Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary

Yulong Liu¹, Michelle E Kossack¹†, Matthew E McFaul¹†, Lana N Christensen¹, Stefan Siebert¹, Sydney R Wyatt¹, Caramai N Kamei², Samuel Horst¹, Nayeli Arroyo¹, Iain A Drummond², Celina E Juliano¹, Bruce W Draper¹*

¹Department of Molecular and Cellular Biology, University of California, Davis, Davis, United States; ²Mount Desert Island Biological Laboratory, Bar Harbor, United States

Abstract: Zebrafish are an established research organism that has made many contributions to our understanding of vertebrate tissue and organ development, yet there are still significant gaps in our understanding of the genes that regulate gonad development, sex, and reproduction. Unlike the development of many organs, such as the brain and heart that form during the first few days of development, zebrafish gonads do not begin to form until the larval stage (≥5 days post-fertilization). Thus, forward genetic screens have identified very few genes required for gonad development. In addition, bulk RNA-sequencing studies that identify genes expressed in the gonads do not have the resolution necessary to define minor cell populations that may play significant roles in the development and function of these organs. To overcome these limitations, we have used single-cell RNA sequencing to determine the transcriptomes of cells isolated from juvenile zebrafish ovaries. This resulted in the profiles of 10,658 germ cells and 14,431 somatic cells. Our germ cell data represents all developmental stages from germline stem cells to early meiotic oocytes. Our somatic cell data represents all known somatic cell types, including follicle cells, theca cells, and ovarian stromal cells. Further analysis revealed an unexpected number of cell subpopulations within these broadly defined cell types. To further define their functional significance, we determined the location of these cell subpopulations within the ovary. Finally, we used gene knockout experiments to determine the roles of foxl2l and wnt9b for oocyte development and sex determination and/or differentiation, respectively. Our results reveal novel insights into zebrafish ovarian development and function, and the transcriptome profiles will provide a valuable resource for future studies.

Editor's evaluation

This single-cell transcriptomic analysis of young adult zebrafish ovaries provides important new data to understand gene expression patterns in numerous ovarian cell types that lead to insights into how ovary development works, and most of the principles will likely apply across vertebrates. The work will interest researchers who study gonad development, sex determination, differences (or ‘disorders’) in sex development, and impacts of the environment (including toxic pollutants) on gonad development and function.

Introduction

Over the last several decades, the zebrafish has emerged as a model to study vertebrate gonad development and function. Zebrafish ovaries and testes contain homologs of most cell types present in mammalian gonads (e.g., Sertoli and Leydig cells in testes, and follicle and theca cells in ovaries). As
in mammals, the zebrafish ovary and testis form from a gonad primordium that is initially bipotential (reviewed in Siegfried and Draper, 2020). The bipotential stage in zebrafish lasts until approximately 15–20 days post-fertilization (dpf) when sex is determined. Though the mechanism by which sex is determined in domesticated zebrafish is not known, zebrafish have orthologs of most genes that drive sex differentiation in mammals. In several cases, mutational analysis has revealed these genes play conserved roles in sex determination and differentiation. Examples include the male-promoting gene **dmrt1** and the female-promoting genes **wnt4**, **foxi2a**, and **foxi2b** (Kossack et al., 2019; Webster et al., 2017; Yang et al., 2017). Though the embryonic origin of the bipotential gonad has yet to be determined in zebrafish, orthologs of genes that are expressed in, and required for, the development of the bipotential gonad in mammals, such as **Gata4** and **Wt1**, are also expressed in the bipotential gonad in zebrafish (Leerberg et al., 2017). Together, these data argue that gonad development in vertebrates is regulated by a largely conserved genetic program.

One major difference between zebrafish and mammalian reproduction is that the mammalian ovary produced a finite number of oocytes only during embryogenesis; by contrast, in many teleost (bony fish), such as the zebrafish, adult females can produce new oocytes throughout their lifetime due to the presence of self-renewing germline stem cells (GSCs; Beer and Draper, 2013; Cao et al., 2019b). Thus, the zebrafish adult ovary contains germ cells at all stages of development, from premeiotic germ cells to mature eggs, while all mammalian oocytes arrest in prophase of meiosis I (also called dictyate arrest) around the time of birth (Faddy et al., 1987; Selman et al., 1993). The ability to continuously produce new germ cells throughout their lifetime makes female zebrafish an excellent research organism to study female GSCs and the somatic cells that regulate their development.

The unipotent GSCs can divide both to replenish their numbers while also producing cells that differentiate into sex-specific gametes, either eggs or sperm. In organisms where GSCs have been identified, they localize to a special microenvironment called the GSC niche. This niche is crucial for maintaining most GSCs in an undifferentiated state while also allowing some to differentiate. GSCs and their niche have been previously identified in a limited number of organisms, such as **Drosophila**, **Caenorhabditis elegans**, and male mice (Xie and Spradling, 2000). Identified GSC niches are maintained by specialized somatic cells that produce short-range signals that regulate GSC proliferation and differentiation. For example, in **Drosophila**, somatic cap cells in the ovary maintain GSCs by expressing specific ligands, such as **decapentaplegic** (dpp, mammalian Bmp), to prevent GSC differentiation (Song et al., 2004; Xie and Spradling, 2000), while in the mouse testis, Sertoli cells in the seminiferous tubules express glial-derived neurotrophic factor (GDNF), which is required for GSC self-renewal (Meng et al., 2000).

GSCs have been identified in the zebrafish ovary, but little is known about the niche and mechanism by which GSCs are maintained. In the adult ovary, early germ cells localize to a discrete region on the surface of the ovary, called the germinal zone, which has been proposed to be the GSC niche (Beer and Draper, 2013). The germinal zone contains mitotically dividing GSCs, oocyte progenitor cells, and early meiotic oocytes (Beer and Draper, 2013; Draper et al., 2007). However, neither the GSC-intrinsic gene expression landscape nor the extrinsic GSC niche cells that likely regulate early GSC differentiation in zebrafish have been characterized.

The best-characterized somatic cell type in the zebrafish ovary are the follicle cells. Teleost follicle cells are homologous to mammalian granulosa cells and thus are the major oocyte support cell. Follicle and granulosa cells produce signals that regulate oocyte development and maturation, and also produce nutrients that are transported into the oocyte cytoplasm via gap junctions (Su et al., 2009). Unlike granulosa cells, which form a multi-cell-layered complex surrounding the mammalian oocyte, follicle cells in fish form only a single-cell layer around the oocytes. In mammals, granulosa cells are the major source of ovarian estradiol (E2) production (Ryan, 1979), and this is likely to be the case in zebrafish. In addition to follicle cells, E2 production also requires steroidogenic theca cells that produce androstenedione, the precursor that follicle cells use for E2 synthesis (Young and McNeilly, 2010). Theca cells are found around and between oocytes but are otherwise not well characterized in zebrafish. The final major cell type in the ovary are the stromal cells, which are composed of all other somatic cell types present, including connective tissue, blood vessels, and immune cells. The fibroblast-like interstitial stromal cells play a largely structural role, but the roles of the other stromal cell types in ovarian development and function are not well characterized in any vertebrates. Although these general cell populations have been observed in the zebrafish ovary, there are significant gaps.
in knowledge regarding the identification of the full diversity of cell types present in the ovary or the functional roles they play.

In this study, we characterize the complex cell environment in the zebrafish ovary using single-cell RNA sequencing. This data has allowed us to identify all known stages of germ cells, from GSCs to pre-follicle stage oocytes (stage IA; Selman et al., 1993), as well as the somatic cell populations, including follicle, theca, and stromal cells. Using subcluster analysis, we have defined subpopulations within these broad cell-type classifications and validated these subpopulations by determining where they reside in the ovary using hybridization chain reaction RNA fluorescent in situ hybridization (HCR RNA-FISH). Further, we demonstrated the accuracy of this dataset to identify developmentally relevant genes using gene knockout analysis. Our data provide strong support that orthologs of genes involved in mammalian ovary development and function play parallel roles in zebrafish, further supporting zebrafish as a relevant research animal to understand human ovarian development and diseases. This reference dataset is available in a processed and interactively browsable form through the Broad Institute Single Cell Portal (available here) and will serve as a resource to greatly enhance future studies of ovarian development and function in unprecedented detail.

Results and discussion
Single-cell RNA sequencing identifies all cell types present in the zebrafish juvenile ovary

To identify all cell types and states in the zebrafish ovary, we collected single-cell transcriptomes from ~25K cells isolated from 40-day post-fertilization (dpf) ovaries. We chose 40 dpf ovaries for three major reasons. First, by 30 dpf sex determination is complete and the initial bipotential gonad has committed to differentiating as either an ovary or testis (Kossack and Draper, 2019). Second, by 40 dpf all major ovarian somatic cell types are likely present (Rodríguez-Mari et al., 2005). Finally, as we were particularly interested in understanding how the developmental progression of GSCs is regulated, 40 dpf ovaries have a higher proportion of early-stage germ cells relative to adult ovaries (Draper, 2012).

We prepared three single-cell libraries from dissociated Tg(piwi1:egfp)uc02 transgenic ovaries, where eGFP is expressed in all germ cells (Figure 1—figure supplement 1; Leu and Draper, 2010). Two of these libraries were prepared using a dissociation method that favored isolation of single somatic cells. The third library was prepared from cells isolated using a milder dissociation method that favored germ cell survival, followed by fluorescence-activated cell sorting (FACS) to purify eGFP+ germ cells (Figure 1—figure supplement 1; see ‘Materials and methods’ for details). We used the 10X Genomics platform for library preparation and sequencing. Following quality control, data cleaning, and the removal of red blood cells, doublets, and ambient RNA, we recovered a total of 25,089 single-cell transcriptomes that comprise 10,658 germ cells and 14,431 somatic cells (Figure 1—figure supplement 2; see ‘Materials and methods’). We obtained germ cells with an average of 2510 genes/cell, 11,294 transcripts/cell and somatic cells with an average of 854 genes/cell, 4812 transcripts/cell.

Cell clustering analysis grouped cells into eight distinct populations using uniform manifold approximation and projection (UMAP; Figure 1, Figure 1—figure supplement 3A, and Supplementary file 1). We assigned these clusters provisional cell-type identities using the expression of known cell-type-specific genes as follows: germ cells – deadbox helicase 4 (ddx4; formerly vasa; Yoon et al., 1997); follicle cells – gonadal soma derived factor (gdf; Gautier et al., 2011); theca cells – cytochrome P450 cholesterol side-chain cleaving enzyme (cyp11a2; Parajes et al., 2013); vasculature – Fli-1 proto-oncogene (fli1a; Brown et al., 2000); neutrophils – myeloid-specific peroxidase (mpx; Lieschke et al., 2001); macrophages – macrophage-expressed gene 1 (mpeg1.1; Zakrzewska et al., 2010); NK cells – NK-lysin tandem duplicate 2 (inl.2; Carmona et al., 2017; Tang et al., 2017; Figure 1, Figure 1—figure supplement 3B). Except for blood vessels and blood cells, the remaining stromal cells are the most ill-defined cell population in the teleost gonad, and no genes have previously been identified that specifically mark this population. Ovarian stromal cells are generally defined as the cells that are not components of the ovarian follicle (i.e., germ and associated follicle and theca cells) and play structural roles (Kinnear et al., 2020). We determined that expression of the gene encoding collagen, type I, alpha 1a (col1a1a; Fisher et al., 2003), is highly enriched in the remaining unidentified population of cells in the UMAP graph. Collagen is well known for its structural role in tissues and therefore...
reasoned that these cells are ovarian stromal cells (Figure 1B, Figure 1—figure supplement 3B). Thus, we were able to identify all previously observed and expected cell types in our dataset. A list of genes with enriched expression in each cluster can be found in Supplementary file 1.

Our initial analysis revealed genes that are differentially expressed between major cell types but did not distinguish possible subpopulations. For example, ddx4 and gsdf were correctly identified as genes expressed specifically in germ cells and follicle cells, respectively; however, specific genes expressed in discrete developmental stages of these populations were not revealed. To gain a more refined view of each of the distinct cell populations, we extracted and reclustered each population.

**Developmental trajectory of female germ cells**

Unlike mammals, female zebrafish can produce new oocytes throughout their life span due to the presence of GSCs (Beer and Draper, 2013; Cao et al., 2019b). Zebrafish therefore provide a unique opportunity to study female GSCs in a vertebrate. The genes that regulate GSC maintenance or progression of progenitor cells toward differentiation are not well defined in zebrafish. To gain further
insight into genes that regulate germ cell development, we extracted and reclustered the germ cell population (Figure 2A) and the differentiation trajectory was immediately evident in the UMAP. To determine the organization of the cells within the UMAP, we first asked which cells express known stage-specific germ cell markers. Using nanos2, DNA meiotic recombination 1 (dmc1), and zona pellucida protein 3b (zp3b), which are expressed in GSCs, early meiotic germ cells (stage 1A pre-follicle phase), and early oocytes (stage 1B follicle phase), respectively (Beer and Draper, 2013; Onichtchouk et al., 2003; Selman et al., 1993; Yoshida et al., 1998), we found that nanos2-expression cells clustered to the left end of the UMAP, dmc1-expressing cells were in an intermediate location, and zp3b-expression cells clustered to the right end of the UMAP. Thus, the cells are organized into what appears to be a developmental trajectory (Figure 2A and B). Because our library preparation method excluded cells larger than 40 μm in diameter, it is likely that only stage 1A (pre-follicle stage) or early stage IB (follicle stage) oocytes are present in our dataset (Selman et al., 1993). To further assess this developmental trajectory, we performed a dedicated trajectory analysis and inferred pseudotime using Monocle 3 (Cao et al., 2019a). We found that the directionality and the sequential gene expression along pseudotime precisely correlated with the trajectory determined by our initial germ cell recluster analysis (Figure 2—figure supplement 1).

With the developmental trajectory of the germ cells constructed, we were able to identify genes with enriched expression in subsets of germ cells (Supplementary file 2). Notably missing from the stage-specific genes listed above are genes expressed in the proliferating oocyte progenitor cells, a population that is intermediate between the GSC and germ cells that have entered meiosis (subcluster 2 in Figure 2A). In most organisms, oocyte progenitor cells undergo several rounds of mitotic amplifying division before entering meiosis (Pepling et al., 1999). A common feature of oocyte progenitor cells, including those in zebrafish ovaries, is that they divide with complete nuclear division but incomplete cytoplasmic division, thus creating a multicell cyst with synchronized developmental progression (Marlow and Mullins, 2008; Pepling et al., 1999). To identify novel markers of oocyte progenitor cells, we searched for genes with enriched expression in this subpopulation (subcluster 2 in Figure 2A; Supplementary file 2). Of the top 100 enriched genes, we found only one gene whose expression was both germ cell-specific and restricted to this subpopulation, an unannotated gene, called zgc:194189 (Supplementary file 2). Further sequence analysis revealed that this gene is the zebrafish ortholog of forhead-box protein L2 like (foxl2; Figure 2—figure supplement 2; Ruzicka et al., 2019). FoxL2L is a paralog of FoxL2 and is present in most teleost genomes as well as in sharks, coelacanths, and spotted gar, but is absent in the genomes of land-dwelling vertebrates (Figure 2—figure supplement 2). Interestingly, medaka foxl2l (formerly foxl3; Figure 2—figure supplement 2) is expressed in oocyte progenitors and is required for progenitor cells to commit to the oocyte-fate (Nishimura et al., 2015). The foxl2l-positive cells in our dataset also express higher levels of proliferating cell nuclear antigen (pcna) relative to the GSC population (Figure 2B and C), suggesting that they divide more rapidly than GSCs, providing further support that this subpopulation are zebrafish oocyte progenitors.

To test the hypothesis that foxl2l-expressing cells are oocyte progenitor cells, we used HCR RNA-FISH to determine where these cells are located in the 40 dpf ovary. Specifically, we determined the location of subcluster 2 cells relative to GSCs (subcluster 1) and to cells that are initiating meiosis (subcluster 3). We identified GSCs using nanos2 expression and early meiotic cells using meiotic recombination protein 8a (rec8a) expression. nanos2 encodes an RNA-binding protein that is expressed specifically in GSCs (Beer and Draper, 2013) while rec8a encodes a meiosis-specific member of the Rad21 cohesin family that is loaded onto chromosomes during premeiotic S-phase (Crespo et al., 2019). Our trajectory analysis indicated that rec8a-expressing cells would be more developmentally advanced than those that express foxl2l, as would be expected given its suspected role in premeiotic S-phase (Figure 2B and C). Indeed, we found that foxl2l- and rec8a-expressing cells localize within multicell clusters (Figure 2E–H), consistent with our hypothesis that foxl2l is expressed in oocyte progenitors, whereas nanos2-expressing cells are found predominantly as single cells or doublets, as previously shown (Figure 2D–G; Beer and Draper, 2013). We found cells expressing nanos2, foxl2l, and rec8a often clustered near one another (Figure 2G). Also consistent with our hypothesis, foxl2l-positive cells were found as clusters of ≥4 cells and rec8a-positive cells were found in clusters of ≥8 cells (Figure 2I). Interestingly, we rarely found cells that were double-positive for nanos2 and foxl2l, and those that were had minimal expression of both genes, suggesting that these
Figure 2. Germ cell subcluster analysis reveals developmental transitions of early germ cells. (A) Germ cell subcluster uniform manifold approximation and projection (UMAP) plot, with cells color-coded by computationally determined cell subtypes. (B) Dot plot showing the relative expression of select genes in the germ cell subclusters. Some genes, like ddx4 and piwi1, are expressed in all germ cells, while others, such as nanos2 or rec8a, are only...
cells were transitioning from GSC to oocyte progenitor (Figure 2F). Cells expressing higher levels of foxl2l could be divided into two populations, based on morphology and gene expression. The first consisted of cells within smaller clusters (2–4 cells) that expressed high levels of foxl2l, but not rec8a. These cells are likely early oocyte progenitors (Figure 2E and G). The second consisted of cells within larger clusters (≥4 cells), expressed relatively less foxl2l but also expressed rec8a. These cells were likely late oocyte progenitors, including those that are undergoing premeiotic S-phase (Figure 2G–I). This latter result is consistent with recent studies in medaka that found rec8a is a direct downstream

Figure 2 continued

expressed in distinct subclusters. (C) Gene expression UMAP plots of select genes. (D–H) Triple hybridization chain reaction RNA fluorescent in situ hybridization (HCR RNA-FISH) for nanos2 (green), foxl2l (magenta), and rec8a (blue) in 40-day-old zebrafish whole-mount ovary. Asterisk in (F) indicates a cell double-positive for nanos2 and foxl2l. (I) Cell number quantification of individual cysts that express the genes indicated on Y-axis. n = 70, N = 3. Scale bar in (F), for (D–G) 10 μm; in (H) 10 μm.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Developmental trajectory analysis of the germ cell library using Monocle 3 produced a similar trajectory to that apparent in the SURAT-based cluster analysis.

Figure supplement 2. Phylogenetic analysis of Foxl2, Foxl2l, Foxl3, and Foxl1 proteins.

Figure supplement 3. Expression of nanos and pumilio orthologs in zebrafish germ cells.

Figure supplement 4. Germ cell library statistics and novel zebrafish germline stem cell (GSC) markers.

Figure supplement 5. Gene module analysis and motif enrichment identify putative germline stem cell (GSC)-specific transcription factors.

Figure 3. Mutational analysis of foxl2l. (A) Schematic diagram of the Foxl2l protein showing the DNA-binding forkhead homology domain (blue), the location of the nuclear localization signal (NLS), and the viral-2A-egfp insertion site in the foxl2l(uc91) allele. (B) GFP expression in germ cells from a foxl2l(uc91) knock-in allele heterozygote recapitulates endogenous foxl2l expression (compare to Figure 2E and F). (C) Sex ratios of foxl2l(uc91) heterozygotes and homozygotes. (D–F) Representative light micrographs of fish examined in (C) (n = 274, N = 4, *p=1 × 10^-7, **p=2 × 10^-7). Wild-type adult female zebrafish (D) has characteristic light-yellow pigmentation on ventral belly and a prominent anal papilla (highlighted with red dashed lines) (D'). Wild-type adult male zebrafish (E) has dark yellow pigmentation on ventral belly and lacks an anal papilla (highlighted with red dashed lines) (E'). foxl2l(uc91) homozygous mutant (F) is phenotypically male. IB, stage 1B oocyte.
target of foxl2 (Kikuchi et al., 2020). Thus, we identified gene expression patterns that define the premeiotic germ cell populations in the ovary.

To determine the function of foxl2, we used CRISPR/Cas9-mediated gene editing to recombine a viral-2A-eGFP insert in-frame into the foxl2 locus, resulting in the Tg(foxl2:foxl2-2A-egfp)uc91 allele (Wierson et al., 2020). The resulting fusion protein encodes all but the last 27 of 239 amino acids of the Foxl2 protein (Figure 3A). We found that in heterozygous animals GFP was expressed in a subset of premeiotic germ cells in the ovary (Figure 3B), identical to the pattern determined using HCR RNA-FISH (Figure 2E–H). To assess the role of foxl2 in germ cell development, we crossed heterozygous parents to produce homozygous mutant offspring. We found that heterozygous knock-in animals had normal sex ratios as adults (Figure 3C–E). By contrast, all homozygous knock-in mutant animals were fertile males as adults, indicating that foxl2 is required for female development (Figure 3C and F). In zebrafish, the ability to produce oocytes is a requirement for female development, thus loss of oocyte production, which occurs in foxl2 mutants, results in all-male development (Rodríguez-Mari et al., 2010; Shive et al., 2010). This is homologous to the role of foxl2 in medaka, which is expressed in oocyte progenitor cells and is required for these cells to commit to oogenesis, and upon loss of foxl2 these cells instead commit to spermatogenesis (Nishimura et al., 2015).

Female GSCs have not been well characterized in any vertebrate. To date, only two genes have been shown to play a cell-autonomous role in female GSC development in zebrafish: nanos2 and nanos3 (Beer and Draper, 2013; Cao et al., 2019b; Draper et al., 2007). nanos2 and nanos3 encode conserved RNA-binding proteins. In zebrafish, nanos2 is expressed specifically in GSC in both male and females and loss of nanos2 leads to sterile males (Beer and Draper, 2013; Cao et al., 2019b). By contrast, nanos3 is only expressed in females where it plays two separate roles. First, nanos3 (formerly called nanos1) is a maternally expressed gene whose mRNA localizes to primordial germ cells (PGCs) and loss of maternal function leads to loss of PGCs during early development (Draper et al., 2007; Köprunner et al., 2001). Second, nanos3 mutant females fail to maintain oocyte production as adults due to loss of nanos2-expressing GSCs (Beer and Draper, 2013; Draper et al., 2007). While genetic mosaic analysis established that nanos3 was required cell-autonomously for GSC maintenance (Beer and Draper, 2013), to date nanos3 expression in premeiotic germ cells in the ovary has not been demonstrated. To address this, we identified the nanos3-expressing cells in our dataset. As previously reported, nanos3 is expressed at high levels in early oocytes, but we also detected expression in the apparent GSC subpopulation, confirming that nanos3 is expressed in GSCs as suggested by the nanos3 mutant phenotype (Figure 2B and C).

In other organisms, Nanos proteins function in complex with members of the Pumilio family of RNA-binding proteins (De Keuckelaere et al., 2018; Forbes and Lehmann, 1998). The zebrafish genome encodes three orthologs of Pumilio, called Pum1–3, but it is not known which of these proteins function together with Nanos2 or Nanos3 to maintain GSC or PGC development. We found that pum1 and pum3, but not pum2, are expressed at significant levels in GSCs (Figure 2—figure supplement 3). In addition, pum1 is predominantly expressed in GSCs while pum3 is expressed more uniformly in early oocytes (Figure 2—figure supplement 3). This raises the possibility that Pum1 may partner with Nanos2 while Pum3 partners with Nanos3, though it is also possible that Nanos2 could form complexes with both Pum1 and Pum3. Further mutational analysis is required to test these relationships.

We detected the highest average number of unique transcripts in GSCs (3517 unique transcripts/cell; Figure 2—figure supplement 4) relative to other premeiotic cells (e.g., 2315 unique transcripts/cell in progenitor cells; Figure 2—figure supplement 4). To identify potential regulators of GSC development, in addition to nanos2, we identified genes in our dataset with enriched expression in this cell population (Supplementary file 2). They include inhibitor of DNA binding 4 (id4; Oatley et al., 2011), which encodes a negative regulator of transcription that is also expressed in mouse spermatogonial stem cells, jagged canonical Notch ligand 2b (jag2b; Haddon et al., 1998), dickkopf WNT signaling pathway inhibitor 3b (dkk3b; Hsu et al., 2010), DNA methyltransferase protein 3bb.1 (dnmt3bb.1; Gore et al., 2016), interferon regulatory factor 10 (irf10; Stein et al., 2007), and chromogranin b (chgb; Xie et al., 2008; Figure 2—figure supplement 4B). Using id4 and chgb as representatives, we performed HCR RNA-FISH to verify the expression pattern in ovaries. We found that all id4-expressing cells also express nanos2 (Figure 2—figure supplement 4C), confirming that id4 is expressed in GSCs in the 40 dpf ovary. Interestingly, not all nanos2-expressing cells were id4+.
We obtained similar results with cgbb (Figure 2-figure supplement 4D). Thus, we identified several new genes whose expression is specific to GSC. In future studies, it will be important to determine if these genes are expressed in both male and female GSCs or, like nanos3, are female-specific.

Identification of germ cell stage-specific gene modules and possible transcriptional regulators

We next aimed to identify the potential transcription factors that regulate the development of germ cells. As a first step, we performed nonzero matrix factorization (NMF) to find gene modules, which are sets of genes that are co-expressed within cell clusters and subclusters and may therefore identify co-regulated genes (Figure 2-figure supplement 5; Brunet et al., 2004; Farrell et al., 2018; Siebert et al., 2019). We identified 32 gene modules (note: modules 8, 13, 19, and 29 were determined to be low-quality gene modules and therefore not considered in our analysis; see Materials and methods), and many of the gene modules showed stage-specific enrichment during germ cell development. For example, genes within modules 25, 28, 10, and 9 were specifically enriched in the GSCs, oocyte progenitors, early meiosis, and oocyte clusters, respectively (Figure 2—figure supplement 5A and B). We next performed gene motif enrichment analysis to find potential transcription factor-binding sites enriched within 2 kb 5’ of the transcription start site of top 20% of genes within individual modules (see Material and methods). We then verified our gene motif enrichment using the figla transcription factor as a test case. Folliculogenesis-specific basic helix-loop-helix (bHLH) transcription factor (Figla) is known to regulate folliculogenesis-specific gene expression in mice and zebrafish (Liang et al., 1997; Qin et al., 2018; Soyal et al., 2000). We found figla is expressed in late meiotic germ cells and early oocytes (Figure 2—figure supplement 5C). We searched for modules that contained genes with enriched putative Figla-binding sites and identified module 9, a module that contains genes with enriched expression in early oocytes (e.g. zp3; Supplementary file 3; Figure 2—figure supplement 5A, C). Therefore, we were confident that NMF and gene motif enrichment analysis could accurately identify important regulatory motifs in germ cells. To identify putative GSC-specific transcription factors, we focused our analysis on modules 7 and 25, which contained genes whose expression was specifically enriched in the GSC cluster (Figure 2—figure supplement 5). We found that genes within these two modules were enriched for putative binding motifs of transcription factors Bhlhe23, early growth response 4 (Egr4), and pre-B-cell leukemia homeobox 3b (Pbx3b) (Figure 2—figure supplement 5A, D). Importantly, GSCs also had enriched expression of Bhlhe23, Egr4, and Pbx3b (Figure 2—figure supplement 5D). Previous studies of Egr4-deficient mice have shown reduced fertility, but the activities of Bhlhe23 and Pbx3b in GSC development and function have not been reported (Tourtellotte et al., 1999). We performed HCR RNA-FISH to confirm the expression of bhlhe23 and found expression specifically in a subset of nanos2-expressing GSCs (Figure 2—figure supplement 5E). Finally, the motif enrichment analysis also identified multiple genes with known functions in germ cell regulation that contain Bhlhe23 binding motifs, such as cnot1, which encoded a component of the CCR4/Not deadenylase complex that functions with Nanos proteins to regulate translation of select targets in GSCs (Suzuki et al., 2012) In addition, dnmt3bb.1, tmb4x, and mcm6 are expressed in GSC and have putative Bhlhe23 sites within 2 kb of their transcription start sites, suggesting that Bhlhe23 may regulate the expression of multiple genes (Figure 2—figure supplement 5F). The information from this analysis will allow us to build more accurate comparative gene regulatory networks between zebrafish and humans, which can be experimentally tested in the genetically tractable zebrafish.

Identification of three follicle cell subpopulations

Follicle cells form a single-layered epithelium that encases developing oocytes once they arrest at the diplotene stage of meiotic prophase I. Teleost follicle cells are homologous to mammalian granulosa cells as evidenced by shared functions and gene expression. For example, zebrafish follicle cells express orthologs of many of the core granulosa cell-expressed genes, such as forkhead-box protein L2 (foxl2a and foxl2b; Crespo et al., 2013), anti-Mullerian hormone (amh; Rodriguez-Mari et al., 2005), follicle-stimulating hormone receptor (fshr; Kwok et al., 2005), aromatase/cytochrome P450 19a1a (cyp19a1a; Dranow et al., 2016), and notch receptor 3 (notch3; Prasasya and Mayo, 2018; Figure 4B, Figure 4-figure supplement 1). Oocyte and follicle cell development is coordinated through bidirectional oocyte-follicle cell-cell interactions (Kidder and Vanderhyden, 2010). Soon
relative expression of select genes in the follicle cell subclusters. Some genes, like gsdf, are expressed in all follicle cells, while others, such as lhx9, are only expressed in distinct subclusters. (C) Gene expression UMAP plots of select cell types. (D) gpc1a expression (pink) is detected in follicle cells surrounding stage IB oocytes, but not stage II oocytes. (E) gja11 expression (pink) is detected in follicle cells surrounding stage II oocytes, but not stage IB oocytes. Ddx4 indirect immunofluorescence (green) labels all germ cells. (F) Triple HCR RNA-FISH shows lhx9-expressing cells (blue) form tracts on the surface of the ovary (arrows) that colocalize with nanos2 (green) and dmc1 (pink)-expressing germ cell stem cells and early meiotic cells, respectively. Lateral edge of the ovary is indicated with a dashed line. (F', F'') Higher-magnification views of regions boxed in (F) showing that lhx9+ cells (arrows) surround the germ cells. (G) Triple HCR RNA-FISH shows gpc1a-expressing cells (blue) are concentrated near the lateral edge of the ovary, a region that contains nanos2 (green) and dmc1 (pink)-expressing germ cell stem cells (arrows) and early meiotic cells, respectively (inset in G shows nanos2 and DNA only for dashed boxed region). Asterisks identify autofluorescent red blood cells. (G') bmp6 channel only. Scale bars in (D, E, G) for (G, G') 50 μm, (F) 100 μm, (F') for (F' and F'') and (G, inset), 20 μm. IB, stage IB oocyte, II, stage II oocyte.

The online version of this article includes the following figure supplement(s) for figure 4:

**Figure supplement 1.** Gene expression uniform manifold approximation and projection (UMAP) plots of select follicle cell-enriched genes.

**Figure supplement 2.** Subcluster analysis of pre-follicle cells (lhx9+).

**Figure supplement 3.** Gene expression uniform manifold approximation and projection (UMAP) plots for genes that function in the retinoic acid signaling pathway.
after an oocyte has formed, it is surrounded by follicle cells and becomes dependent on these cells for factors it cannot produce, such as pyruvate, cholesterol, and select amino acids. In some cases, these compounds are transported to the oocyte via connexin-mediated gap junctions (Su et al., 2008). Mammalian granulosa cells are divided into two major cell types, the cumulus cells that are in direct contact with the oocyte and provide nutrients through gap junctions and the steroidogenic mural cells that form an outer layer and produce estradiol (Gilchrist et al., 2008). By contrast, teleost follicle cells form only a single-cell layer and are the likely source of both estradiol production and select nutrients required for oocyte development (Devlin and Nagahama, 2002). It can be presumed that zebrafish follicle cell maturation is coordinated with oocyte developmental progression, but little is known about stage-specific transcriptional changes that may occur in these cells. Another major difference between mammals and teleosts is that in mammals new oocytes are only formed during embryogenesis, while in many teleosts, including zebrafish, new oocytes are formed continuously from a population of GSCs. Thus, the need to produce new follicle cells extends throughout the reproductive life of the fish. However, pre-follicle cells have not been molecularly identified in any teleost (Beer and Draper, 2013; Cao et al., 2019b).

We extracted and reclustered the cells identified by gsdf expression to further characterize follicle cells (Figure 4). Because we used 40 dpf ovaries as our source of cells, the follicle cells in our dataset likely represent pre-follicle cells and those associated with stage IB and early stage II oocytes. Consistent with this, cyp19a1a, a gene known to be expressed in follicle cells surrounding mid-stage II and older oocytes, is only detected in a very small number of follicle cells (Figure 4B, Figure 4—figure supplement 1; Dronow et al., 2016; Rodriguez-Mari et al., 2005). We found that the majority of genes detected in follicle cells are expressed uniformly throughout the cluster (e.g., gsdf, fox12a/b, and amh; Figure 4A and B, Figure 4—figure supplement 1). However, our recluster analysis distinguished five cell subclusters based on a subset of genes that had nonuniform expression (Figure 4A and Supplementary file 3). Upon further analysis, we concluded that these subclusters likely represent three developmentally relevant cell populations (e.g., subcluster 4 represents a mitotic subpopulation of otherwise developmentally similar cells in subclusters 0 and 2; Figure 4A).

Stage IB follicle cells (subcluster 1)
Cells within subcluster 1 have enriched expression of the heparan sulfate proteoglycan-encoding gene, glypican 1a (gpc1a; Gupta and Brand, 2013), RERG/RAS-like b (rerglb), and Htra serine protease 1b (htra1b; Figure 4A–C). Using gpc1a as a representative gene of subcluster 1, we performed HCR RNA-FISH and determined that gpc1a+ cells are found only surrounding stage IB oocytes (Figure 4D; Selman et al., 1993), and thus represent the primary follicle cells that first enclose oocytes to form the functional follicle. We did not observe significant expression of genes required for cell division in subcluster 1, such as proliferating cell nuclear antigen (pcna) and dihydrofolate reductase (dhhf), suggesting that stage IB follicle cells are not highly proliferative (Figure 4B and C).

Stage II follicle cells (subclusters 0, 2, and 4)
Subcluster 0 contains most of the follicle cells in our dataset, which are characterized by the enriched expression of several genes, including gap junctional protein alpha 11 (gja11; formerly cx34.5; Mikalsen et al., 2020; Figure 4A–C). To determine the location of this cell population within the ovary, we performed HCR RNA-FISH for gja11. These results show that follicle cells expressing gja11 are restricted to stage II follicles (Figure 4E). Gene expression analysis for subcluster 2 cells suggests that these cells are a transitional stage from stage IB to stage II follicle cell as they express several genes, such as sidekick cell adhesion molecule 1a (sdk1a; Galicia et al., 2018), which are expressed in both populations (Figure 4B). Gene expression analysis argues that cells within subcluster 4 are a mitotic subpopulation of subcluster 0 cells as these cells have nearly identical gene expression with cells in subcluster 0 but also express high levels of genes necessary for mitosis, such as pcna (Figure 4B and C). Follicle cell proliferation is not surprising given that the surface area of the zebrafish oocyte increases 1400-fold during oogenesis. In mammals, there is evidence that granulosa cell proliferation is regulated by pituitary-produced follicle-stimulating hormone (Fsh; Goldenberg et al., 1972), and consistent with this, we found the gene encoding the Fsh receptor, called fshr, is expressed in zebrafish follicle cells, including those in subcluster 4 (Laan et al., 2002; Figure 4—figure supplement 1). Finally, we detect only a small percentage of cells expressing cyp19a1a in subcluster 0, arguing that
cells within this subcluster are predominantly composed of follicle cells from early stage II oocytes, which do not express cyp19a1a (Dranow et al., 2016; Figure 4B, Figure 4—figure supplement 1).

Pre-follicle cells (subcluster 3)
Subcluster 3 cells are part of the follicle cell lineage because they express gsdf (Figure 4A–C), yet these cells also express many genes that are unique to this subpopulation, arguing that they are distinct from the stage IIB and II follicle cells identified above. Interestingly, these cells express orthologs of the core genes associated with early undifferentiated somatic gonad cells in mammals, including the transcription factors LIM homeobox 9 (lhx9; Birk et al., 2000), Wilms tumor 1a (wt1a; Kreidberg et al., 1993), empty spiracles homeobox 2 (emx2; Miyamoto et al., 1997), and nuclear receptor 5a1b (nr5a1b; formerly sfl1; Luo et al., 1994; Figure 4B and C, Figure 4—figure supplement 2C and D). Notably, subcluster 3 cells also uniquely express iroquois homeobox 3a and 5a (irx3a and irx5a; Dildrop and Rüther, 2004) whose orthologs in mammals are required for pre-granulosa cell development in the embryonic mouse ovary (Figure 4B and C, Figure 4—figure supplement 2C and D; Fu et al., 2018; Kim et al., 2011). Subcluster 3 cells do not express gata4, a gene required for the earliest stage of somatic gonad cell specification in mice and that is also expressed in early somatic gonad cells during the bipotential phase in zebrafish (10 dpf; Hu et al., 2013; Leerberg et al., 2017). This argues that these cells are not developmentally similar to the somatic gonad precursors present in the bipotential gonad. Finally, subcluster 3 cells express lower levels of genes associated with follicle cell differentiation and function, such as foxl2a/b, amh, fshr, and notch3, than do the stage IIB and stage II follicle cells (Figure 4B and C, Figure 4—figure supplement 1). Together, these results argue strongly that subcluster 3 cells are pre-follicle cells.

We noticed that many of the genes with enriched expression in subcluster 3 did not appear uniformly expressed in all cells within this subcluster. For example, lhx9 appeared to have higher expression in cells distal to the subcluster 1 in the UMAP while gsdf and bone morphogenetic protein 6 (bmp6; Smith et al., 2006) appeared higher in cells proximal to subcluster 1 (Figure 4A and C). To explore this further, we performed additional recluster analysis using only those cells derived from follicle cell subclusters 1 and 3 (Figure 4—figure supplement 2A and B and Supplementary file 4). This analysis confirmed that subcluster 3 cells can be partitioned into two distinct subclusters that, for simplicity, we have designated subclusters 3.1 and 3.2 (Figure 4—figure supplement 2B). While many genes, such as emx2, irx3a, and irx5a, appeared to be expressed uniformly in subclusters 3.1 and 3.2, our analysis confirmed that lhx9 and wt1a have enriched expression in subcluster 3.1, while gsdf and bmp6 appear enriched in subcluster 3.2 (Figure 4—figure supplement 2C and D). Another intriguing difference between these two cell populations is that subcluster 3.2 cells have enriched expression of genes that encode cell signaling components, such as bmp6, fibroblast growth factor receptors 3 and 4 (fgfr3 and fgfr4; Sleptsova-Friedrich et al., 2001; Thisse et al., 1995), and the Fgf-responsive genes ETS variant transcription factor 4 (etv4; Brown et al., 1998) and sprouty homolog 4 (spry4; Fürthauer et al., 2001), while subcluster 3.1 cells have enriched expression of genes that encode cell signaling attenuators, such as the Bmp ligand antagonists noggin1 and 3 (nog1 and nog3; Fürthauer et al., 1999), the Wnt ligand antagonist dkk1b (dkk1b; Hashimoto et al., 2000), and the retinoic acid-degrading enzyme encoded by cytochrome P450 26a1 (cyp26a1; White et al., 1996). In mammals, retinoic acid signaling promotes germ cells to enter meiosis (Bowles et al., 2006; Koubova et al., 2006). Therefore, it is possible that inhibition of Bmp, Wnt, and retinoic acid signaling is necessary to keep these cells in an undifferentiated, stem cell-like state, or to influence other surrounding cells.

To determine where these cells reside in the ovary, we performed HCR RNA-FISH using lhx9 and bmp6 as representative genes for clusters 3.1 and 3.2, respectively. Consistent with our hypothesis, we found that lhx9+ cells do not localize around oocytes, but instead formed cord-like structures that are located on the surface of the ovary (arrows in Figure 4F). Triple HCR RNA-FISH for lhx9, GSC-expressed nanos2, and early meiosis-expressed dmc1 revealed that clusters of early-stage germ cells (premeiotic and stage 1A pre-follicle phase germ cells) were always located adjacent to lhx9+ pre-follicle cells (n = 31 clusters of nanos2 or dmc1-expressing cells; N = 4; Figure 4F, F’ and F”). This analysis also revealed that lhx9+ cells appear to surround these early-stage germ cells (Figure 4F’ and F”). Interestingly, there appear to be two populations of lhx9+ cells, based on expression level and location. The first population appears to express high levels of lhx9 and localize to the chord-like structures, which contact one side of the early-stage germ cell clusters (arrows in Figure 4F, F’ and
The second population appears to express lower levels of *lhx9* and envelop the early-stage germ cell clusters on sides that do not contact the cords (arrowheads in Figure 4F and F’). It is possible that the high-expressing cells correspond to the 3.1 subcluster cells while the low-expressing cells correspond to the 3.2 subcluster cells. In the medaka ovary, sox9b-expressing cells have been identified as pre-follicle cells and similarly formed cord-like structures on the ovarian surface that associate with early-stage germ cells, including GSCs (Nakamura et al., 2008; Nakamura et al., 2010). Indeed, we found the zebrafish sox9b ortholog is expressed in a subset of *lhx9* expression cells, indicating that these cells are likely functional homologs of the medaka *sox9b*-expressing cells (Figure 4—figure supplement 2). Finally, we determined the location of *bmp6* expressing cells together with *os2* and *dmc1*. *bmp6*-expressing cells also appear to colocalize to regions containing early-stage germ cells, but unlike *lhx9*, the *bmp6* staining was more diffuse, though concentrated towards the lateral edge of the ovary (Figure 4G and G’).

In mammals, retinoic acid signaling is required to induce germ cell entry into meiosis (Bowles et al., 2006; Koubouva et al., 2006). Previous studies have argued that retinoic acid is likely produced in the zebrafish ovary by both follicle cells and interstitial cells as these cells express *aldh1a2/neckless*, which encodes the enzyme that converts retinaldehyde to retinoic acid (Figure 4—figure supplement 3A; Rodríguez-Marí et al., 2013). Given that the retinoic acid-degrading enzyme Cyp26a1 appears to be produced by the *lhx9*-expressing pre-follicle cells (Figure 4—figure supplement 2C and D, Figure 4—figure supplement 2B), and the retinoic acid receptor-encoding genes *rxrba* and *rxrb* are expressed in premeiotic germ cells (Figure 4—figure supplement 3C), we speculate that the close association of premeiotic germ cells with these cells may be important for regulating GSC maintenance.

**Follicle cell-expressed wnt9b is required for female sexual development**

In addition to their role in support and regulation of oocyte development during oogenesis, follicle cells also play key roles in promoting female sex determination and differentiation in vertebrates. Our lab previously showed that the follicle cell-expressed gene encoding the Wnt4 ligand was required for female sex differentiation, similar to the role of WNT4 in mammals (Kossack et al., 2019; Vainio et al., 1999). However, while 95% of wnt4 mutants develop as males, we found that ~5% can develop as females. We reasoned that the partial penetrance of the phenotype could be due to the function of an additional Wnt ligand(s). To further test the usefulness of this scRNA-seq dataset for functional gene discovery, we asked if the expression of other Wnt ligands was detected in the dataset, and if so, in what cell type(s)? Our analysis confirmed that wnt4 is expressed in follicle cells in zebrafish ovary (Figure 5A; Kossack et al., 2019). Additionally, there is significant expression of wnt8a (Kelly et al., 1995) in oocytes, wnt9a in stromal cells, wnt9b (Crotwell and Mabee, 2007) in follicle cells, and wnt11 (Makita et al., 1998) in stromal and theca cells (Figure 5—figure supplement 1). We focused our attention on wnt9b because its apparent expression in follicle cells was strikingly similar to that of wnt4 (Figure 5A). To determine if wnt9b was involved in sex determination and/or differentiation, we produced wnt9b mutants using CRISPR/Cas9. The wnt9b(fb207) allele contains a 57 bp deletion of exon1, which includes the splice donor sequence, and is therefore predicted to be a loss-of-function mutation (Figure 5B).

Domesticated zebrafish do not have sex chromosomes, instead sex is determined by a combination of genetic and environmental factors (Kossack and Draper, 2019). Because sex ratios in any cross can vary, it is important to compare sex ratios in mutants to their wild-type siblings. We crossed wnt9b(fb207) heterozygous parents and determined the genotype of their offspring at 50 dpf. We then determined the phenotype by dissecting their gonads to determine whether they had ovaries (females) or testes (males; Figure 5C–G). We found that wild-type animals developed similar numbers of males and females, while wnt9b mutants were majority male (Figure 5H). In addition, we found a subset of wnt9b mutants that had gonads that appeared to be transitioning from an ovary to a testis as they contained stage IB oocytes surrounded by germ cells undergoing early spermatogenesis (Figure 5G). Because oocytes are never observed in wild-type testes at 50 dpf (Figure 5D), the presence of oocyte in these wnt9b mutants suggests that these animals initiated female development before sex reversing to males. Consistent with this hypothesis, wild-type testes at 50 dpf contained germ cells at all stages of spermatogenesis, including early spermatocytes and mature spermatzoa.
Figure 5. Expression and mutational analysis of wnt9b. (A) Expression plots of wnt4a and wnt9b show that wnt9b is expressed only in follicle cells, in a pattern nearly identical to wnt4a. (B) Schematic of the wnt9b genomic locus. Boxes are exons, UTR sequences are shaded. Arrow indicates approximate location of the 57 bp deletion in the wnt9b(fb207) allele. (C–G) Representative regions of gonads stained for Ddx4 protein (green) to identify germ cells. wnt9b(+/+) ovaries (C) and testis (D). wnt9b(-/-) ovary (E) and testis (F). (G) wnt9b(-/-) gonad that contains mostly germ cells that have characteristics of spermatogenesis, but also has a few stage IB oocytes. (H) Sex ratios of offspring produced from wnt9b(fb207) heterozygous (+/-) parents (n = 384, N = 3, *p=0.016, **p=0.037). Trans., transitioning; IB, Stage IB oocyte; Sc, spermatocyte; Sz, spermatozoa. Scale bar in (C) for (C–G) 50 µm.

The online version of this article includes the following figure supplement(s) for figure 5:

Figure supplement 1. Gene expression uniform manifold approximation and projection (UMAP) plots for Wnt ligand-encoding genes in the 40-day post-fertilization (dpf) ovary.
fibroblast-like connective tissue (Subcluster 0 is the largest of the stromal cell population and likely represents the collagen-producing Interstitial cells (clusters 0 and 4)). Of the somatic ovarian cells, the stromal cell cluster is the most transcriptionally diverse. This is also supported by Gene Ontology (GO) analysis (Nicol and Guiguen, 2011). By contrast, Wnt9b in mammals is not expressed in granulosa cells and there is no evidence that it plays a role in sex determination. Instead, Wnt9b is involved in the formation of the male and female reproductive ducts, the Wolffian and Müllerian ducts, respectively (Carroll et al., 2005).

Identification of five stromal cell subpopulations

Ovarian stromal cells can be broadly defined as all ovarian cell types other than germ, follicle, or theca cells. Given this broad definition, it is no surprise that stromal cells are the least characterized cell types in the ovary. Known stromal cell types include interstitial cells, which are collagen-producing connective tissue (fibroblast) that provide structure integrity to the ovary, surface epithelial cells that surround the ovary, vasculature and associated perivascular cells, and resident blood cell types (Kinnear et al., 2020). To date, few of these cell types have been characterized in the zebrafish ovary. From our initial clustering, we identified a large population of cells that had enriched expression of several collagen encoding genes, such as col1a1a (Fisher et al., 2003; Figure 1A) and reasoned that these were likely the stromal cell population. This is further supported by the expression of decorin (dcn; Shintani et al., 2000), which encodes a proteoglycan component of the extracellular matrix, and nuclear receptor subfamily 2, group F, member 2 (nr2f2, formerly coup-tfl; Fjose et al., 1995), both of which are expressed in mammalian ovarian stromal cells (Figure 6A; Pereira et al., 1995; Wagner et al., 2020). To further characterize this cell population, we performed recluster analysis and identified five probable cell subpopulations, which are described below (Figure 6A and Supplementary file 5). Of the somatic ovarian cells, the stromal cell cluster is the most transcriptionally diverse. This is also supported by Gene Ontology (GO) analysis (Figure 6—figure supplement 1).

Interstitial cells (clusters 0 and 4)

Subcluster 0 is the largest of the stromal cell population and likely represents the collagen-producing fibroblast-like connective tissue (Figure 6B). In addition to the expression of the collagen-encoding genes, like col1a1a, notable genes expressed in these cells include the chemokine (C-X-C motif) ligand 12a (cxcl12a, formerly known as sdf1a; Crosier et al., 2001), the nuclear receptor nr2f2, the cell signaling molecule bmp4, and cyp19a1a (Figure 6A and B). Cells within subclusters 0 and 4 have nearly identical gene expression profiles, with the exception that subcluster 4 cells also have strong expression of genes associated with cell cycle progression, such as pcna (Figure 6A and B). Given that the ovary continues to increase in size as the fish grows, it is likely that subcluster 4 cells serve as a source for new interstitial cells similar to follicle cell subcluster 4 cells. Previous studies have shown that cxcl12a is essential for directing migrating PGCs to the forming gonad during embryogenesis (Doitsidou et al., 2002). However, we only detect expression of the cxcl12a receptor, cxcr4b, in neutrophiles and macrophage cells (Supplementary file 1) raising the possibility that Cxcl12/Cxcr4b signaling axis may instead be involved in recruiting these cells into the ovary (Petit et al., 2007). Interestingly, macrophage have been shown to play an important role in regulating spermatogenesis in the mouse testis (DeFalco et al., 2015). Regardless, cxcl21a provides a convenient marker for determining the location of this cell population in the ovary. Using FISH, we found that cxcl12a-expressing cells are enriched in the interstitial regions between early follicles, but not late follicles (<stage II; Figure 6C–C'). Together, these data argue that subclusters 4 and 0 are proliferative and nonproliferative interstitial cells, respectively.
Figure 6. Stromal cell subcluster analysis reveals five main cell subtypes. (A) Dot plot showing the relative expression of select genes in the stromal cell subclusters. Some genes, like col1a1a, are expressed in all stromal cells, while others, such as stm, are only expressed in a specific subcluster. Uniform manifold approximation and projection (UMAP) plots of genes in bold are shown in (B). Gene expression UMAP plots of select genes. Top-left panel shows cells color-coded by computationally determined cell subtype. Cells expressing the indicated gene are colored purple, and the relative intensity...
Figure 6 continued

indicates relative expression levels (intensity scale for each plot is on the right). (C–G) Hybridization chain reaction RNA fluorescent in situ hybridization (HCR RNA-FISH) on whole-mount 40-day post-fertilization (dpf) ovaries reveals the location of cell subtypes. In all panels, DNA is gray. (C) cxcl12a-expressing interstitial cells localize around early-stage oocytes (stage IB), but not around stage II oocytes. Ddx4 indirect immunofluorescence (green) labels all germ cells. (C') Higher magnification of the region boxed in (C). (D) plp1b-expression pericytes colocalize with fltl:egfp-expressing blood vessels (green). (D') Higher magnification of region boxed in (D). (E) acta3:egfp expression vascular smooth muscle cells (red) surround kdrl:dsRed expression blood vessels (green). (E') Higher magnification of the region boxed in (E). (F) smt (blue)-expressing cells localize to the lateral margin of the ovary and colocalize with nanos2 (green) and dmc1 (pink) germine stem cells and early meiotic cells, respectively. (F') Higher magnification of the region boxed in (F). Scale bar in (C–E) 20 µm; (C') 10 µm, (D', E', F') 5 µm. IB, stage IB oocyte; II, stage II oocyte.

The online version of this article includes the following figure supplement(s) for figure 6:

Figure supplement 1. Gene Ontology (GO) terms associated with the follicle and stromal cell subclusters.

Figure supplement 2. Stromal cell subclusters 1 and 2: pericytes and vascular smooth muscle cells.

Figure supplement 3. Stromal cell subcluster 3: stromal progenitor cells.

Figure supplement 4. Max projection view of a confocal stack showing smt expression (blue) and DNA (gray) in a 40-day post-fertilization (dpf) ovary.

Perivascular mural cells: Vascular smooth muscle (subcluster 1) and pericytes (subcluster 2)

Stromal cell subclusters 1 and 2 represent the perivascular mural cells that coat the endothelial-derived blood vessels and provide stability and integrity to the vasculature. Similar to their function in the establishment of the blood–brain barrier, these cells have also been implicated in the formation of a blood–follicle barrier (Siu and Cheng, 2012). Pericytes are generally solitary cells associated with small diameter blood vessels (arterioles, capillaries, and venules) while vascular smooth muscle cells (vSMCs) are associated with large blood vessels where they form a continuous coating (Gaengel et al., 2009). In zebrafish, pericytes express platelet-derived growth factor receptor b (pdgfrb) and notch3, two genes whose expression are enriched in our stromal cell subcluster 2 (Figure 6A, Figure 6—figure supplement 2; Wang et al., 2014). In addition, we find that subcluster 2 cells are enriched for proteolipid protein 1b (plp1b; Brösamle and Halpern, 2002; Figure 6A and B), T-box transcription factor 2b (tbx2b; Dheen et al., 1999), melanoma cell adhesion molecule B (mcamb; Chan et al., 2005), and regulator of G protein signaling 5a and 5b (rgs5a and rgs5b; Wang et al., 2014), all of which have orthologs expressed in human ovarian pericytes (Figure 6A, Figure 6—figure supplement 2A; Wagner et al., 2020). vSMC are characterized by expression of α-smooth muscle actin (acta2) and transgelin (tagln, formally known as SM22a; Bahrami and Childs, 2018). We found that both acta2 and tagln are enriched in stromal cell subcluster 1 (Figure 6A and B, Figure 6—figure supplement 2B and Supplementary file 5). In addition, we found that subcluster 1 has enriched expression of myosin light chain kinase b (mylk;b; Tournoij et al., 2010), tropomyosin 2 (tpm2; Davidson et al., 2013), and cysteine and glycine-rich protein 1b (csrp1b), further supporting the conclusion that these cells are vascular smooth muscle cells (Figure 6—figure supplement 2B).

To determine the location of pericytes in the ovary, we used plp1b as a marker because, in contrast to notch3 and pdgfrb, our data indicate that plp1b is specific to pericytes. We performed HCR RNA-FISH for plp1b using ovaries isolated from Tg(fli1:egfp) that express eGFP in vascular endothelial cells (Lawson and Weinstein, 2002). We found that plp1b+ cells were generally solitary, but always adjacent to blood vessels, consistent with pericytes (Figure 6D–D'). To determine the localization of vSMC, we imaged ovaries from Tg(kdrl:dsRed); Tg(acta2:egfp) double transgenic animals. Tg(kdrl:dsRed) is expressed in vascular endothelial cells, while acta2:egfp is expressed in vSMC (Kikuchi et al., 2011; Whitesell et al., 2014). We found that the majority of the acta2+ vSMC were either adjacent to or wrapped around the kdrl+ vascular endothelial cells (Figure 6E–E'). Together, our data argue strongly that subclusters 2 and 1 represent pericytes and vSMC, respectively.

Stromal progenitor cells (subcluster 3)

In addition to expressing high levels of the general stromal cell genes col1a1a and dcn (Figure 6A), subcluster 3 cells also express many genes that are known to play roles in early mesoderm and stromal cell development (Figure 6A and B, Figure 6—figure supplement 3B and C and Supplementary file 5). For example, the transcription factors odd-skipped related 1 (osr1; Tena et al., 2007) and transcription factor 21 (tcf21; Knight et al., 2008) are expressed specifically in stromal cell subcluster...
3 cells. In mice, Osr1 is one of the earliest genes known to be expressed in lateral plate mesoderm, derivatives of which form the kidney and gonads, and Osr1 mutants fail to form either of these organs (So and Danielian, 1999; Wang et al., 2005). Tcf21-expressing cells in the mouse ovary and testis were identified through lineage labeling as multipotent mesenchymal/stromal progenitor cells that can produce all somatic cell types during development and could contribute to somatic cell turnover during aging or cell replacement following injury in the testis (Cui et al., 2004; Shen et al., 2021). Additionally, subcluster 3 cells and mouse Tcf21-expressing mesenchymal progenitors also share common expression of osr1, procollagen C-endopeptidase enhancer a (pcolcea), slit homolog 3 (slit3), platelet-derived growth factor receptor a (pdgfra), dcn, col1a1a, col1a2, and insulin-like growth factor 2 (igf2a; Figure 6—figure supplement 3B and C; Shen et al., 2021). We therefore propose that these cells are stromal progenitor cells. As we have yet to identify where these cells reside in the ovary, we do not know if these cells are a stable population that persist into adulthood or if they are a transient population present only during the juvenile and early adult stages while the ovary is increasing in size.

Ovarian cavity epithelium (subcluster 5)

Stromal cell subcluster 5 has enriched expression of fgf24, a gene that is expressed in the outer epithelial layer of the early bipotential gonad (10–12 dpf) and is required for early gonad development in zebrafish (Leerberg et al., 2017). Subcluster 5 cells also express the cell adhesion protein encoded by claudin 3d (cldn3d; formerly called cldn3e; Kollmar et al., 2001 #486) and starmaker (stm), a gene required for otolith formation in the ear (Figure 6A and B; Söllner et al., 2003). Given the expression of fgf24 in surface epithelial cells of the early gonad (Leerberg et al., 2017) and the role of claudin proteins in forming tight junctions between epithelial cells (Tsukita et al., 2019), it was plausible that subcluster 5 cells corresponded to the ovarian surface epithelium. To test this, we performed HCR RNA-FISH using a probe for stm. We chose stm because it is not expressed in any other cell type in the ovary and would therefore allow us to specifically identify this population of cells. We found that stm was specifically expressed in cells that localize to the medial and lateral edges of the 40 dpf ovary, a region that also contained nanos2+ GSCs and dmc1+ meiotic germ cells (Figure 6F–F’, Figure 6—figure supplement 4). The colocalization of stm-expressing somatic cells with early-stage germ cells, including GSC’s, raises the interesting possibility that the lateral and medial edges of the 40 dpf ovary are the precursors to the germinal zones in the adult ovary.

To explore the relationship between the lateral and medial edges of the juvenile ovary and the germinal zone in the adult ovary, we performed HCR RNA-FISH for stm on transverse sections from 3-month-old adult ovaries. We found that stm-expressing cells localize to the membrane that forms the ovarian cavity, a fluid-filled space that forms on the dorsal side of the ovary and into which mature eggs are ovulated (Figure 7; Takahashi, 1977). We therefore designate this the ovarian cavity epithelium (OCE). The OCE is contiguous with the reproductive duct located at the posterior end of the ovary and that functions as the conduit between the OCE and genital papilla (Kossack et al., 2019). Interestingly, the OCE appears to adhere to the medial and lateral edges of the ovary at a location that correlates with that of the germinal zone (Figure 7B, C and H). To investigate this further, we prepared transverse histological sections from 3-month-old adult ovaries. We found that premeiotic and early meiotic germ cells are adjacent to where the OCE adheres to the ovary at both the lateral and medial sides (Figure 7D–G). Finally, we found that in addition to stm and fgf24, subcluster 5 cells also express the fgf24-related ligand encoded by fgf18b, transcription factors paired box 2b (pax2b; Pfeffer et al., 1998) and empty spiracles homeobox 2 (emx2; Morita et al., 1995), and wnt7aa (Norton et al., 2005; Figure 7A). Studies in medaka have shown that the transcription factor Pax2 is a probable direct regulator of the stm ortholog, starmaker-like (Bajoghli et al., 2009), raising the possibility that Pax2a may also regulate the expression of stm in OCE cells. Orthologs of pax2, emx2, and wnt7aa are expressed in the Müllerian ducts in mammals (Roly et al., 2020), raising the intriguing possibility that the OCE and mammalian Müllerian ducts share a common evolutionary origin. Altogether, these data argue that stm-expressing cells form the OCE and may also play a role in GSC niche formation and/or function.
Figure 7. Stromal cell subcluster 3: ovarian cavity epithelium (OCE). (A) Gene expression uniform manifold approximation and projection (UMAP) plots for select genes whose expression is enriched in stromal cell subcluster 3. (B) Hybridization chain reaction RNA fluorescent in situ hybridization (HCR RNA-FISH) on transverse sections from a 3-month-old ovary showing that stm (magenta) is expressed in the epithelium that lines the ovarian cavity (OC). DNA is gray. (C) Higher magnification of region boxed in (B) showing that early-stage germ cells localize to the region subjacent to where the OCE
Theca cells and production of the 17ß-estradiol precursor, androstenedione

In addition to producing mature gametes, the other major role of the ovary is to produce the sex steroid hormone 17ß-estradiol (E2), which induces female-specific sex characteristics throughout the body. E2 is the major female sex hormone in vertebrates, including zebrafish (Devlin and Nagahama, 2002). In mammals, the production of E2 requires two cell types: theca and granulosa cells (Ryan, 1979). The primary role of theca cells is to produce the intermediates for estrogen production, using cholesterol as the starting substrate and ending with the precursor androstenedione. Androstenedione is then transferred to granulosa cells for conversion to E2 by two additional enzymes, most notably the Cyp19a1 aromatase (Miller and Auchus, 2011). Although many studies have investigated steroid hormone production in zebrafish, there are still significant gaps in our understanding of the biosynthetic pathway that led to E2 production. For example, while enzyme orthologs necessary for E2 production in mammals have been identified in zebrafish, the specific cell type(s) that produce E2 in zebrafish has not been determined, nor have theca cells been definitively identified. To explore
this further, we first extracted and reclustered the cells identified in our preliminary cluster analysis as theca cells based on their expression of cyp11a2 (Figures 1 and 8).

The first and rate-limiting step in E2 synthesis in mammals is the transfer of cholesterol from the outer to inner mitochondrial membrane by steroidogenic acute regulatory protein, StAR, where it is converted to pregnenolone by the cytochrome P450 cholesterol side-chain-cleaving enzyme, Cyp11a1 (Miller and Auchs, 2011). The expression of both star and cyp11a1 has been detected in zebrafish ovaries (Bauer et al., 2000; Ings and Van Der Kraak, 2006). Unlike other vertebrates, teleost have two paralogs of cyp11a1, called cyp11a1 and cyp11a2, with cyp11a2 sharing the highest degree of similarity with other Cyp11a1 orthologs (Parajes et al., 2013). Our dataset demonstrates that cyp11a1 and cyp11a2 are both expressed in a distinct population of somatic cells, thus identifying these as theca cells (Figure 1, Figure 8—figure supplement 1). Additionally, we found that cyp11a2 is expressed at higher levels than cyp11a1 and therefore may encode the primary cholesterol side chain-cleaving enzyme required for E2 synthesis in the ovary (Figure 8—figure supplement 1). By contrast, the majority of cyp11a1 expression cells in the dataset are stage IA oocytes (Figure 8—figure supplement 1), consistent with previous results showing that in zebrafish cyp11a1 is maternally expressed (Parajes et al., 2013). Finally, expression of the three remaining genes encoding enzyme orthologs required for androstenedione production by theca cells, star, cyp17a1 (Wang and Ge, 2004) and hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1 (hsd3b1; Lin et al., 2015), are also detected in theca cells (Figure 8—figure supplement 1 and Supplementary file 1). As theca cells had only been previously localized by their morphology, we performed HCR RNA-FISH using cyp17a1 to determine the location of theca cells relative to oocytes. As expected, we found that cyp17a1-expressing cells localize to the interstitial spaces between developing oocytes (Figure 8B and C). Thus, we have molecularly characterized the theca cell population in the zebrafish ovary and confirmed that, as in mammals, theca cells are the probable source of the E2 precursor, androstenedione.

Further analysis of gene expression following reclustering of only the theca cell population suggests that theca cells are not a homogeneous cell population (Figure 8A and Supplementary file 6). We found several genes that had nonuniform expression within theca cells in zebrafish, leading to the identification of four theca cell subpopulations (Figure 8A and Supplementary file 6). Examples of genes expressed nonuniformly in theca cells include star (subcluster 1), cyp19a1a and nidogen 1b (nid1b), which encodes a basement membrane-associated protein (subcluster 3), v-ets avian erythroblastosis virus E26 oncogene homolog 2 (ets2; subclusters 0 and 1), and glutamate ammonia ligase A (glula), which encodes an enzyme involved in glutamine synthesis (subcluster 0; Figure 8A). Nonuniformity of theca cells has been previously noted in medaka based on the expression of select genes (Nakamura et al., 2009), providing support that some of these subpopulations have biological relevance that warrants further study. However, nonuniform expression of star was unexpected as StAR is thought to catalyze the rate-limiting step in steroid synthesis – the transfer of cholesterol from the outer to the inner mitochondrial membrane (Miller and Auchs, 2011). This may indicate that the theca cells in our dataset represent at least two developmental states: star- immature and star+ mature theca cells. Alternatively, StAR may not be required for androstenedione production by theca cells and thus E2 production. Two lines of evidence support this hypothesis. First, cyp19a1a mutant zebrafish all develop as males, indicating that E2 production is necessary for female development (Dranow et al., 2016). Second, and in contrast to cyp19a1a mutants, star mutant zebrafish have a normal sex ratio but females are infertile due to defects in oocyte maturation (Shang et al., 2019). These results argue that StAR is not required for the synthesis of all sex steroids in zebrafish, including E2.

Identification of the probable pathway for estrogen production in zebrafish ovary

There are two possible biosynthetic pathways for the conversion of androstenedione to E2, which we will refer to as pathways 1 and 2 (Figure 9A). Both pathways require two enzymatic steps, and the function of the Cyp19a1a aromatase. However, the pathways differ in the intermediate produced and at what step Cyp19a1a functions. Specifically, pathway 1 produces testosterone as an intermediate while pathway 2 produces estrone (E1; Figure 9A). Pathway choice is therefore not dictated solely by the expression of Cyp19a1a, but instead by the presence or absence of the two additional enzymes encoded by 17ß-hydroxysteroid dehydrogenase types 1 and 3, called Hsd17b1 and Hsd17b3. While
Figure 9. Pathway for 17β-estradiol (E2) synthesis in the zebrafish ovary. (A) Two possible pathways for E2 synthesis starting with androstenedione. Colors correspond to the region of the molecules being modified and to the enzyme that catalyzes the modification. (B) Gene expression uniform manifold approximation and projection (UMAP) plots of select genes. Top-left panel shows cells color-coded by computationally determined cell subtype. Cells expressing the indicated gene are colored purple, and the relative intensity indicates relative expression levels (intensity scale for each plot is on the right). (C) Proposed pathway for E2 synthesis in the zebrafish ovary, starting with the 17α-hydroxyprogesterone intermediate precursor, together with the cell types where each reaction occurs.

Figure 9 continued on next page
these enzymes catalyze similar reactions, in zebrafish Hsd17b3 prefers androstenedione as a substrate while Hsd17b1 prefers estrone (Mindnich et al., 2004; Mindnich et al., 2005; Figure 9A, Figure 9—figure supplement 1). By contrast, current evidence argues that the aromatase Cyp19a1a does not have preference for androstenedione over testosterone (Guiguen et al., 2010; Figure 9A). It therefore follows that the expression of Hsd17b1 vs. Hsd17b3 will determine pathway preference. Though it has been proposed that testosterone is the normal intermediate for E2 production in many fish (Devlin and Nagahama, 2002), the pathway by which E2 is produced in the zebrafish ovary remains unknown. Thus, knowing which genes are expressed and in what cells is key to understanding how E2 production is regulated in the zebrafish ovary.

Our lab previously reported that cyp19a1a was likely expressed in theca cells in the early ovary and then was later upregulated in follicle cells surrounding oocytes that had progressed past mid-stage II, a stage that is underrepresented in 40 dpf ovaries (Dranow et al., 2016). Our scRNA-seq data verified that cyp19a1a is expressed in theca cells (Figure 9B) but not in early follicle cells (Figure 9B, see also Figure 4B and Figure 4—figure supplement 1). Unexpectedly we found significant expression of cyp19a1a in interstitial stromal cells. Thus, it is likely that three district ovarian cell populations contribute to E2 production in zebrafish: theca cells, follicle cells, and stromal cells (Figure 9B and C). Finally, we determined in what cell types hsd17b1 and hsd17b3 were expressed. We found that hsd17b1 is expressed strongly in follicle cells while hsd17b3 expression appears to be only weakly expressed in premeiotic germ cells (Figure 8—figure supplement 1). Three major conclusions can be drawn from this analysis: (1) the major pathway for E2 biosynthesis in the zebrafish ovary occurs via pathway 2, (2) Hsd17b1, not Cyp19a1a, catalyzes the final enzymatic step in E2 synthesis, and (3) follicle cells are the major source of E2 production in the zebrafish ovary (Figure 9).

Conclusions
We have presented here the transcriptome of the 40 dpf zebrafish ovary at single-cell resolution, analysis of which has allowed us to define the major cell types present in the juvenile ovary. Further recluster analysis, supported by molecular and mutational evidence, revealed an unexpected level of cell subtype heterogeneity within the somatic cells that likely play important roles in regulating germ cell development as well as the development of female sexual characteristics through the production of estrogen. Our data provides strong support that orthologs of genes involved in mammalian ovary development and function likely play conserved roles in zebrafish, further validating zebrafish as a relevant model for understanding vertebrate ovarian development and function. Similarly, the major cell types required for mammalian ovary development and function have clearly identifiable homologs in the zebrafish ovary. However, unlike mammals, zebrafish females can produce new follicles throughout adult life from a population of self-renewing GSCs and pre-follicle progenitor cells, thus allowing the study of these unique cell types in a tractable vertebrate research animal. Importantly, this reference dataset will facilitate the use of zebrafish to study female reproductive diseases and disorders of sexual development.

Materials and methods

| Key resources table |
|---------------------|
| Reagent type (species) or resource | Designation | Source or reference | Identifiers | Additional information |
| Strain, strain background (Danio rerio) | AB (wild-type) | University of Oregon, Eugene | ZIRC Cat ID: ZL1 | Originally obtained from the University of Oregon, Eugene |
| Genetic reagent (D. rerio) | Tg(piwili1:egfp)uc02 | PMID:20737503 |  |
| Genetic reagent (D. rerio) | Tg(fli1:egfp)y1^{cr} | PMID:12167406 | ZIRC Cat ID: ZL1085 |  |

Continued on next page
Reagent type (species) or resource | Designation | Source or reference | Identifiers | Additional information
--- | --- | --- | --- | ---
Genetic reagent (\textit{D. rerio}) | \textit{Tg}(kdrl:dsRed)\textsuperscript{pd27} | PMID:21397850 | ZIRC Cat ID: ZL4589 | 
Genetic reagent (\textit{D. rerio}) | \textit{Tg}(acta2:eGFP)\textsuperscript{ca7} | PMID:24594685 | ZIRC Cat ID: ZL9948 | 
Genetic reagent (\textit{D. rerio}) | fox\textit{l2}(uc91) | This paper | | CRISPR/Cas9-induced in-frame insertion of V2A-eGFP
Genetic reagent (\textit{D. rerio}) | wnt\textit{9b}(fs209) | This paper | | CRISPR/Cas9-induced deletion mutation
Gene (\textit{D. rerio}) | fox\textit{l2} | PMID:28193729 | ZDB-GENE-081022-71 | 
Gene (\textit{D. rerio}) | wnt\textit{9b} | PMID:17948314 | ZDB-GENE-080201-1 | 
Recombinant DNA reagent | pGTag-fox\textit{l2}-eGFP-\beta\textit{actin} (plasmid) | This paper | | Fox\textit{l2} version of pGTag-eGFP-\beta\textit{actin}
Antibody | Anti-Ddx4 (chicken polyclonal) | PMID:30653507 | IF (1:1500) | 
Antibody | Anti-Ddx4 (rabbit polyclonal) | PMID:10811882 | IF (1:1500) | 
Chemical compound, drug | Type II collagenase | Worthington | Cat# NC9870009 | 
Chemical compound, drug | Trypsin | Worthington | Cat# LS003708 | 
Chemical compound, drug | Trypsin inhibitor | MP Biomedicals | Cat# 100612 | 
Chemical compound, drug | Collagenase I | Sigma-Aldrich | Cat# C0130 | 
Chemical compound, drug | Collagenase II | Sigma-Aldrich | Cat# C6885 | 
Chemical compound, drug | Hyaluronidase | Sigma-Aldrich | Cat# H4272 | 
Chemical compound, drug | 10X TrypLE (trypsin) | Thermo Fisher | Cat# A1217701 | 
Commercial assay or kit | Chromium Single-cell 3’ Library and Gel Bead kit V3 | 10X Genomics | | 
Software, algorithm | Cell Ranger (v3.0.2) | 10X Genomics | | 
Software, algorithm | SoupX (v0.3.1) | PMID:33367645 | | 
Software, algorithm | DoubletFinder (v2.0.2) | PMID:30954475 | | 
Software, algorithm | Seurat (v3.1.0) | PMID:31178118 | | 
Software, algorithm | Monocle 3 (v0.1.3) | PMID:30787437 | | 
Software, algorithm | MEME (v5.1.0) | PMID:25953851 | | 

**Animals**

Zebrafish used for scRNAseq libraries are \textit{Tg}(\textit{piwil1:egfp})\textsuperscript{uc02} transgensics (formerly \textit{ziwi:egfp}; \textit{Leu and Draper, 2010}) in the AB strain. Dissected ovaries for whole-mount immunohistochemistry, RNA in situ hybridization, FISH, and HCR RNA FISH were using the AB strain. The following transgenic lines were also used: \textit{Tg}(fl\textit{i1:egfp})\textsuperscript{y1} (\textit{Lawson and Weinstein, 2002}); \textit{Tg}(kdrl:dsRed)\textsuperscript{pd27} (\textit{Kikuchi et al., 2011}); and \textit{Tg}(acta2:eGFP)\textsuperscript{ca7} (\textit{Whitesell et al., 2014}). All animal work was carried out with approval of the UC Davis IACUC.

**Zebrafish husbandry**

Zebrafish husbandry was performed essentially as previously described (\textit{Westerfield, 2000}), with the following specific details. All zebrafish used in this study were raised in a recirculating aquatic system (Aquaneering, Inc, San Diego, CA), with the following water parameters: water temperature, 28–29°C; pH, 7.0–7.2 (maintained with Sodium Bicarbonate); conductivity, 600–700 µS (maintained with Crystal Sea Marine Mix, Aquaneering, CA); ammonia, 0 ppm; nitrate, 40–60 ppm; water exchange, 10% total volume/day; water source, deionized water. At 5 dpf, zebrafish larvae are placed into 1 l aquaria containing 250 ml 4 parts per thousand (ppt) seawater (Crystal Sea), at no more than 30 fish/tank. At
8 dpf, an additional 250 ml 4 ppt seawater is added to each tank. From days 10 to 30, the tanks are maintained with a slow continuous flow of system water. From 5 to 10 dpf, larvae are fed rotifers ad libitum (Brachionus plicatilis, L-type; Reed Mariculture, Campbell, CA). From 10 to 12 dpf, each tank of larvae is fed both rotifers and one drop concentrated brine shrimp nauplii, twice daily (Aquaneering, Inc). From day 12 to 20 dpf, each tank of larvae is fed two drops concentrated brine shrimp nauplii, twice daily. From 20 to 30 dpf, each tank of larvae is fed three drops concentrated brine shrimp nauplii twice daily. At 30 dpf, juvenile fish are moved to 9 l tanks at 30 fish/tank with continuous system water flow. From 30 to 40 dpf, juvenile fish are fed 500 µl concentrated artemia nauplii and 100 mg Gemma Micro 300 Zebrafish Food (Skretting, France)/tank, twice daily. Starting at 40 dpf, fish are fed 100 mg Zebrafish Select Diet (Aquaneering, Inc)/tank, twice daily.

Zebrafish ovary dissociation for single-cell RNA sequencing

Germ cell library cell dissociation

Germ cells were dissociated from the somatic gonad using a modification of Blokhina et al., 2019. 40 pairs of 40 dpf ovaries were dissected from Tg(piwil1:egfp)uc02 transgenic fish (Leu and Draper, 2010) and stored in a LoBind tube (Cat# 0030108302; Eppendorf) containing 2 ml of L15 medium (Cat# L5520; Sigma-Aldrich). The tissue was minced with small scissors into <1 mm pieces. 200 µl of 20 mg/ml of type 2 collagenase in L15 (Cat# NC9870009; Worthington) were added and incubated on an orbital rotator at 28°C for 35 min. The cell suspension was then gently passed through a 23G needle five times to breakup large cell clumps. 200 µl of 7 mg/ml trypsin (Cat# LS003708; Worthington) in L15 were added and incubated on an orbital rotator for 10 min or until a minimal amount of cell clumps was observed. The trypsin reaction was stopped by adding 500 µl of 20 mg/ml trypsin inhibitor (Cat# 100612; MP Biomedicals) in L15. The supernatant carefully removed. The cell pellet was then resuspended and washed two times with 5 ml of L15 using a P1000 pipette tip and subsequently centrifuged for 3 min at 300 × g. Cells were resuspended in 1 ml L15 and filtered first through a 100 µm nylon filter (Cat# 431752; Corning) and then through a 40 µm nylon filter (Cat# 431750; Corning) to remove cell clumps. The filtrate was centrifuged for 3 min at 300 × g and resuspended in 1 ml of 50 mg/ml BSA (Cat# A8806; Sigma-Aldrich) in phosphate-buffered saline (PBS). Cell viability and number were determined using propidium iodine (Cat# P1304MP; Thermo Fisher) and Hoechst 33342 (Cat# H3570; Thermo Fisher) staining on a Fuchs-Rosenthal hemocytometer (Cat# DHC-F01; Incyto).

Whole ovary library cell dissociation

Somatic ovary cells were dissociated using a modification of Elkouby and Mullins, 2017. 40 pairs of 40 dpf ovaries were dissected from Tg(piwil1:egfp) transgenic fish (Leu and Draper, 2010) and stored in a LoBind tube containing 2 ml of L15. Tissues were minced with microdissection scissors into <1 mm pieces. After the tissue had settled to the bottom of the tube, the media were replaced with 5 ml of digestive enzyme mixture, 3 mg/ml collagenase I (Cat# C0130; Sigma-Aldrich), 3 mg/ml collagenase II (Cat# C6885; Sigma-Aldrich), and 1.6 mg/ml hyaluronidase (Cat# H4272; Sigma-Aldrich) in L15 and incubated on an orbital rotator at room temperature. The suspension was monitored every 10 min until no or a minimal number of cell clumps were observed (~30 min). Cells were centrifuged for 3 min at 300 × g and resuspended in 5 ml of 5X TrypLE (Cat.No. A1217701; Thermo Fisher) in L15 and incubated on an orbital rotator at room temperature for 15 min. The trypsin reaction was stopped by adding 500 µl of 2.8 mg/ml trypsin inhibitor in L15 and incubated on an orbital rotator at room temperature for 1 min. The cell suspension was then added to 25 ml of L15 in a 50 ml conical tube to dilute the trypsin and centrifuged for 3 min at 300 × g. The cell pellet was resuspended and washed two times with 5 ml of L15 using a P1000 pipette and centrifuged for 3 min at 300 × g. The cell pellet was then resuspended and filtered as described above.

Isolation of germ cells by FACS

Following cell dissociation of Tg(piwil1:egfp) transgenic ovaries, GFP+ germ cells were sorted using a MoFlo Astrios EQ Cell Sorter (Beckman Coulter) with a 70 µm nozzle. Cells were sorted using side scatter and GFP purify to identify single germ cells.
Single-cell RNA library preparation and sequencing
Single-cell RNA sequencing libraries were prepared by the UC Davis DNA Technologies core. Briefly, barcoded 3' single-cell libraries were prepared from dissociated cell suspensions or sorted cells using the Chromium Single-Cell 3' Library and Gel Bead kit V3 (10X Genomics) for sequencing according to the manufacturer’s recommendations. All libraries were targeted at 10,000 cell recovery and were amplified using 11 cycles. The cDNA and library fragment size distribution were verified via micro-capillary gel electrophoresis on a Bioanalyzer 2100 (Agilent). The libraries were quantified by fluorometry on a Qubit instrument (Life Technologies) and by qPCR with a Kapa Library Quant kit (Kapa Biosystems) prior to sequencing. Libraries were sequenced on a HiSeq 4,000 sequencer (Illumina) with paired-end 100 bp reads.

Cell Ranger genome reference and gene annotation file generation
A General Transfer Format (GTF) gene annotation file (release 96) for the GRCz11 zebrafish genome was downloaded from Ensembl Genome Browser and filtered using the ‘mkgtf’ function in Cell Ranger (v3.0.2; 10X Genomics) to retain the following attributes: protein_coding, lincRNA, and antisense. A genome reference file was generated with Cell Ranger’s ‘mkref’ function using the GRCz11 zebrafish genome obtained from the Ensembl Genome Browser with alternative loci scaffolds removed and the filtered GTF file.

Count file generation
GRCz11 zebrafish genome FASTQ file from Ensembl Genome Browser with alternative loci scaffolds removed and the filtered GTF file described above were used to generate the count file using the ‘count’ function in Cell Ranger (v3.0.2; 10X Genomics). ‘Expect-cells” was set to 10,000 based on estimated cell recovery.

Cell cluster analysis
Analysis scripts and data for the cell cluster analysis are available on GitHub (https://github.com/yulongliu68/zeb_ov_ssRNAseq, copy archived at swr:1:rev:3430147079ab3840afd7b25b01652f-caeda5f78d; Liu, 2022), and the final clustering can be explored online at the Single Cell Portal (The Broad Institute of MIT and Harvard, https://singlecell.broadinstitute.org/single_cell/study/SCP928).

Briefly, the expression matrices generated from Cell Ranger (germ cell library zx1_40gc, whole ovary libraries zx2_40ov and zx4_40ov) were first processed with SoupX (v0.3.1; Young and Behjati, 2020) to remove ambient RNA. During dissociation and library generation, lysed cells can lead to ambient RNA levels that interfere with subsequent analyses. Despite several washing steps post dissociation, the initial data exploration revealed the presence of oocyte RNA across a variety of cell types. SoupX is designed to identify and remove ambient RNA contaminations. We used non-oocyte cells and inferred top oocyte-specific genes to quantify the extent of the contamination.

The adjusted datasets were then processed with DoubletFinder (v2.0.2; McGinnis et al., 2019) to determine and remove doublets that are expected in any large-scale single-cell RNA-sequencing datasets. DoubletFinder uses cell expression proximity and artificially generated doublets to determine potential doublets. Prior to assessing doublets, we performed quality control on the dataset. For the sorted germ cell library, we retained cells with 200–6000 genes, less than 150,000 unique transcripts, and less than 5% mitochondrial transcripts. For whole ovary libraries, we retained the cells with 200–8000 genes, less than 200,000 unique transcripts, and less than 20% mitochondrial transcripts. These cutoffs were determined by the gene and transcript distributions of those libraries. Blood cells and a small number of germ cells were also removed from the whole ovary libraries. DoubletFinder was run using a conservative 5% estimated cell doublet cutoff (Figure 1—figure supplement 1).

Clustering analyses were conducted using Seurat (v3.1.0; Stuart et al., 2019) and the ‘SCTransform’ workflow to normalize, identify variable genes, and perform scaling of the data. Principal component analysis (PCA) was performed using the top 3000 variable genes, and principal components considered were chosen based on the standard deviation of the elbow plot, p-value of the jackstraw plot, and biological knowledge of the genes in individual components. UMAP analysis and plots are generated based on selected principal components. Gene expression tables for all cells, germ cell subset, pre-follicle cell subset, follicle cell subset, stromal cell subset, and theca cell subset are available for downloaded at Dryad (https://doi.org/10.25338/B8FH12).
Trajectory analysis

Monocle 3 (v0.1.3) was used for trajectory analysis (Cao et al., 2019a). The expression matrix was exported from the Seurat object and used as Monocle 3 input. UMAP was used for dimensionality reduction. We used GSC-specific nanos2 expression to identify the root of the trajectory and measured pseudo-time.

Gene module analysis and motif enrichment

Non-negative matrix factorization (NMF) as a dimensionality reduction strategy was used to identify groups of co-expressed genes as previously described (Brunet et al., 2004; Farrell et al., 2018; Siebert et al., 2019). Expression data for the top 3000 variable genes as identified in the Seurat analyses were used as input to the NMF analysis. To identify the optimal number of gene modules that can describe the dataset, we tested a broad range of K values from 10 to 100 in increments of 5 and then reduced to an increment of 1 in a range between 30 and 40. The optimal K was identified as 36 for the germ cell dataset. Identified gene modules were then filtered to remove low-quality modules if a reproducibility score was lower than 0.6 or the gene module consisted of less than 10 genes.

For motif enrichment within 5’ regions of co-expressed genes, we extracted 2 kb upstream of the transcription start site of the top 20% of the genes within a module based on gene scores. Gene score is a metric describing how well a particular gene reflects the expression of the associated gene module. All sequences were from Ensembl and extracted using biomaRt (v2.40.4; Cunningham et al., 2019). We used MEME (v5.1.0; Bailey et al., 2015; McLeay and Bailey, 2010) for motif enrichment analysis with the following parameter: "--scoring avg --method fisher --hit-lo-fraction 0.25 --value-report-threshold 10.0 --control--shuffle----kmer 2 sequence"; and the following databases: jolma2013.meme, JASPAR2018_CORE_vertebrates_non-redundant.meme, and uniprobe_mouse.meme.

GO analysis

The top 25% of the cluster genes (sorted by p-value) were used for GO analysis inputs. Go enrichment was assessed using g:Profiler2 (v0.2.0; Raudvere et al., 2019). The zebrafish biological process database was used, and enriched GO terms with the significance of <0.05 were considered.

RNA in situ hybridization

All ovary samples were fixed in situ with 4% paraformaldehyde overnight at 4°C, followed by dissection. Samples were dehydrated with 100% methanol for 10 min at room temperature and stored at −20°C at a minimum overnight with fresh 100% methanol. Each probe was independently assayed a minimum of two times (N ≥ 2) using five 40 dpf ovary lobes isolated from five females for each assay (n = 10).

Single-molecule whole-mount RNA in situ hybridization

HCR RNA-FISH probes were ordered from Molecular Instruments. Accession numbers supplied to the manufacturer for designing the probes and the lot numbers for reordering are listed in the following table. Hybridization was performed following the method detailed in Molecular Instruments protocol (MI-Protocol-HCRv3-Zebrafish version 5; Molecular Instruments) with the following modification: fixed ovaries stored in methanol were rehydrated to PBS in a graded series of MeOH:PBS (2:1, 1:1, 1:2), incubating 5 min in each. Ovaries were then washed 4 × 5 min in PBSTw (PBS + 1% Tween 20) before proceeding. Ovaries were then permeabilized with proteinase K at 50 µg/ml PBSTw for 15 min. Following in situ hybridization, samples were cleared in glycerol-PBS gradient (30, 50, and 70%) for 1 hr in each, and then mounted with ProLong Diamond Antifade Mountant (Cat# P36961, Invitrogen).
Probes used for HCR fluorescence RNA in situ hybridization

| Target gene | Organism                  | Lot number | Order number | Amplifier |
|-------------|---------------------------|------------|--------------|-----------|
| nanos2      | Zebrafish (Danio rerio)   | PRB905     | d6f5d8c-b888-4bf7-b9a1-d6d95db66d24 | B2        |
| chgb        | Zebrafish (Danio rerio)   | PRB908     | d6f5d8c-b888-4bf7-b9a1-d6d95db66d24 | B1        |
| id4         | Zebrafish (Danio rerio)   | PRB909     | d6f5d8c-b888-4bf7-b9a1-d6d95db66d24 | B1        |
| fox2l (zgc:194189) | Zebrafish (Danio rerio)   | PRD331     | 165f08fe-24a8-40e3-acb3-02767bf38a37 | B1        |
| lhx9        | Zebrafish (Danio rerio)   | PRD333     | 165f08fe-24a8-40e3-acb3-02767bf38a37 | B3        |
| dmc1        | Zebrafish (Danio rerio)   | PRD334     | 165f08fe-24a8-40e3-acb3-02767bf38a37 | B1        |
| rec8a       | Zebrafish (Danio rerio)   | PRE326     | 20502418-92db-49c1-801c-fc808aa3940f | B1        |
| bhlhe23     | Zebrafish (Danio rerio)   | PRE327     | 20502418-92db-49c1-801c-fc808aa3940f | B3        |
| gpc1a       | Zebrafish (Danio rerio)   | PRF573     | 984cd5dd-9cdc-4496-94eb-c722f6c027ae | B3        |
| plp1b       | Zebrafish (Danio rerio)   | PRF574     | 984cd5dd-9cdc-4496-94eb-c722f6c027ae | B3        |
| cyp17a1     | Zebrafish (Danio rerio)   | PRF576     | 984cd5dd-9cdc-4496-94eb-c722f6c027ae | B3        |
| stm         | Zebrafish (Danio rerio)   | PRF578     | 984cd5dd-9cdc-4496-94eb-c722f6c027ae | B3        |
| bmp6        | Zebrafish (Danio rerio)   | PRG855     | db0a04b4-ef3d-4944-9061-d90ec2d1c956 | B3        |

Standard whole-mount fluorescent RNA in situ hybridization

RNA in situ hybridization was performed essentially as described (Thisse and Thisse, 2008). Template DNA for production of a gja11 RNA in situ probe was generated by RT-PCR. mRNA was first isolated from 40 dpf wild-type AB strain ovaries using TRI reagent (Cat# T9424; Sigma-Aldrich) and then cDNAs were synthesized with RETROScript Reverse Transcription Kit (Cat# AM1710; Thermo Fisher). A gja11-specific template was then generated by PCR using Phusion polymerase (Cat# M0530L; New England Biolabs) and the following primer pair: fwd, 5′-CCCTGAGCATTTTTCCGAC-3′; rev, 5′-taattattacactatatga-GTCTTAAAGCCAGGCTCA-3′. Reverse primers contained a T7 RNA polymerase promoter shown in the lowercase letters. An RNA in situ probe for cxcl12a was generated as previously described (Knaut et al., 2005). For both probes, DNA templates were transcribed with T7 RNA polymerase (Cat# 10881775001; Roche) to yield digoxigenin-labeled antisense RNA probes. Probes were G-50 column purified (Cat# 45-001-398; GE Healthcare), and then diluted to 0.5–2 μg/ml in the hybridization solution with 5% dextran sulfate (Cat# S4030; Thermo Fisher). Fluorescence RNA in situ hybridization procedure was performed following Thisse and Thisse, 2008 with the following modifications: fixed ovaries were permeabilized with proteinase K at 50 μg/ml PBSTw for 15 min, and the in situ was developed with FastRed (Cat# F4648; Sigma-Aldrich).

RNA in situ hybridization on sections

Ovaries were isolated from 3-month-old females the afternoon after spawning. We reasoned that the ovarian cavity and its associated membrane would be more pronounced when the body cavity was devoid of mature oocytes. Whole fish were fixed overnight in neutral alcoholic formalin fixative (29.5% ethanol, 10% formalin, 50 mM sodium phosphate buffer, pH 7.2). Fish were then decalcified by incubating in 0.5 M EDTA, pH 8.0, as previously described (Moore et al., 2002). Fish were then processed and embedded in paraffin, and 7 μm sections were cut. RNA in situ hybridization with stm probes was performed as described in the manufacturer’s protocol (MI-Protocol-HCRv3-Zebrafish version 5; Molecular Instruments). Prior to imaging, sections were stained 1 hr with 1 μg/ml 4′,6-diamidino-2-phenylindole (DAPI) in PBS, cleared in glycerol-PBS gradient (30, 50, and 70%) for 10 min each, and then mounted with ProLong Diamond Antifade Mountant.

Histology

Ovaries were isolated from 3-month-old females the afternoon after spawning. Whole fish were fixed overnight in Bouin’s fixative (Cat# 112016; Fisher Scientific). Fish were then washed 4 × 4 hr in 70% ethanol before processing for paraffin embedding and 7 μm sections cut and stained with hematoxylin
and eosin (H&E) as previously described (Siegfried and Steinfeld, 2021). Imagers were acquired using a Zeiss Axiophot microscope equipped with a Leica DFC 500 camera.

### Immunofluorescence staining

Whole-mount immunofluorescence staining of ovaries was performed as previously published (Leerberg et al., 2017). Primary chicken anti-Ddx4 antibody (Blokhina et al., 2019) or rabbit anti-Ddx4 antibody (Knaut et al., 2000) were used at 1:1500. Secondary goat anti-chicken IgG Alexa Fluor 488 (Cat# A-11039; Thermo Fisher) or goat anti-rabbit IgG Alexa Fluor 488 (Cat# A-11008; Thermo Fisher) were used at 1:300.

### Imaging

Fluorescence RNA in situ hybridization, HCR fluorescence RNA in situ hybridization, and immunohistochemical staining samples were mounted with ProLong Diamond Antifade Mountant and imaged with either Olympus FV1000 laser scanning confocal microscope or Zeiss LSM 880 Airyscan microscope. Histological images were collected using a Zeiss Axiophot microscope equipped with a Leica DFC 500 camera. Images of whole zebrafish were collected using a Leica MZ 16F dissecting microscope equipped with a Leica DFC 500 camera.

### Germ cell quantification

Numbers of early GCs expressing nanos2, nanos2+ foxl2l, foxl2l, or foxl2l+ rec8a in individual clusters were counted in Z-stack images with 1 µm steps through the entire cluster. Individual cells must have more than 20 RNA molecules/puncta after HCR fluorescence RNA in situ to be counted as an expressing cell. 70 germ cell clusters were counted from three independent experiments with three different 40 dpf ovaries. The cell count graph was generated with ggplot2 (v3.3.2).

### Foxl2l phylogenetic analysis

Phylogenetic analysis of Foxl proteins was performed using Clustal Omega with the following protein sequences:

- Carp (Cyprinus carpio): Foxl2l XP_018932368.1; Foxl3 XP_018951720.1; chicken (Gallus gallus): FOXL1 XP_040537252.1, FOXL2 NP_001012630.1; coelacanth (Latimeria chalumnae): Foxl1 XP_005995948.1, Foxl2 XP_006001344.1, Foxl2l XP_005986027.1, Foxl3 XP_005997033.1; elephant shark (Callorhinchus milii): Foxl1 XP_007887429.1; Foxl2l XP_007900370.2; human (Homo sapiens): FOXL1 NP_005241.1, FOXL2 NP_075555.1; killifish (Nothobranchius furzeri): Foxl2l XP_015822944.1; medaka (Oryzias latipes): Foxl1 NP_001116391.1, Foxl2 NP_001098358.1, Foxl2l XP_004070713, Foxl3 XP_011486175.1; Mexican tetra (Astyanax mexicanus): Foxl1 XP_015461274.2, Foxl2l XP_007232357.2, Foxl2b XP_007241719.1, Foxl3 XP_007251255.2; mouse (Mus musculus): FOXL1 NP_032050.2; FOXL2 NP_036150.1, FOXL3 NP_001182057.1; spotted gar (Lepisosteus oculatus): Foxl1 XP_015223442.1, Foxl2 XP_006637658.1, Foxl2l XP_015192516.1, Foxl3 XP_006637427.1; sterlet sturgeon (Acipenser ruthenus): Foxl2l XP_033909703.2; Xenopus tropicalis: Foxl1 XP_012817795.1, Foxl2l XP_004917868.1; zebrafish (Danio rerio): Foxl1 NP_957278.1, Foxl2a, NP_001038717.1, Foxl2b, NP_001304690.1, Foxl2l NP_001122282, Foxl3 NP_001182055.1.

### Mutant production

Generation of foxl2l:egfp knock-in mutant

We used the GeneWeld homology-directed repair technique to generate an in-frame insertion of a viral 2A peptide-eGFP fusion cassette downstream of amino acid 222 (Wierson et al., 2020). A pGTag- foxl2l- eGFP-B-actin-targeted integration plasmid was designed using the pGTag vector series (Wierson et al., 2020). 48 bp foxl2l-specific homology arms were cloned into the pGTag-eGFP-B-actin vector. Upstream homology arm 5′-GCGGgGggAAACGCTTCTGACCTGACTCATGGCCGACGCTTGGGATC3′; downstream homology arm 5′-AAACGGAAGTCCTTACCACTCCCAGAGATGGGATAG3′; Targeted plasmid integration into the foxl2l locus was done using CRISPR/Cas9 using a https://crisprscan.org/-designed sgRNA: 5′-GGGAGATGGGAGCTTGGCCGC3′. Injected F0 fish were outcrossed to wild-type and a sample of F1 embryos was screened for germline transmission of precisely integrated plasmid by PCR using the following primer sets: 5′-integration primers fwd 5′-3′, rev 5′-3′; 3′-integration primers fwd 5′-3′, rev 5′-3′. Precise 5′ integration of the construct was confirmed for the foxl2l(uc91) allele using Sanger
sequencing, and expression of eGFP in germ cells was confirmed at 1.5–2 months post-fertilization. Three pairs of foxl2(uc91) heterozygous fish were incrossed and their offspring genotyped at 50 dpf as follows. The wild-type and knock-in alleles were identified with independent PCR assays using the following primers: Universal_fwd: 5′-GCACATCTCGACTAGGC-3′; wild-type_rev: 5′-CACCAGGTTTGCATTAGT-3′; ampicon = 217 bp; gfp_rev: 5′-CTCTGCTTGTCCGACATGATAG-3′; ampicon = 724 bp. Following genotyping, fish were phenotyped for sex using standard criteria (Kossack and Draper, 2019). Significance values were determined using one-way ANOVA.

Generation of wnt9b mutant
One-celled zebrafish embryos were injected with Cas9 mRNA (1.4 ng/embryo) and a gRNA (58 pg/embryo) (5′-CGTGAGAGAGAAGATGAGGG-3′; target underlined) directed against the wnt9b locus. The target sequence was identified by using CHOPCHOP (Montague et al., 2014). Target site DNA oligos were annealed and cloned into the Bsal site of pDR274 (Hwang et al., 2013). Vector linearized with Dral was used as template for Maxi Script T7 RNA transcription (Cat# AM1312; Thermo Fisher) to generate gRNAs that were purified on a RNeasy Plus column (Cat# 74134; QIAGEN). Mutagenesis was verified by sequencing individual injected embryo colony PCR products covering the targeted locus. Founder heterozygotes were isolated by outcrossing grown G0 fish to TuAB wild-types and genotyping F1 offspring. Primers for genomic PCR: forward, 5′-CCCCCTTTAGAAGTGGACCTG T-3′; reverse, 5′-CAATGAGGCTTATAGGCCAGT-3′. An allele consisting of a 57 bp deletion was identified, which removed the 3′ end of the first exon and 5′ end of the first intron, leading to missplicing and a frameshift leading to a premature stop codon. To determine sex ratios of wnt9b mutant fish, 50 dpf fish produced from wnt9b+/- incross were fixed in 4% paraformaldehyde-1× PBS overnight at 4°C and stored in methanol at –20°C at least overnight. Fish were rehydrated in a methanol:1× PBS gradient and washed three times in 1× PBS prior to dissection. Sex was determined by examining dissected gonads for ovary or testis characteristics. Tissue was collected from each fish during dissection for genotyping post-dissection. Genomic DNA from tissue collected during dissection was extracted and genotyped by PCR. The primers used for genotyping this allele are 5′-CCCTCTTAAAGGCTGCTGAAT-3′ (forward) and 5′-CAATGAGGCATTAGGCCAGT-3′ (reverse). Wild-type allele yields a 238 bp product, wnt9b(fb209) allele yields a 181 bp product. Significance values were determined using one-way ANOVA.

Acknowledgements
The authors wish to thank Jack Cazet for his help with NMF analysis; Members of the Juliano and Burgess labs, Blanche Capel and Jennifer McKey for thoughtful discussions; Florence Marlow and Blanche Capel for critical reading of the manuscript; Jeffery Essner for GeneWeld plasmids; UC Davis Genome Center Sequencing Core; MCB Light Microscope Facility. R01 HD-081551 and NSF/IOS-1456737 to BWD; R35 GM133689 to CEJ; IAD and CNK were supported by NIH grants DK126021 and GM10431 to IAD; MEF and SRW were supported in part by the NIH T32 predoctoral training program in Molecular and Cellular Biology (GM-007377); MEK was supported in part by the NIH T32 predoctoral training program in Environmental Health Science (ES-0070599); SRW was supported in part by the NSF graduate research fellowship program (2036201); YL was supported in part by the UC Davis Dissertation Year Fellowship.

Additional information

Funding

| Funder                      | Grant reference number | Author       |
|-----------------------------|------------------------|--------------|
| National Institutes of Health | R01 HD-081551          | Yulong Liu   |
|                            |                        | Matthew E McFaul |
|                            |                        | Lana N Christensen |
|                            |                        | Bruce W Draper |
Author contributions

Yulong Liu, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing; Michelle E Kossack, Formal analysis, Investigation, Methodology, Visualization, Writing – review and editing; Matthew E McFaul, Formal analysis, Investigation, Methodology, Writing – review and editing; Lana N Christensen, Investigation, Methodology; Stefan Siebert, Methodology, Software, Writing – review and editing; Sydney R Wyatt, Investigation, Visualization, Writing – review and editing; Caramai N Kamei, Formal analysis, Investigation, Writing – review and editing; Samuel Horst, Nayeli Arroyo, Investigation; Iain A Drummond, Celina E Juliano, Funding acquisition, Supervision, Writing – review and editing; Bruce W Draper, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review and editing.

Author ORCIDs

Yulong Liu [http://orcid.org/0000-0003-0775-9258]
Michelle E Kossack [http://orcid.org/0000-0002-3400-5104]
Matthew E McFaul [http://orcid.org/0000-0002-5847-2938]
Sydney R Wyatt [http://orcid.org/0000-0003-2245-7813]
Iain A Drummond [http://orcid.org/0000-0003-4222-0987]
Celina E Juliano [http://orcid.org/0000-0003-3734-1231]

Ethics

This study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All of the animals were handled according to approved institutional animal care and use committee (IACUC) protocols (#20200 and #20201) of the University of California, Davis.

Decision letter and Author response

Decision letter [https://doi.org/10.7554/eLife.76014.sa1]
Author response [https://doi.org/10.7554/eLife.76014.sa2]

Additional files

Supplementary files
- Supplementary file 1. Gene expression tables from the whole dataset.
- Supplementary file 2. Gene expression tables from reclustered germ cells.
- Supplementary file 3. Gene expression tables from the reclustered follicle cells.
• Supplementary file 4. Gene expression tables from the reclustered stage 1A follicle cells and pre-follicle cells.
• Supplementary file 5. Gene expression tables from the reclustered stromal cells.
• Supplementary file 6. Gene expression tables from the reclustered theca cells.
• Transparent reporting form

Data availability
The raw and processed data reported in this paper are archived at NCBI GEO https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE191137 and in an interactively browsable forms at the Broad Institute Single-Cell Portal (https://singlecell.broadinstitute.org/single_cell/study/SCP928/40dpf-ovary-all-cells). Analysis code is archived at github (https://github.com/yulongliu68/zeb_ov_ssRNAseq, copy archived at swh:1:rev:3430147079ab3840afdb725b01652fcaeda5f78d). Gene expression tables for the cell clusters identified and the analysis R objects are archived at Dryad: (https://doi.org/10.25338/B8FH12).

The following datasets were generated:

| Author(s) | Year | Dataset title | Dataset URL | Database and Identifier |
|-----------|------|---------------|-------------|-------------------------|
| Liu Y, Kossack ME, McFaul ME, Christensen L, Siebert S, Wyatt SR, Kamei C, Horst S, Arroyo N, Drummond I, Juliano CE, Draper BW | 2021 | Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary | https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE191137 | NCBI Gene Expression Omnibus, GSE191137 |
| Liu Y, Kossack ME, McFaul ME, Christensen L, Siebert S, Wyatt SR, Kamei C, Horst S, Arroyo N, Drummond I, Juliano CE, Draper BW | 2021 | Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary | https://singlecell.broadinstitute.org/single_cell/study/SCP928/40dpf-ovary-all-cells | Broad Institute Single Cell Portal, SCP928/40dpf-ovary-all-cells |
| Liu Y, Kossack ME, McFaul ME, Christensen L, Siebert S, Wyatt SR, Kamei C, Horst S, Arroyo N, Drummond I, Juliano CE, Draper BW | 2021 | Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary | https://doi.org/10.25338/B8FH12 | Dryad Digital Repository, 10.5061/dryad.25338/B8FH12 |

References
Bahrami N, Childs SJ. 2018. Pericyte Biology in Zebrafish. Advances in Experimental Medicine and Biology 1109:33–51. DOI: https://doi.org/10.1007/978-3-030-02601-1_4, PMID: 30523588
Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME Suite. Nucleic Acids Research 43:W39–W49. DOI: https://doi.org/10.1093/nar/gkv416, PMID: 25953851
Bajoghli B, Ramialison M, Aghaallaei N, Czerny T, Wittbrodt J. 2009. Identification of starmaker-like in medaka as a putative target gene of Pax2 in the otic vesicle. Developmental Dynamics 238:2860–2866. DOI: https://doi.org/10.1002/dvdy.22093, PMID: 19795520
Bauer MP, Bridgham JT, Langenau DM, Johnson AL, Goetz FW. 2000. Conservation of steroidogenic acute regulatory (StAR) protein structure and expression in vertebrates. Molecular and Cellular Endocrinology 168:119–125. DOI: https://doi.org/10.1016/s0303-7207(00)00316-6, PMID: 11064158
Beer RL, Draper BW. 2013. nanos3 maintains germline stem cells and expression of the conserved germline stem cell gene nanos2 in the zebrafish ovary. Developmental Biology 374:308–318. DOI: https://doi.org/10.1016/j.ydbio.2012.12.003, PMID: 23228893
Birk OS, Casiano DE, Wassif CA, Cogliati T, Zhao L, Zhao Y, Grinberg A, Huang S, Kreidberg JA, Parker KL, Porter FD, Westphal H. 2000. The LIM homeobox gene Lhx9 is essential for mouse gonad formation. Nature 403:909–913. DOI: https://doi.org/10.1038/35002622, PMID: 10706291
Blokhina YP, Nguyen AD, Draper BW, Burgess SM. 2019. The telomere bouquet is a hub where meiotic double-strand breaks, synapsis, and stable homolog juxtaposition are coordinated in the zebrafish, Danio rerio. PLOS Genetics 15:e1007730. DOI: https://doi.org/10.1371/journal.pgen.1007730, PMID: 30653507
Bowles J, Knight D, Smith C, Wilhelm D, Richman J, Mamiya S, Yashiro K, Chawengsaksophak K, Wilson MJ, Rossant J, Hamada H, Koopman P. 2006. Retinoid signaling determines germ cell fate in mice. Science (New York, N.Y.) 312:596–600. DOI: https://doi.org/10.1126/science.1125691, PMID: 16574820

Brösmilie C, Halpern ME. 2002. Characterization of myelination in the developing zebrafish. Glia 39:47–57. DOI: https://doi.org/10.1002/glia.10088, PMID: 12112375

Brown LA, Amores A, Schilling TF, Jowett T, Baert JL, de Launoit Y, Sharrocks AD. 1998. Molecular characterization of the zebrafish PEA3 ETS-domain transcription factor. Oncogene 17:93–104. DOI: https://doi.org/10.1038/sj.onc.1201911, PMID: 9671138

Brown LA, Rodaway AR, Schilling TF, Jowett T, Ingham PW, Patient RK, Sharrocks AD. 2000. Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Flt-1 in wild-type and mutant zebrafish embryos. Mechanisms of Development 90:237–252. DOI: https://doi.org/10.1016/s0925-4773(99)00256-7, PMID: 10640707

Brunet JP, Tamayo P, Golub TR, Mesirov JP. 2004. Metagenes and molecular pattern discovery using matrix factorization. PNAS 101:4164–4169. DOI: https://doi.org/10.1073/pnas.0308531101, PMID: 15016911

Cao Z, Mao X, Luo L. 2019a. Germline Stem Cells Drive Ovary Regeneration in Zebrafish. Cell Reports 26:1709–1717. DOI: https://doi.org/10.1016/j.celrep.2019.01.061

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J. 2019b. The single-cell transcriptomic landscape of mammalian organogenesis. Nature 566:496–502. DOI: https://doi.org/10.1038/s41586-019-0969-x

Carmona SJ, Teichmann SA, Ferreira L, Macaulay IC, Stubbington MJT, Cvejic A, Gfeller D. 2017. Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. Genome Research 27:451–461. DOI: https://doi.org/10.1101/gr.207704.116, PMID: 28087841

Crespo B, Lan-Chow-Wing O, Rocha A, Zanuy S, Gómez A. 2013. foxl2 and foxl3 are two ancient paralogs that beyond. Developmental Dynamics 232:232–244. DOI: https://doi.org/10.1002/dvdy.20220, PMID: 15580611

Crespo B, Lan-Chow-Wing O, Rocha A, Zanuy S, Gómez A. 2013. foxl2 and foxl3 are two ancient paralogs that remain fully functional in teleosts. General and Comparative Endocrinology 194:81–93. DOI: https://doi.org/10.1016/j.ygcen.2013.08.016, PMID: 24045113

Crespo D, Assis LHC, van de Kant HJG, de Waard S, Safian D, Lemos MS, Bogerd J, Schulz RW. 2019. Endocrine and local signaling interact to regulate spermatogenesis in zebrafish: follicle-stimulating hormone, retinoic acid and androgens. Development (Cambridge, England) 146:dev178665. DOI: https://doi.org/10.1242/dev.178665, PMID: 31597660

Crosier PS, Bardsley A, Horsfield JA, Krasowska AK, Lavalie ER, Collins-Racie LA, Postlethwait JH, Yan YL, McCoy JM, Crosier KE. 2001. In situ hybridization screen in zebrafish for the selection of genes encoding secreted proteins. Developmental Dynamics 222:637–644. DOI: https://doi.org/10.1002/dvdy.1218, PMID: 11748832

Crotwell PL, Mabee PM. 2007. Gene expression patterns underlying proximal-distal skeletal segmentation in late-stage zebrafish, Danio rerio. Developmental Dynamics 236:3111–3128. DOI: https://doi.org/10.1002/dvdy.21352, PMID: 17948314

Cui S, Ross A, Stallings N, Parker KL, Capel B, Quaggin SE. 2004. Disrupted gonadogenesis and male-to-female sex reversal in Pod1 knockout mice. Development (Cambridge, England) 131:4095–4105. DOI: https://doi.org/10.1242/dev.01266, PMID: 15289436

Cunningham F, Achuthan P, Akanni W, Allen J, Amores A, Armode MJ, Bennett R, Bhai J, Billis K, Boddu S, Cummins C, Davidson C, Dodiya KJ, Gall A, Girón CG, Gil L, Grego T, Haggerty L, Haskell E, Hourlier T, et al. 2019. Ensembl 2019. Genome Research 29:D745–D751. DOI: https://doi.org/10.1093/nar/gky1113, PMID: 30407521

Davidson AE, Siddiqui FM, Lopez MA, Lunt P, Carlson HA, Moore BE, Love S, Born DE, Roper H, Majumdar A, Jayadev S, Underhill HR, Smith CO, von der Hagen M, Hubner A, Jardine P, Merrison A, Curtis E, Ellwagen G, Jungbluth H, et al. 2013. Novel deletion of lysine 7 expands the clinical, histopathological and genetic spectrum of TPM2-related myopathies. Brain 136:508–521. DOI: https://doi.org/10.1093/brain/awss344, PMID: 23413262

De Keuckelaere E, Hulpiau P, Saeyes Y, Berx G, van Roy F. 2018. Nanos genes and their role in development and beyond. Cellular and Molecular Life Sciences 75:1929–1946. DOI: https://doi.org/10.1007/s00018-018-2766-3, PMID: 29397397

DeFalco T, Potter SJ, Williams AV, Waller B, Kan MJ, Capel B. 2015. Macrophages Contribute to the Spermatogonial Niche in the Adult Testis. Cell Reports 12:1107–1119. DOI: https://doi.org/10.1016/j.celrep.2015.07.015, PMID: 26257171

Devlin RH, Nagahama Y. 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture 208:191–364. DOI: https://doi.org/10.1016/S0044-8486(02)00057-1

Dheen T, Sleptsova-Friedrich I, Xu Y, Clark M, Lehrae H, Gong Z, Korzh V. 1999. Zebrafish tbx-5 functions during formation of midline structures. Development (Cambridge, England) 126:2703–2713. DOI: https://doi.org/10.1242/dev.126.12.2703, PMID: 10331981
Hashimoto H, Itoh M, Yamanaka Y, Yamashita S, Shimizu T, Solnica-Krezel L, Hibi M, Hirano T. 2000. Zebrafish Dkk1 functions in forebrain specification and axial mesendoderm formation. Developmental Biology 217:138–152. DOI: https://doi.org/10.1006/dbio.1999.9537, PMID: 10625541

Hsu RJ, Lin CY, Hsu HS, Zheng SK, Lin CC, Tsai HJ. 2010. Novel intronic microRNA represses zebrafish myf5 promoter activity through silencing dickkopf-3 gene. Nucleic Acids Research 38:4384–4393. DOI: https://doi.org/10.1093/nar/gkq146, PMID: 20236986

Hu YC, Okamura LM, Page DC. 2013. Gata4 is required for formation of the genital ridge in mice. PLOS Genetics 9:e1003629. DOI: https://doi.org/10.1371/journal.pgen.1003629, PMID: 23874227

Hwang WY, Fu Y, Reydon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh J-RJ, Joung JK. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature Biotechnology 31:227–229. DOI: https://doi.org/10.1038/nbt.2501, PMID: 23360946

Ings JS, Van Der Kraak GJ. 2006. Characterization of the mRNA expression of StAR and steroidogenic enzymes in zebrafish ovarian follicles. Molecular Reproduction and Development 73:943–954. DOI: https://doi.org/10.1002/mrd.20490, PMID: 16700073

Kelly GM, Greenstein P, Erezylilmaf DF, Moon RT. 1995. Zebrafish wnt8 and wnt8b share a common activity but are involved in distinct developmental pathways. Development (Cambridge, England) 121:1787–1799. DOI: https://doi.org/10.1242/dev.121.6.1787, PMID: 760994

Kidder GM, Vanderhyden BC. 2010. Bidirectional communication between oocytes and follicle cells: ensuring oocyte developmental competence. Canadian Journal of Physiology and Pharmacology 88:399–413. DOI: https://doi.org/10.1139/y10-009, PMID: 20555408

Kikuchi K, Holdway JE, Major RJ, Blum N, Dahn RD, Begemann G, Poss KD. 2011. Retinoic acid production by endocardiac and epicardiac is an injury response essential for zebrafish heart regeneration. Developmental Cell 20:397–404. DOI: https://doi.org/10.1016/j.devcel.2011.01.010, PMID: 21397850

Kikuchi M, Nishimura T, Ishishita S, Matsuda Y, Tanaka M. 2020. foxl3, a sexual switch in germ cells, initiates two independent molecular pathways for commitment to oogenesis in medaka. PNAS 117:12174–12181. DOI: https://doi.org/10.1073/pnas.1918556117, PMID: 32409601

Kim B, Kim Y, Cooke PS, Rüeth U, Jorgensen JS. 2011. The fused toes locus is essential for somatic-germ cell interactions that foster germ cell maturation in developing gonads in mice. Biology of Reproduction 84:1024–1032. DOI: https://doi.org/10.1093/biolreprod.110.088559, PMID: 21293032

Kinnear HM, Tomaszewski CE, Chang FL, Moravek MB, Xu M, Padmanabhan V, Shikanov A. 2020. The ovarian stroma as a new frontier. Reproduction (Cambridge, England) 160:R25–R39. DOI: https://doi.org/10.1530/REP-19-0501, PMID: 32716007

Knaut H, Pelegri F, Bohmann K, Schwarz H, Nüsslein-Volhard C. 2000. Zebrafish vasa RNA but not its protein is a component of the germ plasm and segregates asymmetrically before germline specification. The Journal of Cell Biology 149:875–888. DOI: https://doi.org/10.1083/jcb.149.4.875, PMID: 10811828

Knaut H, Blader P, Strähle U, Schier AF. 2005. Assembly of trigeminal sensory ganglia by chemokine signaling. Neuron 47:653–666. DOI: https://doi.org/10.1016/j.neuron.2005.07.014, PMID: 16129936

Knight RD, Mebus K, Roehl HH. 2008. Mandibular arch muscle identity is regulated by a conserved molecular process during vertebrate development. Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution 310:355–369. DOI: https://doi.org/10.1002/jsexb.21215, PMID: 18338789

Kollmar R, Nakamura SK, Kappler JA, Hudspeth AJ. 2001. Expression and phylogeny of claudins in vertebrate primordia. PNAS 98:10196–10201. DOI: https://doi.org/10.1073/pnas.171325898, PMID: 11517306

Köprunner M, Thisse C, Thisse B, Raz E. 2001. A zebrafish nanos-related gene is essential for the development of primordial germ cells. Genes & Development 15:2877–2885. DOI: https://doi.org/10.1101/gad.212401, PMID: 11691838

Kossack ME, Draper BW. 2019. Genetic regulation of sex determination and maintenance in zebrafish (Danio rerio). Current Topics in Developmental Biology 134:119–149. DOI: https://doi.org/10.1016/btb.2019.02.004, PMID: 30999793

Kossack ME, High SK, Hopton RE, Yan YL, Postlethwait JH, Draper BW. 2019. Female Sex Development and Reproductive Duct Formation Depend on Wnt4a in Zebrafish. Genetics. 219:211–233. DOI: https://doi.org/10.1534/genetics.118.301620, PMID: 30446521

Koubova J, Menke DB, Zhou Q, Capel B, Griswold MD, Page DC. 2006. Retinoic acid regulates sex-specific timing of meiotic initiation in mice. PNAS 103:2474–2479. DOI: https://doi.org/10.1073/pnas.0510813103, PMID: 16461896

Kreidberg JA, Sariola H, Loring JM, Maeda M, Pelletier J, Housman D, Jaenisch R. 1993. WT-1 is required for early kidney development. Cell 74:679–691. DOI: https://doi.org/10.1016/0026-0495(93)90515-r, PMID: 8395349

Kwok HF, So WK, Wang Y, Ge W. 2005. Zebrafish Gonadotropins and Their Receptors: I. Cloning and Characterization of Zebrafish Follicle-Stimulating Hormone and Luteinizing Hormone Receptors—Evidence for Their Distinct Functions in Follicle Development I. Biology of Reproduction 72:1370–1381. DOI: https://doi.org/10.1095/biolreprod.104.038190, PMID: 15728795

Laan M, Richmond H, He C, Campbell RK. 2002. Zebrafish as a Model for Vertebrate Reproduction: Characterization of the First Functional Zebrafish (Danio rerio) Gonadotropin Receptor. General and Comparative Endocrinology 125:349–364. DOI: https://doi.org/10.1016/s0016-6480(02)00422-7, PMID: 12167406

Lawson ND, Weinstein BM. 2002. In vivo imaging of embryonic vascular development using transgenic zebrafish. Developmental Biology 248:307–318. DOI: https://doi.org/10.1006/dbio.2002.0711, PMID: 12167406
Leerberg DM, Sano K, Draper BW. 2017. Fibroblast growth factor signaling is required for early somatic gonad development in zebrafish. PLOS Genetics 13:e1006993. DOI: https://doi.org/10.1371/journal.pgen.1006993, PMID: 28873404

Leu DH, Draper BW. 2010. The zwi1 promoter drives germline-specific gene expression in zebrafish. Developmental Dynamics 239:2714–2721. DOI: https://doi.org/10.1002/dvdy.22404, PMID: 20737503

Liang L, Soyal SM, Dean J. 1997. FiGalpah, a germ cell specific transcription factor involved in the coordinate expression of the zona pellucida genes. Development (Cambridge, England) 124:4939–4947. DOI: https://doi.org/10.1242/dev.124.4939, PMID: 9262457

Lieschke GJ, Oates AC, Crowhurst MO, Ward AC, Layton JE. 2001. Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. Blood 98:3087–3096. DOI: https://doi.org/10.1182/blood.v98.10.3087, PMID: 11698295

Lin J-C, Hu S, Ho P-H, Hsu H-J, Postlethwait JH, Chung B. 2015. Two Zebrafish hsd3b Genes Are Distinct in Function, Expression, and Evolution. Endocrinology 156:2854–2862. DOI: https://doi.org/10.1210/en.2014-1584, PMID: 25974401

Liu Y. 2022. zeb_ov_ssRNAseq. swh:1:rev:3430147079ab3840afdb725b01652fcaeda5f78d. GitHub. https://github.com/yulongliu68/zeb_ov_ssRNAseq

Luo X, Ikeda Y, Parker KL. 1994. A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. Cell 77:481–490. DOI: https://doi.org/10.1016/0092-8674(94)90211-9, PMID: 8187173

Makita R, Mizuno T, Koshida S, Kuroiwa A, Takeda H. 1998. Zebrafish wnt11: pattern and regulation of the expression by the yolk cell and No tail activity. Mechanisms of Development 71:165–176. DOI: https://doi.org/10.1016/s0925-4773(98)00013-6, PMID: 9507106

Marlow FL, Mullins MC. 2008. Bucky ball functions in Balbiani body assembly and animal-vegetal polarity in the oocyte and follicle cell layer in zebrafish. Developmental Biology 321:40–50. DOI: https://doi.org/10.1016/j.ydbio.2008.05.557, PMID: 18582455

McGinnis CS, Murrow LM, Gartner ZJ. 2019. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. BMC Bioinformatics 11:165. DOI: https://doi.org/10.1186/s12859-019-2887-2, PMID: 30954475

McLeay RC, Bailey TL. 2010. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. BMC Bioinformatics 11:165. DOI: https://doi.org/10.1186/1471-2105-11-165, PMID: 20356413

Meng X, Lindahl M, Hyvönen ME, Parvinen M, de Rooij DG, Hess MW, Raatikainen-Ahokas A, Sainio K, Rauvala H, Lakso M, Pichel JG, Westphal H, Saarma M, Sariola H. 2000. Regulation of cell fate decision of undifferentiated spermatogonia by GDNF. Science (New York, N.Y.) 287:1489–1493. DOI: https://doi.org/10.1126/science.287.5457.1489, PMID: 10688798

Mikalsen S-O, Tausen M, i Kongststovu S. 2020. Phylogeny of teleost connexins reveals highly inconsistent intra- and interspecies use of nomenclature and misassemblies in recent teleost chromosome assemblies. BMC Genomics 21:223. DOI: https://doi.org/10.1186/s12862-020-06620-2

Miller WL, Auchus RJ. 2011. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. Endocrine Reviews 32:81–151. DOI: https://doi.org/10.1210/er.2010-0013, PMID: 21051590

Mindnich R, Deluca D, Adamski J. 2004. Identification and characterization of 17 beta-hydroxysteroid dehydrogenases in the zebrafish, Danio rerio. Molecular and Cellular Endocrinology 215:19–30. DOI: https://doi.org/10.1016/j.mce.2003.11.010, PMID: 15026171

Mindnich R, Haller F, Halbach F, Moeller G, Hrabé de Angelis M, Adamski J. 2005. Androgen metabolism via 17beta-hydroxysteroid dehydrogenase type 3 in mammalian and non-mammalian vertebrates: comparison of the human and the zebrafish enzyme. Journal of Molecular Endocrinology 35:305–316. DOI: https://doi.org/10.1677/jme.0.01853, PMID: 16216911

Miyamoto N, Yoshida M, Kuratani S, Matsuo I, Aizawa S. 1994. Defects of urogenital development in mice lacking Emx2. Development (Cambridge, England) 121:1643–1664. DOI: https://doi.org/10.1242/dev.121.1643, PMID: 9165114

Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. 2014. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Research 42:W401–W407. DOI: https://doi.org/10.1093/nar/gku410, PMID: 24861617

Moore JL, Aros M, Steudel KG, Cheng KC. 2002. Fixation and decalcification of adult zebrafish for histological, immunocytochemical, and genotypic analysis. BioTechniques 32:296–298. DOI: https://doi.org/10.2144/02322203, PMID: 11848405

Morita T, Nitta H, Kiyama Y, Mori H, Mishina M. 1995. Differential expression of two zebrafish emx homeoprotein mRNAs in the developing brain. Neuroscience Letters 198:131–134. DOI: https://doi.org/10.1016/0304-3940(95)11988-9, PMID: 8592638

Nakamura S, Aoki Y, Saito D, Kuroki Y, Fujiyama A, Naruse K, Tanaka M. 2008. Sox9b/sox9a2-EGFP transgenic medaka reveals the morphological reorganization of the gonads and a common precursor of both the female and male supporting cells. Molecular Reproduction and Development 75:472–476. DOI: https://doi.org/10.1002/mrd.20764, PMID: 17474097

Nakamura S, Kurokawa H, Asakawa S, Shimizu N, Tanaka M. 2009. Two distinct types of theca cells in the medaka gonad: germ cell-dependent maintenance of cyp19a1-expressing theca cells. Developmental Dynamics 238:2652–2657. DOI: https://doi.org/10.1002/ddy.22068, PMID: 19705448
Nakamura S, Kobayashi K, Nishimura T, Higashijima S, Tanaka M. 2010. Identification of germine stem cells in the ovary of the teleost medaka. Science (New York, N.Y.) 328:1561–1563. DOI: https://doi.org/10.1126/science.1185473, PMID: 20488987

Nicol B, Guiguen Y. 2011. Expression profiling of Wnt signaling genes during gonadal differentiation and gametogenesis in rainbow trout. Sexual Development 5:318–329. DOI: https://doi.org/10.1002/dvdy.10392, PMID: 14579378

Nishimura T, Sato T, Yamamoto Y, Watakabe I, Ohkawa Y, Suyama M, Kobayashi S, Tanaka M. 2015. Sex determination: foxl3 is a germ cell-intrinsic factor involved in sperm-egg fate decision in medaka. Science (New York, N.Y.) 349:326–331. DOI: https://doi.org/10.1126/science.aaa2657, PMID: 26067255

Norton WHJ, Ledin J, Grandel H, Neumann CJ. 2005. HSPG synthesis by zebrafish Ext2 and Extl3 is required for Fgf10 signalling during limb development. Development (Cambridge, England) 132:4963–4973. DOI: https://doi.org/10.1242/dev.02084, PMID: 16221725

Oatley MJ, Kaucher AV, Racicot KE, Oatley JM. 2011. Inhibitor of DNA binding 4 is expressed selectively by spermatogonia in the male germ line and regulates the self-renewal of spermatogonial stem cells in mice. Biology of Reproduction 85:347–356. DOI: https://doi.org/10.1095/biolreprod.111.091330, PMID: 21543770

Onichtchouk D, Aduroja K, Beltng HG, Gunige L, Driever W. 2003. Transgene driving GFP expression from the promoter of the zona pellucida gene zpc is expressed in oocytes and provides an early marker for gonadal differentiation in zebrafish. Developmental Dynamics 228:393–404. DOI: https://doi.org/10.1002/dvdy.10392, PMID: 14579378

Parajes S, Griffin A, Taylor AE, Rose IT, Miguel-Escalada I, Hadzhevy Y, Airlt W, Shackleton C, Müller F, Krone N. 2013. Redefining the initiation and maintenance of zebrafish interrenal steroidogenesis by characterizing the key enzyme cyp11a2. Endocrinology 154:2702–2711. DOI: https://doi.org/10.1210/en.2013-1145, PMID: 23671259

Pepling ME, de Cuevas M, Spradling AC. 1999. Germline cysts: a conserved phase of germ cell development? Trends in Cell Biology 9:257–262. DOI: https://doi.org/10.1016/s0962-8924(99)01594-9, PMID: 10370240

Pereira FA, Qiu Y, Tasi MJ, Tsi SY. 1995. Chicken ovalbumin upstream promoter transcription factor (COUP-TF): expression during mouse embryogenesis. The Journal of Steroid Biochemistry and Molecular Biology 53:503–508. DOI: https://doi.org/10.1016/0960-0760(95)00097-J, PMID: 7626501

Petit I, Jin D, Rafii S. 2007. The SDF-1-CXCR4 signaling pathway: a molecular hub modulating neo-angiogenesis. Trends in Immunology 28:299–307. DOI: https://doi.org/10.1016/j.it.2007.05.007, PMID: 17560169

Pfeffer PL, Gerster T, Lun K, Brand M, Busslinger M. 1998. Characterization of three novel members of the zebrafish Pax2/5/8 family: dependency of Pax5 and Pax8 expression on the Pax2.1 (noi) function. Development (Cambridge, England) 125:3063–3074. DOI: https://doi.org/10.1242/dev.125.16.3063, PMID: 9671580

Prasasya RD, Mayo KE. 2018. Notch Signaling Regulates Differentiation and Steroidogenesis in Female Mouse Ovarian Granulosa Cells. Endocrinology 159:184–198. DOI: https://doi.org/10.1210/endo.2017-00677, PMID: 29126263

Qin M, Zhang Z, Song W, Wng QW-L, Chen W, Shigaonkar N, Ge W. 2018. Roles of Figla/figla in Juvenile Development and Follicle Formation During Zebrafish Gonadogenesis. Endocrinology 159:3699–3722. DOI: https://doi.org/10.1210/en.2018-00648, PMID: 30184072

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Research 47+W191–W198. DOI: https://doi.org/10.1093/nar/gkz639, PMID: 3106453

Rodriguez-Mari A, Yan Y-L, Bremiller RA, Wilson C, Cañestro C, Postlethwait JH. 2005. Characterization and expression pattern of zebrafish Anti-Müllerian hormone (Amh) relative to sox9a, sox9b, and cyp19a1a, during gonad development. Gene Expression Patterns 5:655–667. DOI: https://doi.org/10.1016/j.modgep.2005.02.008, PMID: 15939378.

Rodriguez-Mari A, Cañestro C, Bremiller RA, Nguyen-Johnson A, Asakawa K, Kawakami K, Postlethwait JH. 2010. Sex reversal in zebrafish fancl mutants is caused by Tp53-mediated germ cell apoptosis. PLOS Genetics 6:e1001034. DOI: https://doi.org/10.1371/journal.pgen.1001034, PMID: 20661450

Rodriguez-Mari A, Cañestro C, Bremiller RA, Catchen JM, Yan Y-L, Postlethwait JH. 2013. Retinoic acid metabolic genes, meiosis, and gonadal sex differentiation in zebrafish. PLOS ONE 8:e73951. DOI: https://doi.org/10.1371/journal.pone.0073951, PMID: 24040125

Roly ZY, Godini R, Estermann MA, Major AT, Pocock R, Smith CA. 2020. Transcriptional landscape of the embryonic chicken Müllerian duct. BMC Genomics 21:688. DOI: https://doi.org/10.1186/s12864-020-07106-8, PMID: 33008304

Ruzicka L, Howe DG, Ramachandran S, Toro S, Van Slyke CE, Bradford YM, Eagle A, Fashena D, Frazer K, Kalita P, Mani P, Martin R, Moxon ST, Paddock H, Pich C, Schaper K, Shao X, Singer A, Westerfield M. 2019. The Zebrafish Information Network: new support for non-coding genes, richer Gene Ontology annotations and the Alliance of Genome Resources. Nucleic Acids Research 47:D867–D873. DOI: https://doi.org/10.1093/nar/gky1090, PMID: 30407545

Ryan KJ. 1979. Granulosa-thecal cell interaction in ovarian steroidogenesis. Journal of Steroid Biochemistry 11:799–800. DOI: https://doi.org/10.1016/0022-4731(79)90014-1, PMID: 497143

Selman K, Wallace RA, Sarka A, Qi X. 1993. Stages of oocyte development in the zebrafish, Brachydanio rerio. Journal of Morphology 218:203–224. DOI: https://doi.org/10.1002/jmor.1052180209, PMID: 29865471

Shang G, Peng X, Ji C, Zhai G, Ruan Y, Lou Q, Jin X, He J, Wang H, Yin Z. 2019. Steroidogenic acute regulatory protein and luteinizing hormone are required for normal ovarian steroidogenesis and oocyte maturation in...
Thiese B, Thiese C, Weston JA. 1995. Novel FGF receptor (Z-FGFR4) is dynamically expressed in mesoderm and neurectoderm during early zebrafish embryogenesis. Developmental Dynamics 203:377–391. DOI: https://doi.org/10.1002/aja.1002030309, PMID: 8589434

Thiese C, Thiese B. 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. Nature Protocols 3:59–69. DOI: https://doi.org/10.1038/nprot.2007.514, PMID: 18193022

Tourtelotte WG, Nagarajan R, Auyeung A, Muller C, Milbrandt J. 1999. Infertility associated with incomplete spermatogenic arrest and oligozoospermia in Egr4-deficient mice. Development (Cambridge, England) 126:5061–5071. DOI: https://doi.org/10.1242/dev.126.22.5061, PMID: 10529423

Tsukita S, Tanaka H, Tamura A. 2019. The Claudins: From Tight Junctions to Biological Systems. Trends in Biochemical Sciences 44:141–152. DOI: https://doi.org/10.1016/j.tibsc.2018.09.008, PMID: 30665499

Vainio S, Heikilä M, Kispert A, Chinn N, McMahon AP. 1999. Female development in mammals is regulated by Wnt-4 signalling. Nature 397:405–409. DOI: https://doi.org/10.1038/17068, PMID: 9984040

Wagner M, Yoshihara M, Douagi I, Damdimopoulos A, Panula S, Petropoulos S, Lu H, Pettersson K, Palm K, Katayama S, Hovatta O, Kere J, Lanner F, Damdimopoulos P. 2020. Single-cell analysis of human ovarian cortex identifies distinct cell populations but no oogonial stem cells. Nature Communications 11:1147. DOI: https://doi.org/10.1038/s41467-020-14936-3, PMID: 32123174

Wang Y, Ge W. 2004. Cloning of zebrafish ovarian P450sc17 (CYP17, 17alpha-hydroxylase/17, 20-lyase) and characterization of its expression in gonadal and extra-gonadal tissues. General and Comparative Endocrinology 135:241–249. DOI: https://doi.org/10.1016/j.ygcen.2003.09.015, PMID: 14697311

Wang Q, Lan Y, Cho ES, Maltby KM, Jiang R. 2005. Odd-skipped related 1 (Odd 1) is an essential regulator of heart and urogenital development. Developmental Biology 288:582–594. DOI: https://doi.org/10.1016/j.ydbio.2005.09.024, PMID: 16223478

Wang Y, Pan L, Moens CB, Appel B. 2014. Notch3 establishes brain vascular integrity by regulating pericyte number. Development 141:307–317. DOI: https://doi.org/10.1242/dev.096107

Webster KA, Schach U, Ordaz A, Steinfeld JS, Draper BW, Siegfried KR. 2017. Dmrt1 is necessary for male sexual development in zebrafish. Developmental Biology 422:33–46. DOI: https://doi.org/10.1016/j.ydbio.2016.12.008, PMID: 27940159

Westefield, M. (2000). The zebrafish book. A guide to the laboratory use of zebrafish (Danio rerio). 4th edn (Univ. of Oregon Press, Eugene OR).

White JA, Guo YD, Baetz K, Beckett-Jones B, Bonasoro J, Hsu KE, Dilworth FJ, Jones G, Petkovich M. 1996. Identification of the retinoic acid-inducible all-trans-retinoic acid 4-hydroxylase. The Journal of Biological Chemistry 271:29922–29927. DOI: https://doi.org/10.1074/jbc.271.47.29922, PMID: 8939936

Whitesell TR, Kennedy RM, Carter AD, Rollins EL, Georgijevic S, Santoro MM, Chae SJ. 2014. An α-smooth muscle actin (acta2/asm) zebrafish transgenic line marking vascular mural cells and visceral smooth muscle cells. PLOS ONE 9:e90590. DOI: https://doi.org/10.1371/journal.pone.0090590, PMID: 24594685

Wiersen WA, Welker JM, Almeida MP, Mann CM, Webster DA, Torrie ME, Weiss TJ, Kambakam S, Vollbrecht MK, Lan M, Li X, McKeighan KC, Levey J, Ming Z, Wehmeier A, Mikelson CS, Haltom JA, Kwan KM, Chien C-B, Balcıunas D, Ekker SC, et al. 2020. Efficient targeted integration directed by short homology in zebrafish and mammalian cells. Cell 180:299–309. DOI: https://doi.org/10.1016/j.cell.2020.01.008, PMID: 31998681

Wixted JR, Kennedy RM, Carter AD, Rollins EL, Georgijevic S, Santoro MM, Chae SJ. 2014. An α-smooth muscle actin (acta2/asm) zebrafish transgenic line marking vascular mural cells and visceral smooth muscle cells. PLOS ONE 9:e90590. DOI: https://doi.org/10.1371/journal.pone.0090590, PMID: 24594685

Wixted JR, Kennedy RM, Carter AD, Rollins EL, Georgijevic S, Santoro MM, Chae SJ. 2014. An α-smooth muscle actin (acta2/asm) zebrafish transgenic line marking vascular mural cells and visceral smooth muscle cells. PLOS ONE 9:e90590. DOI: https://doi.org/10.1371/journal.pone.0090590, PMID: 24594685

Wixted JR, Kennedy RM, Carter AD, Rollins EL, Georgijevic S, Santoro MM, Chae SJ. 2014. An α-smooth muscle actin (acta2/asm) zebrafish transgenic line marking vascular mural cells and visceral smooth muscle cells. PLOS ONE 9:e90590. DOI: https://doi.org/10.1371/journal.pone.0090590, PMID: 24594685

Wixted JR, Kennedy RM, Carter AD, Rollins EL, Georgijevic S, Santoro MM, Chae SJ. 2014. An α-smooth muscle actin (acta2/asm) zebrafish transgenic line marking vascular mural cells and visceral smooth muscle cells. PLOS ONE 9:e90590. DOI: https://doi.org/10.1371/journal.pone.0090590, PMID: 24594685

Wixted JR, Kennedy RM, Carter AD, Rollins EL, Georgijevic S, Santoro MM, Chae SJ. 2014. An α-smooth muscle actin (acta2/asm) zebrafish transgenic line marking vascular mural cells and visceral smooth muscle cells. PLOS ONE 9:e90590. DOI: https://doi.org/10.1371/journal.pone.0090590, PMID: 24594685

Xie J, Wang WQ, Liu TX, Deng M, Ning G. 2008. Spatio-temporal expression of chromogranin A during zebrafish embryogenesis. The Journal of Endocrinology 198:451–458. DOI: https://doi.org/10.1677/JOE-08-0221, PMID: 18586978

Yang YJ, Wang Y, Li Z, Zhou L, Gui JF. 2017. Sequential, Divergent, and Cooperative Requirements of Foxl2a and Foxl2b in Ovary Development and Maintenance of Zebrafish. Genetics 205:1551–1572. DOI: https://doi.org/10.1534/genetics.116.199133, PMID: 28193729

Yoon C, Kawakami K, Hopkins N. 1997. Zebrafish vasa homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. Development (Cambridge, England) 124:3157–3165. DOI: https://doi.org/10.1242/dev.124.16.3157, PMID: 9272956

Yoshida K, Kondoh G, Matsuda Y, Habe T, Nishimune Y, Morita T. 1998. The mouse RecA-like gene Dmc1 is required for homologous chromosome synopsis during meiosis. Molecular Cell 1:707–718. DOI: https://doi.org/10.1016/s1097-2765(00)00707-2, PMID: 9660954

Young JM, McNeilly AS. 2010. Theca: the forgotten cell of the ovarian follicle. Reproduction (Cambridge, England) 140:489–504. DOI: https://doi.org/10.1530/REP-10-0094, PMID: 20628033

Young MD, Behjati S. 2020. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. GigaScience 9:giaa151. DOI: https://doi.org/10.1093/gigascience/giaa151, PMID: 33367645

Zakrzewska A, Cui C, Stockhammer OW, Benard EL, Spaink HP, Meijer AH. 2010. Macrophage-specific gene functions in Sp1-directed innate immunity. Blood 116:e1–e11. DOI: https://doi.org/10.1182/blood-2010-01-262873, PMID: 20424185