13.
691 Pilpel, Noam, Ifat Nezer, Shalom W. Applebaum, and Yael Heifetz. 2008. "Mating-Increases
692   Trypsin in Female Drosophila Hemolymph." *Insect Biochemistry and Molecular Biology* 38
693   (3): 320–30.
694 Prokupek, A. M., S. D. Kachman, I. Ladunga, and L. G. Harshman. 2009. "Transcriptional
695   Profiling of the Sperm Storage Organs of Drosophila Melanogaster." *Insect Molecular*
696   *Biology* 18 (4): 465–75.
697 Reinhardt, Josie A., Bryan Kolaczkowski, Corbin D. Jones, David J. Begun, and Andrew D.
698   Kern. 2014. "Parallel Geographic Variation in Drosophila Melanogaster." *Genetics* 197 (1):
699   361–73.
700 Rezával, Carolina, Hania J. Pavlou, Anthony J. Dornan, Yick-Bun Chan, Edward A. Kravitz, and
701   Stephen F. Goodwin. 2012. "Neural Circuitry Underlying Drosophila Female Postmating
702   Behavioral Responses." *Current Biology: CB* 22 (13): 1155–65.
703 Rice, W. R. 1996. "Sexually Antagonistic Male Adaptation Triggered by Experimental Arrest of
704   Female Evolution." *Nature* 381 (6579): 232–34.
705 Robinson, Scott W., Pawel Herzyk, Julian A. T. Dow, and David P. Leader. 2013. "FlyAtlas:
706   Database of Gene Expression in the Tissues of Drosophila Melanogaster." *Nucleic Acids*
707   *Research* 41 (Database issue): D744–50.
708 Rubinstein, C. Dustin, and Mariana F. Wolfner. 2013. "Drosophila Seminal Protein Ovulin
709   Mediates Ovulation through Female Octopamine Neuronal Signaling." *Proceedings of the*
710   *National Academy of Sciences of the United States of America* 110 (43): 17420–25.
711 Sanchez-Lopez, Javier Arturo, Shai Twena, Ido Apel, Shani Chen Kornhaeuser, Michael
712   Chasnitsky, Andras G. Miklosi, Perla J. Vega-Dominguez, Alex Shephard, Amir Hefetz, and
713   Yael Heifetz. 2022. "Male-Female Communication Enhances Release of Extracellular
714   Vesicles Leading to High Fertility in Drosophila." *Communications Biology* 5 (1): 815.
715 Satija, Rahul, Jeffrey A. Farrell, David Gennert, Alexander F. Schier, and Aviv Regev. 2015.
716   "Spatial Reconstruction of Single-Cell Gene Expression Data." *Nature Biotechnology* 33 (5):
717   495–502.
718 Saudan, Philippe, Klaus Hauck, Matthias Soller, Yves Choffat, Michael Ottiger, Michael Spörri,
719   Zhaobing Ding, et al. 2002. "Ductus Ejaculatorius Peptide 99B (DUP99B), a Novel
720   Drosophila Melanogaster Sex-Peptide Pheromone." *European Journal of Biochemistry /*
721   *FEBS* 269 (3): 989–97.
722 Scheunemann, L., A. Lampin-Saint-Amaux, J. Schor, and T. Preat. 2019. "A Sperm Peptide

723   Enhances Long-Term Memory in Female Drosophila." *Science Advances* 5 (11): eaax3432.
724 Schnakenberg, Sandra L., Wilfredo R. Matias, and Mark L. Siegal. 2011. "Sperm-Storage
725   Defects and Live Birth in Drosophila Females Lacking Spermathecal Secretory Cells."
726   *PLoS Biology* 9 (11): e1001192.
727 Sepil, Irem, Ben R. Hopkins, Rebecca Dean, Marie-Laëtitia Thézénas, Philip D. Charles,
728   Rebecca Konietzny, Roman Fischer, Benedikt M. Kessler, and Stuart Wigby. 2019.
729   "Quantitative Proteomics Identification of Seminal Fluid Proteins in Male Drosophila
730   Melanogaster." *Molecular & Cellular Proteomics: MCP* 18 (Suppl 1): S46–58.
731 Sirot, Laura K., Geoffrey D. Findlay, Jessica L. Sitnik, Dorina Frasheri, Frank W. Avila, and
732   Mariana F. Wolfner. 2014. "Molecular Characterization and Evolution of a Gene Family
733   Encoding Both Female- and Male-Specific Reproductive Proteins in Drosophila." *Molecular
734   Biology and Evolution* 31 (6): 1554–67.
735 Sun, Jianjun, and Allan C. Spradling. 2012. "Female Reproductive Glands Play Essential Roles
736   in Reproduction That May Have Been Conserved During Evolution." *Biology of
737   Reproduction* 87 (Suppl_1): 347–347.
738 ———. 2013. "Ovulation in Drosophila Is Controlled by Secretory Cells of the Female
739   Reproductive Tract." *eLife* 2 (April): e00415.
740 Svetec, Nicolas, Julie M. Cridland, Li Zhao, and David J. Begun. 2016. "The Adaptive
741   Significance of Natural Genetic Variation in the DNA Damage Response of Drosophila
742   Melanogaster." *PLoS Genetics* 12 (3): e1005869.
743 Waring, G. L., and A. P. Mahowald. 1979. "Identification and Time of Synthesis of Chorion
744   Proteins in Drosophila Melanogaster." *Cell* 16 (3): 599–607.
745 Wigby, Stuart, Nora C. Brown, Sarah E. Allen, Snigdha Misra, Jessica L. Sitnik, Irem Sepil,
746   Andrew G. Clark, and Mariana F. Wolfner. 2020. "The Drosophila Seminal Proteome and
747   Its Role in Postcopulatory Sexual Selection." *Philosophical Transactions of the Royal
748   Society of London. Series B, Biological Sciences* 375 (1813): 20200072.
749 Wigby, Stuart, and Tracey Chapman. 2004. "Female Resistance to Male Harm Evolves in
750   Response to Manipulation of Sexual Conflict." *Evolution; International Journal of Organic
751   Evolution* 58 (5): 1028–37.
752 Woitowich, Nicole C., Annaliese Beery, and Teresa Woodruff. 2020. "A 10-Year Follow-up
753   Study of Sex Inclusion in the Biological Sciences." *eLife* 9 (June).
754   https://doi.org/10.7554/eLife.56344.
755 Xi, Nan Miles, and Jingyi Jessica Li. 2021. "Benchmarking Computational Doublet-Detection
756   Methods for Single-Cell RNA Sequencing Data." *Cell Systems* 12 (2): 176–94.e6.
757 Yapici, Nilay, Young-Joon Kim, Carlos Ribeiro, and Barry J. Dickson. 2008. "A Receptor That
758   Mediates the Post-Mating Switch in Drosophila Reproductive Behaviour." *Nature* 451
759   (7174): 33–37.
760 Young, Matthew D., and Sam Behjati. 2020. "SoupX Removes Ambient RNA Contamination
761   from Droplet-Based Single-Cell RNA Sequencing Data." *GigaScience* 9 (12).
762   https://doi.org/10.1093/gigascience/giaa151.
763 Zappia, Luke, and Alicia Oshlack. 2018. "Clustering Trees: A Visualization for Evaluating
764   Clusterings at Multiple Resolutions." *GigaScience* 7 (7).
765   https://doi.org/10.1093/gigascience/giy083.
766 Zhao, Li, Janneke Wit, Nicolas Svetec, and David J. Begun. 2015. "Parallel Gene Expression
767   Differences between Low and High Latitude Populations of Drosophila Melanogaster and
768   D. Simulans." *PLoS Genetics* 11 (5): e1005184.

769