

## MIT Open Access Articles

*A single-cell atlas reveals unanticipated cell type complexity in Drosophila ovaries*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Slaidina, Maija, Gupta, Selena, Banisch, Torsten U and Lehmann, Ruth. 2021. "A single-cell atlas reveals unanticipated cell type complexity in Drosophila ovaries." Genome Research, 31 (10).

**As Published:** 10.1101/GR.274340.120

**Publisher:** Cold Spring Harbor Laboratory

**Persistent URL:** <https://hdl.handle.net/1721.1/146886>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of use:** Creative Commons Attribution NonCommercial License 4.0



# A single-cell atlas reveals unanticipated cell type complexity in *Drosophila* ovaries

Maija Slaidina,<sup>1</sup> Selena Gupta,<sup>1</sup> Torsten U. Banisch,<sup>1</sup> and Ruth Lehmann<sup>1,2</sup>

<sup>1</sup>Department of Cell Biology, NYU Grossman School of Medicine, HHMI, Skirball Institute of Biomolecular Medicine, New York, New York 10016, USA

Organ function relies on the spatial organization and functional coordination of numerous cell types. The *Drosophila* ovary is a widely used model system to study the cellular activities underlying organ function, including stem cell regulation, cell signaling and epithelial morphogenesis. However, the relative paucity of cell type-specific reagents hinders investigation of molecular functions at the appropriate cellular resolution. Here, we used single-cell RNA sequencing to characterize all cell types of the stem cell compartment and early follicles of the *Drosophila* ovary. We computed transcriptional signatures and identified specific markers for nine states of germ cell differentiation and 23 somatic cell types and subtypes. We uncovered an unanticipated diversity of escort cells, the somatic cells that directly interact with differentiating germline cysts. Three escort cell subtypes reside in discrete anatomical positions and express distinct sets of secreted and transmembrane proteins, suggesting that diverse micro-environments support the progressive differentiation of germ cells. Finally, we identified 17 follicle cell subtypes and characterized their transcriptional profiles. Altogether, we provide a comprehensive resource of gene expression, cell type-specific markers, spatial coordinates, and functional predictions for 34 ovarian cell types and subtypes.

[Supplemental material is available for this article.]

Most organs are composed of numerous cell types. Their spatial organization determines the cellular interactions that ensure sustainable organ function, which lasts for the duration of an organism's life. The fly ovary is an attractive and widely used model system to study the general principles of organ function, due to a combination of genetic tractability, small size, and cell type complexity. Oogenesis, or egg production, is the major function of the ovary. Adult ovaries comprise cells of two distinct lineages: the somatic cells of mesodermal origin; and the germ cells (GCs) that arise from primordial germ cells and differentiate into eggs (Boyle and DiNardo 1995; Moore et al. 1998; Riechmann et al. 1998). Ovaries are composed of 16 to 20 units called ovarioles, each of which serves as an egg production line (Fig. 1C; Spradling 1993). At the anterior tip of an ovariole is a stem cell compartment, called the germarium, which houses germline stem cells (GSCs) and, at a more posterior position, the somatic follicle stem cells (FSCs) and their support cells (Fuller and Spradling 2007; Kirilly and Xie 2007; Duhart et al. 2017). More posteriorly, follicles, or egg chambers, of increasingly advanced stages are lined up, with the mature oocytes located distally. Each follicle contains a germline cyst surrounded by a follicular epithelium and will produce an egg. Specialized follicle cells, called stalk cells, separate follicles from one another, whereas polar cells at each pole of the follicle serve as organizers (King 1970; Mahowald and Kambyzellis 1980; Margaritis et al. 1980; Duhart et al. 2017).

Both stem cell populations are excellent models to study stem cell maintenance and differentiation, and stem cell-niche interactions. GSC self-renewal has been extensively studied and relies on

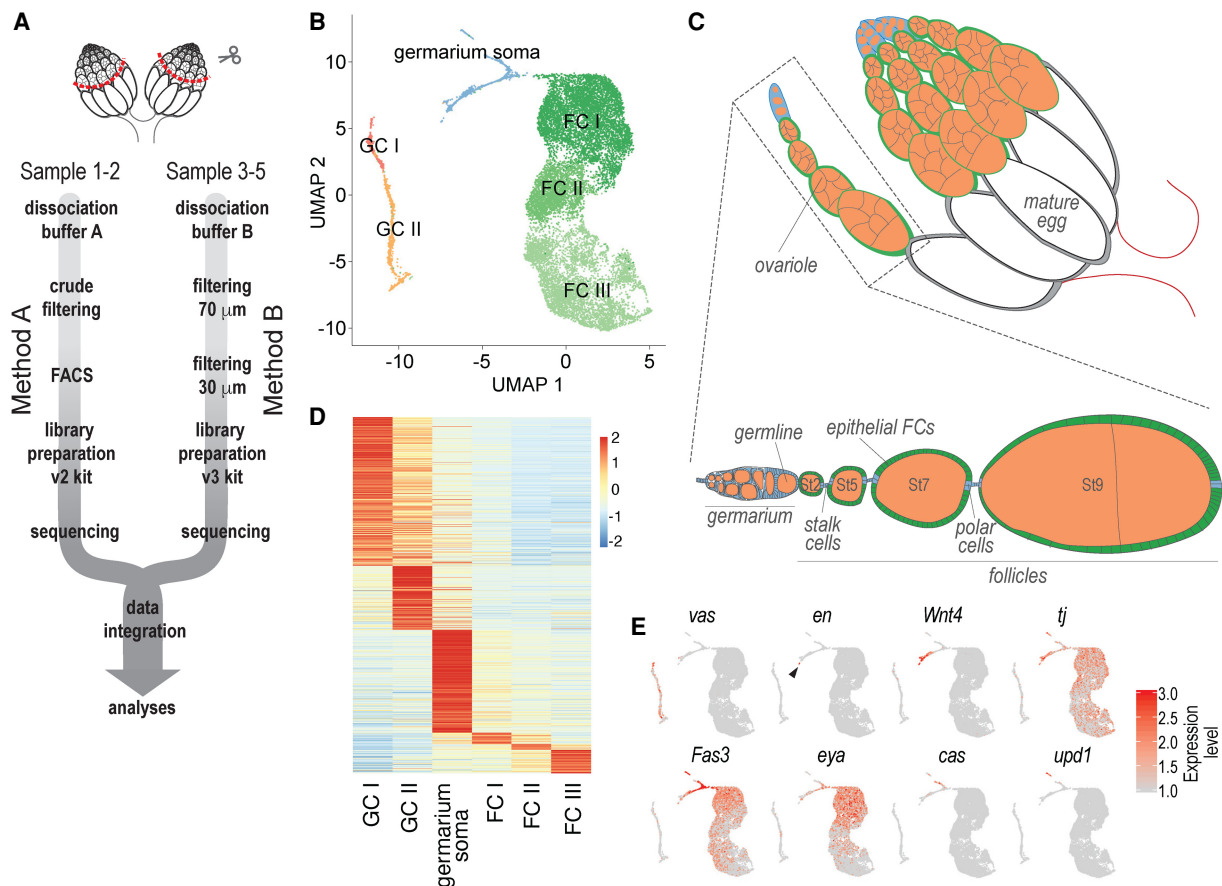
signals from and adhesion to the niche, which comprise terminal filament (TF) and cap cells (CCs) (Fig. 3B, below; Sahut-Barnola et al. 1995; Xie and Spradling 1998, 2000; Song et al. 2002). On the other hand, the mechanisms governing FSC self-renewal and differentiation are poorly understood (Margolis and Spradling 1995; Reilein et al. 2017; Rust and Nystul 2020).

Numerous somatic cell types support the multistage process of GC differentiation into mature eggs. Specifically, escort cells (ECs), also called inner germarium sheath cells (IGS), send protrusions around GCs and have a dual role—to promote both GSC renewal and GC differentiation (Kirilly et al. 2011; Rojas-Ríos et al. 2012; Liu et al. 2015; Wang and Page-McCaw 2018). This regulation relies on large surface contacts between the two cell types, and it is unclear whether distinct EC populations execute each function (Schulz et al. 2002; Kirilly et al. 2011; Banisch et al. 2017). Oogenesis is further supported by follicle cells (FCs) that arise from FSCs (Margolis and Spradling 1995). FCs form an epithelial layer around the germline-derived oocyte and nurse cells, thus delimiting the follicles. FCs are essential for egg development; they prepattern the body axes of the future embryo and secrete yolk proteins and the eggshell. Numerous signaling pathways are involved in specification and morphogenesis of follicle cell subtypes (Wu et al. 2008; Duhart et al. 2017; Rust and Nystul 2020). Therefore, the fly ovary also serves as a powerful experimental paradigm to dissect the molecular basis of cell–cell signaling. Despite the identification of various genes regulating oogenesis, the persistent dearth of cell type-specific markers and tools has hindered progress in determining the exact role of these factors with the appropriate cellular resolution. For example, specific markers for GSCs, FSCs, and other somatic cell types remain to be identified. To bridge this gap, we set out to use a single-cell RNA sequencing

<sup>2</sup>Present address: Whitehead Institute and Department of Biology at MIT, Cambridge, MA 02142, USA  
Corresponding author: [lehmann@wi.mit.edu](mailto:lehmann@wi.mit.edu)

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.274340.120>. Freely available online through the *Genome Research* Open Access option.

© 2021 Slaidina et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.



**Figure 1.** scRNA-seq of *Drosophila* adult ovaries. (A) Schematic of the scRNA-seq sample preparation. (B) UMAP plot of the entire data set. Each dot represents a transcriptome of a single cell and is color-coded according to cluster membership. (C) Schematic drawing of an adult ovary. (D) A heat map visualizing average gene expression levels of all marker genes in each cluster. Red indicates highest and blue lowest expression. (E) Expression of major cell type markers *vas*, *en*, *Wnt4*, *tj*, *Fas3*, *eya*, *cas*, and *upd1*. Red indicates highest and light gray lowest expression. Arrowhead indicates *en*-expressing cells.

(scRNA-seq) approach to systematically identify and characterize all cell types that compose the stem cell compartment and early-to mid-stage egg chambers.

## Results

### scRNA-seq using two distinct methods produces equivalent results

We aimed to obtain and characterize single-cell transcriptomes from the germarium and early follicles to identify cell type-specific markers and thus enable further functional studies. We used a nanoliter droplet-based 10x Genomics Chromium system for single-cell RNA sequencing (Fig. 1A). We performed the experiment in five replicates and selected high-quality transcriptomes (Supplemental Text). Despite using two different methods, all five sequencing experiments produced similar results (Fig. 1A; Supplemental Text; Supplemental Fig. S1A,C,D). Data from all replicates were aligned and merged using the Seurat 3 algorithm for library integration (Supplemental Fig. S1A).

Our initial coarse clustering revealed six clusters, visualized on a UMAP (Uniform Manifold Approximation and Projection) plot, where each dot represents a single-cell transcriptome (Fig. 1B). Two clusters expressed the GC marker *vas* (Lasko and

Ashburner 1988) and were thus annotated as germ cells (clusters GC I and II) (Fig. 1B,E). The three largest clusters expressed the follicle cell markers *Fas3* and *eya* (Bai and Montell 2002; Nystul and Spradling 2007) and therefore corresponded to epithelial follicle cells (clusters FC I, II, and III) (Fig. 1B,E). Finally, the last cluster contained somatic cell types of the germarium—TF and CC that express *en*, ECs that express *Wnt4* and *tj* (Forbes et al. 1996; Kawashima et al. 2003; Li et al. 2003; Mottier-Pavie et al. 2016), early follicle cell lineages that express *Fas3*, *tj* and *eya* (Bai and Montell 2002; Kawashima et al. 2003; Li et al. 2003; Nystul and Spradling 2007), and stalk and polar cell lineages that express *cas* (Chang et al. 2013) and *upd1* (cluster germarium soma, Fig. 1B,E; Silver and Montell 2001). We did not detect a separate cluster containing the muscle cells that ensheath the ovarioles (epithelial sheath cells) or peritoneal sheath, suggesting that those cells were lost during dissections or dissociations.

To compare the transcriptional signatures of these clusters, we computed markers for each cluster and visualized their expression in a heat map (Fig. 1D; Supplemental Table S1). GC I and GC II clusters shared a fraction of their marker genes, but there was minimal overlap between marker genes of other clusters, indicating that these cell types are transcriptionally distinct despite shared mesodermal origins. Thus, we produced a high-quality single-cell RNA sequencing data set of 15,227 cell transcriptomes, which we

used to generate ovarian cell type gene expression profiles, transcriptional signatures, and function predictions.

### Identification of nine distinct steps of GC differentiation

Next, we sought to characterize the transcriptional dynamics underlying GC differentiation. GSCs are in close contact with the niche and are marked by a round GC-specific organelle called the spectrosome (Lin and Spradling 1995). GSCs divide asymmetrically to give rise to a cystoblast (CB), which subsequently undergoes four rounds of synchronous divisions with incomplete cytokinesis to produce a 16-cell cyst. Within the cyst, the spectrosome evolves into a fusome that spans all 16 cells (Huynh 2006). The germarium is divided into four regions organized along the anterior-posterior axis—1, 2a, 2b, and 3 (Fig. 2B). Incomplete divisions of GCs occur in region 1, the anterior part of the germarium. Initially, the 16-cell cysts are rounded in shape, and multiple cysts can reside next to each other in region 2a of the germarium (Spradling 1993). During these initial stages of GC differentiation, ECs send protrusions around GCs and regulate their differentiation (Schulz et al. 2002; Kirilly et al. 2011; Banisch et al. 2017). As 16-cell cysts transition into region 2b, they flatten to a disc shape and become surrounded by follicle cells. Cysts grow larger and rounder as they reach region 3, where stage 1 of oogenesis begins. Finally, a follicle containing a cyst surrounded by epithelial FCs pinches off from the germarium and transitions through 14 stages of oogenesis to give rise to a mature egg (Figs. 1C, 2B). One of the 16 cells is specified as an oocyte and is destined to become the egg, whereas the other 15 cells become nurse cells that support oocyte growth by providing mRNAs and proteins. Toward the end of oogenesis during stage 11, nurse cells transfer all their content to the oocyte and subsequently undergo apoptosis (Spradling 1993).

To identify gene expression profiles specific to distinct stages of GC differentiation, we subclustered the GC I and GC II clusters and obtained nine clusters (Fig. 2A). To determine the identity of each cluster, we identified cluster-specific marker genes (Fig. 2C; Supplemental Table S1), selected markers that are not or are lowly expressed in other cell types (Supplemental Fig. S2), and visualized their expression patterns using a hybridization chain reaction (HCR), a highly sensitive fluorescence in situ hybridization (FISH) method (Choi et al. 2018). For gene description and brief functional annotation of genes described, refer to Supplemental Text.

The cluster expressing *CG2887*, *CG6628*, and *CG15398* corresponded to the earliest stages of GC differentiation and contained GSCs, CBs, and two-cell cysts (cluster GSC/CB/2-cc) (Fig. 2B–E; Supplemental Fig. S3A,B). *bam*, a master regulator of GC differentiation onset (Xie and Spradling 1998, 2000; Chen and McKearin 2003), was lowly expressed in two-cell cysts, displayed the strongest expression in four-cell cysts (cluster 4-cc), and lower levels at later stage cysts (Fig. 2B,C,E,G; Supplemental Fig. S3B,D). *HP6* was expressed in GSCs and early cysts (clusters GSC/CB/2-cc and 4-cc), whereas *Orc1* showed the strongest expression in four- and eight-cell cysts (clusters 4-cc and 8-cc) and low expression in GSCs, two-cell and 16-cell cysts (Fig. 2C,F; Supplemental Fig. S3C). Thus, *HP6* and *Orc1* coexpression in four-cell cysts demarcated cluster 4-cc. *CG15628* displayed the strongest expression starting from the eight-cell cyst stage demarcating cluster 8-cc and slowly fades after 16-cell cysts reached region 2b (Fig. 2B,C,G; Supplemental Fig. S3D). The next five clusters corresponded to 16-cell cysts at various stages. Strong expression of *CG3691* and

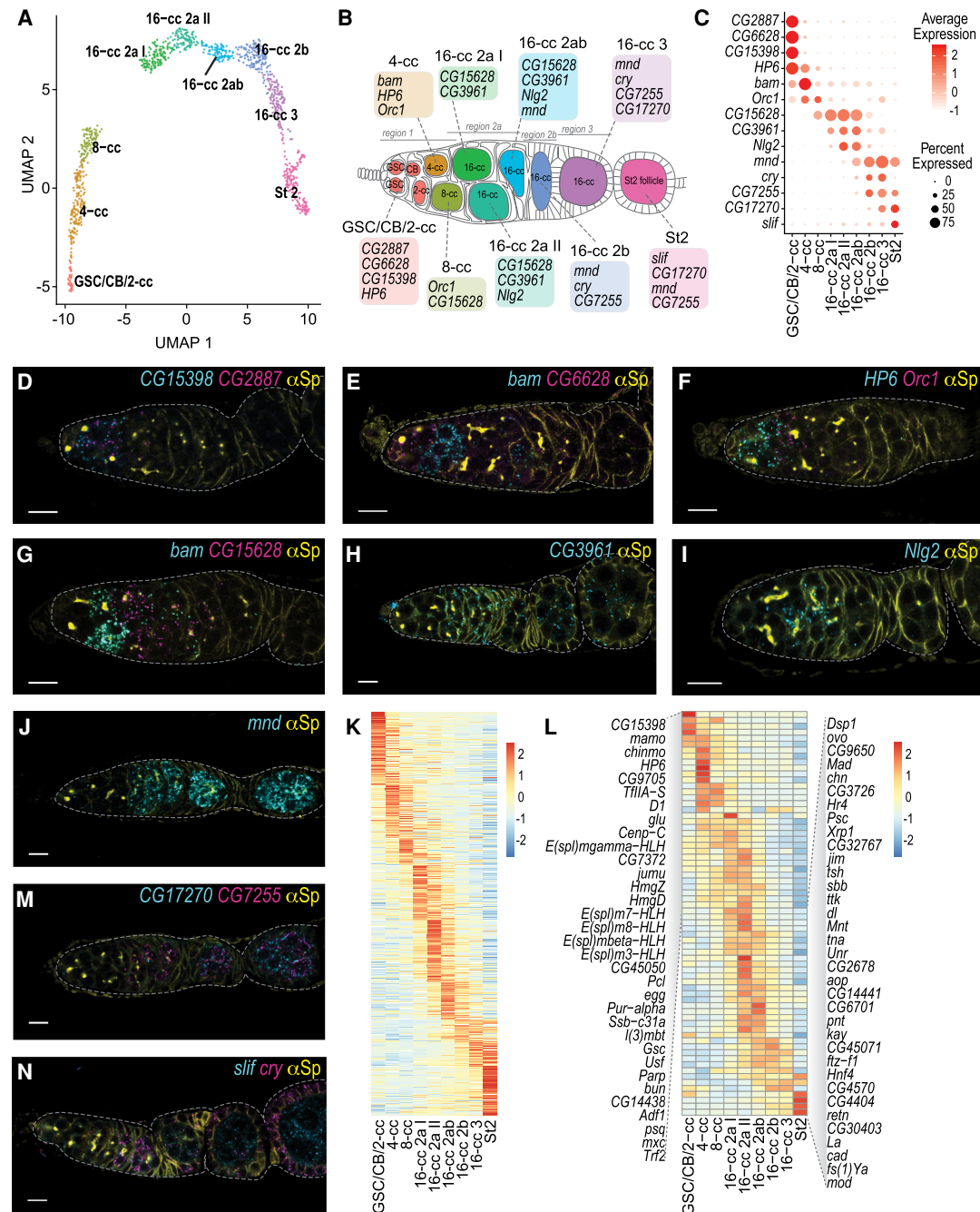
*Nlg2* was first detected in 16-cell cysts in region 2a, thus demarcating clusters 16-cc 2a I and 16-cc 2a II (Fig. 2B,C,H,I; Supplemental Fig. S3E,F). The next cluster contained 16-cell cysts from region 2a and 2b (16-cc 2ab). It was marked by the onset of strong *mnd* expression, which was detected in 16-cell cysts in region 2a, and by *CG3961* whose expression extends into 16-cell cysts in region 2b (Fig. 2B,C,H,J; Supplemental Fig. S3E,G). Strong expression of *CG7255* was first detected in 16-cell cysts in region 2b marking cluster 16-cc 2b (Fig. 2B,C,M; Supplemental Fig. S3H). Finally, onset of strong *CG17270* expression in 16-cell cysts in region 3 marked cluster 16-cc 3, and onset of *slif* expression marked stage 2 follicles (cluster St2) (Fig. 2B,C,M,N; Supplemental Fig. S3H,I). *cry* was predominantly expressed in 16-cell cysts in regions 2b and 3 (clusters 16-cc 2b and 16-cc 3) and was down-regulated in stage 2 follicles (cluster St 2) (Fig. 2B,C,N; Supplemental Fig. S3I). Thus, by exploring the expression of 14 marker genes in situ, we have assigned all nine clusters to specific stages of germ cell cyst development. GCs from later stages of oogenesis were not present in our data set due to their larger size and thus are incompatible with our cell suspension preparation and cell capture methods. The clusters were arranged on UMAP in the order of their differentiation status, with undifferentiated cells at the bottom left corner and differentiated cells on the right side (Fig. 2A). This indicates that differentiation is a predominant source of transcriptome variation in our data set. The data did not allow us to differentiate nurse cells from oocytes by clustering, and it remains unclear if our data set contains oocytes.

We were unable to clearly separate GSCs from CBs and two-cell cysts by clustering. This is likely due to the high degree of similarity in the gene expression profiles of GSCs, CBs, and two-cell cysts, the reliance on translational rather than transcriptional control for early GSC differentiation (Slaidina and Lehmann 2014), and the low abundance of these cell types. An observational study using wild-type ovaries estimated that, in each germarium, on average there were two GSCs, 1.3 CBs, 1.6 two-cell, 0.9 four-cell, 0.8 eight-cell cysts, and 8.4 16-cell cysts. Of the 16-cell cysts, 5.5 were found in region 2a, 1.9 in region 2b, and one in region 3 (Drummond-Barbosa and Spradling 2001). We used these observational numbers to estimate the coverage obtained for the different cell types (Supplemental Fig. S1B). Coverage varied between cyst stages, suggesting that some cluster populations may not be pure. Our analysis revealed distinct transcriptomes for nine steps of GC differentiation from GSC to stage 2 of oogenesis and identified markers for most (Fig. 2C), suggesting that the transcriptomes identified correspond to the developmental stages described previously.

### Gene expression dynamics during GC differentiation

Next, we characterized gene expression dynamics during GC differentiation more closely. Eight hundred thirteen genes were differentially expressed between GC clusters, indicating that even though the first steps of differentiation rely heavily on translational regulation, there are significant changes in mRNA levels (Slaidina and Lehmann 2014). We visualized the mean expression levels of differentially expressed genes on a heat map (Fig. 2K; Supplemental Tables S1, S2). Sixteen-cell cyst clusters displayed distinct expression signatures. Differences observed between the five germarial 16-cell cyst states reveal distinct mRNA expression patterns as 16-cell cysts transition from cyst cell division to differentiation. In region 2a, 16-cell cysts are enwrapped by ECs, whereas in region 2b and 3, GCs are surrounded by pre-FCs, which start





**Figure 2.** GC subclustering in nine differentiation states. (A) UMAP plot of GCs subclustered in nine differentiation states. Each dot represents a transcriptome of a single cell and is color-coded according to cluster membership. (B) Schematic drawing of a germarium and stage 2 follicle with GCs highlighted in the colors corresponding to the UMAP plot in A. Germ cell differentiation stages (GSC, CB, and cyst stage) are indicated in each cell. TFs, CCs, ECs, and FC lineages are colored in white. Germarium regions are indicated above the cartoon. Color-coded boxes indicate select marker genes for each cluster, whose mRNAs are visualized in D–J, M, and N. Note: Contrary to the schematic drawing, in vivo germ cell differentiation stages are not represented equally frequently. (C) A dot plot visualizing expression of 14 marker genes in nine GC clusters. Dot diameter represents the fraction of cells expressing each gene in each cluster, as shown in scale. Color intensity represents the average normalized expression level. (D–J, M, N) Marker gene mRNA in situ hybridization using HCR (cyan, magenta). Spectrosomes, fusomes, and somatic cell membranes are labeled by anti- $\alpha$ -Spectrin (yellow). Scale bars: 10  $\mu$ m. (D) *CG15398* (cyan) and *CG2887* (magenta) are expressed in GSCs, CBs, and two-cell cysts. (E) *bam* (cyan) is predominantly expressed in four-cell cysts, *CG6628* (magenta) is expressed in GSCs, CBs, and two-cell cysts. (F) *HP6* (cyan) is expressed in GSCs, CBs, two-cell and four-cell cysts, *Orc1* (magenta) is expressed in four-cell and eight-cell cysts. (G) *bam* (cyan) is predominantly expressed in four-cell cysts, with lower expression in two-cell and later stage cysts, *CG15628* (magenta) is expressed in eight-cell and 16-cell cysts in germarium regions 2a and 2b. (H) *CG3961* (cyan) is expressed predominantly in 16-cell cysts in germarium regions 2a and 2b. (I) *Nlg2* (cyan) is expressed in 16-cell cysts in germarium regions 2a and 2b. (J) *mnd* (cyan) is expressed in 16-cell cysts in germarium regions 2a, 2b and 3, and in stage 2 follicles. (K, L) Heat maps visualizing average gene expression levels of all GC marker genes (K) and dynamically expressed transcription factors (L) in each cluster. Red indicates highest and blue lowest expression. (M) *CG17270* (cyan) is expressed in 16-cell cysts in germarium region 3 and in stage 2 follicle, *CG7225* (magenta) is expressed in 16-cell cysts in germarium regions 2b and 3, in stage 2 follicle and beyond. (N) *slif* (cyan) is expressed in stage 2 follicle and beyond, *cry* (magenta) is expressed in 16-cell cysts in germarium regions 2b and 3 and in follicle cells.

forming the follicular epithelium. Thus, transcriptional differences may also reflect the changing signals GC cysts receive from various bordering somatic cell types.

We explored the functional predictions for these genes using Gene List Annotation for *Drosophila* (GLAD) (Supplemental Fig. S4A; Supplemental Table S3; Hu et al. 2015). GC stages in contact with ECs (up to 16-cell cysts in region 2a) were enriched for transmembrane proteins, some of which may be involved in signaling between GCs and ECs and others in formation of the fusome, which is composed of ER-derived vesicles (Huynh 2006). Furthermore, the GSCs/CB/2-cc cluster was enriched for chaperones and heat shock proteins, which have a protective function. Consistent with divisions of early cysts, GSC/CB/2-cc and 4-cc clusters were enriched for cytoskeletal proteins, in particular, those involved in cytokinesis. In contrast, most late stage cysts were enriched for mitochondria-related genes, which coincide with rapid expansion of the mitochondrial pool and their selection based on fitness (Hill et al. 2014; Lieber et al. 2019). Finally, transcription factors and DNA-binding proteins were significantly enriched at multiple stages over the course of GC differentiation, and we visualized their expression in a heat map (Fig. 2L). A few were expressed in GSC/CB/2-cc cluster, including, *CG15398* (Fig. 1D; Supplemental Fig. S3A) and *chinmo*, which regulates cyst stem cell renewal in *Drosophila* testis (Flaherty et al. 2010), suggesting that it may have a conserved role in stem cell renewal across many tissues. Five Notch-responsive *Enhancer of split Complex (E [spl]-C)* transcription factors were expressed in 4-cc, 8-cc, and 16-cc 2a I clusters, suggesting that the Notch signaling pathway may regulate early steps in GC differentiation.

Altogether, we have identified specific transcriptional signatures for nine distinct steps of GC differentiation. Further analyses of gene expression dynamics and genetic requirements between the clusters may reveal novel aspects of GC differentiation regulation.

### *Drosophila* germaria contain three EC subtypes

Next, we aimed to identify transcriptional signatures for somatic cells of the germarium. There are five major somatic cell types in germaria. At the most anterior tip of the germarium reside the somatic terminal filament and cap cells of the GSC niche (Fig. 3B; Sahut-Barnola et al. 1995). Directly posterior to them in regions 1 and 2a are ECs, which extend protrusions that enwrap GCs (Schulz et al. 2002). Posterior to the ECs at the region 2a and 2b boundary are FSCs, which divide asymmetrically to give rise to a transit amplifying FC population called prefollicle cells (pre-FC) (Margolis and Spradling 1995). Pre-FCs reside in regions 2b and 3 where they proliferate and further differentiate into polar, stalk, and epithelial FC lineages.

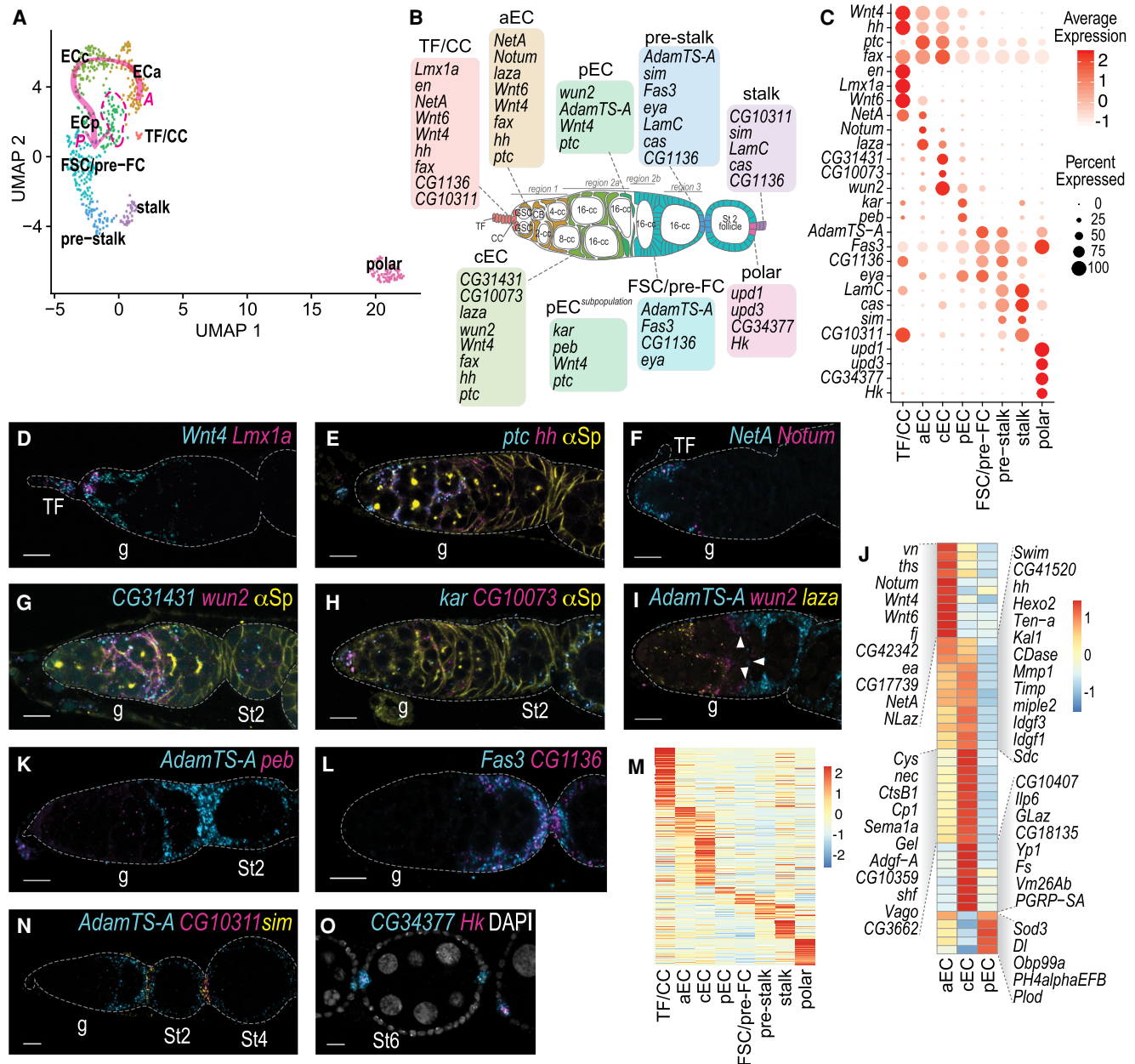
To assign transcriptomes to individual cell types, we subclustered the ‘germarium soma’ cluster and initially obtained seven clusters. A small number of cells located separately on UMAP and were included in one of the larger clusters by the clustering algorithm. These cells expressed *en* and *Lmx1a*, well known TF and CC markers (Fig. 3A–D; Supplemental Figs. S1E, S3J; Bolívar et al. 2006; Allbee et al. 2018). Therefore, we manually annotated this cell population as TF/CCs (see Supplemental Text). Each ovariole contains only about eight TFs and six CCs (Godt and Laski 1995; Xie and Spradling 2000). In total, our data set contained only 10 such cells, and thus we were unable to distinguish TFs from CCs. We computed markers for all eight clusters (Supplemental Table S1). Of the remaining seven clusters, four expressed *Fas3*, *cas*, or

both, suggesting they represent FC subtypes (Fig. 3L; Zhang and Kalderon 2001; Chang et al. 2013), and the remaining three expressed high levels of *fax* and *Wnt4* (Decotto and Spradling 2005; Mottier-Pavie et al. 2016), suggesting that they correspond to three distinct subtypes of ECs (Fig. 3A–D; Supplemental Fig. S3J). EC subtypes may correspond to ECs that are functionally distinct or that are in three distinct anatomical positions. Such subdivision of ECs along the anterior-posterior axis in the germarium has been suggested on the basis of morphological differences and limited functional analyses (Rojas-Ríos et al. 2012; Eliazar et al. 2014; Banisch et al. 2017). To distinguish between these possibilities, we explored marker gene expression of each cluster. *hh*, *Wnt4* and *ptc* are expressed in an anterior to posterior gradient, and their expression differed between EC clusters; suggesting the three clusters might correspond to distinct anatomical positions—anterior (aEC), central (cEC), and posterior (pEC) (Fig. 3A–E). Moreover, *Wnt6* functions in anterior ECs (Wang and Page-McCaw 2018) and was enriched in the aEC compared to other EC clusters (Fig. 3C). We selected a few EC subtype markers and assessed their mRNA expression in the ovary. *NetA* and *Notum* were specifically expressed in the aEC cluster (Fig. 3C) and, indeed, labeled anterior ECs in region 1 (Fig. 3F; Supplemental Fig. S3L). *CG31431* and *CG10073* were expressed in the cEC cluster (Fig. 3C), and the strongest expression of their mRNAs was detected in the central region of the germarium, in close contact to eight- and 16-cell cysts (Fig. 3G,H; Supplemental Fig. S3N,O). We used a combination of two markers to identify the remaining EC population. *wun2* is predominantly expressed in cEC and pEC clusters, whereas *AdamTS-A* is expressed in the pEC and FC lineage (Fig. 3C); therefore, cells coexpressing *wun2* and *AdamTS-A* correspond to pEC. We observed that *wun2* is predominantly expressed in ECs in region 2a, and *AdamTS-A* in the early FC lineage, and expression of both genes overlapped near the region 2a and 2b boundary (Fig. 3I; Supplemental Fig. S3M, arrowheads). These results suggest that the aEC, cEC, and pEC clusters correspond to the most anterior, central, and posterior ECs. EC transcriptomes are aligned from anterior to posterior, as indicated by a pink arrow in the UMAP plot (Fig. 3A, pink arrow).

Two additional pEC markers, *kar* and *peb*, were expressed in a fraction of pECs that lacked expression of *wun2* and *AdamTS-A* (Fig. 3A[dotted line in magenta],C; Supplemental Fig. S1E). In the germarium, *kar* and *peb* displayed sparse weak expression throughout the germarium (Fig. 3H,K; Supplemental Fig. S3O,P), suggesting that there might be additional complexity within the EC population that has not been revealed by clustering. Indeed, it has been suggested that ECs can migrate toward the anterior to possibly replace lost ECs (Reilein et al. 2017).

To test whether we captured all EC subtypes, we labeled *laza*, which is expressed in aEC and cEC together with *wun2* and *AdamTS-A*, and observed that all regions of the germarium were labeled, except the most anterior tip (TF/CC) (Fig. 3I; Supplemental Fig. S3M). Thus, we have, indeed, captured all EC subtypes.

Altogether, we have identified three transcriptionally distinct EC subtypes arranged along the anterior-to-posterior axis of the germarium. The number of cells recovered corresponded well to previous observations (Supplemental Fig. S1B). To explore the potential functional differences between EC subtypes in our data set, we compared their transcriptomes. Four hundred five genes were differentially expressed between EC clusters. We visualized their expression in a heat map together with other somatic cell markers (Fig. 3M) and performed GLAD analyses on EC marker genes (Supplemental Fig. S4B; Supplemental Table S3). Consistent with



**Figure 3.** Identification of three transcriptionally distinct EC subtypes. (A) UMAP plot of somatic cells of the germarium subclustered in eight clusters. Each dot represents a transcriptome of a single cell and is color-coded according to cluster membership. (B) Schematic drawing of a germarium and stage 2 follicle with somatic cells highlighted in the colors corresponding to the UMAP plot in A. GCs are colored in white, and their differentiation stages (GSC, CB, and cyst stage) are indicated in each cell. Germarium regions are indicated above the cartoon. Color-coded boxes indicate select marker genes for each cluster. Note: Contrary to the schematic drawing, in vivo germ cell differentiation stages are not represented equally frequently. (C) A dot plot visualizing expression of select 28 marker genes. Dot diameter represents the fraction of cells expressing each gene in each cluster, as shown in the scale. Color intensity represents the average normalized expression level. (D–I, K, L, N, O) Marker gene mRNA in situ hybridization using HCR (cyan, magenta, and yellow). Scale bars: 10  $\mu$ m. (g) Germaria, (TF) terminal filament, (St2, St4, St6) oogenesis stages for follicles. (E, G, H) Spectrosomes, fusomes, and somatic cell membranes are labeled by anti- $\alpha$ -Spectrin (yellow). (D) *Wnt4* (cyan) is expressed in TF, CC, and ECs, *Lmx1a* (magenta) is expressed in TF and CC. (E) *ptc* (cyan) is expressed in ECs, *hh* (magenta) is expressed in ECs and pre-FCs. (F) *NetA* (cyan) and *Notum* (magenta) are expressed in anterior ECs in germarium region 1. (G) *CG31431* (cyan) is predominantly expressed in ECs in germarium region 2a with some expression in region 1 and 2b, *wun2* (magenta) is predominantly expressed in ECs in germarium region 2a. (H) *kar* (cyan) is sparsely and lowly expressed throughout the germarium, *CG10073* (magenta) is predominantly expressed in region 2a ECs. (I) *AdamTS-A* (cyan) is expressed in FSCs, pre-FCs, and the most posterior ECs, *wun2* (magenta) is predominantly expressed in ECs in germarium region 2a, *laza* (yellow) is expressed in anterior ECs in germarium region 1 and some ECs in germarium region 2a. (J) A heatmap visualizing average expression levels of EC enriched secreted proteins and ECM components. Red indicates highest and blue lowest expression. (K) *AdamTS-A* (cyan) is expressed in FSCs, pre-FCs, and the most posterior ECs, *peb* (magenta) is sparsely and lowly expressed throughout the germarium. (L) *Fas3* (cyan) and *CG1136* (magenta) are expressed in FC lineage including FSCs, pre-FCs, prestalk, and epithelial FCs. (M) A heatmap visualizing average gene expression levels of all somatic cell marker genes in each cluster. Red indicates highest and blue lowest expression. (N) *AdamTS-A* (cyan) is expressed in FC lineage including FSCs, pre-FCs, prestalk, and epithelial FCs, *CG10311* (magenta) is expressed in stalk cells, *sim* (yellow) expressed in prestalk and stalk cells. (O) *CG34377* (cyan) and *Hk* (magenta) are expressed in polar cells, DAPI (white) labels cell nuclei.



their role as GC differentiation regulators, ECs were strongly enriched for gene classes related to cell-cell communication, such as signaling pathway components, secreted proteins, receptors, transmembrane proteins, and extracellular matrix (matrisome) proteins. Moreover, to facilitate discovery of transcription factors regulating distinct EC fates, we have visualized expression of differentially expressed transcription factors (Supplemental Fig. S4C).

Because of their anatomical position, the three EC populations are in contact with GCs at different states of differentiation, suggesting that each EC subtype might send a distinct set of signals to regulate different stages of GC self-renewal or differentiation. To identify such potential regulators, we visualized the expression of all secreted proteins and ECM components that are differentially expressed among EC subtypes (Fig. 3J).

*Fs* (*Follistatin*), an inhibitor of activin and TGF- $\beta$  signaling (Bickel et al. 2008), was expressed specifically in cECs. Because the signaling of the TGF- $\beta$  homolog *dpp* promotes GSC self-renewal and impedes GC differentiation, it is plausible that *Fs* expression in the cECs may facilitate GC differentiation by suppressing the *dpp* signal emanating from the GSC niche. Thus, *Fs* expression in cECs may provide an additional layer of BMP signaling regulation in the germarium (Supplemental Fig. S4D; Song et al. 2004).

aECs expressed Wnt signaling pathway components and regulators, including ligands *Wnt4* and *Wnt6*, and *Notum*, which regulates Wnt ligand activity (Kakugawa et al. 2015). aECs also expressed EGF and FGF ligands *vn* and *ths*. Indeed, we observed higher levels of pERK, a EGFR and FGFR pathway readout, in the anterior and central part of the germarium (region 1 and 2) than in region 3 (Supplemental Fig. S4E). *Insulin-like peptide 6* (*Ilp6*) was specifically expressed in cECs, and two regulators of insulin/IGF signaling *Idgf1* and *Idgf3* were expressed in aECs and cECs. Thus, a number of signaling pathway activators and regulators were expressed in distinct regions of the germarium and may regulate GC differentiation. Moreover, ECs may also signal to each other and to additional somatic cells of the germarium, such as TFs and CCs, or FSCs. Indeed, it has been suggested that ECs regulate FSC maintenance (Vied et al. 2012; Sahai-Hernandez and Nystul 2013). cECs also expressed a number of secreted peptidases (*Cys*, *CtsB1*, *Cp1*), which may be involved in signaling molecule activation or degradation. Moreover, peptidase inhibitors *nec*, *CG17739*, and *Kal1* were expressed in adjacent EC clusters and thus may refine the peptidase activity domain. pECs were enriched for genes required for collagen production and secretion (*PH4aEFB*, *Plod*), whereas aECs and cECs expressed matrix metalloprotease (*Mmp1*) and its inhibitor (*Timpp*), which regulate ECM structure, suggesting that there is a distinct ECM structure and composition in different parts of the germarium. Moreover, a number of adhesion molecules and axon guidance molecules (*Neta*, *Ten-a*, *Sema1a*) were differentially expressed between the EC subtypes. Altogether, these results indicate that each EC subgroup secretes a different set of molecules, which may create distinct microenvironments for GC development and differentiation along the anterior-posterior axis of the germarium. Because signaling between ECs and GCs relies at least in part on large surface contacts between GCs and EC protrusions (Schulz et al. 2002; Kirilly et al. 2011; Banisch et al. 2017), we sought to search for physical interactor pairs (such as transmembrane–transmembrane and secreted–transmembrane) that may mediate signaling between GCs and ECs (Guruharsha et al. 2011). We selected secreted ligands, extracellular matrix proteins (ECM), and transmembrane proteins enriched in ECs (158 genes) or GC clusters adjacent to ECs (GSC/CB/2-cc, 4-cc, 8-cc, 16-cc 2a I, 16-cc 2a II, and 16-cc

2ab; 200 genes). Forty-nine genes (16 expressed in GCs, 23 in ECs, and 10 in both) formed 119 predicted protein:protein interaction pairs between transmembrane/secreted proteins in GCs and ECs (Supplemental Table S4). In 30 pairs, both genes had GO terms associated with plasma membrane or extracellular space, and in 65 pairs, at least one of the genes was uncharacterized (in 17 both were uncharacterized). Further analyses of the physical interaction pairs may reveal novel mechanisms for GC differentiation regulation by ECs.

Altogether, our data, together with recent reports by Rust et al. (2020), Tu et al. (2021), and Shi et al. (2021), identify multiple EC subtypes and suggest functional specialization among ECs. Subtype-specific markers, transcriptomes, and potential signaling molecules and interactors provide a resource for the functional characterization of EC subtypes and may reveal novel signaling mechanisms between GCs and soma.

### Identification of early FC lineages

We next sought to characterize the FC lineages present in the germarial somatic cell clusters. One cluster was rather distant from other clusters and expressed polar cell markers *upd1* and *upd3* (Fig. 3A–C). Visualization of two newly identified markers, *CG34377* and *Hk*, confirmed that the cluster corresponded to polar cells (cluster polar) (Fig. 3C,O; Supplemental Fig. S3R).

Stalk cells express *LamC* and *cas* (Song and Xie 2003; Chang et al. 2013). Two adjacent clusters expressed both markers, albeit at different levels (Fig. 3A,C). To determine the cluster identity, we searched for genes that were enriched and differentially expressed between the clusters. *sim* was expressed in both, whereas *CG10311* and *AdamTS-A* were each enriched in one of the clusters (Fig. 3C). *sim* was expressed in FCs at the region 3 and stage 2 follicle interface and in stalk cells (Fig. 3N; Supplemental Fig. S3Q), suggesting that the two clusters correspond to stalk cells and FCs differentiating into stalk cells (prestalk). *CG10311* was strongly expressed in stalk and weakly expressed in prestalk cells, whereas *AdamTS-A* was expressed in prestalk (and other pre-FCs) but absent from stalk cells, indicating that the *AdamTS-A*-expressing cluster corresponds to the prestalk cells and the *CG10311*-expressing cluster corresponds to the stalk cells (Fig. 3C,N; Supplemental Fig. S3Q).

Finally, the remaining cluster expressed high levels of *AdamTS-A* and *CG1136*, which are both expressed in pre-FCs in germarium regions 2b and 3 (Fig. 3K,L). Therefore, we concluded that this cluster contains FSCs and pre-FCs (FSC/pre-FC). Consistently, this cluster also expressed high levels of the pre-FC marker *eya* (Fig. 3C; Bai and Montell 2002). Our clustering analyses did not reveal a clear FSC population, likely because the low numbers of FSCs per germarium resulted in too few FSC transcriptomes present for our analyses (Reilein et al. 2017; Fadiga and Nystul 2019), or their transcription profiles are highly similar to pre-FCs (Nystul and Spradling 2010; Hartman et al. 2015; Reilein et al. 2017; Rust and Nystul 2020).

Aiming to identify the FSCs and their markers, we performed pseudotime analyses on follicle cell lineages from germarium soma clusters (FSC/pre-FC, prestalk, stalk, polar) and early epithelial FC (St 2-4I, St 2-4 II) using Monocle, an excellent tool to align cells on a “pseudotime” axis and study gene expression dynamics over the course of differentiation (Qiu et al. 2017). These analyses revealed three branches of FC differentiation, one corresponding to epithelial FCs, one to polar cells, and the third to stalk cells, where FCs transition through a prestalk state before acquiring stalk

cell identity (Supplemental Fig. S5A,B; Assa-Kunik et al. 2007). *eya* was down-regulated as FCs committed to polar and stalk cell fate, and remained highly expressed in epithelial FCs (Supplemental Fig. S5C). In contrast, *CG10311* was up-regulated and *AdamTS-A* was down-regulated in stalk cells over time, and *upd1* was up-regulated in polar cells (Supplemental Fig. S5D–F). Thus, our Monocle analysis agrees well with literature reports (Tworoger et al. 1999; Beccari et al. 2002) and with our cluster assignment to cell types. We, therefore, used the results of this analysis for identification of putative FSCs and their markers.

Monocle analyses suggested two putative markers for FSCs, *if* and *CG6044* (Supplemental Fig. S5G–I), and we visualized their expression using HCR. To discern the precise location of the region 2a/2b boundary where FSCs are positioned, we labeled with *Fas3* mRNA, which labels the follicle cell lineage, and DAPI, which labels cell nuclei. We found that *if* was expressed in the region where FSCs reside; however, its expression extended to pre-FCs and ECs, where it was weakly expressed (Supplemental Fig. S5J). *CG6044* was predominantly expressed in the pre-FCs, with weak expression at the germarium region 2a and 2b boundary where FSCs reside (Supplemental Fig. S5K). Therefore, neither *if* nor *CG6044* are FSC-specific markers. We were unable to identify FSC-specific markers, likely because the FSC gene expression profile is highly similar to the pre-FC gene expression profiles (Nystul and Spradling 2010; Hartman et al. 2015; Reilein et al. 2017), and therefore they are indistinguishable by bioinformatics in our data set. Increasing the number of cells of interest might increase the statistical power to alleviate these limitations. Rust et al. (2020) used a similar approach to identify FSCs and their markers. Although their markers are, indeed, enriched in our putative FSC population, their mRNA expression extends to ECs and pre-FCs similarly to the genes we identified as marking FSC populations (Supplemental Text; Supplemental Fig. S10C).

Three hundred ninety-nine genes were differentially expressed between the FC clusters, and we visualized their expression in a heat map (Fig. 3M). The polar cell cluster displayed the most distinct signatures, whereas prestalk and stalk cells shared a fraction of their marker genes. We performed GLAD analyses on all marker genes (Supplemental Fig. S4B). Major signaling pathway components were enriched in FC lineages, consistent with high signaling activity during FC lineage specification. FC lineage clusters were also enriched for cytoskeletal genes. The FSC/pre-FSC cluster was particularly enriched for cell division-related cytoskeletal proteins, consistent with active proliferation. A number of TF/CC-specific genes were also enriched in stalk cells. Both cell types are postmitotic and share a similar flattened shape, suggesting that many of the shared genes may comprise cell shape regulators. Moreover, we have visualized expression of transcription factors that are differentially expressed in follicle cell lineages (Supplemental Fig. S4C). Further studies will determine whether they are required for acquisition of distinct FC fates. Altogether, we determined FC lineage gene expression profiles and markers. Exploration of the gene expression dynamics as FSCs transition into specialized FC fates may reveal novel regulatory mechanisms for specification of FC lineages.

### Identification of epithelial FC subtypes

Epithelial follicle cells together with GCs transition through 14 stages of oogenesis and are the most abundant ovarian cell type. FC and GC progression through oogenesis is highly coordinated, because FCs play an essential role in egg development by deter-

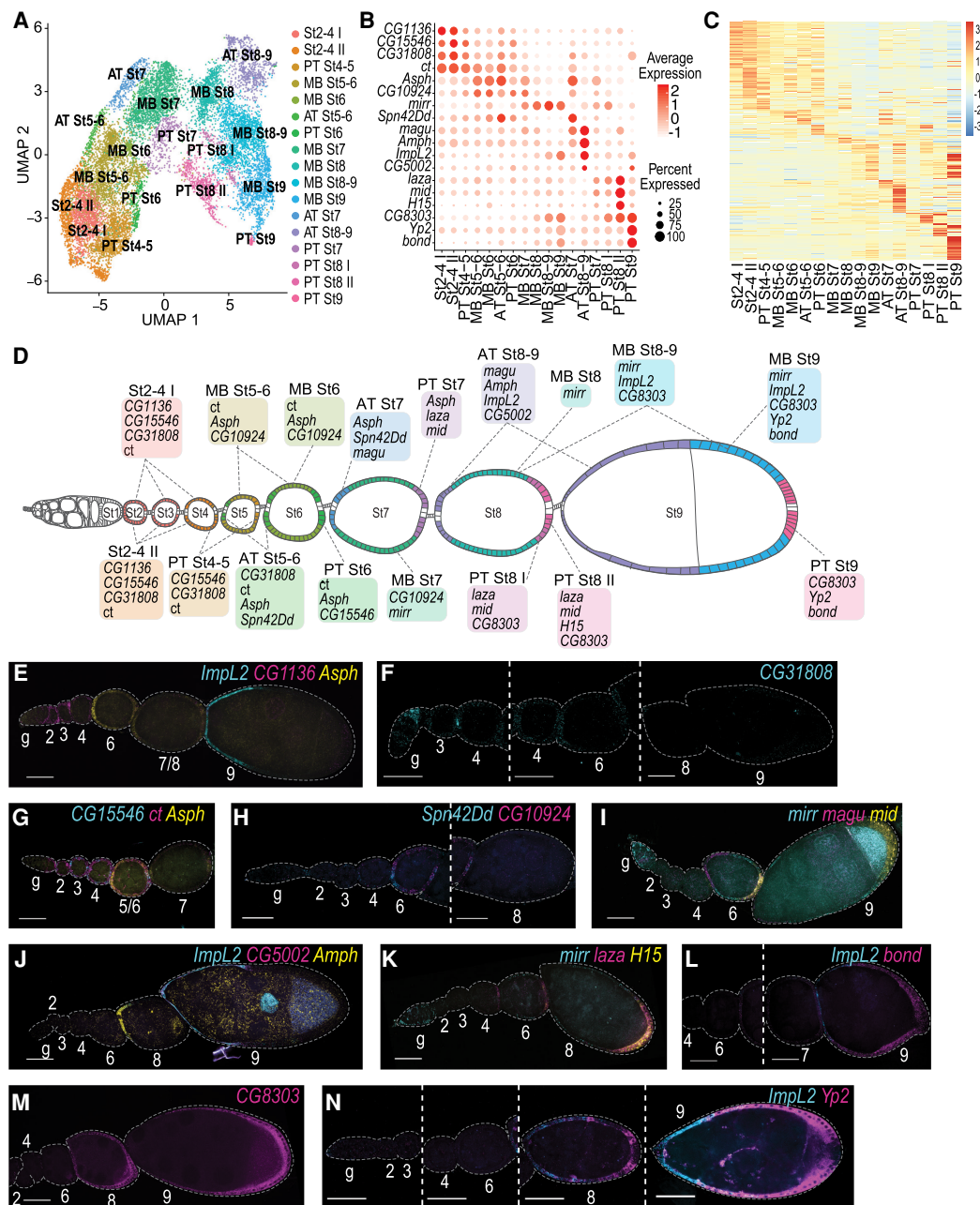
mining the body axes of the future embryo and producing essential yolk proteins (Wu et al. 2008). Determining the transcriptional differences between FC subtypes will strengthen our understanding of how this coordination is achieved. Because we manually removed late stage follicles during dissections (see Methods), we expected to only identify FC transcriptomes from early- and mid-stage follicles. Indeed, the stage 9/10A FC marker *Yp2* was well expressed, but the stage 10B marker *Cp36* expression was barely detectable in our data set (Supplemental Fig. S6A; Tootle et al. 2011), indicating that our data set contains FC transcriptomes up to stage 9/10A. Therefore, we focused on FC development up to that stage.

Follicle cells progress from mitotic (stage 1–6) to transitioning (stage 7) to endocycling states (stage 7–9) (Fig. 4D; Deng et al. 2001). Starting from stage 5, FCs differentiate into multiple subtypes determined by external cues, which largely depend on the anatomical position within the follicle (Wu et al. 2008; Duhart et al. 2017). FCs at the anterior and posterior tip of the follicle receive JAK/STAT signal (*upd1* and *upd3*) from polar cells, and assume terminal fate, while the rest become mainbody (MB) follicle cells (MBFCs) (Xi et al. 2003). Anterior terminal (AT) follicle cells (ATFCs) assume three distinct fates determined by the distance from the tip—border cells, stretched cells, and centripetal cells (Xi et al. 2003). Posterior terminal cells receive an additional EGF signal (*grk*) emanating from the oocyte nucleus starting from stage 6 and assume the posterior terminal (PT) follicle cell (PTFC) fate (González-Reyes et al. 1995). Initially, epithelial FCs have cuboidal shape, and, at stage 9, the PTFCs transition into columnar-shaped FCs. Simultaneously, border cells together with anterior polar cells start migrating posteriorly between the nurse cells. Stretched cells start producing a TGF- $\beta$  ligand (*Dpp*), and in response to it, flatten in a wave from the anterior to posterior covering the rapidly growing nurse cells (Brigaud et al. 2015). The MBFCs cover the growing oocyte and assume columnar shape as well. At stage 9/10A, the oocyte nucleus migrates to the anterior dorsal corner of the oocyte, and by activating EGFR signaling, specifies the dorsal FCs (Spradling 1993; González-Reyes et al. 1995). Centripetal cells, which initially are located between stretched cells and MBFCs, start migrating between the nurse cells and oocyte at stage 10B (Wu et al. 2008). Further follicle cell morphogenesis events occur beyond this stage, and each follicle cell subtype has a specific role in forming various eggshell structures (Duhart et al. 2017).

To find gene expression profiles of epithelial follicle cell subpopulations, we performed in-depth analyses on FC I, II, and III clusters. We first assessed expression of known markers to assign transcriptomes to distinct stages of oogenesis and follicle cell subtypes.

For staging, we assessed the expression of *ct*, *stg*, *CycB*, and *peb*. *ct* and *CycB* are expressed in mitotic FCs up to stage 6, whereas *stg* expression extends into stage 7 (Fig. 4D,G; Supplemental Figs. S6A, S7B; Jackson and Blochlinger 1997; Deng et al. 2001; Sun and Deng 2005). Transcription factor *peb* (also known as *Hindsight*) is expressed sporadically starting at stage 6 and its expression is strongly up-regulated during stage 7 and beyond (Supplemental Fig. S6A; Sun and Deng 2007). As mentioned above, *Yp2* marks stage 9 follicles (Fig. 4D,N; Supplemental Figs. S6A, S7H; Tootle et al. 2011). Therefore, the end of stage 6 is marked by down-regulation of *CycB* and *ct*, stage 7 is marked by overlapping expression of *stg* and *peb* (Supplemental Fig. S6A), *peb* expression in absence of *stg* marks FCs starting from stage 8, and *Yp2* marks FCs starting from stage 9. Therefore, in the UMAP plot (Fig. 4A), transcriptomes of early stage FCs are located in the





**Figure 4.** Characterization of 17 FC subpopulations. (A) UMAP plot of FCs subclustered in 17 clusters. Each dot represents a transcriptome of a single cell and is color-coded according to cluster membership. (B) A dot plot visualizing expression of select 18 marker genes. Dot diameter represents the fraction of cells expressing each gene in each cluster, as shown in the scale. Color intensity represents the average normalized expression level. (C) A heatmap visualizing average gene expression levels of all FC marker genes in each cluster. Red indicates highest and blue lowest expression. (D) Schematic drawing of an ovariole with FC subtypes highlighted in the colors corresponding to the UMAP plot in A. Note: Contrary to the schematic drawing, in vivo follicles of all stages are not represented equally frequently. (E–N) Marker gene mRNA in situ hybridization using HCR (cyan, magenta, yellow). Follicle stages (2–9) and germarium (g) are indicated above or below each follicle. Scale bars: 30 μm. (E) *Impl2* (cyan) is expressed in stretch cells at stage 9, *CG1136* (magenta) is expressed in pre-FCs, and stage 2–4 FCs, *Asph* (yellow) expressed in stage 6 FCs, and restricted to both termini at stage 7. (F) *CG31808* (cyan) is faintly expressed in stage 3–4 FCs and is restricted to ATFCs at stage 6. (G) *CG15546* (cyan) has patchy expression in stage 2–4 follicles and is restricted to PTFs by stage 5–6, *ct* (magenta) is expressed in all FCs from pre-FC to stage 6, *Asph* (yellow) expressed in stage 5–6 follicles. (H) *Spn42Dd* (cyan) is expressed in ATFCs at stage 6, *CG10924* (magenta) is expressed in MBFCs at stage 6. (I) *mirr* (cyan) is expressed in MBFC from stage 6–7 onward, *magu* (magenta) is expressed in ATFCs from stage 6 and is down-regulated by stage 9, *mid* (yellow) expressed in PTFs starting from stage 6. (J) *Impl2* (cyan) is expressed in ATFCs excluding the most anterior tip at stage 8 and in stretch and border cells at stage 9, *CG5002* (magenta) is expressed in stretch cells at stage 9, *Amph* (yellow) expressed in ATFCs at stage 8. (K) *mirr* (cyan) is expressed in MBFCs from stage 6–7 onward, *laza* (magenta) is weakly expressed in PTFs starting from stage 4 and strongly expressed in a broad band of PTFs at stage 8, *H15* (yellow) is expressed in a narrow band of the most posterior PTFs at stage 8; weak expression is detectable in PTFs at stage 6. (L) *Impl2* (cyan) is expressed in stretch cells at stage 9, *bond* (magenta) is expressed in a posterior-to-anterior gradient in PTFs and MBFCs at stage 9. (M) *CG8303* (magenta) is expressed in a posterior-to-anterior gradient in PTFs and MBFCs at stages 8 and 9. (N) *Impl2* (cyan) is expressed in ATFCs at stage 8, extending into MBFCs, and in stretched cells at stage 9, *Yp2* (magenta) is expressed in a posterior-to-anterior gradient in PTFs and MBFCs at stages 8 and 9. Its expression is excluded from the most posterior tip.

bottom left corner and transition in an arch toward the right where FCs differentiate into multiple subtypes.

Next, we sought to assess whether the anatomical position correlated with specific UMAP regions. We assessed expression patterns of a number of MB and terminal follicle cell markers. *mirr* marks MBFCs and was expressed in a band in the middle of the UMAP (Fig. 4A,I,K; Supplemental Figs. S6A, S7D,F; Jordan et al. 2000). *mid*, *H15*, and *pnt* are expressed at the posterior terminus at stage 6 and onward, and we observed their expression in a band on the bottom of the UMAP (Fig. 4A,I,K; Supplemental Figs. S6A, S7D,F; Morimoto et al. 1996; Lomas et al. 2013). Following dorsal migration of the oocyte nucleus at stage 9, *pnt*, but not *mid* or *H15*, expression is induced in anterior dorsal cells (Morimoto et al. 1996; Lomas et al. 2016). In our data set, *pnt* was always coexpressed with *mid* and *H15*, indicating that the anterior dorsal stage 9/10A cells are absent in our data set. *dpp* is expressed in stretched cells, and we observed its expression in the top right corner of the UMAP (Supplemental Fig. S6A; Xi et al. 2003). The border cell and centripetal cell marker *slbo* was detected in a very small number of cells adjacent to *dpp*-expressing cells (Supplemental Fig. S6A; Murphy et al. 1995). Therefore, AT follicle cells are on the top, MB in the middle, and PT on the bottom of the UMAP plot (Fig. 4A), indicating that the transcriptome variation in FCs is driven by both the anatomic position in the follicle and the oogenesis stage.

To obtain precise subdivision of FCs by stage and subtype, we subclustered the transcriptomes into 15 clusters and manually split two clusters (Fig. 4A; Supplemental Text). We identified specific markers for each cluster and visualized their expression along with a few previously described markers (Fig. 4B,E-N; Supplemental Fig. S7A–H).

*CG1136* was predominantly expressed in FCs of very early follicles up to stage 4 (clusters St2-4 I, St2-4 II) (Fig. 4B,D,E; Supplemental Figs. S6A, S7A). *CG15546* expression was patchy in stage 2–4 follicles (clusters St2-4 I, St2-4 II, PT St4-5), and slightly enriched at the termini prior to enriching to the PT at stage 6 (cluster PT St6) (Fig. 4B,D,G; Supplemental Figs. S6A, S7B), whereas *CG31808* had weak patchy expression in stage 2–4 follicles and restricted predominantly to ATFCs at stage 5–6 (clusters St2-4 I, St2-4 II, PT St4-5, and AT St5-6) (Fig. 4B,D,F; Supplemental Fig. S6A). *ct* was expressed in FCs up to stage 6 as previously reported (clusters St2-4 I, St2-4 II, PT St4-5, MB St5-6, MB St6, AT St5-6, PT St6) (Fig. 4B,D,G; Supplemental Figs. S6A, S7B; Jackson and Blochlinger 1997). *Asph* was ubiquitously expressed starting from stage 5 and enriched at both termini by stage 7 (clusters MB St5-6, MB St6, AT St5-6, PT St6, AT St7, PT St7) (Fig. 4B,D,E,G; Supplemental Figs. S6A, S7A,B). MBFCs were labeled by patchy expression of *CG10924* at stage 6–7 (clusters MB 5-6, MB St6, MB St7), and the previously described transcription factor *mirr* from stage 6 onward (clusters MB St7, MB St8, MB St8-9, MB St9) (Fig. 4B,D,H,I,K; Supplemental Figs. S6A, S7C,D,F; Jordan et al. 2000).

We visualized expression of four genes expressed in AT. *Spn42Dd* was predominantly expressed in the ATFCs at stages 5–7 (clusters AT St5-6, AT St7) (Fig. 4B,D,H; Supplemental Figs. S6A, S7C). *magu* was expressed in AT and lowly at PT starting at stage 6, peaking at stage 8, and declining at stage 9 (clusters AT St7, AT St8-9) (Fig. 4B,D,I; Supplemental Figs. S6A, S7D). *Amph* was expressed in ATFCs at stage 8 (cluster AT St8-9), and its expression was undetectable in stretched cells at stage 9 (Fig. 4B,D,J; Supplemental Figs. S6A, S7E), whereas *CG5002* was expressed in stretched cells at stage 9 (cluster AT St8-9) (Fig. 4B,D,J; Supplemental Figs. S6A, S7E). *ImpL2* expression started in ATFCs

at stage 8 and reached far posterior in MBFCs but was excluded from the very anterior tip (clusters AT St8-9, MB St8-9, MB St9). In stage 9, *ImpL2* was expressed in flattened stretched cells and border cells (Fig. 4B,D,E,J,L,N; Supplemental Figs. S6A, S7A,E,G,H).

PTFCs were faintly labeled by *laza* and *mid* at early stages, reaching maximum at stage 8 (clusters PT St7, PT St8 I, PT St8 II) (Fig. 4B,D,I,K; Supplemental Figs. S6A, S7D,F). *CG8303* was enriched in PTFCs at stages 8 and 9 (clusters PT St8 I, PT St8 II, PT St9), *Yp2* and *bond* at stage 9 (cluster PT St9), but all three were expressed in MBFCs as well (clusters MB St9, and MB St8-9 for *CG8303* only) (Fig. 4B,D,L,M,N; Supplemental Figs. S6A, S7G,H). Stage 8 PT FCs were divided in two subpopulations: cluster PT St8 II corresponded to the most posterior tip and expressed *H15* along with *laza* and *mid*, and cluster PT St8 I corresponded to a band of cells adjacent to PT St8 II cells, which expressed *laza* and *mid* and only low levels of *H15* (Fig. 4B,D,I,K; Supplemental Figs. S6A, S7D,F).

### Characterization of FC transcriptional signatures

Five hundred eighteen genes were differentially expressed between FC subtypes, and we visualized their expression in a heat map (Fig. 4C). We observed temporal signatures (stage 2–6, and stage 8–9) and cell fate-dependent (ATFC, MBFC, PTFC) signatures. The gene expression signature of each cluster was composed of blending of both the temporal and cell fate signatures, especially during earlier stages of oogenesis.

To gain insights of functional differences between FC subtypes, we explored gene class enrichment in the signature of each cluster (Supplemental Fig. S6B–D; Supplemental Table S3; Hu et al. 2015). Transcription factors/DNA-binding proteins were enriched in multiple clusters, highlighting the role of transcriptional regulation in follicle cell subtype diversification. Cytoskeletal proteins were enriched in multiple early stage follicle cell subtypes and in ATFCs at later stages. The sets of genes, however, were distinct. The early proliferative FCs were enriched for cell division associated cytoskeletal proteins, whereas stage 7 and 8–9 ATFCs were predominantly enriched for actin binding proteins and actin regulators, many of which are known to be involved in cell morphogenesis events such as axon targeting. In FC cells, they are likely involved in stretched cell flattening.

Transporters, in particular, vacuolar ATPase (V-ATPase) subunits were enriched in late stage follicle cells, especially in cluster ATFCs St8-9. V-ATPase is a large multimolecular complex that uses the energy of ATP hydrolysis to create proton gradients, for example, to lower the pH of particular cell compartments, such as lysosomes or endosomes (Nelson et al. 2000). V-ATPases locate to the stretch cell plasma membrane at stage 13 of oogenesis to acidify nurse cells and facilitate their death and engulfment by stretch cells (Timmons et al. 2016). It is plausible that FCs begin up-regulating V-ATPase subunit genes as early as stage 8–9 in preparation for this acidification. Indeed, the first signs of nurse cell death can be observed already at stage 10 (Cooley et al. 1992), suggesting that V-ATPases may start executing their function well prior to stage 13.

Furthermore, we visualized expression of transcription factors that are differentially expressed among epithelial FC subtypes (Supplemental Fig. S6E). In addition, numerous other gene classes were enriched in individual clusters, and functional studies may reveal new insight into specific roles of follicle cell subtypes. Moreover, in-depth studies of differentially expressed transcription factors will help understand the transcriptional networks governing complex FC subtype specification and morphogenesis.

## Discussion

We generated a single-cell atlas of the stem cell compartment and early differentiating egg chambers of adult ovaries of *Drosophila melanogaster*. We characterized cell type-specific transcriptional signatures and identified novel markers. In only a few cases does a single marker gene uniquely identify a specific cell type, but in combination and by intensity of gene expression, these genes can be used as cell type markers that distinguish between related cell types and developmental transitions. With these transcriptional profiles, we generated functional predictions for 34 cell types and subtypes—nine states of GC differentiation, and 25 somatic cell types including GSC niche cells, three EC subtypes, FSC/pre-FCs, three clusters corresponding to polar and stalk cell lineages, and 17 epithelial FC subtypes. This extensive annotation will bolster future studies, for example, by capitalizing on cluster-specific markers to develop cell type-specific genetic tools and for the analysis of stage and cell type-specific functions.

We were unable to distinguish GSCs and FSCs from their daughters by clustering. This is likely due to the relatively low numbers of stem cells in our data set, but it also reflects the high similarity between the transcriptomes of stem cells and their daughters (Kai et al. 2005; Nystul and Spradling 2010; Slaidina and Lehmann 2014; Hartman et al. 2015; Reilein et al. 2017; Rust and Nystul 2020). This perdurance of the stem cell transcriptome might have a functional relevance. Indeed, CBs and germline cysts can dedifferentiate and compete with GSCs for niche occupancy (Xie and Spradling 1998; Liu et al. 2015). Likewise, FSC daughters migrate across the germarium and possibly compete for niche occupancy with other FSCs (Nystul and Spradling 2007; Reilein et al. 2017). Thus, stem cell daughters initially retain the ability to revert to a stem cell state, possibly because they do not extensively remodel their transcriptome immediately after the asymmetric division.

ECs have a dual role—they promote GSC self-renewal and GC differentiation forming a domain termed the differentiation niche (Schulz et al. 2002; Kirilly et al. 2011; Rojas-Ríos et al. 2012; Liu et al. 2015; Wang and Page-McCaw 2018). We and Rust et al. (2020) identified three EC subtypes—anterior, central, and posterior ECs. Similarly, Shi et al. (2021) and Tu et al. (2021) report multiple EC subtypes. Each EC subtype interacts with GCs of a particular differentiation state and likely sends and receives distinct signals (Tu et al. 2021). Likewise, EC morphologies differ between locations; cEC and pEC protrusions are longer than aEC protrusions, as they interact with increasingly larger germline cysts (Banisch et al. 2017). We uncovered that a number of secreted proteins, adhesion molecules, and ECM components are differentially expressed between EC subtypes, suggesting that each subtype creates a distinct microenvironment. Thus, as GCs progress through differentiation and move posteriorly along the germarium, their immediate microenvironment changes. These observations open the possibility that the spatial organization of distinct EC microenvironments supports progressive GC differentiation and that maturing GC may feed-back on their microenvironment to define and stabilize its pattern.

Highly granular clustering of GC transcriptomes and precise cluster identity assignments allowed us to identify transcription factors that are dynamically expressed over the course of differentiation. We uncovered that a number of Notch signaling responsive transcription factors (*Enhancer of split Complex*) are enriched in four- and eight-cell cysts, raising a possibility that Notch signaling regulates early steps of GC differentiation.

Two recent studies have generated similar adult ovary atlases (Jevitt et al. 2020; Rust et al. 2020). Two additional studies have focused mainly on escort cells (Shi et al. 2021; Tu et al. 2021). Each study had a distinct focus and approach (for a detailed comparative analysis of these studies and discussion, refer to [Supplemental Text](#); [Supplemental Figs. S8–S10](#)). Overall, our studies and those of Rust et al. (2020) and Jevitt et al. (2020) produced similar results ([Supplemental Figs. S8, S9](#)), despite technical differences in fly lines, sample preparation, and analytical methods. The majority of markers identified showed the expected expression in our data set ([Supplemental Fig. 10](#); [Supplemental Text](#)). The exact cluster boundaries occasionally differ between the studies—for example, for EC subtypes. Shi et al. (2021) and Tu et al. (2021) focused more specifically on ECs to understand their functional diversity. Our study provides a more granular subclustering of ovarian cell types and precise cluster mapping to cell types and differentiation stages by direct visualization of mRNAs in situ. Thus, we deliver a comprehensive resource of gene expression profiles and markers for each cluster and provide gene class annotations for transcriptional signatures and functional predictions.

Despite being one of the most extensively studied adult organs, our analysis of the *Drosophila* ovary reveals higher cell type diversity than previously anticipated. These findings suggest that numerous, yet unidentified, cell subpopulations with distinct functions exist even in the most thoroughly studied organs. The ongoing Human Cell Atlas and similar projects in model organisms will start revealing this complexity, although focused studies are needed to uncover the interplay between the subpopulations and their functions, finally allowing us to fully comprehend organ function in homeostasis and disease.

## Methods

### Ovary dissociation for scRNA-seq

For each sample, 30–40 pairs of adult ovaries were dissected one by one in ice-cold DPBS, no calcium, no magnesium (Thermo Fisher Scientific #14190136). Using forceps and a dissection needle, the anterior tip of the ovary was cut and transferred to a new well. We used two dissociation methods: Method A: Dissociation by type I collagenase and trypsin and FACS for purification (see Slaidina et al. 2020); Method B: Dissociation by type I collagenase and elastase and filtration for purification. Detailed description of the methods can be found in the [Supplemental Text](#).

### scRNA-seq data analysis

We used Seurat v3 for sc-RNA seq analyses following standard guidelines for data processing, identification of variable genes, dimensionality reduction, and cell clustering. For coarse clustering, we used resolution parameter 0.2. For step-wise clustering, we selected cell groups from initial coarse clustering—GCs, follicle cells, or somatic cells of germaria. The data were scaled, new PCAs were computed, and new UMAP plots were generated, followed by clustering with higher resolution parameters (GC; 0.6, somatic cells of germarium; 1, and FCs; 0.6). For marker identification, we used Wilcoxon test and visualized marker expression in heat maps using the *ph heatmap* package.

We used the GLAD online tool to align marker genes for each cluster with particular gene categories and assessed significant enrichment in each cluster's transcriptional signature. For physical interaction analyses, we selected genes from the following GLAD categories: GPCRs, Ion channels, Matrisome (ECM), Receptors, Secreted proteins, Transmembrane proteins, and Transporters



(Hu et al. 2015) from EC cluster markers, and select GC cluster (GSC/CB/2-cc, 4-cc, 8-cc, 16-cc 2a I, 16-cc 2a II, and 16-cc 2ab) markers. We created a putative GC surface molecule list and an EC surface molecule list. Next, we scanned the database generated by Guruharsha et al. (2011) to identify which entries from the GC surface molecule list are predicted to form physical protein:protein interaction pairs with entries from the EC surface molecule list. Next, we explored GO terms associated with genes forming pairs. In [Supplemental Table S4](#), we shaded entries with the GO term ‘extracellular space’ in yellow, entries with the GO term ‘plasma membrane’ in blue, and in green if both GO terms are associated with the gene.

### Monocle pseudotime analyses

We used Monocle 2 for pseudotime analyses of FSC differentiation (Qiu et al. 2017) using the standard workflow. Detailed description of the methods can be found in the [Supplemental Text](#).

### Data set comparison

To compare the five data sets, we used Seurat v4 (Hao et al. 2021) data integration with SCTransform. For marker comparison, we first computed positive markers using a Wilcoxon test on all data sets individually (Rust et al. 2020; Jevitt et al. 2020; Shi et al. 2021; Tu et al. 2021; and ours). For our data sets, we merged all data into a single data set. We omitted markers with an adjusted  $p$ -value  $>0.01$ , and selected 50 markers with the lowest  $p$ -values for each cluster. Note that some clusters had fewer than 50 markers. Next, we performed a pairwise comparison of marker overlap between the clusters and used Fisher’s exact test to calculate corrected  $p$ -values. On the dot plot, we displayed percentage of marker overlap for each cluster and  $-\log_{10}$  of the corrected  $p$ -value. For a detailed discussion of data set comparisons, see [Supplemental Figures S8–S10 and Supplemental Text](#).

### RNA in situ hybridization and immunofluorescence

The HCR protocol was adapted from Choi et al. (2018). Custom designed probes, probe hybridization buffer, probe wash buffer, and amplification buffer were procured from Molecular Instruments, Inc. Detailed description of the methods can be found in the [Supplemental Text](#).

### Data access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO); <https://www.ncbi.nlm.nih.gov/geo/> under accession number GSE162192.

### Competing interest statement

The authors declare no competing interests.

### Acknowledgments

The alpha-spectrin antibody developed by D. Branton and R. Dubreuil was obtained from the Developmental Studies Hybridoma Bank, created by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) of the National Institutes of Health (NIH) and maintained at the University of Iowa, Department of Biology, Iowa City, Iowa. pMad antibody was from Dr. Edward Laufer. We thank Drs. Katja Rust and Amy Reilein for help identifying follicle stem

cells in stained germaria. We thank Drs. Katja Rust and Todd Nystul for sharing their results prior to publication. We thank Drs. Katja Rust and Todd Nystul, Deeptiman Chatterjee and Wu-Min Deng, Shiyuan Chen and Ting Xie, and Rongwen Xi for sharing annotated data sets. We thank Drs. Sherilyn Grill and Lionel Christiaen for critical comments on the manuscript. We would like to thank the Genome Technology Center (GTC) for expert library preparation and sequencing. GTC is a shared resource partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center. Cell sorting technologies were provided by NYU Langone’s Cytometry and Cell Sorting Laboratory, which is supported in part by grant P30CA016087 from the NIH/National Cancer Institute. S.G. was supported by Dean’s Undergraduate Research Fund Grant, R.L. was supported by the Simons Foundation, NIH R37HD041900, and was a Howard Hughes Medical Institute (HHMI) investigator.

**Author contributions:** M.S. and R.L. conceived and designed the study; M.S., S.G., and T.U.B. performed experiments; M.S. collected and analyzed the data; and M.S. and R.L. wrote the manuscript.

### References

- Allbee AW, Rincon-Limas DE, Biteau B. 2018. Lmx1a is required for the development of the ovarian stem cell niche in *Drosophila*. *Development* **145**: dev163394. doi:10.1242/dev.163394
- Assa-Kunik E, Torres IL, Schejter ED, Johnston DS, Shilo B-Z. 2007. *Drosophila* follicle cells are patterned by multiple levels of Notch signaling and antagonism between the Notch and JAK/STAT pathways. *Development* **134**: 1161–1169. doi:10.1242/dev.02800
- Bai J, Montell D. 2002. Eyes absent, a key repressor of polar cell fate during *Drosophila* oogenesis. *Development* **129**: 5377–5388. doi:10.1242/dev.00115
- Banisch TU, Maimon I, Dadosh T, Gilboa L. 2017. Escort cells generate a dynamic compartment for germline stem cell differentiation via combined Stat and Erk signalling. *Development* **144**: 1937–1947. doi:10.1242/dev.143727
- Beccari S, Teixeira L, Rørth P. 2002. The JAK/STAT pathway is required for border cell migration during *Drosophila* oogenesis. *Mech Dev* **111**: 115–123. doi:10.1016/S0925-4773(01)00615-3
- Bickel D, Shah R, Gesualdi SsC, Haerry TE. 2008. *Drosophila* Follistatin exhibits unique structural modifications and interacts with several TGF- $\beta$  family members. *Mech Dev* **125**: 117–129. doi:10.1016/j.mod.2007.09.013
- Bolívar J, Pearson J, López-Onieva L, González-Reyes A. 2006. Genetic dissection of a stem cell niche: the case of the *Drosophila* ovary. *Dev Dyn* **235**: 2969–2979. doi:10.1002/dvdy.20967
- Boyle M, DiNardo S. 1995. Specification, migration and assembly of the somatic cells of the *Drosophila* gonad. *Development* **121**: 1815–1825. doi:10.1242/dev.121.6.1815
- Brigaud I, Duteyrat J-L, Chlasta J, Bail SL, Couderc J-L, Grammont M. 2015. Transforming Growth Factor  $\beta$ /activin signalling induces epithelial cell flattening during *Drosophila* oogenesis. *Biol Open* **4**: 345–354. doi:10.1242/bio.201410785
- Chang Y-C, Jang ACC, Lin C-H, Montell DJ. 2013. Castor is required for Hedgehog-dependent cell-fate specification and follicle stem cell maintenance in *Drosophila* oogenesis. *Proc Natl Acad Sci* **110**: E1734–E1742. doi:10.1073/pnas.1300725110
- Chen D, McKearin DM. 2003. A discrete transcriptional silencer in the *bam* gene determines asymmetric division of the *Drosophila* germline stem cell. *Development* **130**: 1159–1170. doi:10.1242/dev.00325
- Choi HMT, Schwarzkopf M, Fornace ME, Acharya A, Artavanis G, Stegmaier J, Cunha A, Pierce NA. 2018. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**: dev165753. doi:10.1242/dev.165753
- Cooley L, Verheyen E, Ayers K. 1992. Chickadee encodes a profilin required for intercellular cytoplasm transport during *Drosophila* oogenesis. *Cell* **69**: 173–184. doi:10.1016/0092-8674(92)90128-Y
- Decotto E, Spradling AC. 2005. The *Drosophila* ovarian and testis stem cell niches: similar somatic stem cells and signals. *Dev Cell* **9**: 501–510. doi:10.1016/j.devcel.2005.08.012
- Deng WM, Althausen C, Ruohola-Baker H. 2001. Notch- $\Delta$  signaling induces a transition from mitotic cell cycle to endocycle in *Drosophila* follicle cells. *Development* **128**: 4737–4746. doi:10.1242/dev.128.23.4737

- Drummond-Barbosa D, Spradling AC. 2001. Stem cells and their progeny respond to nutritional changes during *Drosophila* oogenesis. *Dev Biol* **231**: 265–278. doi:10.1006/dbio.2000.0135
- Duhart JC, Parsons TT, Raftery LA. 2017. The repertoire of epithelial morphogenesis on display: progressive elaboration of *Drosophila* egg structure. *Mech Dev* **148**: 18–39. doi:10.1016/j.mod.2017.04.002
- Eliazer S, Palacios V, Wang Z, Kollipara RK, Kittler R, Buszczak M. 2014. Lsd1 restricts the number of germline stem cells by regulating multiple targets in escort cells. *PLoS Genet* **10**: e1004200. doi:10.1371/journal.pgen.1004200
- Fadiga J, Nystul TG. 2019. The follicle epithelium in the *Drosophila* ovary is maintained by a small number of stem cells. *eLife* **8**: 1161. doi:10.7554/eLife.49050
- Flaherty MS, Salis P, Evans CJ, Ekas LA, Marouf A, Zavadil J, Banerjee U, Bach EA. 2010. *chimmo* is a functional effector of the JAK/STAT pathway that regulates eye development, tumor formation, and stem cell self-renewal in *Drosophila*. *Dev Cell* **18**: 556–568. doi:10.1016/j.devcel.2010.02.006
- Forbes AJ, Spradling AC, Ingham PW, Lin H. 1996. The role of segment polarity genes during early oogenesis in *Drosophila*. *Development* **122**: 3283–3294. doi:10.1242/dev.122.10.3283
- Fuller MT, Spradling AC. 2007. Male and female *Drosophila* germline stem cells: two versions of immortality. *Science* **316**: 402–404. doi:10.1126/science.1140861
- Godt D, Laski FA. 1995. Mechanisms of cell rearrangement and cell recruitment in *Drosophila* ovary morphogenesis and the requirement of bric à brac. *Development* **121**: 173–187. doi:10.1242/dev.121.1.173
- González-Reyes A, Elliott H, Johnston DS. 1995. Polarization of both major body axes in *Drosophila* by *gurken-torpedo* signalling. *Nature* **375**: 654–658. doi:10.1038/375654a0
- Guruharsha KG, Rual J-F, Zhai B, Mintseris J, Vaidya P, Vaidya N, Beekman C, Wong C, Rhee DY, Cenaj O, et al. 2011. A protein complex network of *Drosophila melanogaster*. *Cell* **147**: 690–703. doi:10.1016/j.cell.2011.08.047
- Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. 2021. Integrated analysis of multimodal single-cell data. *Cell* **184**: 3573–3587.e29. doi:10.1016/j.cell.2021.04.048
- Hartman TR, Ventresca EM, Hopkins A, Zinshteyn D, Singh T, O'Brien JA, Neubert BC, Hartman MG, Schofield HK, Stavrides KP, et al. 2015. Novel tools for genetic manipulation of follicle stem cells in the *Drosophila* ovary reveal an integrin-dependent transition from quiescence to proliferation. *Genetics* **199**: 935–957. doi:10.1534/genetics.114.173617
- Hill JH, Chen Z, Xu H. 2014. Selective propagation of functional mitochondrial DNA during oogenesis restricts the transmission of a deleterious mitochondrial variant. *Nat Genet* **46**: 389–392. doi:10.1038/ng.2920
- Hu Y, Comjean A, Perkins LA, Perrimon N, Mohr SE. 2015. GLAD: an online database of Gene List Annotation for *Drosophila*. *J Genomics* **3**: 75–81. doi:10.7150/jgen.12863
- Huynh J-R. 2006. Fusome as a cell-cell communication channel of *Drosophila* ovarian cyst. In *Cell-cell channels*, pp. 217–235. Springer, New York.
- Jackson SM, Blochlinger K. 1997. Cut interacts with Notch and protein kinase A to regulate egg chamber formation and to maintain germline cyst integrity during *Drosophila* oogenesis. *Development* **124**: 3663–3672. doi:10.1242/dev.124.18.3663
- Jevitt A, Chatterjee D, Xie G, Wang X-F, Otwell T, Huang Y-C, Deng W-M. 2020. A single-cell atlas of adult *Drosophila* ovary identifies transcriptional programs and somatic cell lineage regulating oogenesis. *PLoS Biol* **18**: e3000538. doi:10.1371/journal.pbio.3000538
- Jordan KC, Clegg NJ, Blasi JA, Morimoto AM, Sen J, Stein D, McNeill H, Deng WM, TwoRoger M, Ruohola-Baker H. 2000. The homeobox gene *mirror* links EGF signalling to embryonic dorso-ventral axis formation through Notch activation. *Nat Genet* **24**: 429–433. doi:10.1038/74294
- Kai T, Williams D, Spradling AC. 2005. The expression profile of purified *Drosophila* germline stem cells. *Dev Biol* **283**: 486–502. doi:10.1016/j.ydbio.2005.04.018
- Kakugawa S, Langton PF, Zebisch M, Howell S, Chang T-H, Liu Y, Feizi T, Bineva G, O'Reilly N, Snijders AP, et al. 2015. Notum deacylates Wnt proteins to suppress signalling activity. *Nature* **519**: 187–192. doi:10.1038/nature14259
- Kawashima T, Nakamura A, Yasuda K, Kageyama Y. 2003. Dmef, a novel member of Maf transcription factor family is expressed in somatic gonadal cells during embryonic development and gametogenesis in *Drosophila*. *Gene Expr Patterns* **3**: 663–667. doi:10.1016/S1567-133X(03)00093-0
- King RC. 1970. *Ovarian development in Drosophila melanogaster*. Academic, New York.
- Kirilly D, Xie T. 2007. The *Drosophila* ovary: an active stem cell community. *Cell Res* **17**: 15–25. doi:10.1038/sj.cr.7310123
- Kirilly D, Wang S, Xie T. 2011. Self-maintained escort cells form a germline stem cell differentiation niche. *Development* **138**: 5087–5097. doi:10.1242/dev.067850
- Lasko PF, Ashburner M. 1988. The product of the *Drosophila* gene *vasa* is very similar to eukaryotic initiation factor-4A. *Nature* **335**: 611–617. doi:10.1038/335611a0
- Li MA, Alls JD, Avancini RM, Koo K, Godt D. 2003. The large Maf factor Traffic Jam controls gonad morphogenesis in *Drosophila*. *Nat Cell Biol* **5**: 994–1000. doi:10.1038/ncb1058
- Lieber T, Jeedigunta SP, Palozzi JM, Lehmann R, Hurd TR. 2019. Mitochondrial fragmentation drives selective removal of deleterious mtDNA in the germline. *Nature* **570**: 380–384. doi:10.1038/s41586-019-1213-4
- Lin H, Spradling AC. 1995. Fusome asymmetry and oocyte determination in *Drosophila*. *Dev Genet* **16**: 6–12. doi:10.1002/dvg.1020160104
- Liu Z, Zhong G, Chai PC, Luo L, Liu S, Yang Y, Baeg G-H, Cai Y. 2015. Coordinated niche-associated signals promote germline homeostasis in the *Drosophila* ovary. *Journal of Cell Biology* **211**: 469–484. doi:10.1083/jcb.201503033
- Lomas MF, Hails J, Lachance J-FB, Nilson LA. 2013. Response to the dorsal anterior gradient of EGFR signaling in *Drosophila* oogenesis is prepatterned by earlier posterior EGFR activation. *Cell Rep* **4**: 791–802. doi:10.1016/j.celrep.2013.07.038
- Lomas MF, Vito SD, Lachance J-FB, Houde J, Nilson LA. 2016. Determination of EGFR signaling output by opposing gradients of BMP and JAK/STAT activity. *Curr Biol* **26**: 2572–2582. doi:10.1016/j.cub.2016.07.073
- Mahowald A, Kambysellis M. 1980. Oogenesis. In *The genetics and biology of Drosophila* (ed. Ashburner M, Wright TRF), Vol. 2d, pp. 141–224. Academic Press, New York.
- Margaritis LH, Kafatos FC, Petri WH. 1980. The eggshell of *Drosophila melanogaster*. I. Fine structure of the layers and regions of the wild-type eggshell. *J Cell Sci* **43**: 1–35. doi:10.1242/jcs.43.1.1
- Margolis J, Spradling A. 1995. Identification and behavior of epithelial stem cells in the *Drosophila* ovary. *Development* **121**: 3797–3807. doi:10.1242/dev.121.11.3797
- Moore LA, Broihier HT, Van Doren M, Lunsford LB, Lehmann R. 1998. Identification of genes controlling germ cell migration and embryonic gonad formation in *Drosophila*. *Development* **125**: 667–678. doi:10.1242/dev.125.4.667
- Morimoto AM, Jordan KC, Tietze K, Britton JS, O'Neill EM, Ruohola-Baker H. 1996. Pointed, an ETS domain transcription factor, negatively regulates the EGF receptor pathway in *Drosophila* oogenesis. *Development* **122**: 3745–3754. doi:10.1242/dev.122.12.3745
- Mottier-Pavie VI, Palacios V, Eliazer S, Scoggin S, Buszczak M. 2016. The Wnt pathway limits BMP signaling outside of the germline stem cell niche in *Drosophila* ovaries. *Dev Biol* **417**: 50–62. doi:10.1016/j.ydbio.2016.06.038
- Murphy AM, Lee T, Andrews CM, Shilo BZ, Montell DJ. 1995. The breathless FGF receptor homolog, a downstream target of *Drosophila* C/EBP in the developmental control of cell migration. *Development* **121**: 2255–2263. doi:10.1242/dev.121.8.2255
- Nelson N, Perzov N, Cohen A, Hagai K, Padler V, Nelson H. 2000. The cellular biology of proton-motive force generation by V-ATPases. *J Exp Biol* **203**: 89–95. doi:10.1242/jeb.203.1.89
- Nystul T, Spradling A. 2007. An epithelial niche in the *Drosophila* ovary undergoes long-range stem cell replacement. *Cell Stem Cell* **1**: 277–285. doi:10.1016/j.stem.2007.07.009
- Nystul T, Spradling A. 2010. Regulation of epithelial stem cell replacement and follicle formation in the *Drosophila* ovary. *Genetics* **184**: 503–515. doi:10.1534/genetics.109.109538
- Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* **14**: 979–982. doi:10.1038/nmeth.4402
- Reilein A, Melamed D, Park KS, Berg A, Cimet E, Tandon N, Vunjak-Novakovic G, Finkelstein S, Kalderon D. 2017. Alternative direct stem cell derivatives defined by stem cell location and graded Wnt signalling. *Nat Cell Biol* **19**: 433–444. doi:10.1038/ncb3505
- Riechmann V, Rehorn KP, Reuter R, Leptin M. 1998. The genetic control of the distinction between fat body and gonadal mesoderm in *Drosophila*. *Development* **125**: 713–723. doi:10.1242/dev.125.4.713
- Rojas-Ríos P, Guerrero I, González-Reyes A. 2012. Cytoneme-mediated delivery of hedgehog regulates the expression of bone morphogenetic proteins to maintain germline stem cells in *Drosophila*. *PLoS Biol* **10**: e1001298. doi:10.1371/journal.pbio.1001298
- Rust K, Nystul T. 2020. Signal transduction in the early *Drosophila* follicle stem cell lineage. *Curr Opin Insect Sci* **37**: 39–48. doi:10.1016/j.cois.2019.11.005
- Rust K, Byrnes LE, Yu KS, Park JS, Sneddon JB, Tward AD, Nystul TG. 2020. A single-cell atlas and lineage analysis of the adult *Drosophila* ovary. *Nat Commun* **11**: 5628–5617. doi:10.1038/s41467-020-19361-0



- Sahai-Hernandez P, Nystul TG. 2013. A dynamic population of stromal cells contributes to the follicle stem cell niche in the *Drosophila* ovary. *Development* **140**: 4490–4498. doi:10.1242/dev.098558
- Sahut-Barnola I, Godt D, Laski FA, Couderc JL. 1995. *Drosophila* ovary morphogenesis: analysis of terminal filament formation and identification of a gene required for this process. *Dev Biol* **170**: 127–135. doi:10.1006/dbio.1995.1201
- Schulz C, Wood CG, Jones DL, Tazuke SI, Fuller MT. 2002. Signaling from germ cells mediated by the *rhomboid* homolog *stet* organizes encapsulation by somatic support cells. *Development* **129**: 4523–4534. doi:10.1242/dev.129.19.4523
- Shi J, Jin Z, Yu Y, Zhang Y, Yang F, Huang H, Cai T, Xi R. 2021. A progressive somatic cell niche regulates germline cyst differentiation in the *Drosophila* ovary. *Curr Biol* **31**: 840–852.e5. doi:10.1016/j.cub.2020.11.053
- Silver DL, Montell DJ. 2001. Paracrine signaling through the JAK/STAT pathway activates invasive behavior of ovarian epithelial cells in *Drosophila*. *Cell* **107**: 831–841. doi:10.1016/S0092-8674(01)00607-9
- Slaidina M, Lehmann R. 2014. Translational control in germline stem cell development. *J Cell Biol* **207**: 13–21. doi:10.1083/jcb.201407102
- Slaidina M, Banisch TU, Gupta S, Lehmann R. 2020. A single-cell atlas of the developing *Drosophila* ovary identifies follicle stem cell progenitors. *Genes Dev* **34**: 239–249. doi:10.1101/gad.330464.119
- Song X, Xie T. 2003. *wingless* signaling regulates the maintenance of ovarian somatic stem cells in *Drosophila*. *Development* **130**: 3259–3268. doi:10.1242/dev.00524
- Song X, Zhu C-H, Doan C, Xie T. 2002. Germline stem cells anchored by adherens junctions in the *Drosophila* ovary niches. *Science* **296**: 1855–1857. doi:10.1126/science.1069871
- Song X, Wong MD, Kawase E, Xi R, Ding BC, McCarthy JJ, Xie T. 2004. Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, *bag of marbles*, in germline stem cells in the *Drosophila* ovary. *Development* **131**: 1353–1364. doi:10.1242/dev.01026
- Spradling AC. 1993. Developmental genetics of oogenesis. In *The development of Drosophila melanogaster* (ed. Bate M, Martinez Arias A), Vol. 1, pp. 1–70. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sun J, Deng W-M. 2005. Notch-dependent downregulation of the homeo-domain gene *cut* is required for the mitotic cycle/endocycle switch and cell differentiation in *Drosophila* follicle cells. *Development* **132**: 4299–4308. doi:10.1242/dev.02015
- Sun J, Deng W-M. 2007. Hindsight mediates the role of Notch in suppressing Hedgehog signaling and cell proliferation. *Dev Cell* **12**: 431–442. doi:10.1016/j.devcel.2007.02.003
- Timmons AK, Mondragon AA, Schenkel CE, Yalonetskaya A, Taylor JD, Moynihan KE, Etchegaray JI, Meehan TL, McCall K. 2016. Phagocytosis genes nonautonomously promote developmental cell death in the *Drosophila* ovary. *Proc Natl Acad Sci* **113**: E1246–E1255. doi:10.1073/pnas.1522830113
- Tootle TL, Williams D, Hubb A, Frederick R, Spradling A. 2011. *Drosophila* eggshell production: identification of new genes and coordination by Pxt. *PLoS One* **6**: e19943. doi:10.1371/journal.pone.0019943
- Tu R, Duan B, Song X, Chen S, Scott A, Hall K, Blanck J, DeGraffenreid D, Li H, Perera A, et al. 2021. Multiple niche compartments orchestrate step-wise germline stem cell progeny differentiation. *Curr Biol* **31**: 827–839.e3. doi:10.1016/j.cub.2020.12.024
- TwoRoger M, Larkin MK, Bryant Z, Ruohola-Baker H. 1999. Mosaic analysis in the *Drosophila* ovary reveals a common Hedgehog-inducible precursor stage for stalk and polar cells. *Genetics* **151**: 739–748. doi:10.1093/genetics/151.2.739
- Vied C, Reilein A, Field NS, Kalderon D. 2012. Regulation of stem cells by intersecting gradients of long-range niche signals. *Dev Cell* **23**: 836–848. doi:10.1016/j.devcel.2012.09.010
- Wang X, Page-McCaw A. 2018. Wnt6 maintains anterior escort cells as an integral component of the germline stem cell niche. *Development* **145**: dev158527. doi:10.1242/dev.158527
- Wu X, Tanwar PS, Raftery LA. 2008. *Drosophila* follicle cells: morphogenesis in an eggshell. *Semin Cell Dev Biol* **19**: 271–282. doi:10.1016/j.semdb.2008.01.004
- Xi R, McGregor JR, Harrison DA. 2003. A gradient of JAK pathway activity patterns the anterior-posterior axis of the follicular epithelium. *Dev Cell* **4**: 167–177. doi:10.1016/S1534-5807(02)00412-4
- Xie T, Spradling AC. 1998. *decapentaplegic* is essential for the maintenance and division of germline stem cells in the *Drosophila* ovary. *Cell* **94**: 251–260. doi:10.1016/S0092-8674(00)81424-5
- Xie T, Spradling AC. 2000. A niche maintaining germ line stem cells in the *Drosophila* ovary. *Science* **290**: 328–330. doi:10.1126/science.290.5490.328
- Zhang Y, Kalderon D. 2001. Hedgehog acts as a somatic stem cell factor in the *Drosophila* ovary. *Nature* **410**: 599–604. doi:10.1038/35069099

Received November 13, 2020; accepted in revised form August 9, 2021.



## A single-cell atlas reveals unanticipated cell type complexity in *Drosophila* ovaries

Maija Slaidina, Selena Gupta, Torsten U. Banisch, et al.

*Genome Res.* 2021 31: 1938-1951 originally published online August 13, 2021

Access the most recent version at doi:[10.1101/gr.274340.120](https://doi.org/10.1101/gr.274340.120)

---

**Supplemental Material** <http://genome.cshlp.org/content/suppl/2021/09/20/gr.274340.120.DC1>

**References** This article cites 84 articles, 37 of which can be accessed free at:  
<http://genome.cshlp.org/content/31/10/1938.full.html#ref-list-1>

**Open Access** Freely available online through the *Genome Research* Open Access option.

**Creative Commons License** This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

Affordable, Accurate  
Sequencing.



---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

---