1   **Distinct gene expression dynamics in germ line and somatic tissue during ovariole**

2   **morphogenesis in *Drosophila melanogaster***

3   Shreeharsha Tarikere[1a], Guillem Ylla[1a] and Cassandra G. Extavour[1, 2]*

4

5   1. Department of Organismic and Evolutionary Biology, Harvard University

6   2. Department of Molecular and Cellular Biology, Harvard University

7

8   * Author for correspondence: extavour@oeb.harvard.edu

9   a. These authors contributed equally to this work

10

11

Abstract

The survival and evolution of a species is a function of the number of offspring it can produce. In insects the number of eggs that an ovary can produce is a major determinant of reproductive capacity. Insect ovaries are made up of tubular egg-producing subunits called ovarioles, whose number largely determines the number of eggs that can be potentially laid. Ovariole number is directly determined by the number of cellular structures called terminal filaments, which are stacks of cells that assemble in the larval ovary. Elucidating the developmental and regulatory mechanisms of terminal filament formation is thus key to understanding the regulation of insect reproduction through ovariole number regulation. We systematically measured mRNA expression of all cells in the larval ovary at the beginning, middle and end of terminal filament formation. We also separated somatic and germ line cells during these stages and assessed their tissue-specific gene expression during larval ovary development. We found that the number of differentially expressed somatic genes is highest during late stages of terminal filament formation and includes many signaling pathways that govern ovary development. We also show that germ line tissue, in contrast, shows greater differential expression during early stages of terminal filament formation, and highly expressed germ line genes at these stages largely control cell division and DNA repair. We provide a tissue-specific and temporal transcriptomic dataset of gene expression in the developing larval ovary as a resource to study insect reproduction.

KEY WORDS: Ovary, FACS, RNA-Seq, Terminal filament, Germ line, Stem cell niche.

Tarikere, Ylla & Extavour Page 2

INTRODUCTION

34

35 Healthy reproductive organs are among the most important factors that determine the

36 fertility of an individual, and more importantly, continuity of the species itself. Reproductive

37 fitness, including fecundity, is determined by the number of progenies an organism can

38 produce. In insects, egg-producing subunits of ovaries are called ovarioles (BÜNING

39 1994). In flies of the genus *Drosophila*, the number of ovarioles predicts the peak egg

40 laying potential of the females of the species (DAVID 1970), and is negatively correlated

41 with egg size but positively correlated with reproductive output (CHURCH *et al.* 2021). The

42 number of ovarioles varies widely across insects and is in the range of 18-24 ovarioles

43 per ovary in wild type North American populations of *Drosophila melanogaster* (HONEK

44 1993; MARKOW AND O'GRADY 2007; HODIN 2009). In *Drosophila*, adult ovariole number is

45 established in the larval stages through the development of a species-specific number of

46 linear somatic cell stacks called terminal filaments (KING *et al.* 1968). The genetic

47 mechanisms governing ovary morphogenesis, which includes the process of regulation

48 of terminal filament number and assembly during larval ovary development, remain poorly

49 understood.

50 Ovary morphogenesis is orchestrated by interactions of the cell types of somatic

51 and germ line tissues. Somatic ovarian tissue is principally made up of five cell types -

52 sheath cells, swarm cells, terminal filaments, cap cells, and intermingled cells. The

53 anterior most cells of the ovary are the sheath cells, and a sub-population of these apically

54 positioned cells undergo two cell migration events during larval ovary development. First,

55 a population of sheath cells called swarm cells migrates from the anterior to the posterior

56 of the ovary to form the basal region in the mid third larval instar stage (COUDERC *et al.*

57 2002; GREEN II AND EXTAVOUR 2012). Secondly, in the late third instar and early pupal

58 stages, sheath cells migrate from the apical to the basal region, traversing in between

59 terminal filaments (KING *et al.* 1968). These sheath cells lay down basement membrane

60 in their path, which encapsulates developing ovarioles (KING 1970).

61 Terminal filaments are stacks of cells located just below the sheath cells in the

62 anterior larval ovary. They are formed by a process of progressive intercalation of

63 flattened cells into stacks, and stack formation occurs in a "wave" that proceeds from the

64 medial to the lateral side in the larval ovary (**Figure 1A**) (SAHUT-BARNOLA *et al.* 1995).

65 Morphogenesis in larval ovary and the mechanisms controlling the process are not

66 completely understood.

67 The genes *bric á brac 1 (bab1), bric á brac 2 (bab2)* and *engrailed (en)* are

68 expressed in the terminal filaments and essential for terminal filament cell differentiation

69 and terminal filament assembly (GODT AND LASKI 1995; SAHUT-BARNOLA *et al.* 1995;

70 COUDERC *et al.* 2002; BOLÍVAR *et al.* 2006). We previously showed that the Hippo signaling

71 pathway controls the regulation of cell proliferation in somatic cells, thereby affecting the

72 number of terminal filaments and their constituent terminal filament cells (SARIKAYA AND

73 EXTAVOUR 2015). During early terminal filament formation, Actin and Armadillo (arm)

74 proteins deposited in the region between terminal filaments make a scaffold to flatten and

75 intercalate terminal filament cells (GODT AND LASKI 1995; SAHUT-BARNOLA *et al.* 1995;

76 CHEN *et al.* 2001). Expression of the protein cofilin (*twinstar*) is required in terminal

77 filament and apical cells for actin-based change in cell shape, and loss of cofilin causes

78 a reduction in terminal filament and apical cell numbers (CHEN *et al.* 2001).

79   Normal growth of an ovary depends on the homeostatic of proliferation of the

80   somatic and germ line tissues (GILBOA AND LEHMANN 2006; GILBOA 2015). This balance

81   between somatic and germ line tissue populations is achieved by regulation of

82   proliferation, differentiation and apoptosis of stem cell populations of somatic and germ

83   cell lineages (SAHUT-BARNOLA *et al.* 1995; SAHUT-BARNOLA *et al.* 1996). Somatic cells

84   called intermingled cells interact with the germ cells and control their proliferation (LI *et*

85   *al.* 2003; GILBOA AND LEHMANN 2006; SARIKAYA AND EXTAVOUR 2015; LAI *et al.* 2017;

86   PANCHAL *et al.* 2017; LI *et al.* 2019). *Notch, hedgehog, Mitogen Activated Protein Kinase*

87   *(MAPK)* and *Epidermal growth factor receptor* (*EGFR)* signaling pathways, as well as the

88   transcription factor *traffic jam* maintain the germ line stem cell niche (BESSE *et al.* 2005;

89   SONG *et al.* 2007; MATSUOKA *et al.* 2013; SARIKAYA AND EXTAVOUR 2015), which is

90   established at the posterior of each terminal filament.

91   Recent work by Slaidina and colleagues used single-cell transcriptomics to

92   describe the gene expression profiles of the various cell types of the late third instar larval

93   ovary (SLAIDINA *et al.* 2020). They sub-divided terminal filament cells into anterior or

94   posterior cell types, and sheath cells into migratory or non-migratory cell types, based on

95   gene expression patterns of the single cell sub-populations. (SLAIDINA *et al.* 2020). While

96   this study examined a single time point of ovary development, given that ovary

97   morphogenesis is a temporal process, we hypothesize that changes in gene expression

98   patterns over the course of development may be important to regulate morphogenesis.

99   Thus, a gene expression study across the developing stages of larval ovary would

100   advance our understanding of the transcriptomic regulation of ovarian morphogenesis.

101   Although all major conserved animal signaling pathways are known to be involved

102   in ovarian morphogenesis (TWOMBLY *et al.* 1996; COHEN *et al.* 2002; HUANG *et al.* 2005;

103   SONG *et al.* 2007; GANCZ AND GILBOA 2013; GREEN AND EXTAVOUR 2014; SARIKAYA AND

104   EXTAVOUR 2015; KUMAR *et al.* 2020), a systematic gene expression profile of a developing

105   ovary is lacking. Such system-wide gene expression data for the ovary throughout

106   terminal filament morphogenesis, including the potentially distinct transcriptional profiles

107   of germ cells and somatic cells, could shed light on the processes involved in the

108   maintenance of cell types necessary to shape the ovary and control the number of

109   ovarioles.

110   To this end, we measured gene expression during the development of the larval

111   ovary by systematically staging and sequencing mRNA from whole ovaries before, during

112   and after terminal filament formation. Furthermore, we separated somatic and germ line

113   tissue types at each of these stages to analyze tissue-specific gene expression. We

114   compared the gene expression profiles across tissues and also across stages of ovary

115   development. We then employed functional enrichment analysis to determine the different

116   biological functions active in the three larval developmental stages and two tissue types

117   that could yield information on ovary morphogenesis. This dataset is an important

118   temporal and tissue specific gene expression resource for the insect developmental

119   biology community to understand early ovary development.

120

121   **RESULTS**

122   **Staging larval ovary development during terminal filament formation**

Tarikere, Ylla & Extavour Page 6

123   We divided the developing *Drosophila* larval ovary into three stages during terminal

124   filament formation and used RNA-seq to quantify gene expression at these stages

125   (**Figure 1A**). First, we considered an early stage of terminal filament formation at the

126   early third instar larva (72 hours After Egg Laying, 72h AEL), when terminal filament

127   assembly is initiating (GODT AND LASKI 1995) (**Figure 1A-i)**. Second, we assigned the

128   middle (mid) stage (96h AEL) as 24 hours after the early stage, at the midway point of

129   terminal filament assembly (GODT AND LASKI 1995) (**Figure 1A-ii**). Third, the late stage

130   (120h AEL) was defined as the time point of white pupa formation (when the larvae

131   become immobile at the larval to pupal transition (ASHBURNER *et al.* 2005)), which

132   occurs 24 hours after the middle stage (**Figure 1A-iii**). At the white pupa stage, terminal

133   filament assembly is complete and the number of terminal filaments reflects the number

134   of adult ovarioles (HODIN AND RIDDIFORD 2000) (**Figure 1B**).

135   We dissected these three stages of developing ovaries from larvae obtained from

136   synchronized eggs and sequenced the transcripts present at each stage from pools of

137   30-100 ovaries **(Supplementary Table S1)**. We aligned reads to the *Drosophila*

138   *melanogaster* genome (FlyBase v6.36), which yielded between 88.49% and 98.06% of

139   reads aligned per sample (**Supplementary Figure S1, Supplementary Table S1**).

140   Clustering analysis based on the variance-stabilizing transformation (VST) of the gene

141   counts of each sample confirmed that the three biological replicates of each stage

142   clustered together, and that the three stages were well separated, as reflected by the

143   principal component analysis (PCA) and the dendrogram of the hierarchical analysis

144   (**Figure 2A-B**). Furthermore, the dendrogram visualization of the hierarchical clustering

145   results revealed that the mid stage was more similar in expression profile to the early

146   stage than to the late stage. This indicates a more pronounced transcriptomic change at

147   the transition from mid to late, than from early to mid, despite the fact that the same

148   chronological amount of time had elapsed between each stage.

149

150   **Differential gene expression analysis of whole ovary samples at different stages**

151   We analyzed the transcriptional differences between each stage and the successive one,

152   thus performing a differential expression analysis comparing early to mid and mid to late

153   transitions, using DESeq2 (LOVE *et al.* 2014) with a threshold of p<0.01 (see Methods).

154   We found a significantly higher number of genes differentially expressed in the mid to late

155   transition (2,727 genes), than in the early to mid transition (685) (**Figure 2C,**

156   **Supplementary Table S2**). Interestingly, from early to mid stages twice as many genes

157   were down-regulated (480) as upregulated (206), while from mid to late stages

158   approximately the same proportion of genes were upregulated (1,264) and

159   downregulated (1,463). We then identified the genes that were differentially expressed in

160   one stage as compared to the other two stages, with the aim of revealing genes with

161   stage-specific over- or under-expression. We found that early and late stages had many

162   more over-expressed genes (1,434 and 1,626 respectively) than the mid stage (538)

163   (**Figure 2D, Supplementary Table S3**). A heatmap representing the expression levels

164   of the stage-specific overexpressed genes clearly separates the three groups of genes

165   (**Figure 2E**). The first group in the heatmap contains the 1,478 genes that are highly

166   expressed specifically at early stages, with less expression at mid stages and very low

167   expression at the late stage. Another large group of 1,618 genes are highly expressed

168   specifically at late stages, and show low expression at early and mid stages. Finally, we

Tarikere, Ylla & Extavour Page 8

169   identified a third and smallest group of 202 genes that are highly expressed at mid stages,

170   with some detectable expression at early stages, but little detectable expression at the

171   late stage (**Figure 2E**). These results are consistent with our previous observation that

172   there is a high gene expression similarity in early and mid-stages, and an increased

173   transcriptomic change from mid to late stages.

174

175   **Separation of somatic and germ line tissues in the developing ovary**

176   Given our ultimate interest in gene regulatory functions and dynamics during terminal

177   filament formation, we wished to understand the predicted functions of the many

178   differentially expressed genes across stages. We reasoned, however, that given the

179   different developmental numbers, roles and behaviors of germ line and somatic cells in

180   this developing organ, considering functional categories of differentially expressed genes

181   in these whole ovary samples would be only minimally informative. We therefore designed

182   an experimental strategy that allowed us to consider the transcriptional dynamics of the

183   germ line and soma separately, described below.

184   To understand the gene expression differences between the somatic and germ

185   line tissues of the ovary during terminal filament morphogenesis, we drove somatic and

186   germ line tissue-specific GFP expression using the UAS-GAL4 system (BRAND AND

187   PERRIMON 1993), using the drivers *bab:GAL4* and *nos:GAL4* respectively (see Methods).

188   We dissociated ovaries at the three stages described above, and isolated the GFP-

189   positive cells using Fluorescence-Activated Cell Sorting (FACS). Cellular debris was

190   eliminated with gate R1, non-singlets were eliminated by gate R2, and the R3 gate

191   selected for GFP positive cells. A combination of the three gates yielded singlet GFP

192   positive cells, minimizing the possibility of tissue contamination by undissociated cells.

193   When similar number of ovaries were used to obtain sorted cells for somatic and germ

194   line tissue-types, we found larger number of somatic cells as compared to germ cells as

195   expected, indicating a successful separation of the desired tissue type. (**Supplementary**

196   **Figure S2**).

197   With this method, we obtained tissue-specific transcriptomes of somatic and germ

198   line tissues at the same three stages of terminal filament development used to generate

199   the whole ovary dataset. We sequenced three biological replicates for all datasets, and

200   retained replicates that had at least 10 million reads. The number of reads aligned to the

201   genome ranged from 11 to 81 million. Greater than 94% of reads aligned in all datasets,

202   with the single exception of one dataset with 88% of aligned reads **(Supplementary**

203   **Table S1, Supplementary Figure S3**). The PCA analysis based on the counts

204   normalized by variance-stabilization transformation (VST) shows a clear separation of

205   somatic and germ cell libraries along the first principal component, suggesting a

206   successful separation of cell types by FACS (**Figure 3A**). For the somatic samples, the

207   three biological replicates cluster closely together (**Figure 3A-B**) while the different stages

208   are separated from each other in the second principal component. The structure of the

209   dendrogram for the somatic samples resembles that of the whole ovary, in which early

210   and mid-stages are closer to each other than either is to the late stage. As for the germ

211   cell libraries, unlike the biological replicates of the early and late stages, the mid-stage

212   replicates do not cluster together. A possible explanation is the low number of reads from

213   the sample Mid-1 (**Supplementary Table S1**).

214   To further assess the successful separation of somatic and germ cells, we checked

215   the expression of well-known tissue-type-specific markers. The genes *nanos* and *vasa*

216   are two genes known to be specifically expressed in germ cells in the ovary (SCHUPBACH

217   AND WIESCHAUS 1986; LEHMANN AND NUSSLEIN-VOLHARD 1991). Both genes show higher

218   expression in the germ cell libraries than in the somatic cell libraries (mean $\log_2$(Fold

219   Change) of 8.37 for *nanos,* and 8.23 for *vasa*) (**Figure 3C**), confirming that the preparation

220   and sequencing of the germ cell libraries successfully captured the germ cells and their

221   RNAs, and suggesting that germ cells were not present (or present only at very low levels)

222   in the somatic cell libraries. *bab1*, *bab2*, and *tj* are considered somatic gene markers

223   (SAHUT-BARNOLA *et al.* 1995; COUDERC *et al.* 2002) These three somatic markers display

224   higher expression levels in our somatic libraries than in the germ cell libraries at each

225   stage (mean $\log_2$(Fold Change) -0.31 for *bab1*, -1.3 for *bab2*, and -1.7 for *tj*) (**Figure 3D**).

226   However, in four of the 18 libraries, either *bab1* or *bab2* (but not *tj*) showed higher

227   expression levels in a specific germ cell library than in the somatic libraries. These specific

228   cases were as follows: (1) one early stage germ cell replicate had higher *bab1* levels than

229   one of the early somatic replicates; (2) two mid stage germ cell replicates had higher *bab1*

230   levels than the somatic replicates; (3) one late stage germ cell replicate had higher *bab1*

231   levels than the somatic replicates; (4) one mid stage germ cell replicate had higher *bab2*

232   levels than the somatic replicates. This could indicate that some somatic cells might have

233   been included in these particular germ cell libraries. Nonetheless, despite this putative

234   small amount of contamination, we can clearly differentiate both tissue types based on

235   their expression profiles as shown in the PCA (**Figure 3A**), suggesting that we captured

236   the transcriptional differences between cell types (**Figure 3A**) sufficiently to allow us to

237   achieve our goal of successfully retrieving the genes that are highly and differentially

238   expressed in each of these two tissues.

239

240   **Differential expression analysis of somatic and germ line tissues across all stages**

241   The differential expression analysis between the somatic and germ line tissues across all

242   three stages revealed 1,880 genes significantly upregulated (adjusted p-value<0.01) in

243   germ cells and 1,585 genes significantly upregulated in the somatic cells (**Figure 4A;**

244   **Supplementary Table S4**).

245   Among the 20 most significant genes (with the lowest adjusted p-value)

246   overexpressed in germ cells relative to somatic cells, we detected known germ line-

247   specific genes including piRNA biogenesis genes *Argonaute3 (AGO3), krimper (krimp)*,

248   and *tejas (tej)*, along with *Aubergine (aub)*, (BRENNECKE *et al.* 2007; OLIVIERI *et al.* 2010;

249   PATIL AND KAI 2010; SATO *et al.* 2015), *sisters unbound (sunn)* (KRISHNAN *et al.* 2014)*,*

250   *benign gonial cell neoplasm (bgcn)* (OHLSTEIN *et al.* 2000), and uncharacterized genes

251   including *CG32814* and *CG12851* on the chromosome 2R. As for the somatic cells, the

252   most significantly overexpressed gene relative to the germ cells is the cytochrome gene

253   *Cyp4p2*, whose role is unknown in the ovary, followed by cytochrome *Cyp4p1* and the

254   uncharacterized genes CG32581 and CG42329. Some genes known to play roles in the

255   ovary were also among this group, including the regulator of the niche cells and ecdysone

256   receptor *Taiman (tai)* (KÖNIG *et al.* 2011), and the regulator of vitellogenesis *apterous (ap)*

257   (GAVIN AND WILLIAMSON 1976).

258

Tarikere, Ylla & Extavour Page 12

259   **Temporally dynamic expression of genes previously studied in somatic ovary**

260   **development**

261   We explored the expression dynamics of some of the previously studied genes expressed

262   in the *Drosophila* ovary. To our knowledge, temporal gene expression studies in the larval

263   ovary for many these genes have not yet been conducted.

264   First, we considered the temporal expression patterns of some adhesion proteins

265   known to play a role in ovary development. *RanBPM* is an adhesion linker protein

266   expressed in the germ line niche in the adult ovary (DANSEREAU AND LASKO 2008). In our

267   dataset we see opposing trends of expression levels in somatic and germ line tissues,

268   such that in germ line tissue *RanBPM* expression decreases progressively from early to

269   mid to late stages, while in the somatic tissue it increases from early to mid to late stages

270   (**Supplementary Figure S4A**). Cofilin (encoded by the gene *twinstar*) is an adhesion

271   protein required for terminal filament cell rearrangement during terminal filament

272   morphogenesis, as well as for adult border cell migration (CHEN *et al.* 2001). Cofilin shows

273   similar germ line and somatic cell expression trends, with higher levels at early stages

274   that decrease progressively at mid and late stages (**Supplementary Figure S4B**).

275   We then looked at temporal expression of *RhoGEF64C* and *Wnt4,* genes involved

276   in cell motility. RhoGEF64C is a small apically localized RhoGTPase that regulates cell

277   shape and migration in the ovary (SIMOES *et al.* 2006). In our datasets we found

278   *RhoGEF64C* expressed at higher levels in early and late stage somatic cells than at mid

279   stages (**Supplementary Figure S4C**). *Wnt4* is involved in cell motility during ovarian

280   morphogenesis (COHEN *et al.* 2002) and is expressed in the posterior terminal filaments

281   and other somatic cell types of the third instar larval ovary (SLAIDINA *et al.* 2020). We

282   found *Wnt4* to be expressed in lower levels in early and middle stages while the

283   expression increases significantly in the late stage (**Supplementary Figure S4D**).

284   We also examined the temporal expression dynamics of a number of terminal

285   filament cell-type-specific genes previously identified in a single cell sequencing study of

286   the late third larval instar ovary (SLAIDINA *et al.* 2020). For example, *Diuretic hormone 44*

287   *receptor 2 (Dh442)* was identified as highly expressed in terminal filament cells (SLAIDINA

288   *et al.* 2020). In our datasets, we observed a significant increase in expression levels only

289   at the late stage relative to early and mid-stage expression levels (adjusted p-value

290   2.04×10^{-11}) (**Supplementary Figure S4E**). Additional genes known to function in terminal

291   filaments are *engrailed* and *patched* (FORBES *et al.* 1996; BESSE *et al.* 2005; BOLÍVAR *et*

292   *al.* 2006). In our datasets we observed *engrailed* expressed at lowest levels at the early

293   stage, showing a progressive increase in expression levels from mid to late stages.

294   *patched* showed a similar progressive increase across stages, with a significant increase

295   from early to mid-stage (**Supplementary Figure S4F-G**). Finally, we considered

296   members of the Fibroblast Growth Factor (FGF) signaling pathway, which controls sheath

297   cell proliferation in the pupal ovary (IRIZARRY AND STATHOPOULOS 2015). Three key genes

298   of this pathway, the FGF ligand *thisbe,* the FGF scaffolding protein *stumps* and the

299   upstream FGF signaling activator *heartless*, show significantly higher differential

300   expression levels at early to mid-stage than at mid to late stages (**Supplementary Figure**

301   **S4H-J**). These temporal profiles add to our understanding of the roles of these genes in

302   ovarian morphogenesis by suggesting distinct putative critical regulatory periods for

303   different genetic pathways.

304

Functional enrichment analysis of differentially expressed genes in somatic and germ line tissues across all stages

305
306

307 To gain insight into the general functional categories of genes likely involved in ovarian

308 germ cell and somatic behaviors during terminal filament development, we performed a

309 gene ontology (GO) enrichment analysis of the biological processes of differentially

310 expressed genes across cell types and developmental stages (ASHBURNER *et al.* 2000).

311 We found 31 level four GO-terms enriched (adjusted p-value<0.05) within the upregulated

312 genes in germ cells, and 188 level four GO-terms enriched in the upregulated genes in

313 somatic cells (**Supplementary Figure S5**). This analysis highlighted clear differences in

314 the biological functions performed by the genes expressed in each tissue. The GO-terms

315 enriched in the germ cells are primarily related to meiotic processes (9/31 contain the

316 words "meiosis" or "meiotic"), chromosome stability (6/31 contain the words

317 "chromosome" or "karyosome") and cell cycle (12/31 contain "cell cycle"). In contrast, the

318 GO-terms enriched in the somatic cells are principally related to cellular response (21/188

319 contain "response"), development (18/188), growth (16/188), morphogenesis (10/188)

320 cell migration (6/188 contain the word "migration") and signaling pathways (6/188).

321 To complement this GO enrichment analysis, we performed a KEGG pathway

322 enrichment analysis on the same cell-type-specific overexpressed genes. The KEGG

323 pathway database is a manually curated database of molecular interactions used to study

324 enrichment of genetic regulatory pathways in gene lists (KANEHISA AND GOTO 2000). With

325 this analysis, we identified nine KEGG pathways significantly enriched in the germ cells,

326 and 16 significantly enriched pathways in the somatic cells (adjusted p-value<0.05)

327 (**Figure 4B**). The KEGG pathways enriched in the germ cells are generally related to

328   meiosis and genome protection, while upregulated genes in the somatic cells are

329   enriched for pathways involved in cell proliferation and cell death, including the previously

330   identified *Hippo* (BARRY AND CAMARGO 2013; SARIKAYA AND EXTAVOUR 2015) and *MAPK*

331   (SHAUL AND SEGER 2007) signaling pathways.

332

333   **Stage- and tissue-specific differential gene expression analysis**

334   To obtain a finer-grained view of the dynamic regulation of terminal filament development,

335   we also performed differential expression analysis between the somatic and germ line

336   tissue types at each of the three stages. In the somatic cells the number of differentially

337   expressed genes between the early and mid-stages (867 genes) is lower than between

338   the mid and late stages (1,404 genes) (**Figure 5A; Supplementary Table S6**). To identify

339   genes with stage-specific upregulation, we compared each stage to the other two stages.

340   We identified a higher number of stage-specific upregulated genes in early (1,227) and

341   late stages (139) than at mid (1,409) (**Figure 5B; Supplementary Table S7**).

342   The germ cells, in general, display fewer differentially expressed genes between

343   stages than the somatic cells. From early to mid-stages there are twice as many

344   differentially expressed genes (557 genes) as from mid to late stages (248 genes) (**Figure

345   5D; Supplementary Table S8**). In terms of stage-specific upregulated genes, the highest

346   number of such genes are found at early stages (209), followed by mid (186), and late

347   (84) stages (**Figure 5E; Supplementary Table S9**).

348   To explore the functions of the stage-specific upregulated genes in each tissue type, we

349   performed a GO analysis of biological functions and KEGG pathway enrichment analysis

Tarikere, Ylla & Extavour Page 16

350   on these six sets of genes (upregulated at early, mid, and late stages in germ and somatic

351   cells). The GO enrichment analysis of the genes differentially expressed in somatic cells

352   over time (**Supplementary Table S6**) revealed that four key biological processes are

353   consistent throughout all three stages including the mid stage, which has the smallest

354   number of differentially expressed genes across stages. Specifically, these are the GO

355   terms taxis, cell growth, actin filament-based process and cell adhesion. At early and late

356   stages, we additionally observe many key biological processes related to morphogenesis

357   in the somatic cells, including cell proliferation, differentiation and migration. Considering

358   gene expression levels specific to each stage, we identified 1,227, 139, and 1,409

359   upregulated genes at early, mid, and late stages respectively. The upregulated genes at

360   early and late stages were enriched for 97 and 764 GO-terms of biological process, while

361   none were enriched in the mid-stage (**Supplementary Figure S6**). As for the KEGG

362   pathway enrichment analysis, there were two enriched pathways at early stages, one at

363   mid-stage and 17 in late stages (**Figure 5C**). This analysis allowed us to pinpoint the

364   stage(s) at which specific pathways were enriched in somatic cells relative to germ cells,

365   which included Apoptosis, *Hippo* signaling, and *MAPK* signaling. In addition, we detected

366   some signaling pathways enriched only in somatic cells at late stages, such as the

367   *Hedgehog*, *FoxO*, and *Notch* pathways (**Figure 5C**).

368   Given the known role of the Hippo pathway in cell proliferation (BARRY AND

369   CAMARGO 2013), and specifically in terminal filament cell and terminal filament number

370   regulation (SARIKAYA AND EXTAVOUR 2015), we proceeded to analyze the expression

371   patterns of the genes belonging to the core Hippo signaling pathway. We found that most

372   Hippo pathway core genes display increasing expression levels from early to mid to late

373   stages, with the exception of the expression of the core gene *Rae1* which progressively

374   decreases in expression level from early to late stages (**Supplementary Figure S8**).

375   In the germ cells, across stages we find fewer processes directly involved in

376   development and morphogenesis with gene ontology categories belonging to meiosis and

377   cell cycle (**Supplementary Figure S6**). Among the 209, 186, and 84 upregulated genes

378   in germ cells at early, mid, and late stages respectively, only one KEGG pathway

379   (Ribosome,) and one biological process GO-term (cytoplasmic translation) of were found

380   significantly enriched at early-stages. No such enrichment was detected at mid-stages,

381   and three KEGG pathways were enriched at late stages (**Supplementary Figure S7**).

382

383   **Uncharacterized genes**

384   The detection of uncharacterized genes among the top differentially expressed genes in

385   germ cells drove us to ask if there were any differences in the proportion of

386   uncharacterized genes in each set of differentially expressed genes. We found that in the

387   genes significantly upregulated in somatic cells compared to germ cells, 29.15% are

388   categorized as "uncharacterized proteins" in FlyBase (LARKIN *et al.* 2021), while within the

389   significantly upregulated genes in germ cells, the proportion of uncharacterized genes

390   was 39.10%. Within the stage-specific upregulated genes, the proportion of

391   uncharacterized genes remained constant (between 28.96% and 29.63%) in somatic

392   cells, while in germ cells it increased from 29.08% in early stages, to 34.83% in mid

393   stages, and to 37.40% in late stages (**Supplementary Figure S9).**

394

Expression of cell type-specific markers

395

396   A previous single cell RNA-sequencing dataset of the late third stage larval ovary ()

397   SLAIDINA *et al.* (2020) identified transcriptional profile clusters interpreted as indicative of

398   cell types, and suggested gene markers associated with each cell type. To determine

399   whether the cell types identified at this late stage might also be present at earlier

400   developmental stages than that previously assessed, we examined the expression levels

401   of those suggested marker across our datasets. As expected, the majority of the germ

402   cell markers are highly expressed in our germ cell libraries and expressed only at low

403   levels in the somatic cells (**Supplementary Figure S10**). Among the somatic markers

404   detected in our somatic tissue libraries, we do not observe any particular temporal

405   expression pattern specific to a given somatic cell type. Nevertheless, we clearly

406   distinguish two groups of somatic markers (**Supplementary Figure S11**). One group is

407   composed of somatic markers whose expression levels are highest at early and mid-

408   stages, and decay at the late stage, and a larger group of makers that are less strongly

409   expressed at early stages, show increased expression at the mid stage, and show highest

410   expression at late stages. By contrast, the germ cell markers detected in our germ cell

411   libraries do not display any clear temporal expression pattern. Instead, most of these

412   genes were expressed at similar levels across the three studied stages (**Supplementary**

413   **Figure S12**). This is consistent with our previous observations that the germ line dataset

414   is not enriched for any signaling pathway directly implicated in development during these

415   three times points as the somatic cells do (Supplementary Figure S6).

416

417   **DISCUSSION**

Tarikere, Ylla & Extavour Page 19

Temporal gene expression during ovary morphogenesis

418

419   We systematically staged and sequenced entire larval ovaries to generate a gene

420   expression dataset during terminal filament formation. We then separated somatic and

421   germ line tissues during these stages and generated tissue-specific transcriptomes. While

422   the development of the *Drosophila* ovary has been studied for the last several decades,

423   and progress has been made on identifying the roles of some signaling pathways in its

424   morphogenesis (COHEN *et al.* 2002; BESSE *et al.* 2005; GILBOA AND LEHMANN 2006; GANCZ

425   *et al.* 2011; GANCZ AND GILBOA 2013; MATSUOKA *et al.* 2013; GILBOA 2015; IRIZARRY AND

426   STATHOPOULOS 2015; LENGIL *et al.* 2015; MENDES AND MIRTH 2016; PANCHAL *et al.* 2017)

427   to our knowledge, there are no publicly available transcriptomes of developing larval

428   ovaries of *Drosophila*. Recent articles have reported single cell RNA-sequencing for

429   *Drosophila* ovaries, focusing either on a single larval time point or on adult ovaries (JEVITT

430   *et al.* 2020; RUST *et al.* 2020; SLAIDINA *et al.* 2020; SLAIDINA *et al.* 2021). Our stage and

431   tissue-type specific data thus represent a valuable complementary transcriptomic

432   resource on the morphogenesis of the larval ovaries of *Drosophila*, a complex process

433   that ultimately influences reproductive capacity.

434   In the whole ovary dataset, the increased differential gene expression at the mid-

435   late transition and at the late stages, enriched the expression of genes in key signaling

436   pathways that are necessary in ovary development (**Figure 2C-D**). Signaling pathways

437   found exclusively enriched at the late stages suggest that morphogenetic processes of

438   the larval ovary at these stages operate through these key pathways.

439

Separating somatic and germ line tissue

440

Tarikere, Ylla & Extavour Page 20

441   FACS-based tissue-type separation coupled with RNA-seq proved to be a successful way

442   to obtain transcriptomes of somatic and germ line cells during larval ovary development.

443   The overall expression profile of the somatic cells across the three studied stages is

444   similar to the profile of the whole ovary dataset in terms of the proportion of differentially

445   expressed genes across stages(**Figures 2D, 5B**) This may be because since the number

446   of somatic cells in the whole ovary is higher than that of the germ cells at all stages

447   (**Supplementary Figure S2**), the whole ovary gene expression profile is likely dominated

448   by somatic cell expression. Only upon physically separating the germ cells from the

449   somatic cells, could we observe a different gene expression pattern in the germ line. Germ

450   cells showed the highest number of differentially expressed genes at early stages, and

451   the lowest numbers at late stages (**Figure 5E**). In contrast, the highest number of

452   differentially expressed genes of somatic stages occurs at the late stages, and the lowest

453   number at the mid stage (**Figure 5B**).

454   The results of the functional enrichment analyses in somatic and germ line tissue

455   reveal distinct functions and pathways likely operate in these tissues during larval ovary

456   development. Germ cells may be especially sensitive to DNA damage given their role in

457   propagating genetic material, which we speculate may explain the enrichment of

458   processes related to nucleotide replication, recombination and repair in our analysis

459   (Figures 4, S5). Similarly, we observed many genes of the piRNA pathway (e.g., *AGO3*,

460   *aub*, *krimp*, *tej*), which protect the genome from transposable elements (**Supplementary**

461   **Table S6**) (SATO AND SIOMI 2020) among the top significantly enriched genes in germ

462   cells. On the other hand, the somatic tissue is enriched for different signaling pathways

463   including Hippo*,* MAPK, and apoptosis (**Figures 5C, S6**), which are known to play a role

464   in either larval or adult ovary morphogenesis (LYNCH *et al.* 2010; KHAMMARI *et al.* 2011;

465   ELSHAER AND PIULACHS 2015; SARIKAYA AND EXTAVOUR 2015).

466   Detection of a higher number of uncharacterized genes in germ line tissue datasets

467   than in the somatic tissue (**Supplementary Figure S9**), suggest that our understanding

468   of the genetic regulation of the germ line in developing ovaries is still incomplete. Datasets

469   like the one herein provided help to identify new genes that could be important for ovary

470   morphogenesis.

471

472   **Cell adhesion and migration during ovary morphogenesis**

473   We assessed the temporal dynamics of genes expressed in specific cell types during

474   development to serve as generators of new hypotheses to understand the role of genes

475   and pathways during morphogenesis. *RhoGEF64C* is a RhoGTPase with some role in

476   regulating control cell shape changes that lead to epithelial cell invagination (SIMOES *et*

477   *al.* 2006; TORET AND LE BIVIC 2021). In a genome-wide association study on ovariole

478   number phenotypes in natural populations of *Drosophila*, *RhoGEF64C* driven in somatic

479   tissue had a significant effect on adult ovariole number (LOBELL *et al.* 2017). The

480   significant increase in expression of *RhoGEF64C* we observed in early and late stages

481   (Figure S4C) suggests its role in somatic cell shape and migration in both early and late

482   stages. We see the GO processes of adhesion and migration enriched in somatic cells

483   but not in germ cells (**Supplementary Figure S5**), which could mean that these

484   morphogenetic processes in germ cells are minimal compared to somatic cells.

485   GO-terms related to cell adhesion, motility and taxis were enriched in all three

486   stages in somatic cells (**Supplementary Figure S6**). Previous studies have shown

487   signaling pathways involved in ovary development to affect cell adhesion and migration

488   processes (COHEN *et al.* 2002; LI *et al.* 2003; BESSE *et al.* 2005; LAI *et al.* 2017). Migratory

489   events in mid to late stages of the larval ovary have been described for two ovarian cell

490   types, swarm cells and sheath cells (SAHUT-BARNOLA *et al.* 1995; SAHUT-BARNOLA *et al.*

491   1996; GREEN II AND EXTAVOUR 2012; SLAIDINA *et al.* 2020). Migrating sheath cells in the

492   late third instar larvae lay the basement membrane in between terminal filaments to form

493   ovarioles (KING *et al.* 1968).

494   The FGF signaling pathway supports terminal filament cell differentiation in the

495   early larval stages through *thisbe (ths)* and upstream activator *heartless (htl)*, and also

496   controls sheath cell proliferation in late larval and pupal stages (IRIZARRY AND

497   STATHOPOULOS 2015). In our dataset, we observe that in somatic cells these FGF pathway

498   genes show a significant progressive upregulation from early to mid and from mid to late

499   stages (**Supplementary Figure S4I-J**). Consistently, *ths* and *stumps* were identified as

500   markers of a distinct migratory ovarian cell population, the sheath cells (SLAIDINA *et al.*

501   2020). *stumps* is expressed in stages corresponding to our "late" stage in the

502   differentiating terminal filament cells and at later stages (144h AEL), also in migratory

503   sheath cells (IRIZARRY AND STATHOPOULOS 2015).

504

505   **Functional enrichment analysis and signaling pathways**

506   Our results show that in the late stage of somatic cells there is an increase in expression

507   of genes involved in multiple signaling pathways, including the Wnt, MAPK, Hippo,

508   Hedgehog, FoxO, TGF and Notch pathways (**Figure 5C**). The molecular mechanisms of

509   all these signaling pathways during larval ovary development have not yet been

510   extensively studied, but all of them have been functionally implicated in ovariole number

511   determination by a large-scale genetic screen (KUMAR *et al.* 2020).

512   We previously showed that Hippo signaling pathway controls proliferation of

513   somatic cells, which affects terminal filament number (SARIKAYA AND EXTAVOUR 2015).

514   Our differential gene expression data show that members of the Hippo pathway are

515   significantly differentially expressed in the somatic tissue (**Figure 4B; Supplementary**

516   **Figure S8**), and all of its core genes except one show a progressive increase in

517   expression levels across the three studied stages (**Supplementary Figure S8**). Loss of

518   function mutations in Yki, an effector of the Hippo signaling pathway, cause increased

519   growth and reduced apoptosis through an increase in the levels of the cell cycle protein

520   Cyc E and the apoptosis inhibitor *Diap1* (HARVEY *et al.* 2003; HUANG *et al.* 2005). In our

521   somatic cell datasets, we observe *Diap1* transcript levels significantly increase from early

522   to late stages, and those of CycE increase from mid to late stages (**Supplementary**

523   **Figure S13A-B**). However, the Apoptosis KEGG pathway appears significantly enriched

524   in the somatic late stage (**Figure 5C; Supplementary Figure S6**) . Furthermore,

525   apoptosis-related genes *Dronc* and *Dark*, which form the apoptosome (**Supplementary**

526   **Figure 13C, D**) (YUAN *et al.* 2011), are also significantly upregulated in the late stage, as

527   are the caspases *Dcp-1*, *Drice*, and *Dredd* (**Supplementary Figure 13C, E-G**) (HARVEY

528   *et al.* 2001). Thus, we observe both an upregulation of apoptosis and an upregulation of

529   the apoptosis inhibition genes in late stage somatic cells. This could mean that genes

530   controlling apoptosis both positively and negatively are acting to exert tight control of this

531   process. Alternatively, our observations may reflect that each process is upregulated

532   within different somatic cell types.

533   Cap cells and intermingled cells are somatic cells that interact with the germ cells

534   for the maintenance of germ line stem cell niches (LI *et al.* 2003; SONG *et al.* 2007). The

535   Notch signaling pathway, enriched in the late-stage somatic dataset (**Figure 5C;**

536   **Supplementary Figure S6**), is required for cap cell fate (PANCHAL *et al.* 2017). We

537   observed an expression level increase in Notch pathway components at late stages,

538   suggesting that the role of the Notch pathway in cap cell fate determination may be

539   particularly important at mid to late stages of larval ovary development.

540   Components of the TGFβ pathway, enriched in late stage somatic cells in our

541   dataset (**Figure 5C**), are known to contribute to ovarian development. These include the

542   TGFβ component *decapentaplegic* (*dpp)*, previously documented as expressed in all

543   larval ovarian somatic cells (SATO *et al.* 2010) and also in the larval-pupal stage cap cells

544   and intermingled cells of the germ line stem cell niche, where it promotes proliferation

545   and represses differentiation of primordial germ cells (GILBOA AND LEHMANN 2004;

546   MATSUOKA *et al.* 2013). The activin pathway, a branch of the TGFβ pathway (PANGAS AND

547   WOODRUFF 2000), controls terminal filament cell proliferation and differentiation (LENGIL

548   *et al.* 2015). We find that the activin receptor *baboon* shows a significant expression level

549   increase in the late stage somatic cells (adjusted p-value of 0.002602) and could indicate

550   its role in terminal filament cell differentiation in late stage.

551

552   **Conclusions**

553   Here we provide a dataset that explores gene expression during larval ovary development

554   and morphogenesis, which is crucial to understand how the ovary is shaped in early

555   stages to develop into a functional adult organ. This study advances our understanding

556   of the process of building an ovary and regulating the morphogenesis processes. More

557   importantly, this work offers a dataset for the developmental biology community to probe

558   the genetic regulation of larval ovarian morphogenesis.

559

560   **MATERIALS AND METHODS**

561   **Fly Stocks**

562   Flies were reared at 25°C at 60% humidity with food containing yeast and in uncrowded

563   conditions. The following two fly lines were obtained from the Bloomington Stock Center:

564   *w[*]; P{w[+mW.hs]=GawB}bab1[Pgal4-2]/TM6B, Tb[1]* (abbreviated herein as *bab:GAL4*;

565   stock number 6803), *P{w[+mC]=UAS-Dcr-2.D}1, w[1118]; P{w[+mC]=GAL4-nos.NGT}40*

566   (abbreviated herein as *nos:GAL4*; stock number 25751). *w[1118], P[UAS Stinger]*

567   (abbreviated herein as *UAS:Green Stinger I*, (Barolo *et al.* 2000) used for GFP

568   expression was a gift from Dr. James Posakony (University of California, San Diego).

569   Crosses were set with 100-200 virgin UAS females and 50-100 GAL4 males in a 180 ml

570   bottle containing 50ml standard fly media one day prior to egg laying.

571

572   **Staging larvae**

573   To obtain uniformly staged larvae for the experiments, a protocol was devised to collect

574   eggs that were near-synchronously laid, from which the larvae were then collected. To

575   obtain a desired genotype, crosses were set as described above. The cross was set at

576   25°C at 60% humidity and left overnight to mate. Hourly egg collections were set up on

577   60 mm apple juice-agar plates (9 g agar, 10 g sugar, 100 ml apple juice and 300 ml water)

578 with a pea-sized spread of fresh yeast paste (baker's yeast granules made into a paste

579 in a drop of tap water). Eggs were collected hourly for eight hours. The first two collection

580 plates were discarded to remove asynchronously laid eggs that may have been retained

581 inside the females following fertilization. Staged first instar larvae were collected into vials

582 24 hours after egg collection. Larvae at 72h AEL (hours After Egg Laying) were

583 designated as early stage, at 96h AEL as middle stage and at 120h AEL as late stage of

584 Terminal Filament development. For a step-by-step detailed protocol see

585 **Supplementary File 1**.

586

587 **Dissection and dissociation of larval ovary**

588 Staged larvae were collected for dissection every hour. The head of the larva was

589 removed with forceps and the cuticle and gut were carefully pulled with one forceps while

590 holding the fat body with another forceps. This process left just the fat bodies in the

591 dissection dish as long as the larvae were well fed and fattened with yeast. Ovaries

592 located in the center of the length of each fat body were then dissected free of the fat

593 body using an insulin syringe needle (BD 328418). Ovaries dissected clear of fat body

594 were collected in DPBS (Thermo Fisher 14190144) and batches of 20-30 ovaries in DPBS

595 were kept on ice until dissociation. Ovaries were harvested hourly at the appropriate

596 times, placed on ice immediately following dissection, and maintained on ice for a

597 maximum of four hours before dissociation and subsequent FACS processing.

598 Dissociation of the larval ovary required two enzymatic steps. After seven hours of

599 dissection, batches of dissected ovaries were placed in 0.25% Trypsin solution (Thermo

600 Fisher 25200056) for ten minutes at room temperature in the cavity of a glass spot plate

601 (Fisher Scientific 13-748B). They were then transferred to another cavity containing 2.5

602 % Liberase (5 g Liberase reconstituted in 2ml nuclease free water; Sigma 5401119001)

603 and teased apart with tungsten needles until most of the clumps were separated and left

604 (without agitation) at room temperature for ten minutes. Using a 200µl pipette with a filter

605 tip (pre rinsed in 1X PBS), the dissociated cells in Liberase were pipetted up and down

606 gently ten times to uniformly mix and separate the cells. The cell suspension was then

607 transferred to an RNA Lobind tube (Eppendorf 8077-230) and placed on a vortexer for 1

608 minute. Meanwhile the well was rinsed in 1.4 ml of PBS by pipetting repeatedly. This PBS

609 was then mixed with the Liberase mixture and vortexed for another minute, and the entire

610 sample was then was placed on ice. This sample was then taken directly to the FACS

611 facility on ice along with an RNA Lobind collection tube containing 100-200µl Trizol

612 (Thermo Fisher 15596026). For a step-by-step detailed protocol see **Supplementary File**

613 **1**.

614

615 **Flow Sorting GFP-positive cells**

616 The dissociated tissue sample was sorted in a MoFlo Astrios EQ Cell sorter (Beckman

617 Coulter) run with Summit v6.3.1 software. The dissociated cell solution was diluted and a

618 flow rate of 200 events per second was maintained with high sorting efficiency (< 98%)

619 during the sorting process. A scatter gate (R1) was employed to eliminate debris

620 (**Supplementary Figure S6**) and a doublet gate (R2) was used to exclude non-singlet

621 cells. A 488 nm emission Laser was used to excite the GFP and the collection was at 576

622 nm. The GFP-positive cells were designated in gate R3 and sorted directly into Trizol.

623 The resulting cells collected in Trizol were frozen immediately by plunging the tube in

624   liquid nitrogen and then stored at -80°C until RNA extraction. A single replicate consisted

625   of at least 1000 cell counts pooled from FACS runs.

626

RNA extraction

628   Flow-sorted cells were stored at -80°C were thawed at room temperature. Trizol contents

629   were lysed with a motorized pellet pestle (Kimble 749540-0000). Zymo RNA Micro-Prep

630   kit (Zymo Research R2060) was used to isolate RNA from the Trizol preparations. Equal

631   amounts of molecular grade ethanol (Sigma E7023) were added to Trizol and mixed well

632   with a pellet pestle, then pipetted onto a spin column. All centrifugation steps were done

633   at 10,000g for one minute at room temperature. The column was washed with 400µl Zymo

634   RNA wash buffer and then treated with Zymo DNase (6U/µl) for 15 minutes at room

635   temperature. The column was then washed twice with 400µl Zymo RNA Pre-wash buffer

636   and once with Zymo RNA wash-buffer. The RNA was eluted from the column in 55 µl of

637   Nuclease-free water (Thermo Fisher 10977015). The RNA obtained was quantified first

638   using a NanoDrop (Model ND1000) spectrophotometer and then using a high sensitivity

639   kit (Thermo Fisher Q32852) on a Qubit 3.0 Fluorometer (Thermo Fisher Q33216). It was

640   also checked for integrity on a high sensitivity tape (Agilent 5067-5579) with an electronic

641   ladder on an Agilent Tapestation 2200 or 4200. RNA extraction from staged whole ovaries

642   was carried out by crushing entire ovaries in Trizol and following the same protocol

643   described above. For a step-by-step detailed protocol see **Supplementary File 1**.

644

Library Preparation

646   cDNA libraries were prepared using the Takara Apollo library preparation kit (catalogue

647   # 640096). Extracted RNA samples were checked for quality using Tapestation tapes.

648   50µl of RNA samples were pipetted into Axygen PCR 8-strip tubes (Fisher Scientific 14-

649   222-252) and processed through PrepX protocols on the Apollo liquid handling system.

650   mRNA was isolated using PrepX PolyA-8 protocol (Takara 640098). The mRNA samples

651   were then processed for cDNA preparation using PrepX mRNA-8 (Takara 640096)

652   protocol. cDNA products were then amplified for 15 cycles of PCR using longAmp Taq

653   (NEB M0287S). During amplification PrepX RNAseq index barcode primers were added

654   for each library to enable multiplexing. The amplified library was then cleaned up using

655   PrepX PCR cleanup-8 protocol with magnetic beads (Aline C-1003). The final cDNA

656   libraries were quantified using a high sensitivity dsDNA kit (Thermo Fisher Q32854) on a

657   Qubit 3.0 Fluorometer (Thermo Fisher Q33216). cDNA content and quality were

658   assessed with D1000 (Agilent 5067-5582) or High sensitivity D1000 tape (Agilent 5067-

659   5584, when cDNA was in low amounts) on an Agilent Tapestation 2200 or 4200. For a

660   step-by-step detailed protocol see **Supplementary File 1**.

661

662   **Sequencing cDNA libraries**

663   Libraries were sequenced on an Illumina HiSeq 2500 sequencer. Single end-50bp reads

664   were sequenced on a high-throughput flow cell. Libraries of varying concentrations were

665   normalized to be equimolar, the concentrations of which ranged between 2-10nM per

666   lane. All the samples in a flow cell were multiplexed and later separated on the basis of

667   unique prepX indices to yield at least 10 million reads per library. The reads were

668   demultiplexed and trimmed of adapters using the bcl2fastq2 v2.2 pipeline to yield final

669   fastq data files.

670

RNA-seq data processing

672   The *D. melanogaster* genome assembly and gene annotations were obtained from

673   FlyBase version dmel_r6.36_FB2020_05 (LARKIN *et al.* 2021). The reads were aligned

674   with RSEM v1.3.3 (LI AND DEWEY 2011) and using STAR v2.7.6a as read aligner (DOBIN

675   *et al.* 2013) we obtained the gene counts in each library. Because some of the tissue-

676   specific biological samples were sequenced in more than one lane or run, and therefore

677   the reads were split into multiple fastq files, the gene counts belonging to the same

678   biological sample were summed. Gene counts in each dataset were normalized with the

679   variance stabilizing transformation (VST) method implemented in the DESeq2 v1.26.0

680   (LOVE *et al.* 2014) R package. Further analyses, such as principal component analysis,

681   hierarchical clustering, and differential expression analysis, were performed in R using

682   the VST-normalized counts.

683

Differential Expression (DE) analysis

685   The differential expression analyses were performed with DESeq2 v1.26.0 (LOVE *et al.*

686   2014). On the whole ovary dataset, the contrasts tested were early vs mid, and mid vs

687   late stages. For the tissue-specific datasets, three different comparisons were performed.

688   First, to identify differentially expressed genes independently of the stage, all stages of

689   somatic cells were compared to all stages of germ cells. Second, to identify genes up-

Tarikere, Ylla & Extavour Page 31

690   regulated in a stage-specific manner within each tissue, we compared the expression

691   level at each stage to the mean expression level of the other two stages. Third, we

692   compared germ cells and somatic cells independently at each stage. Genes with a

693   Benjamini-Hochberg (BH) adjusted p-value lower than 0.01 were selected as differentially

694   expressed in the corresponding contrast.

695

696   **Functional analysis**

697   The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG)

698   pathways enrichment analyses were performed on the differentially expressed genes with

699   the enrichGO and enrichKEGG functions of the clusterProfiler package (v3.14.3) for R

700   (YU *et al.* 2012). The GO terms were obtained using the R package AnnotationDbi

701   (CARLSON 2015) with the database org.Dm.eg.db v3.10.0. The GO overrepresentation

702   analysis of biological process (BP) was performed against the gene universe of all *D.*

703   *melanogaster* annotated genes in org.Dm.eg.db, adjusting the p-values with the

704   Benjamini-Hochberg method (BH), adjusted p-value and q-value cutoff of 0.01, and a

705   minimum of 30 genes per term. For the KEGG enrichment analysis, p-values were

706   adjusted by the BH procedure, and an adjusted p-value cutoff of 0.05 was used.

707

708   **Data availability**

709   All the raw data are publicly available at NCBI-Gene Expression Omnibus (GEO)

710   database under the accession code GSE172015. The scripts used to process and

711 analyze   the   data   are   available   at   GitHub   repository

712 https://github.com/guillemylla/Ovariole_morphogenesis_RNAseq.

713

Competing Interests

715 The authors have no competing interests to declare.

716

Acknowledgements

724 **FIGURE LEGENDS**

725

726 **Figure 1**: **Experimental scheme for generating stage-specific transcriptomes of**
727 **germ cells and somatic cells of larval ovaries during terminal filament formation.**
728 **A)** Location of the larval ovaries (white circles within the larva), and illustration of larval
729 ovary development divided into three stages during terminal filament formation (colored
730 in black). **B)** Left to right: location of the ovaries in an adult female abdomen; a single
731 adult ovary containing multiple ovarioles; an individual ovariole; anterior tip of an ovariole
732 enlarged to show the germarium and terminal filament (black) at the tip. **C)**
733 Representation of the three stages of whole larval ovaries chosen for library preparation
734 and sequencing (yellow: early stage, green: mid, blue: late). **D)** Somatic cells and **E)** germ
735 cells from developing ovaries at the three chosen stages were labelled with GFP using
736 tissue-specific GAL4 lines and **F)** GFP-positive cells were separated using FACS. **G)**
737 Schematics of representative plot layouts of somatic and germ line tissue separation
738 using FACS. Y axis: autofluorescence, 488-576/21 Height Log; X axis: GFP fluorescence
739 intensity, 488-513/26 Height Log (see Supplementary Figure S2 for actual data plots). **H)**
740 Separated cells or whole ovaries were processed for mRNA extraction and cDNA library
741 preparation followed by high throughput sequencing.

742

743 **Figure 2: Whole ovary RNA-seq dataset overview**. **A)** hierarchical clustering
744 dendrogram and **B)** PCA of the whole ovary RNA-seq dataset, both showing that
745 biological replicates are similar to each other and that early and mid-stages are more
746 similar to each other than either of them is to late stage. **C)** Number of differentially
747 expressed genes between early and mid stages, and between mid and late stages
748 (adjusted p-value<0.01; black: upregulated genes; white: downregulated genes). **D)**
749 Number of significantly upregulated stage-specific genes (adjusted p-value<0.01). **E)**
750 Heatmap showing the expression of all the stage-specific upregulated genes as a row-
751 wise z-score. Genes are clustered hierarchically and separated into three groups using
752 the function "cutree", and greyscale row labels immediately to the right of the tree are
753 colored based on the stage in which the gene was detected to be significantly upregulated
754 (x axis categories).

755

756 **Figure 3: Cell type-specific RNA-seq dataset concordance and positive controls**. **A)**
757 PCA Plot and **B)** hierarchical clustering dendrogram of germ cell and somatic cell RNA-
758 seq libraries. Expression in normalized counts by variance stabilization transformation
759 (VST) in each of the cell-type-specific RNA-seq libraries of **C)** known germ cell markers
760 *nanos* and *vasa,* and **D)** known terminal filament markers *bric a brac 1*, *bric a brac 2*, and
761 *traffic jam*.

762

763 **Figure 4: Transcriptomic differences between germ cells and somatic cells**. **A)**
764 Number of significantly upregulated genes (adjusted p-value<0.01) in germ cells and
765 somatic cells. **B)** Significantly enriched KEGG pathways (adjusted p-value<0.05) within

766   the upregulated genes of each cell type. The circle size is proportional to the number of
767   differentially expressed genes that the indicated KEGG pathway contains, and the color
768   gradient indicates the p-value.

769

770   **Figure 5: Cell type-specific differential expression analysis**. **A)** Number of
771   differentially expressed genes (adjusted p-value<0.01) upregulated (black) and
772   downregulated (white) in somatic cells at each stage compared to the two other stages.
773   **B)** Number of differentially expressed genes (adjusted p-value<0.01) upregulated (black)
774   and downregulated (white) in somatic cells at each examined stage. **C)** Significantly
775   enriched (adjusted p-value<0.05) KEGG pathways within the upregulated genes at each
776   somatic stage. Circle size is proportional to the number of differentially expressed genes
777   it contains, and the color gradient indicates the p-value. **D)** Number of differentially
778   expressed genes (adjusted p-value<0.01) upregulated (black) and downregulated (white)
779   in germ cells at each stage compared to the two other stages. **E)** Number of differentially
780   expressed genes (adjusted p-value<0.01) upregulated (black) and downregulated (white)
781   in germ cells at each examined stage.

REFERENCES

Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler *et al.*, 2000 Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics 25**:** 25-29.

Ashburner, M., K. G. Golic and R. S. Hawley, 2005 *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Barolo, S., L. A. Carver and J. W. Posakony, 2000 GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. Biotechniques 29**:** 726, 728, 730, 732.

Barry, E. R., and F. D. Camargo, 2013 The Hippo superhighway: signaling crossroads converging on the Hippo/Yap pathway in stem cells and development. Current Opinion in Cell Biology 25**:** 247-253.

Besse, F., D. Busson and A. M. Pret, 2005 Hedgehog signaling controls Soma-Germen interactions during Drosophila ovarian morphogenesis. Dev Dyn 234**:** 422-431.

Bolívar, J., J. Pearson, L. López-Onieva and A. González-Reyes, 2006 Genetic dissection of a stem cell niche: the case of the *Drosophila* ovary. Developmental Dynamics 235**:** 2969-2979.

Brand, A. H., and N. Perrimon, 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118**:** 401-415.

Brennecke, J., A. A. Aravin, A. Stark, M. Dus, M. Kellis *et al.*, 2007 Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell 128**:** 1089-1103.

Büning, J., 1994 *The Insect Ovary: ultrastructure, previtellogenic growth and evolution*. Chapman and Hall, London.

Carlson, M., 2015 AnnotationDbi: Introduction To Bioconductor Annotation Packages.

Chen, J., D. Godt, K. Gunsalus, I. Kiss, M. Goldberg *et al.*, 2001 Cofilin/ADF is required for cell motility during Drosophila ovary development and oogenesis. Nat Cell Biol 3**:** 204-209.

Church, S. H., B. A. S. de Medeiros, S. Donoughe, N. L. Marquez Reyes and C. G. Extavour, 2021 Repeated loss of variation in insect ovary morphology highlights the role of developmental constraint in life-history evolution. Proceedings of the Royal Society of London. Series B: Biological Sciences in press.

Cohen, E. D., M. C. Mariol, R. M. Wallace, J. Weyers, Y. G. Kamberov *et al.*, 2002 DWnt4 regulates cell movement and focal adhesion kinase during Drosophila ovarian morphogenesis. Dev Cell 2**:** 437-448.

Couderc, J. L., D. Godt, S. Zollman, J. Chen, M. Li *et al.*, 2002 The *bric a brac* locus consists of two paralogous genes encoding BTB/POZ domain proteins and acts as a homeotic and morphogenetic regulator of imaginal development in *Drosophila*. Development 129**:** 2419-2433.

Dansereau, D. A., and P. Lasko, 2008 RanBPM regulates cell shape, arrangement, and capacity of the female germline stem cell niche in Drosophila melanogaster. J Cell Biol 182**:** 963-977.

David, J., 1970 Le nombre d'ovarioles chez la Drosophila: relation avec la fecondite et valeur adaptive [Number of ovarioles inDrosophila melanogaster: relationship with fertility and adaptive value]. Arch. Zool. Exp. Gen 111**:** 357-370.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski *et al.*, 2013 STAR: ultrafast universal RNA-seq aligner., pp. 15-21 in *Bioinformatics*.

Elshaer, N., and M. D. Piulachs, 2015 Crosstalk of EGFR signalling with Notch and Hippo pathways to regulate cell specification, migration and proliferation in cockroach panoistic ovaries. Biol Cell 107**:** 273-285.

Forbes, A. J., H. Lin, P. W. Ingham and A. C. Spradling, 1996 hedgehog is required for the proliferation and specification of ovarian somatic cells prior to egg chamber formation in Drosophila. Development 122**:** 1125-1135.

Gancz, D., and L. Gilboa, 2013 Insulin and Target of rapamycin signaling orchestrate the development of ovarian niche-stem cell units in *Drosophila*. Development 140**:** 4145-4154.

Gancz, D., T. Lengil and L. Gilboa, 2011 Coordinated regulation of niche and stem cell precursors by hormonal signaling. PLoS Biology 9**:** e1001202.

Gavin, J. A., and J. H. Williamson, 1976 Juvenile hormone-induced vitellogenesis in Apterous4, a non-vitellogenic mutant in Drosophila melanogaster. Journal of Insect Physiology 22**:** 1737-1742.

Gilboa, L., 2015 Organizing stem cell units in the *Drosophila* ovary. Current Opinion in Genetics & Development 32C**:** 31-36.

Gilboa, L., and R. Lehmann, 2004 Repression of primordial germ cell differentiation parallels germ line stem cell maintenance. Curr Biol 14**:** 981-986.

Gilboa, L., and R. Lehmann, 2006 Soma-germline interactions coordinate homeostasis and growth in the *Drosophila* gonad. Nature 443**:** 97-100.

Godt, D., and F. A. Laski, 1995 Mechanisms of cell rearrangement and cell recruitment in *Drosophila* ovary morphogenesis and the requirement of *bric à brac*. Development 121**:** 173-187.

Green, D. A., 2nd, and C. G. Extavour, 2014 Insulin signalling underlies both plasticity and divergence of a reproductive trait in Drosophila. Proc Biol Sci 281**:** 20132673.

Green II, D. A., and C. G. Extavour, 2012 Convergent Evolution of a Reproductive Trait Through Distinct Developmental Mechanisms in *Drosophila*. Developmental Biology 372**:** 120-130.

Harvey, K. F., C. M. Pfleger and I. K. Hariharan, 2003 The *Drosophila* Mst ortholog, *hippo*, restricts growth and cell proliferation and promotes apoptosis. Cell 114**:** 457-467.

Harvey, N. L., T. Daish, K. Mills, L. Dorstyn, L. M. Quinn *et al.*, 2001 Characterization of the Drosophila caspase, DAMM. J Biol Chem 276**:** 25342-25350.

Hodin, J., 2009 Chapter 11: She Shapes Events As They Come: Plasticity in Female Insect Reproduction, pp. 423-521 in *Phenoypic Plasticity of Insects*, edited by D. Whitman and T. N. Ananthakrishnan. Science Publishers, Inc., Enfield, NH.

Hodin, J., and L. M. Riddiford, 2000 Different mechanisms underlie phenotypic plasticity and interspecific variation for a reproductive character in Drosophilids (Insecta: Diptera). Evolution 5**:** 1638-1653.

Honek, A., 1993 Intraspecific Variation in Body Size and Fecundity in Insects - a General Relationship. Oikos 66**:** 483-492.

Huang, J., S. Wu, J. Barrera, K. Matthews and D. Pan, 2005 The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the *Drosophila* Homolog of YAP. Cell 122**:** 421-434.

872   Irizarry, J., and A. Stathopoulos, 2015 FGF signaling supports Drosophila fertility by
873   regulating development of ovarian muscle tissues. Dev Biol 404**:** 1-13.
874   Jevitt, A., D. Chatterjee, G. Xie, X. F. Wang, T. Otwell *et al.*, 2020 A single-cell atlas of
875   adult Drosophila ovary identifies transcriptional programs and somatic cell lineage
876   regulating oogenesis. PLoS Biol 18**:** e3000538.
877   Kanehisa, M., and S. Goto, 2000 KEGG: kyoto encyclopedia of genes and genomes.
878   Nucleic Acids Res 28**:** 27-30.
879   Khammari, A., F. Agnes, P. Gandille and A. M. Pret, 2011 Physiological apoptosis of polar
880   cells during Drosophila oogenesis is mediated by Hid-dependent regulation of
881   Diap1. Cell Death Differ 18**:** 793-805.
882   King, R. C., 1970 *Ovarian development in Drosophila melanogaster*. Academic Press.
883   King, R. C., S. K. Aggarwal and U. Aggarwal, 1968 The Development of the Female
884   *Drosophila* Reproductive System. Journal of Morphology 124**:** 143-166.
885   König, A., A. S. Yatsenko, M. Weiss and H. R. Shcherbata, 2011 Ecdysteroids
886   affectDrosophilaovarian stem cell niche formation and early germline
887   differentiation. The EMBO Journal 30**:** 1549-1562.
888   Krishnan, B., S. E. Thomas, R. Yan, H. Yamada, I. B. Zhulin *et al.*, 2014 Sisters Unbound
889   Is Required for Meiotic Centromeric Cohesion in Drosophila melanogaster.
890   Genetics 198**:** 947-965.
891   Kumar, T., L. Blondel and C. G. Extavour, 2020 Topology-driven analysis of protein-
892   protein interaction networks detects functional genetic sub-networks regulating
893   reproductive capacity. eLife 9**:** e54082.
894   Lai, C. M., K. Y. Lin, S. H. Kao, Y. N. Chen, F. Huang *et al.*, 2017 Hedgehog signaling
895   establishes precursors for germline stem cell niches by regulating cell adhesion.
896   Journal of Cell Biology 216**:** 1439-1453.
897   Larkin, A., S. J. Marygold, G. Antonazzo, H. Attrill, G. Dos Santos *et al.*, 2021 FlyBase:
898   updates to the Drosophila melanogaster knowledge base. Nucleic Acids Res 49**:**
899   D899-D907.
900   Lehmann, R., and C. Nusslein-Volhard, 1991 The maternal gene *nanos* has a central role
901   in posterior pattern formation of the *Drosophila* embryo. Development 112**:** 679-
902   691.
903   Lengil, T., D. Gancz and L. Gilboa, 2015 Activin signaling balances proliferation and
904   differentiation of ovarian niche precursors and enables adjustment of niche
905   numbers. Development 142**:** 883-892.
906   Li, B., and C. N. Dewey, 2011 RSEM: accurate transcript quantification from RNA-Seq
907   data with or without a reference genome., pp. 323 in *BMC Bioinformatics*.
908   Li, M., X. Hu, S. Zhang, M. S. Ho, G. Wu *et al.*, 2019 Traffic jam regulates the function of
909   the ovarian germline stem cell progeny differentiation niche during pre-adult stage
910   in Drosophila. Sci Rep 9**:** 10124.
911   Li, M. A., J. D. Alls, R. M. Avancini, K. Koo and D. Godt, 2003 The large Maf factor Traffic
912   Jam controls gonad morphogenesis in *Drosophila*. Nature Cell Biology 5**:** 994-
913   1000.
914   Lobell, A. S., R. R. Kaspari, Y. L. Serrano Negron and S. T. Harbison, 2017 The Genetic
915   Architecture of Ovariole Number in Drosophila melanogaster: Genes with Major,
916   Quantitative, and Pleiotropic Effects. G3 (Bethesda) 7**:** 2391-2403.

Love, M. I., W. Huber and S. Anders, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15**:** 550.

Lynch, J. A., A. D. Peel, A. Drechsler, M. Averof and S. Roth, 2010 EGF signaling and the origin of axial polarity among the insects. Curr Biol 20**:** 1042-1047.

Markow, T. A., and P. M. O'Grady, 2007 *Drosophila* biology in the genomic age. Genetics 177**:** 1269-1276.

Matsuoka, S., Y. Hiromi and M. Asaoka, 2013 Egfr signaling controls the size of the stem cell precursor pool in the Drosophila ovary. Mech Dev 130**:** 241-253.

Mendes, C. C., and C. K. Mirth, 2016 Stage-Specific Plasticity in Ovary Size Is Regulated by Insulin/Insulin-Like Growth Factor and Ecdysone Signaling in Drosophila. Genetics 202**:** 703-719.

Ohlstein, B., C. A. Lavoie, O. Vef, E. Gateff and D. M. McKearin, 2000 The Drosophila cystoblast differentiation factor, benign gonial cell neoplasm, is related to DExH-box proteins and interacts genetically with bag-of-marbles. Genetics 155**:** 1809-1819.

Olivieri, D., M. M. Sykora, R. Sachidanandam, K. Mechtler and J. Brennecke, 2010 An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. EMBO J 29**:** 3301-3317.

Panchal, T., X. Chen, E. Alchits, Y. Oh, J. Poon *et al.*, 2017 Specification and spatial arrangement of cells in the germline stem cell niche of the Drosophila ovary depend on the Maf transcription factor Traffic jam. PLoS Genet 13**:** e1006790.

Pangas, S. A., and T. K. Woodruff, 2000 Activin signal transduction pathways. Trends Endocrinol Metab 11**:** 309-314.

Patil, V. S., and T. Kai, 2010 Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas. Current Biology 20**:** 724-730.

Rust, K., L. E. Byrnes, K. S. Yu, J. S. Park, J. B. Sneddon *et al.*, 2020 A single-cell atlas and lineage analysis of the adult Drosophila ovary. Nat Commun 11**:** 5628.

Sahut-Barnola, I., B. Dastugue and J.-L. Couderc, 1996 Terminal filament cell organization in the larval ovary of *Drosophila melanogaster*: ultrastructural observations and pattern of divisions. Roux's Archives of Developmental Biology 205**:** 356-363.

Sahut-Barnola, I., D. Godt, F. A. Laski and J.-L. Couderc, 1995 *Drosophila* Ovary Morphogenesis: Analysis of Terminal Filament Formation and Identification of a Gene Required for This Process. Developmental Biology 170**:** 127-135.

Sarikaya, D. P., and C. G. Extavour, 2015 The Hippo pathway regulates homeostatic growth of stem cell niche precursors in the *Drosophila* ovary. PLoS Genetics 11**:** e1004962.

Sato, K., and M. C. Siomi, 2020 The piRNA pathway in Drosophila ovarian germ and somatic cells. Proc Jpn Acad Ser B Phys Biol Sci 96**:** 32-42.

Sato, K., Yuka, A. Shibuya, P. Carninci, Y. Tsuchizawa *et al.*, 2015 Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the Drosophila Germline. Molecular Cell 59**:** 553-563.

Sato, T., J. Ogata and Y. Niki, 2010 BMP and Hh signaling affects primordial germ cell division in Drosophila. Zoolog Sci 27**:** 804-810.

Schupbach, T., and E. Wieschaus, 1986 Maternal-effect mutations altering the anterior-posterior pattern of the Drosophila embryo. Rouxs Arch Dev Biol 195**:** 302-317.

963   Shaul, Y. D., and R. Seger, 2007 The MEK/ERK cascade: from signaling specificity to
964   diverse functions. Biochimica et Biophysica Acta 1773**:** 1213-1226.
965   Simoes, S., B. Denholm, D. Azevedo, S. Sotillos, P. Martin *et al.*, 2006
966   Compartmentalisation of Rho regulators directs cell invagination during tissue
967   morphogenesis. Development 133**:** 4257-4267.
968   Slaidina, M., T. U. Banisch, S. Gupta and R. Lehmann, 2020 A single-cell atlas of the
969   developing Drosophila ovary identifies follicle stem cell progenitors. Genes Dev
970   34**:** 239-249.
971   Slaidina, M., S. Gupta, T. Banisch and R. Lehmann, 2021 A single cell atlas reveals
972   unanticipated cell type complexity in Drosophila ovaries. bioRxiv**:** 427703.
973   Song, X., G. B. Call, D. Kirilly and T. Xie, 2007 Notch signaling controls germline stem
974   cell niche formation in the *Drosophila* ovary. Development 134**:** 1071-1080.
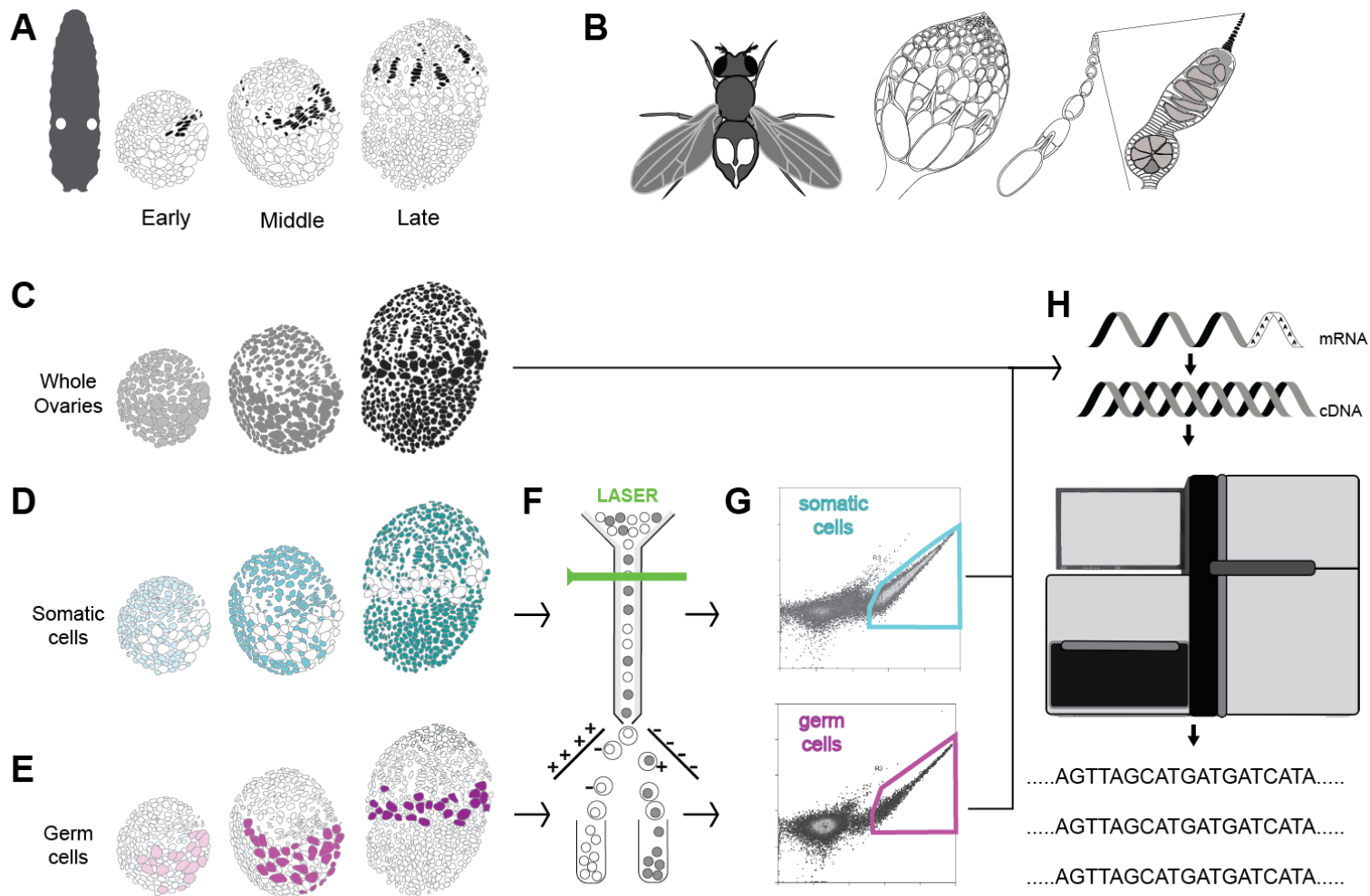975   Toret, C. P., and A. Le Bivic, 2021 A potential Rho GEF and Rac GAP for coupled Rac
976   and Rho cycles during mesenchymal-to-epithelial-like transitions. Small GTPases
977   12**:** 13-19.
978   Twombly, V., R. K. Blackman, H. Jin, J. M. Graff, R. W. Padgett *et al.*, 1996 The TGF-
979   beta signaling pathway is essential for Drosophila oogenesis. Development 122**:**
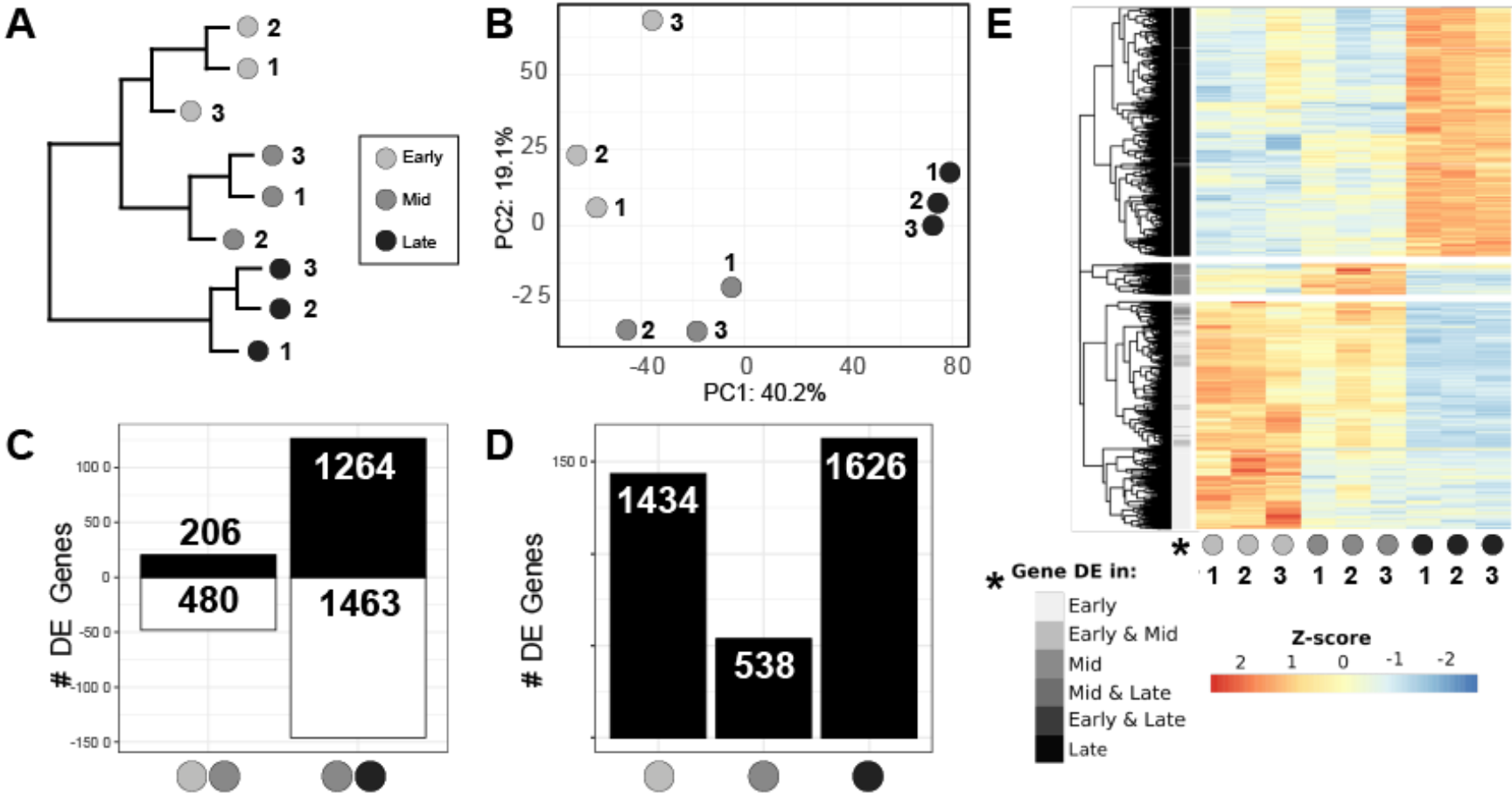980   1555-1565.
981   Yu, G., L. G. Wang, Y. Han and Q. Y. He, 2012 clusterProfiler: an R package for
982   comparing biological themes among gene clusters. OMICS 16**:** 284-287.
983   Yuan, S., X. Yu, M. Topf, L. Dorstyn, S. Kumar *et al.*, 2011 Structure of the Drosophila
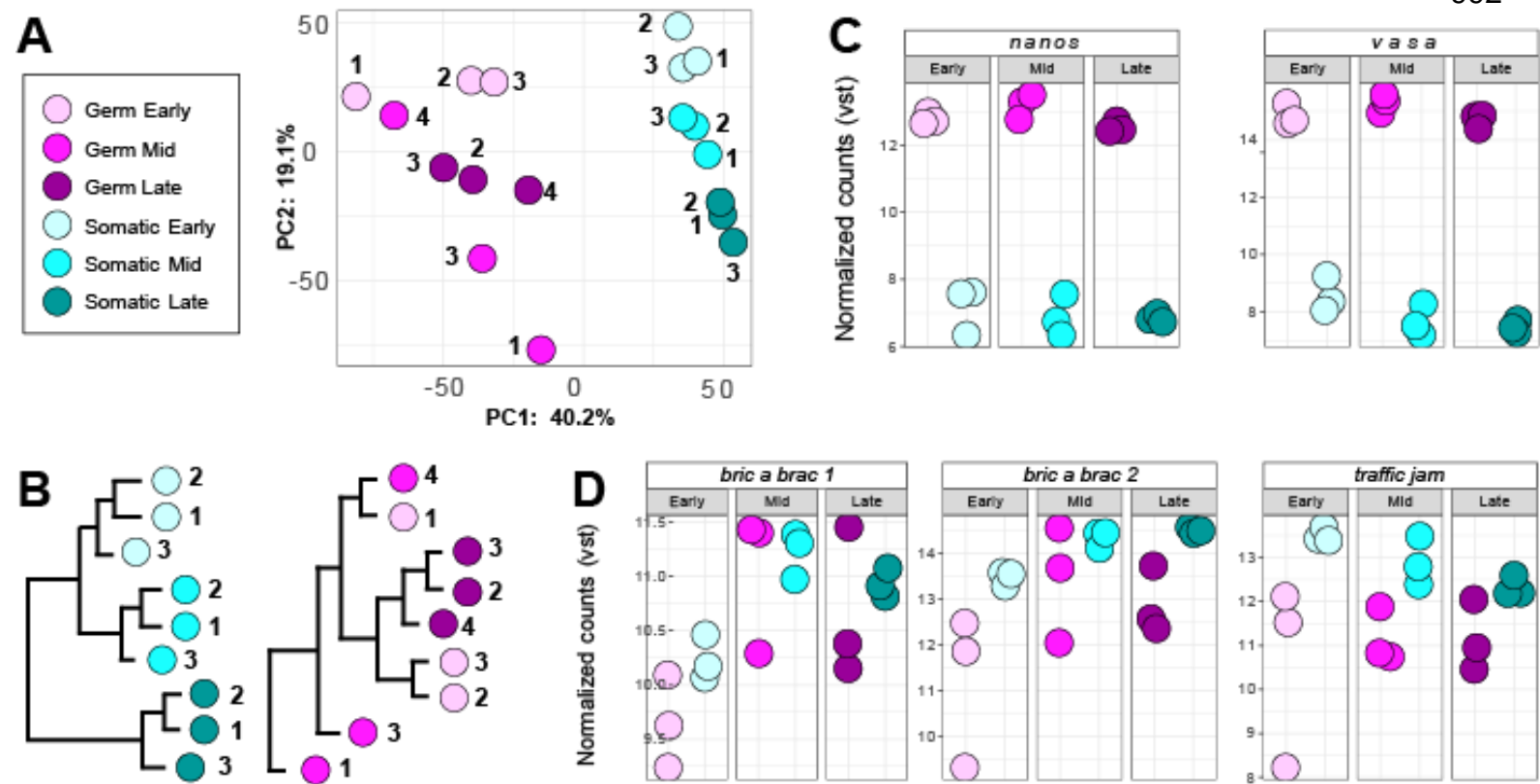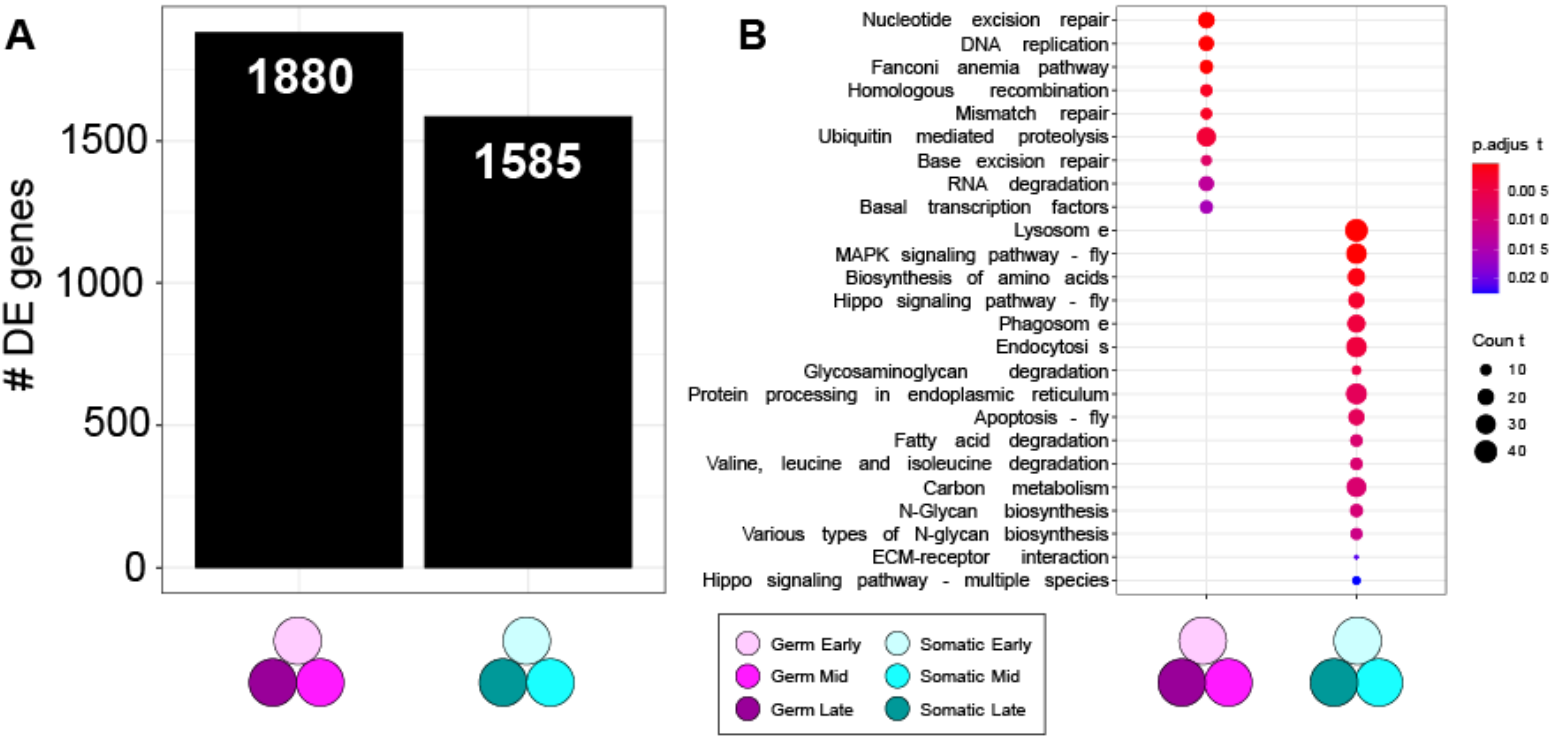984   apoptosome at 6.9 a resolution. Structure 19**:** 128-140.
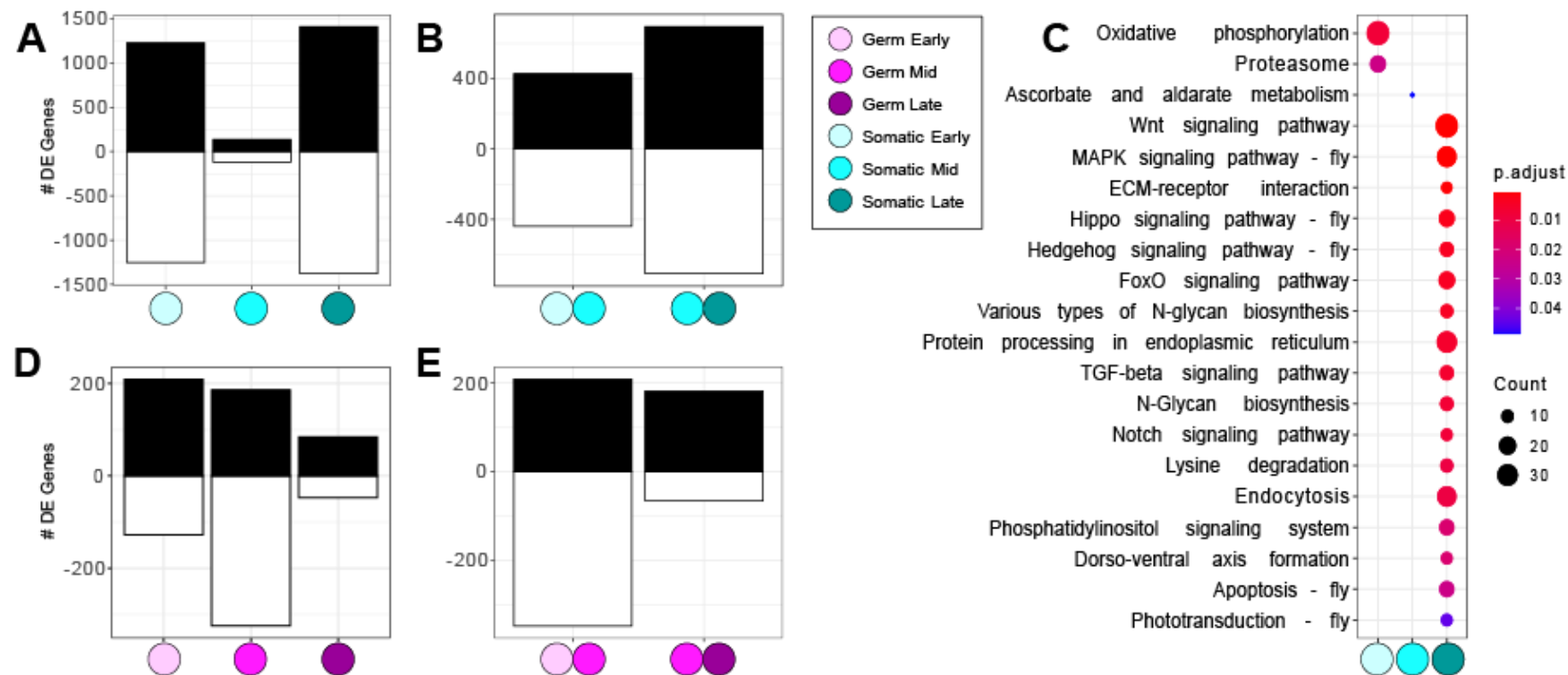
985

Figure 1

A

Early Middle Late

B

C
Whole
Ovaries

H
mRNA
cDNA

D
Somatic
cells

F
LASER

G
somatic
cells

E
Germ
cells

germ
cells

.....AGTTAGCATGATGATCATA.....

.....AGTTAGCATGATGATCATA.....

.....AGTTAGCATGATGATCATA.....

Figure 2

991   **Figure 3**

992

Figure 4

Figure 5