Single-Cell Transcriptome Reveals Insights into the Development and Function of the Zebrafish Ovary

By

YULONG LIU

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry, Molecular, Cellular and Developmental Biology

and a

Designated Emphasis in Biotechnology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Dr. Bruce W. Draper

Celina E. Juliano

Anna La Torre

Committee in Charge

2021

Contents

ACKNOWLEDGEMENT

I like to thank all my mentors, advisors, colleagues, friends, and families for their mentorship and support. Without them, it wouldn't be possible for me to complete my dissertation at UC Davis successfully. Specifically, I like to thank Dr. Bruce Draper for his continuous mentorship as my thesis advisor. He gave me the freedom to perform independent studies and explore new ideas, but he was also always available when I needed advice and took the time to check in on my progress and give feedback. I will never forget when he called me about my acceptance at UC Davis a few days before Christmas, it was the best Christmas present I have ever received. I like to thank both present and past members of the Draper Lab family at UC Davis (please do not be confused with the Draper Lab at Cambridge, MA, because they will threaten to sue our laboratory with their well-paid lawyers if there is any confusion due to the similarity in the name), Dr. Matt McFaul, Dr. Michelle Kossack, Sydney Wyatt, Dr. Dena Leerberg, Rachel (Hanz) Hopton, Lana Christensen, Samuel Horst, and Nayeli Arroyo for all their support and thoughtful discussions.

I like to thank my thesis committee members, Celina Juliano and Anna La Torre, for their valuable advice and support. I like to thank all my past and continuing collaborators, Dr. Stefan Siebert from the Juliano Lab (UC Davis), Dr.Zexia Gao from the Gao Lab (Huazhong Agricultural University), Jenn Cossaboon from the Teh Lab (UC Davis), Dr. Jingxun Chen from the Brunet Lab (Stanford) for either supported me on my project or let me be part of their exciting projects. I like to thank my former lab family at Arizona State University in Cheryl Nickerson Lab, especially

Dr. Cheryl Nickerson and Dr. Aurélie Crabbé, for their early mentorship that kickstarted my career in science.

I like to thank all my colleagues at UC Davis, specifically the most amazing BMCDB cohort, Shanaya Shah, Jack Cazet, Glyn Noguchi, Becka Jane, Adam Miltner, Bryan Teefy, Hannah Petrek, Jisoo Han, Lindsay Fague, Dean Natwick, and Marcus Moreno. I don't think I would have survived the first year without their friendship and support. A special thanks to Prema Karunanithi and Anastasia Berg. They were a year above me, and they have done above and beyond to help me navigate through the early years of graduate school and helped me on many fellowships and award applications. But, of course, above all, their friendship. I like to thank all the other amazing friends that I have met at UC Davis. They were always there to support me during the most challenging time during my graduate career, especially Robin Stewart, Luis Eduardo Contreras-Llano, Bianca Yaghoobi, Dr. Paige Mundy, and Dr. Melissa Bolotaolo. I also like to thank all my friends outside of UC Davis that I met back in Arizona and the bay area, especially Nick Jia and Zipeng Jia, for their dearest friendship, and they have always been there for me when I need help and support since high school.

Finally, and most importantly, I like to thank my family, especially my mother, Jessica Green. She came to the US many years ago by herself in search of a better future for me. She is the most wonderful mother in the world and the strongest person that I know. Also, thank all the extended family members in the US and China, especially stepfather Paul Green, aunt Junzi Guo, and uncle Shouming Zhang. I lived with aunt Junzi and Uncle Shouming for many years when my parents were away during my youth, and they have always treated me as their children.

ABSTRACT

Zebrafish are an established research organism that has been used to make 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 dpf). As such, forward genetic screens have identified only a few genes required for gonad development. In addition, bulk RNA sequencing studies used to 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 used single-cell RNA sequencing to determine the transcriptomes of cells isolated from juvenile zebrafish ovaries. In total, we profiled 10,658 germ cells and 14,431 somatic cells. In **Chapter** 1, I introduced germline stem cells in zebrafish and germline stem cell niches identified in other organisms. In **Chapter 2**, we used our single-cell data to identify the major cell types of the ovary. We then focused on the germ cell cluster, which contains the developmental stages of the germ cells from germline stem cells to early meiotic oocytes. From these data, we discovered a novel progenitor-stage germ cell marker in zebrafish, *foxl2l*, and revealed that it is required for female sex development. Furthermore, we identified multiple transcription factors that potentially have a role in germline stem cell regulation. As described in **Chapter 3**, we explored our somatic cell data, which represents all known somatic cell types, including follicle cells, theca cells, and interstitial stromal cells. Further analysis of these data revealed an unexpected number of somatic 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. Our results reveal novel insights into ovarian development and function and our Atlas will provide a valuable resource for future studies. In **Chapter 4**, I discuss future directions that should be pursued as a result of this dissertation.

Chapter 1: Introduction

Organisms evolve to meet their reproductive needs. In most mammals, such as humans, females are born with a finite number of post-meiotic matured eggs. In contrast, in many teleost fish (bony fish), such as the zebrafish, females have the ability to continuously produce eggs throughout their lifetime. This is evidenced by their profound fecundity and the continuous presence of oocytes in all stages (McMillan, 2007). This difference in reproductive ability is due to the female zebrafish possessing germline stem cells (GSCs)(Draper et al., 2007a). However, female GSCs are not well studied in vertebrates, and little is known about the factors that regulate their function to continuously produce germ cells.

The unipotent GSC can both divide to replenish their numbers and to produce differentiating cells, specifically, the sex-specific gametes eggs and sperm. In organisms where they have been identified, GSCs localize to a special microenvironment, the GSC niche, that is crucial for maintaining them in an undifferentiated state. Similar to other stem cell niches, specialized somatic cells produce short-range signals that regulate GSC proliferation and differentiation. Furthermore, physical anchoring of the GSCs via adherent proteins is also a reoccurring characteristic of the niche (Song et al., 2002).

GSCs and their niches have been well studied in a limited number of organisms, such as *Drosophila*, *C. elegans*, and mice. For example, in *Drosophila,* cap cells in females and hub cells in males maintain GSCs by expressing Dpp/Gbb(BMP) ligands to activate BMP signaling in GSCs (Chen & McKearin, 2003; Song et al., 2004a). Activation of BMP signaling in GSCs represses the expression of differentiation genes, such as *bam*. In the hermaphrodite *C. elegans*, distal tip

cells maintain GSCs by expressing LAG-2(Notch ligand) to activate GLP-1 (Notch receptor) signaling in GSCs. Activation of GLP-1 in the GSCs activates FBF genes that repress the expression of differentiation-promoting genes, such as Gld (Kimble & Crittenden, 2007). In male mice, specialized Sertoli cells maintain GSCs by expressing GDNF to activate Ret and GFRa signaling, which promote self-renewal (Lee et al., 2007; J. M. Oatley & Brinster, 2012). These findings suggest that the interactions between GSCs and their niche is crucial for maintenance of GSCs. Although the overall physical characteristics of GSC niches are similar, their signaling pathways can be unique (T. Xie, 2008).

In teleost fish, the identity of genes that regulate GSCs and their niche are largely unknown. A discrete band of the earliest GCs wrap around the adult ovary, which contains a heterogeneous population of both mitotic GCs and early meiotic GCs. The presence of mitotic cells is evidenced by the presence of Phosphorylated histone H3-positive germ cells or by Bromodeoxyuridine incorporation. The presence of early meiotic cells is evidenced by the expression of the early meiotic genes, *DNA meiotic recombinase 1* (*dmc1*), *synaptonemal complex protein 3* (*sycp3*), and the marker of double-strand breaks, *phosphorylated histone H2AX* (*γH2AX*). Thus, this discrete band, called the "germinal zone," contains both replicating (mitotic) cells and differentiating (early meiotic) GCs, and this is characteristic of a previously studied GSC niches(Beer & Draper, 2013b). In the distantly related medaka, mitotic and early meiotic cells are localized on the surface of the medaka ovary, however, unlike zebrafish, these cells do not localize to a discrete zone but are instead widely dispersed (Nakamura, 2010).

In zebrafish, *nanos2* is specifically expressed in a subset of mitotic cells in the germinal zone, which was hypothesized as zebrafish GSCs, as well as in male GSCs (Beer & Draper,

2013a). Nanos proteins, which are conserved zinc-finger RNA-binding proteins that function as post-transcriptional repressors, play essential roles in GSC maintenance in many organisms such as in *Drosophila*, mice, *C.elegans*, pig, and zebrafish (Beer & Draper, 2013a; Draper et al., 2007a; Forbes & Lehmann, 1998a; Park et al., 2017; Sada et al., 2009; Subramaniam & Seydoux, 1999). There are three Nanos homologs (Nos 1-3) in vertebrates and one in *Drosophila* (Nos). Pumilio (PUM), another RNA binding protein, binds in a sequence-specific manner to the Nanos response elements (NREs) located in the 3' UTR of its targeted RNA. Nanos binds to both the RNA and to PUM to regulate PUM the downstream effect of PUM binding on target mRNAs (Weidmann et al., 2016). Nanos recruits the CCR4-NOT deadenylation complex to the Nos-PUM-RNA complex to remove the poly-A tail of the targeted RNA transcript and, consequently, promotes RNA degradation. In female *Drosophila*, Nanos is required for the maintenance of GSCs by preventing meiosis, and Nanos overexpression results in an overabundance of GSCs (Z. Wang & Lin, 2004). Similarly, in male mice, Nanos2 prevents meiosis entry by inhibiting *stra8* expression, a meiosis promoting gene (Sada et al., 2009; Suzuki et al., 2008). Intriguingly, ectopic expression of Nanos2 in female mice embryonic germ cell also inhibits meiotic progression (Suzuki et al., 2008). Although the function of human NANOS2 has not been determined, NANOS2 mutation are associated with male infertility (Kusz et al., 2009). Both in female *Drosophila* and male mice, Nanos and Nanos2 expression, respectively, are GSC-specific in the adult gonads.

A comprehensive study of the GSC and its niche in the female zebrafish ovary will be a valuable resource for understanding how GSCs are regulated and maintained. In the following

chapters, I will explore my studies utilize single-cell RNA sequencing to identify early germ cell regulation and GSC niche cell candidates, and our other discoveries.

References

Beer, R. L., & Draper, B. W. (2013). Nanos3 maintains germline stem cells and expression of the conserved germline stem cell gene nanos2 in the zebrafish ovary. Developmental Biology, 374(2), 308–318. https://doi.org/10.1016/j.ydbio.2012.12.003

Chen, D., & McKearin, D. (2003). Dpp signaling silences bam transcription directly to establish asymmetric divisions of germline stem cells. Current Biology, 13(20), 1786–1791.

Draper, B. W., McCallum, C. M., & Moens, C. B. (2007). Nanos1 Is Required To Maintain Oocyte Production in Adult Zebrafish. Developmental Biology, 305(2), 589–598. https://doi.org/10.1016/j.ydbio.2007.03.007

Forbes, A., & Lehmann, R. (1998). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. Development, 125(4), 679–690.

Kimble, J., & Crittenden, S. L. (2007). Controls of germline stem cells, entry into meiosis, and the sperm/oocyte decision in Caenorhabditis elegans. Annu. Rev. Cell Dev. Biol., 23, 405–433.

Kusz, K. M., Tomczyk, L., Sajek, M., Spik, A., Latos-Bielenska, A., Jedrzejczak, P., Pawelczyk, L., & Jaruzelska, J. (2009). The highly conserved NANOS2 protein: Testis-specific expression and significance for the human male reproduction. Molecular Human Reproduction. https://doi.org/10.1093/molehr/gap003

Lee, J., Kanatsu-Shinohara, M., Inoue, K., Ogonuki, N., Miki, H., Toyokuni, S., Kimura, T., Nakano, T., Ogura, A., & Shinohara, T. (2007). Akt mediates self-renewal division of mouse spermatogonial stem cells. Development (Cambridge, England), 134(10), 1853–1859. https://doi.org/10.1242/dev.003004

McMillan, D. B. (2007). Fish histology: Female reproductive systems. Springer Science & Business Media.

Nakamura, S. (2010). Identification of Germline Stem Cells. 1561(2004). https://doi.org/10.1126/science.1185473

Oatley, J. M., & Brinster, R. L. (2012). The germline stem cell niche unit in mammalian testes. Physiological Reviews, 92(2), 577–595. https://doi.org/10.1152/physrev.00025.2011

Park, K.-E., Kaucher, A. V., Powell, A., Waqas, M. S., Sandmaier, S. E. S., Oatley, M. J., Park, C.-H., Tibary, A., Donovan, D. M., Blomberg, L. A., Lillico, S. G., Whitelaw, C. B. A., Mileham, A., Telugu, B. P., & Oatley, J. M. (2017). Generation of germline ablated male pigs by CRISPR/Cas9 editing of the NANOS2 gene. Scientific Reports, 7(December 2016), 40176. https://doi.org/10.1038/srep40176

Sada, A., Suzuki, A., Suzuki, H., & Saga, Y. (2009). The RNA-binding protein NANOS2 is required to maintain murine spermatogonial stem cells. Science, 325(5946), 1394–1398.

Song, X., Wong, M. D., Kawase, E., Xi, R., Ding, B. C., McCarthy, J. J., & Xie, T. (2004). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. Development, 131(6), 1353–1364.

Song, X., Zhu, C.-H., Doan, C., & Xie, T. (2002). Germline stem cells anchored by adherens junctions in the Drosophila ovary niches. Science, 296(5574), 1855–1857.

Subramaniam, K., & Seydoux, G. (1999). Nos-1 and nos-2, two genes related to Drosophila nanos, regulate primordial germ cell development and survival in Caenorhabditis elegans. Development (Cambridge, England), 126(21), 4861–4871.

Suzuki, A., Saga, Y., Suzuki, A., & Saga, Y. (2008). Nanos2 suppresses meiosis and promotes male germ cell differentiation service Nanos2 suppresses meiosis and promotes male germ cell differentiation. 430–435. https://doi.org/10.1101/gad.1612708

Wang, Z., & Lin, H. (2004). Nanos maintains germline stem cell self-renewal by preventing differentiation. Science (New York, N.Y.), 303(5666), 2016–2019. https://doi.org/10.1126/science.1093983

Weidmann, C. A., Qiu, C., Arvola, R. M., Lou, T.-F., Killingsworth, J., Campbell, Z. T., Hall, T. M. T., & Goldstrohm, A. C. (2016). Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio. ELife, 5, e17096.

Xie, T. (2008). Germline stem cell niches. StemBook, 1–18. https://doi.org/10.3824/stembook.1.23.1

Chapter 2: scRNA-seq Transcriptomics of Zebrafish Ovarian Germ Cells

2.1 Preface

This chapter included works from the following contributors:

Yulong Liu, Michelle E. Kossack, Matthew E. McFaul, Lana Christensen, Stefan Siebert, Samuel Horst, Nayeli Arroyo, Celina E. Juliano and Bruce W. Draper

2.2 Abstract

The famously fecund female zebrafish can produce hundreds of eggs each week. Unlike female mammals, both male and female zebrafish have the ability to continuously produce new germ cells throughout its lifetime because of a stem cell population called germline stem cell (GSC). Like most known stem cell systems, GSCs are restricted to a specific location, called the "niche," which contains specialized somatic gonad cells that produce specific factors to regulate and maintain GSCs. While genes regulating GSCs in mammalian males have been identified, little is known about the regulation of GSCs and the niche cell type in any female vertebrate. To elucidate the gene regulation of GSCs in female zebrafish, we performed single-cell RNA sequencing on more than 10,000 GCs and 14,000 somatic gonad cells. We characterized stage-specific GC gene expression from GSC differentiation to meiosis and oogenesis and identified covariant gene modules and putative regulatory transcription factors for various GC development stages. This included novel candidate GSC and GSC differentiation marker genes, and we validated the expression with single-molecule RNA in situ hybridization. These findings provide a comprehensive inquiry into GSC regulation, and our transcriptome dataset of the

zebrafish ovary expression landscape offers a valuable resource for female GC differentiation, meiosis, reproduction, and ovarian cancer research.

2.3 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). In addition, the core developmental regulators of mammalian gonad development have orthologs in the zebrafish genome, and in many cases mutational analysis has revealed that these genes have similar function in both mammals and zebrafish (reviewed in (Siegfried & Draper, 2020)). Though zebrafish gonad development is similar to that of mammals, there are also important differences. The testes of adult mammals and zebrafish can produce new sperm throughout life. By contrast, the ability to produce new oocytes as adults is not universal. For example, in mammals, such as humans, females are born with a finite number of oocytes arrested in prophase of meiosis I (also called dictyate arrest) whereas as in many teleost fish (bony fish), such as the zebrafish, adult females can produce new follicles, and therefore mature eggs, throughout their lifetime. Zebrafish can breed year-round with females able to produce hundreds of eggs on a near weekly basis. The zebrafish adult ovary contains germ cells at all stages of development, from premeiotic germ cells to mature eggs (Selman et al., 1993) due to the presence of germline stem cells (GSCs) (Beer & Draper, 2013b; J. Cao et al., 2019). The ability to continuously produce new germ cells throughout their lifetime makes female zebrafish an excellent research organism to study female germline stem cells 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 (T. Xie & 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*, encodes a member of the Bone morphogenetic protein family of ligands), to prevent GSC differentiation (Song et al., 2004b; T. Xie & Spradling, 2000). 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).

Although GSCs have been identified in the zebrafish ovary, very little is known about the niche and the 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. The germinal zone contains mitotically dividing GSCs, oocyte progenitor cells, and early meiotic oocytes (Beer & Draper, 2013b; Draper et al., 2007b). However, neither the GSC-intrinsic gene expression landscape nor the extrinsic GSC niche cells that regulate early GSC differentiation in zebrafish have been identified.

In this study we performed a comprehensive analysis of the GC populations in the developing zebrafish ovary through single-cell RNA sequencing. In the GC populations, we characterized detailed

expression changes from GSC to early oocyte and identified a novel progenitor GC population. We found a forkhead box transcription factor *foxl2l* is highly expressed in the progenitor GCs and *foxl2l* mutation caused all homozygous mutants to become males.

2.4 Materials and Methods

Animal

Zebrafish used for scRNA-seq libraries were *Tg(ziwi:EGFP)* transgenics (Leu and Draper, 2010) generated from the AB strain. Dissected ovaries for whole-mount immunochemistry, and single-molecule hybridization chain reaction RNA fluorescent *in situ* hybridization (smFISH) were from the wild-type AB strain. Zebrafish husbandry was performed as previously described with reduced 25 fish per tank density after 5dpf (Westerfield, n.d.).

Zebrafish ovary dissociation for single-cell RNA sequencing

For the Fluorescence Cell Sorter (FACS) sorted germ cell (GC) library, we performed a dissociation protocol that was optimized for GC dissociation. This protocol resulted in a higher number of undissociated somatic cell clumps during dissociation, but we observed higher retention of single GCs at different stages. The GC library dissociation was adopted from a previously published protocol (Blokhina et al., 2019). 20 pairs of 40dpf ovaries were dissected and stored in a LoBind tube (Cat.No. 0030108302; Eppendorf) containing 2mL of L15 medium (Cat.No. L5520; Sigma-Aldrich). Tissues were minced with microdissection scissors into <1mm pieces. 200µL of 20mg/mL type 2 collagenase in L15 (Cat.No. NC9870009; Worthington) was added and incubated on an orbital rotator at 28°C for 35 min.

The cell suspension was then gently passed through a 23g needle for 5 times to break up large cell clumps. 200µL of 7mg/mL trypsin (Cat.No. LS003708; Worthington) in L15 was 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 500µL of 20mg/mL trypsin Inhibitor (Cat.No. 100612; MP Biomedicals) in L15, the cells were centrifuged for 3 min at 300 x g and the supernatant carefully discarded. The cell pellet was then resuspended and washed with 5 mL of L15 using a P1000 pipette and subsequently centrifuged for 3 min at 300 x g. This washing step was repeated two more times. Cells were resuspended in 1mL L15 and passed first through a 100µm nylon filter (Cat.No. 431752; Corning) and then through a 40µm nylon filter (Cat.No. 431750; Corning). The filtrate was centrifuged for 3 min at 300 x g and resuspended in 1mL of 50mg/mL BSA (Cat.No. A8806; Sigma-Aldrich) in PBS. Cell viability and number were assessed using propidium iodine and Hoechst 33342 (Cat.No. H3570; Thermo Fisher) staining on a Fuchs-Rosenthal hemocytometer (Cat.No. DHC-F01; Incyto).

For whole ovary libraries, we performed a dissociation protocol that was optimized for somatic cell dissociation. This protocol resulted in high recovery of somatic single cells but with a low retention of GCs. The whole ovary library dissociation protocol was adopted from a previously published protocol (Elkouby & Mullins, 2017). 40 pairs of 40dpf ovaries were dissected and stored in a LoBind tube containing 2mL of L15. Tissues were minced with microdissection scissors into <1mm pieces. After the tissue had settled the medium was replaced with 2mL of digestive enzyme mixture (3 mg/mL collagenase I (Cat.No. C0130; Sigma-Aldrich), 3 mg/mL collagenase II (Cat.No. C6885; Sigma-Aldrich), and 1.6 mg/mL hyaluronidase (Cat.No. H4272; Sigma-Aldrich) in L15 and incubated on an orbital rotator at room temperature (RT). The suspension was monitored every 2 minutes until no or a minimal number of cell clumps were observed. Cells were centrifuged for 3 mins at 300 x g and

resuspended in 2mL of 5x TrypLE (Cat.No. A1217701; Thermo Fisher) in L15 and incubated on an orbit rotator at 28°C for 5 min. The trypsin reaction was stopped by adding 500µL of 2.8mg/mL trypsin inhibitor in L15 and incubated on an orbital rotator at 28°C for 1 min. The cell suspension was then added to 25ml of L15 in a 50ml conical tube to dilute the trypsin and centrifuged for 3 min at 300 x g. The cell pellet was resuspended and washed with 5 mL of L15 using a P1000 pipette and centrifuged for 3 mins at 300 x g. This washing step was repeated once. The cell pellet was resuspended in 1mL of L15 and then passed through first a 100µm nylon filter, then a 70µm nylon filter (Cat.No. 431751; Corning), and finally through a 40µm nylon filter. The filtrate was centrifuged for 3 min at 300 x g and resuspended in 1mL of 50mg/mL BSA in PBS. Cell viability and number were assessed using propidium iodide and Hoechst staining on a Fuchs-Rosenthal hemocytometer.

Fluorescence-Activated Cell Sorting (FACS)

40dpf Tg(*ziwi:EGFP*) transgenics zebrafish ovaries were dissociated and the cell suspension was sorted on a MoFlo Astrios EQ Cell Sorter (Beckman Coulter) with a 70 µm nozzle to obtain eGFP+ cells.

Single-cell RNA sequencing 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 (LifeTechnologies) and by qPCR with a Kapa Library Quant kit (Kapa Biosystems) prior to sequencing. Libraries were sequenced on a HiSeq 4000 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 Genomic) to retain the following attributes: protein_coding, lincRNA, and antisense. A genome reference file was generated with Cell Ranger's "mkref" function using the filtered GTF file and GRCz11 zebrafish genome obtained from the Ensembl Genome Browser.

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 Genomic). "expect-cells" was set to 10000 based on estimated cell recovery.

Cell cluster analysis

The expression matrices generated from Cell Ranger were first processed with SoupX (v0.3.1) 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) 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 GC library, we retained cells with 200 to 6000 genes, less than 150,000 unique transcripts, and less than 5% mitochondrial transcripts. For whole ovary libraries, we retained the cells with 200 to 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 GCs were also removed from the whole ovary libraries. DoubletFinder was run using a conservative 5% estimated cell doublet cutoff.

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 were generated based on selected principal components.

Trajectory analysis

Monocle3 (v0.1.3) was used for trajectory analysis (J. Cao et al., 2019).The expression matrix was exported from the Seurat object and used as Monocle3 input. UMAP was used for dimensionality reduction. We used germline stem cell (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 3,000 variable genes as identified in the Seurat analyses were used as input to the NMF analysis. To identify the optimal number of gene modules to describe the dataset we tested a broad range of K values from 10-100 in increments of 5 and then reduced to an increment of 1 in a range between 30-40. The optimal K was identified as 36 for the sorted-GC 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 the 5’ regions of co-expressed genes, we extracted 2kb 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 & Bailey, 2010) for motif enrichment analysis with the following parameter: "--scoring avg --method fisher --hit-lo-fraction 0.25 --evalue-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.

Gene ontology analysis

The top 25% of the cluster markers (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 primers for *in situ* RNA probes were designed with NCBI Primer-BLAST tool (Ye et al., 2012). RNA probes for fluorescence RNA in situ hybridizations were synthesized from mRNA isolated from 40dpf wild-type AB strain dissected ovaries. mRNA was isolated through TRI reagent (Cat. No. T9424; Sigma-Aldrich) extraction, and the cDNAs were synthesized with RETROScript Reverse Transcription Kit (Cat. No. AM1710; Thermo Fisher). Primer-specific amplified DNA templates for the RNA probes were generated using Phusion polymerase (Cat. No. M0530L; New England BioLabs). Reverse primers contained a T7 RNA polymerase promoter shown in the lower-case letters. DNA templates were

reverse transcribed with T7 RNA polymerase (Cat. No. 10881775001; Roche) to yield Digoxigenin labeled antisense RNA probes. Probes were G-50 column purified (Cat. No. 45-001-398; GE Healthcare) and column concentrated (Cat. No. R1015; Zymo). Probes were diluted to 0.5-2 mg/ml in the hybridization solution with 5% dextran sulfate.

All dissected ovary samples were fixed with 4% paraformaldehyde overnight at 4°C. Samples were dehydrated with 100% methanol for 10 mins at room temperature and stored at -20°C at a minimum overnight with fresh 100% methanol.

The smFISH procedure was performed following a published protocol with the following modifications (Thisse and Thisse, 2008): 40 dpf ovaries were permeabilized with proteinase K at 50 µg/mL for 15 minutes, and the fluorescence in situs were developed with FastRed (Cat.No. F4648; Sigma-Aldrich).

smFISH probes were purchased from Molecular Instruments. Hybridization procedures followed previous instructions and the Molecular Instruments protocol (MI-Protocol-HCRv3-Zebrafish version 5; Molecular Instruments) with the following modification: tissue preparation followed the procedure above, tissues were permeabilized with proteinase K at 50 µg/mL for 15 minutes. Prior to imaging samples were cleared with 30%/50%/70% glycerol-PBS gradient for 1 hour in each step, and then mounted with ProLong™ Diamond Antifade Mountant (Cat.No P36961, Invitrogen)

Immunofluorescence staining

Whole ovary immunofluorescence were performed based on a previously published protocol (Leerberg et al., 2017). The anti-DEAD-Box Helicase 4 (Ddx4) antibody (Knaut et al., 2000) was used at 1:1500

dilution, and the Proliferating cell nuclear antigen (PCNA) antibody (Cat.No. ab29, Abcam) was used at 1:100 dilution.

Imaging

Fluorescence RNA in situ hybridization, HCR fluorescence RNA in situ hybridization, and immunohistochemical staining samples were mounted with ProLong Diamond antifade mountant (Cat.No. P36971, Invitrogen) and imaged with either Olympus FV1000 laser scanning confocal microscope or Zeiss LSM 880 Airyscan microscope.

Generation of *foxl2l:egfp* knock-in mutant

Design plasmid – Generation of the *pGTag- foxl2l- eGFP-B-actin* targeted integration plasmid was designed using the pGTag vector series (Wierson et al. 2020). Targeted integration of the eGFP into the foxl2l locus was done using CRISPR/Cas9 cutting followed by integration of the eGFP from the pGTag plasmid via homology directed repair. 48-bp homology arms for foxl2l corresponding to the gRNA CRISPR cut-site were cloned into the *pGTag-eGFP-B-actin* vector following Wierson et al. methods. Upstream homology arm 5’ GCGGgggAAACGCTCTAGTGCCTCTGAGCGGCATGACTCCGCCGGTGAGCCCGGGcGGAT-3’ Downstream homology arm 5’-AAGCGGAAGCTCCATCTCCACCTGCAGTTACGCGCCGCAGAACAGTCACCCcccCCG-3’ CRISPR gRNA 5’-GGGAGATGGAGCTTCCGCCC-3’ (oligos ordered from IDT; gRNA designed using CRISPRscan.org). Homology arms were cloned into the pGTag vector backbone using the Wierson et al.

methods for 1-pot digestion using type II restriction enzymes BspQI and BfuAI (NEB) with alterations in the amount of BfuAI (using 1ul/sample instead of 0.5ul). This is followed by ligation with T4 ligase. Ligation mixtures were then transformed via heat shock into NEB stable E.Coli (C3040H) following the Wierson et al. methods, generating the pGTag-foxl2l-eGFP-B-actin plasmid. The plasmid was then screened and sequenced for precise integration of the homology arms.

Embryo injection – Assembly of gRNA oligo was performed using IDT gRNA oligo, scaffold oligo, and Phusion polymerase in a PCR reaction. Samples were then purified using the Monarch PCR purification kit, followed by IVT using T7 RNA-pol and TURBO DNase clean-up to generate the foxl2l gRNA. Cas9 RNA was generated from pT3TS-nCas9n plasmid using the mMESSAGE mMACHINE kit with T3 polymerase for IVT. 150 pg/nl cas9 mRNA, 50 pg/nl foxl2l gRNA, 0.75 ul of pGTag UgRNA (325 ng/ul), 10 ng/ul of pGTag-foxl2l-eGFP-B-Actin plasmid, and 0.8 ul of 5X Phenol red (1.5M KCl) was mixed and injected into 1-cell WT (AB-line) zebrafish embryos at 2nl per embryo. Injection amounts were calibrated by injecting reaction mixture into oil droplets on a hemocytometer.

Identify inserts: Sort and outcross fish – Injected embryos were grown in a 10cm dish at 37°C. Injected embryos were then screened at 24-hr by whole-embryo DNA extraction followed by PCR flanking the 5' and 3' insert sites. 24-hr PCR screen contains 0.5 ul house-made Taq polymerase, 1 ul 2.5 mM dNTPs, 1.5 ul 10X buffer, 1 ul of each primer, 10 ul water. PCR reaction products were then separated on a 1.5% agarose gel to look for expected insert size bands. If the PCR screen showed bands of the expected size for the correct insertion, then injected embryos were grown to adulthood. Upon sexual maturity, the F0 fish were outcrossed to WT. A sample of the embryos from these were screened as well using the same 24-hr PCR screen to assess germline transmission. These heterozygous F1 fish were

then screened for fluorescence at 1.5-2 mpf to see if any GFP could be seen in the expected location (the gonad). Fish that were GFP-positive were sorted and later in-crossed to generate homozygous pGTag-foxl2l-eGFP-B-actin transgenic fish.

Genotyping *foxl2l:egfp* knock-in mutant

Genomic DNA was extracted from *foxl2l:egfp* knock-in mutant crosses. Used two primer pairs of "LNC 55-F GTACCACCGGCATCC" with "LNC 56-R GCACCGAGGTTTGCCATTAGT", and "LNC 55-F GTACCACCGGCATCC" with "LNC 60-R GCTGAACTTGTGGCCGTTTA" under the following PCR condition: initial denaturation at 95°C for 60s, 35 cycles at 95°C for 15s, 50°C for 15s, and 68°C for 20s, and final extension at 68°C for 30s. Two PCR products were separated on separate 2% agarose gels. Wildtype has only 155bp band, heterozygous mutant has both 155 bp and 217 bp bands, and homozygous mutant has only 217 bp band.

Germ cell marker quantification

The number of early GCs expressing *nanos C2HC-Type Zinc Finger 2* (*nanos2)*, *nanos2* + *forkhead Box L3 foxl3* (*foxl3)*, or *foxl3* + *REC8 meiotic recombination protein a* (*rec8a)* in individual clusters were counted in Z-stack images with 1µm 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 3 independent experiments with 3 different 40dpf ovaries. The cell count graph was generated with ggplot2 (v3.3.2) (Villanueva & Chen, 2019).

2.5 Results

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 differentiate as either an ovary or testis (Kossack & Draper, 2019). Second, by 40 dpf all major ovarian somatic cell types are likely present (Rodriguez-Mari et al., 2005). Finally, as we were particularly interested in understanding how the developmental progression of germline stem cells are 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(piwil1:egfp)*uc02 transgenic ovaries, where eGFP is expressed in all germ cells (Leu & 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 gently dissociated ovaries followed by fluorescence-activated cell sorting (FACS) to purify eGFP+ germ cells; (Figure 2; see 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 of 10,658 germ cells and 14,431 somatic cells (Figure 3; see Methods). We obtained germ cells with an average of 2,510 genes/cell, 11,294 transcripts/cell and somatic cells with an average of 854 genes/cell, 4,812 transcripts/cell.

Cell clustering analysis grouped cells into eight distinct populations using uniform manifold approximation and projection (UMAP; Figure 1, Figure 4A). 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*) (Gautier et al., 2011), follicle cells (*gonadal soma derived factor*, *gsdf*) (Gautier et al., 2011), theca cells (*cytochrome P450 cholesterol side-chain cleaving enzyme*, *cyp11a2*) (Parajes et al., 2013), vasculature (*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), and NK cells (*NK-lysin tandem duplicate 2*, *nkl.2*; Figure 1 and Figure 4B) (Carmona et al., 2017; Tang et al., 2017). 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 had previously been identified that specifically mark this population. 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). We reasoned that the remaining unidentified population of cells in the UMAP graph represented the stromal cells. We determined that expression of the gene encoding *collagen, type I, alpha 1a* (*col1a1a*) (Morvan-Dubois et al., 2003), is highly enriched in this cell population, consistent with the structural role some subtypes of stromal cells likely play in the ovary (Figure 1B and Figure 4B). Thus, we were able to identify all previously observed and expected cell types in our dataset.

The analysis revealed genes that are differentially expressed between major cell types but did not distinguish possible sub-populations. 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 sub-clustered each population.

Developmental trajectory of female germ cells

Unlike mammals, female zebrafish are able to produce new oocytes throughout their lifespan due to the presence of GSCs (Beer & Draper, 2013b; Z. Cao et al., 2019). 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 the genes regulating germ cell development, we performed sub-cluster analysis on the germ cell population (Figure 5A) and the differentiation trajectory was immediately evident in the UMAP. Using the known stage-specific germ cell markers *nanos2*, *DNA meiotic recombinase 1* (*dmc1*) and *zona pellucida protein 3b* (*zp3b*), which are expressed in GSCs, early meiotic germ cells and early oocytes, respectively, we found that the position of a cell within the UMAP corresponded to its developmental stage from GSC to early diplotene-arrested oocyte (dictyate; Figure 5A,B) (Beer & Draper, 2013b; Onichtchouk et al., 2003; Yoshida et al., 1998). Because our cell dissociation method excluded cells greater than 40µm in diameter, it is likely that only Stage IA (pre-follicle stage) or early Stage IB (follicle stage) oocytes are present in our scRNA-seq data set (Selman et al., 1993). To further assess this developmental trajectory, we performed a dedicated trajectory analysis and inferred pseudotime using Monocle 3 (Z. Cao et al., 2019). We found that the directionality and the sequential gene expression along pseudotime precisely correlated with the trajectory determined by our initial analysis (Figure 6).

With the developmental trajectory of the germ cells constructed, we were able to identify genes with enriched expression in subsets of germ cells. Notably missing from the known stage-specific

genes listed above are genes known to be expressed in proliferating oocyte progenitor cells that are intermediate between the GSC and the cells that have entered meiosis (sub-cluster 2 in Figure 5A). In most organisms, oocyte progenitor cells undergo several rounds of mitotic amplifying division before entering meiosis. 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 multi-cell cyst with synchronized developmental progression (Marlow & Mullins, 2008; Pepling et al., 1999). To identify novel markers of oocyte progenitor cells, we analyzed the genes with enriched expression in this sub-population (sub-cluster 2 in Figure 5A). 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*. Further sequence analysis revealed that this gene is the zebrafish ortholog of *forkhead-box protein L2 like* (*foxl2l*; Figure 7) (Ruzicka et al., 2019). FoxL2l is a paralog of FoxL2 that 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 7). Interestingly, medaka *foxl2l* (formerly *foxl3*; Figure 7) is expressed in oocyte progenitors and is required for germ cells to commit to the oocyte-fate (Nishimura et al., 2015). The *foxl2l* positive cells in our data set also express higher levels of *proliferating cell nuclear antigen* (*pcna*) gene relative to the GSC population (Figures 5B,C), indicating they divide more rapidly than GSCs, providing further support that this sub-population are zebrafish oocyte progenitors.

To test the hypothesis that *foxl2l*-expressing cells are oocyte progenitor cells, we used single molecule fluorescent in situ hybridization (smFISH) to determine where these cells are located in the 40 dpf ovary. Specifically, we determined the location of sub-cluster #2 cells relative to GSCs (sub-cluster #1) and to cells that are initiating meiosis (sub-cluster #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 germline stem cells (Beer & Draper, 2013b) while *rec8a* encodes a meiosis-specific member of the Rad21 cohesin family that is loaded onto chromosomes during pre-meiotic S-phase (D. 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 pre-meiotic S-phase (Figures 5B,C). Indeed, we found that *foxl2l*- and *rec8a*-expressing cells localize within multi-cell clusters (Figures 5E-H), consistent with our hypothesis that *foxl2l* is expressed in oocyte progenitors, whereas *nanos2*+ cells are found predominantly as single cells or doublets, as previously shown (Figures 5D-G) (Beer & Draper, 2013b). We found cells expressing *nanos2*, *foxl2l* and *rec8a* often clustered near one another, (Figure 5G). 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 (Figures 5I). Interestingly, we rarely found cells that were double-positive for *nanos2* and *foxl2l*, and when we did, they had minimal expression of either gene, suggesting these cells were differentiating from GSC to oocyte progenitor (Figure 5F). 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 (Figures 5E,G). The second consisted of cells within larger clusters (≤8 cells), expressed relatively less *foxl2l* but also expressed *rec8a*. These cells were likely late oocyte progenitors, including those that are undergoing pre-meiotic S-phase (Figures 5G,H). Thus, we have identified gene expression patterns that define the pre-meiotic germ cell populations in the ovary.

To determine the function of *foxl2l*, we used CRISPR/Cas9-mediated gene editing to recombine a viral-2A-EGFP insert in-frame into the *foxl2l* locus, resulting in the *Tg(foxl2l:foxl2l-2A-egfp)*uc91 allele (Wierson et al., 2020). The resulting fusion protein encodes all but the last 27 of 239 amino acids of the FoxL2l protein (Figure 11A). We found that in heterozygous animals, GFP was expressed in a subset of premeiotic germ cells in the ovary (Figure 11B), a pattern that was identical to that determined using smFISH (Figure 5E-H). To assess the role of *foxl2l* 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 (9 females and 11 males, Figure 11C-E). By contrast, all homozygous mutant animals (0 female, 10 males) were fertile males as adults, indicating that *foxl2l* is required for female development (Figure 11C,F). In zebrafish, the ability to produce oocytes is a requirement for female development, thus loss of oocyte production, which we hypothesize occurs in *foxl2l* mutants, results in all-male development (Rodriguez-Mari et al., 2013). In medaka, *foxl2l* is expressed in oocyte progenitor cells and is required for these cells to commit to oogenesis, and upon loss of *foxl2l* these cells instead commit to spermatogenesis (Nishimura et al., 2015).

Transcriptional characterization of female germline stem cells

Female germline stem cells 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 & Draper, 2013b; J. Cao et al., 2019; Draper et al., 2007b). *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 animals (Beer & Draper, 2013b; Z. Cao et al.,

2019). By contrast, *nanos3* is only expressed in females where it plays two separate roles. First, *nanos3* (formally called *nanos1*) is a maternally expressed gene whose mRNA localizes to primordial germ cells (PGC) and loss of maternal function leads to loss of PGCs during early development (Draper et al., 2007b; Köprunner et al., 2001). Second, *nanos3* mutant females fail to maintain oocyte production as adults due to loss of *nanos2*+ GSCs (Beer & Draper, 2013b; Draper et al., 2007b). While genetic mosaic analysis established that *nanos3* was required cell autonomously for GSC maintenance (Beer & Draper, 2013b), to date *nanos3* expression in premeiotic germ cells has not been demonstrated. To address this, we identified the *nanos3*-positive cells in our data set. 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 (Figures 5B,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 & Lehmann, 1998b). The zebrafish genome encodes three orthologs of Pumilio, called *Pum1-3*, but it is not known which of these functions along 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 8). In addition, *pum3* is also expressed in early oocytes (Figure 8). 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 (3,517 unique transcripts/cell; Figure 9) relative to other premeiotic cells (e.g. 2,315 unique transcripts/cell in

progenitor cells). To identify potential regulators of GSC development, in addition to *nanos2*, we identified genes in our data set with enriched expression in this cell population. They include *inhibitor of DNA binding 4* (*id4*), which encodes a negative regulator of transcription that is also expressed in mouse spermatogonial stem cells*, jagged canonical Notch ligand 2b* (*jag2b*), *dickkopf WNT signaling pathway inhibitor 3b* (*dkk3b*), *DNA methyltransferase protein 3bb.1* (*dnmt3bb.1*), which encodes a protein potentially involved in epigenetic programing, *interferon regulatory factor 10* (*irf10*), and *chromogranin b* (*chgb*), which encodes a protein involved in production of dense core granules in neuroendocrine cells (Figure 9B) (Gore et al., 2016; Haddon et al., 1998; Hsu et al., 2010; M. J. Oatley et al., 2011; J. Xie et al., 2008). Using *id4* and *chgb* as examples, we performed smRNA in situ to verify their expression pattern in ovaries. We found that all *id4*-expressing cells also expressed *nanos2* (Figure 9C), confirming that *id4* is expressed in GSCs in the 40dpf ovary. Interestingly, not all *nanos2*+ cells were *id4*+. We obtained similar results with *chgb* (Figure 9D). Thus, this analysis has identified several new genes with specific expression in GSCs. 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 models and possible transcriptional regulators

We next aimed to identify the potential transcription factors that regulate the development of germ cell stages. As a first step we performed non-zero matrix factorization (NMF) to find gene modules, which are sets of genes that are co-expressed within cell clusters and sub-clusters and may therefore identify co-regulated genes (Brunet et al., 2004; Farrell et al., 2018; Siebert et al., 2019). We identified 32 gene modules (note: module 8, 13, 19 and 29 were determined to be low quality gene

modules and therefore not considered in our analysis; see 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 10A, 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 Methods). We then verified our gene motif enrichment using the *figla* transcription factor as a test case. Folliculogenesis specific bHLH transcription factor (Figla) is known to regulate folliculogenesis-specific gene expression in mice and zebrafish and *figla* is expressed in late meiotic germ cells and early oocytes (Figure 10C) (Qin et al., 2018; Soyal et al., 2000). We identified modules that contained gene with enriched putative Figla binding sites, and found that these motifs were highly enriched in module 9, a module that contains genes with enriched expression in early oocytes (e.g. *zp3*; Figure 10A,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 were specifically enriched in the GSC cluster (Figure 10). We found that genes within these two modules had enrichmed putative binding motifs for transcription factors basic helix-loop-helix family member e23 (Bhlhe23), early growth response 4 (Egr4), and pre-B-cell leukemia homeobox 3b (Pbx3b) (Figure 10A,D), and importantly, the genes encoding these factors had enriched expression in GSCs (Figure 10D). Previous studies of Egr4 deficient mice have shown reduced fertility, but the activities of Bhlhe23 and Pbx3b in GSC development or function have not been reported (Tourtellotte et al., 1999). We performed smRNA in situ hybridization to confirm the expression of *bhlhe23* and found it specifically expressed in a subset of *nanos2*+ GSCs (Figure 10E). We

also found multiple genes that are involved in germ cell regulation and contain the Bhlhe23 binding motifs through the motif enrichment, 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*, *tsmb4x* and *mcm6* are expressed in GSCs and have putative Bhlhe23 sites within 2KB of their transcription start sites, suggesting that Bhlhe23 may regulate the expression of multiple genes (Figure 10F).

2.6 Discussion

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. We transcriptionally profiled over 10,000 germ cells that represent all developmental stages from germline stem cells to early-stage oocytes (Stage IA). Though previous work had identified genes expressed in most of these developmental stages (e.g. *nanos2* for germline stem cells and *dmc1* for early meiosis) (Beer & Draper, 2013b), our data set allowed us to identify markers of progenitor cells and of cells that are undergoing pre-meiotic S-phase (*foxl2l* and *rec8a*, respectively). The ability to identify specific germ cell stages in the ovaries will allow us to more accurately define the structure of the germline stem cell niche and to better characterize germline developmental defects that result from experimental manipulations (e.g. mutant analysis or following chemical perturbation). In addition to identifying new stage-specific markers, the sequencing depth has allowed us to also identify gene modules that may represent co-regulated genes (e.g. genes expressed in germline stem cells), as well as candidate transcriptional regulators. This information should allow us to build more accurate gene

regulatory networks, which can be experimentally tested in the genetically tractable zebrafish. Finally, we provided genetic evidence that the oocyte progenitor cell-expressed *foxl2l* is required for oocyte development in zebrafish, similar to the role of *foxl2l* in medaka (Nishimura et al., 2015).

2.7 Figures

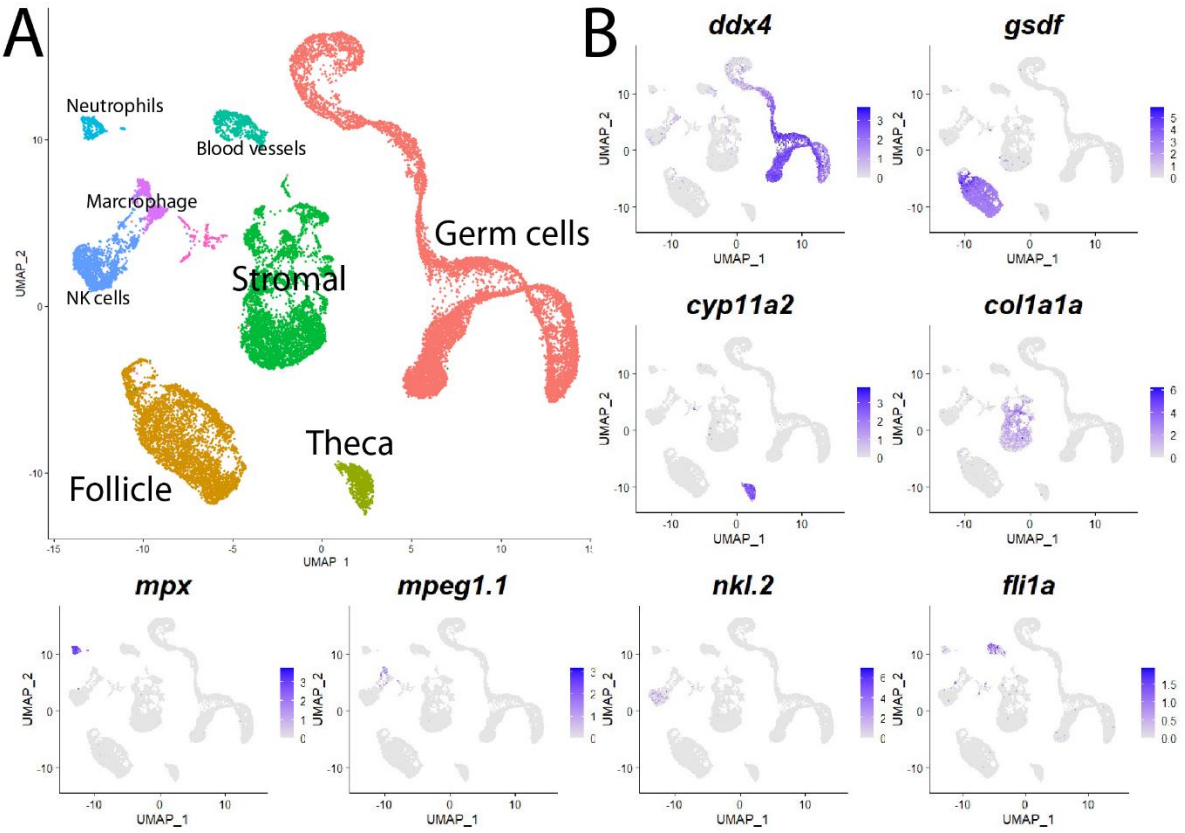


Figure 1. Single-cell RNA sequencing of 25,089 single cells isolated from 40-day old zebrafish ovaries. A. Single-cell UMAP plot of the 40-day old zebrafish ovary. Cells are color-coded by computationally determined cell clusters. B. Gene expression plots of known cell-specific marker genes identify the major cell type (labeled in A) that each cluster corresponds to.

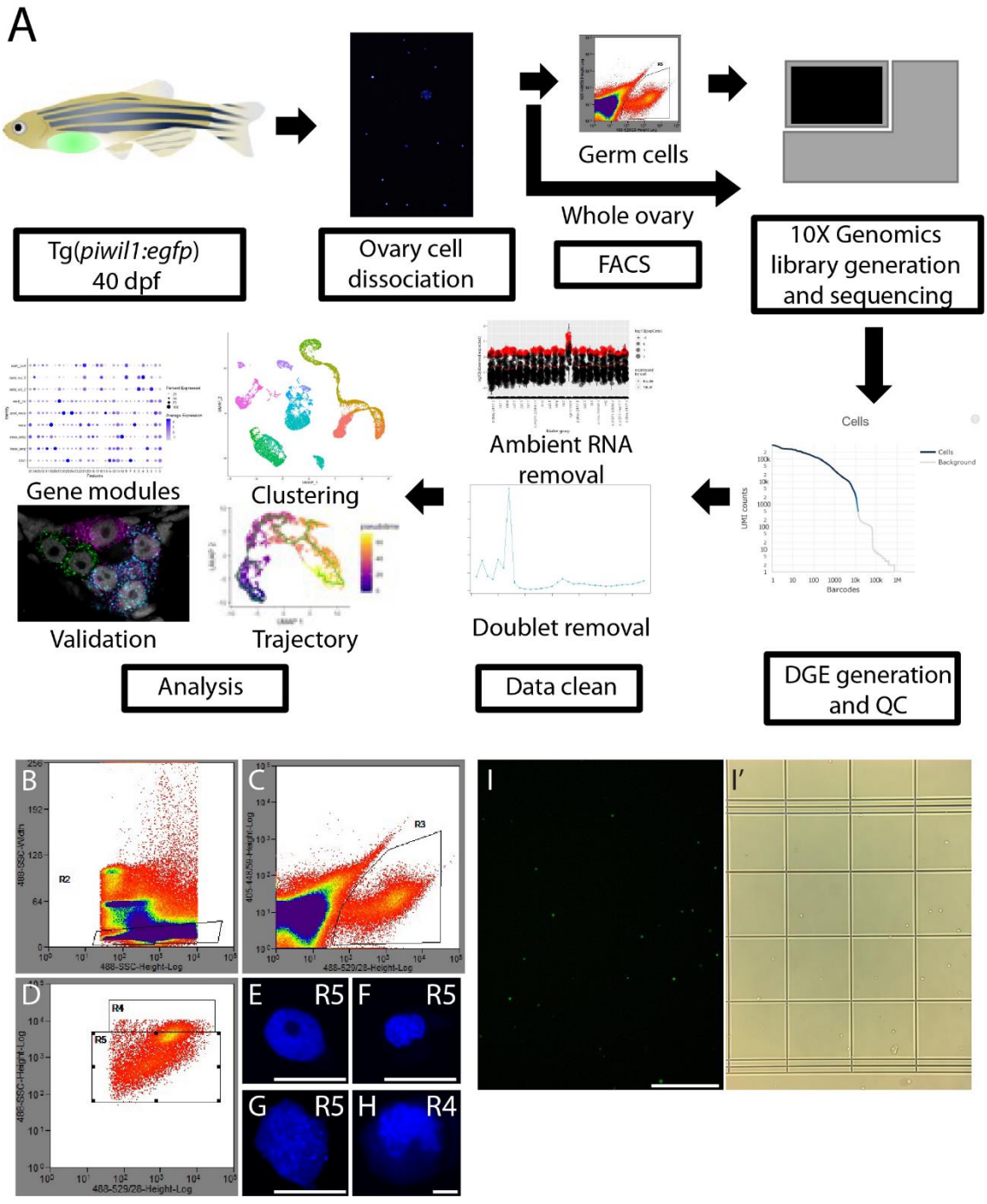


Figure 2. A. Data generation and processing flow chart. Schematic of the experimental workflow, including library generation, data processing, and analysis. GC-specific library cell sorting and validation (**B-D**) Scatter plot and gating perimeter used to sort eGFP+ GCs for the

GC-specific library. R2 gated for single cells (**B**), R3 gated for GFP+ signals (**C**), R4 contains relatively larger oocytes, and R5 are obtained that was enriched for early GCs, including early oocytes (**D**). (**E-H**) Representative images of cell nucleus staining obtained from either R4 (**H**) or R5 (**E-G**) regions. All 10µm scale bars. (**I-I'**) Florescence (**I**) or bright-field (**I'**) microscopy images of cells obtained after sorting shows all cells are eGFP+. 0.25 mm scale bar.

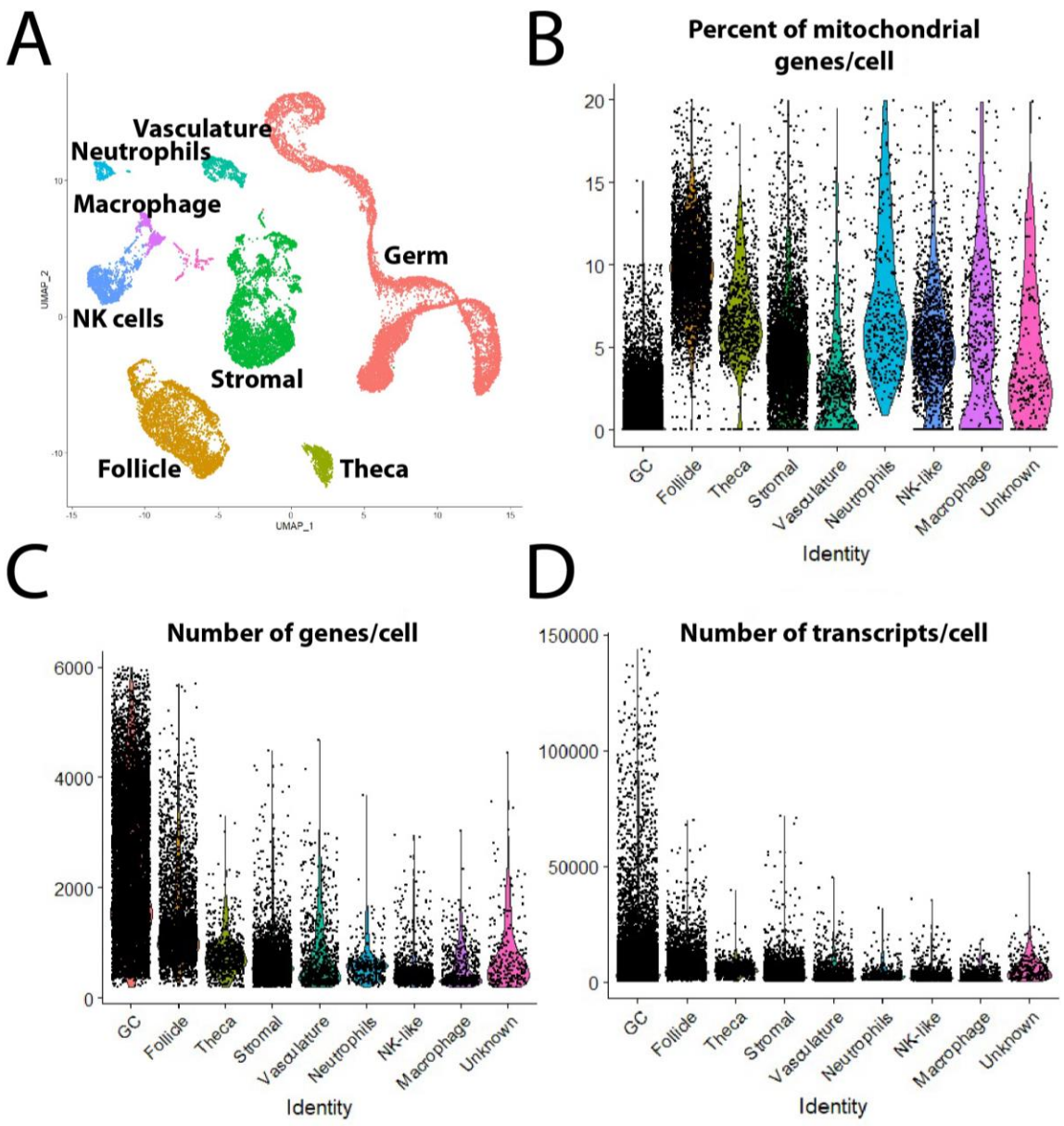
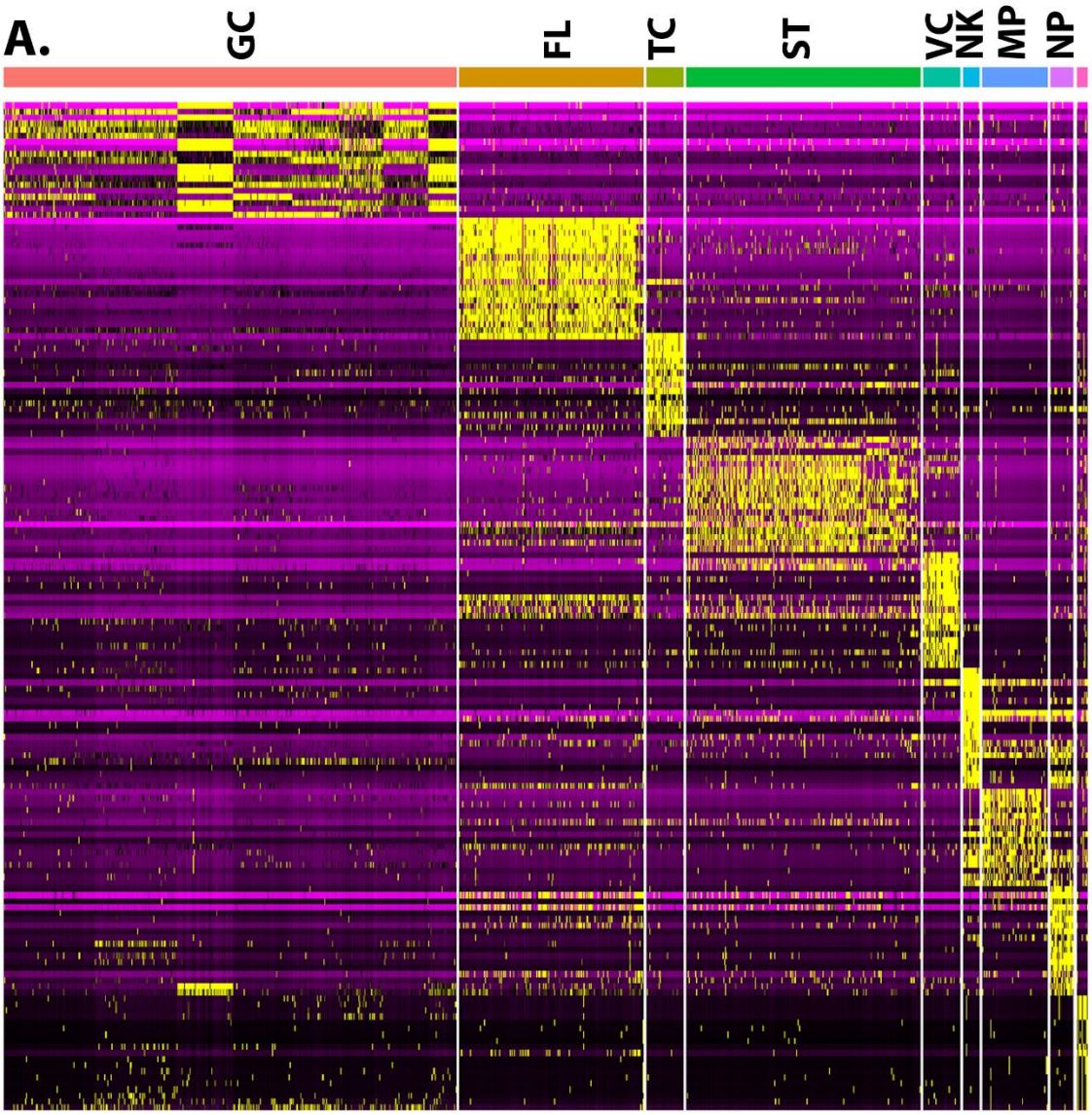


Figure 3. Major cell types statistics **A.** Single-cell UMAP plot of 40 days old zebrafish ovary with combined clusters based on major cell types. **B-D.** Violin plot of percent of mitochondria genes (**B**), number of genes per cell (**C**), and number of transcripts per cell (**D**) in each major cell type in the final dataset.



B.

GC	FL	TC	ST	VC	NP	NK	MP
ddx4	*gsdf*	*cy917a1*	*col1a1a*	*cdh5*	*lyz*	*ccl36.1*	*grn1*
sycp1	*hsd19b1*	*star*	*nid1b*	*etv2*	*lect21*	*nkl.2*	*ccl35.1*
zp3.2	*her2*	*cyp11a2*	*krt15*	*sox7*	*mpx*	*wasb*	*ccr9a*
piwil1	*cldn11a*	*ddx1b*	*tagln*	*aoc2*	*cpa5*	*laptm5*	*c1qb*
nasp	*notch3*	*fgg*	*anxa1a*	*krt8*	*cxcr4b*	*arpc1b*	*lygl1*

Figure 4. Major cell-types differential expression heatmap and top markers **A.** Gene expression heatmap of differentially expressed genes between major cell types. Yellow represents highly expressed genes, purple represents lowly expressed genes, and black represents no expression.

B. Top 5 marker genes computationally identified in each major cell type based on statistical significance.

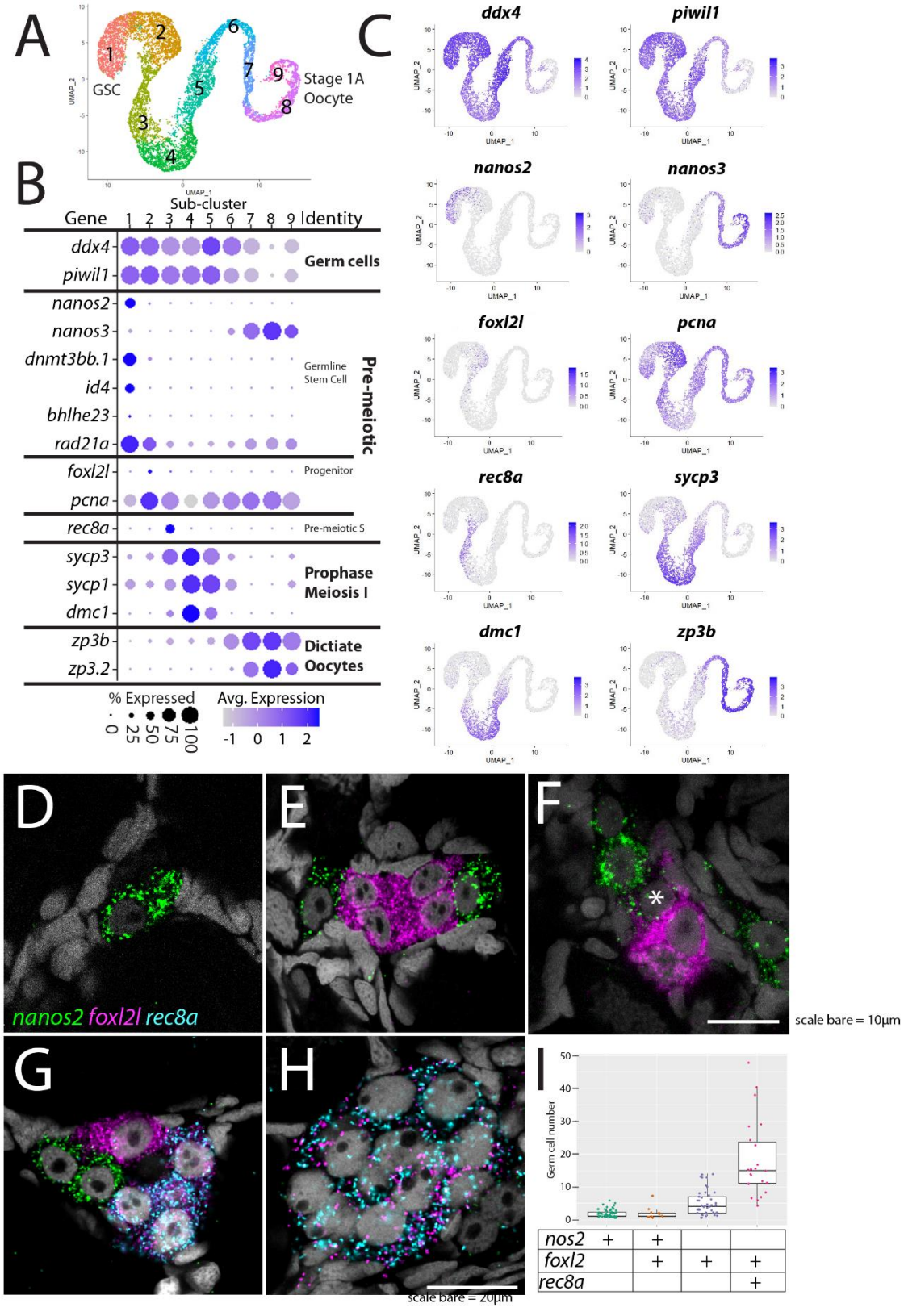

A
GSC
Stage 1A Oocyte
B
Sub-cluster
Gene 1 2 3 4 5 6 7 8 9 Identity
ddx4
piwil1
Germ cells
nanos2
nanos3
dnmt3bb.1
id4
bhlhe23
rad21a
Germline Stem Cell
Pre-meiotic
foxl2l
pcna
Progenitor
rec8a
Pre-meiotic S
sycp3
sycp1
dmc1
Prophase Meiosis I
zp3b
zp3.2
Dictiate Oocytes
% Expressed
Avg. Expression
C
ddx4
piwil1
nanos2
nanos3
foxl2l
pcna
rec8a
sycp3
dmc1
zp3b
D
E
F
nanos2 foxl2l rec8a
scale bare = 10µm
G
H
scale bare = 20µm
I
Germ cell number
nos2 + + +
foxl2 + + +
rec8a + +

Figure 5. Germ cell sub-cluster analysis reveals developmental transitions of early germ cells. A. Germ cell sub-cluster 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 sub-clusters. Some genes, like *ddx4* and *piwil1*, are expressed in all germ cells, while others, such as *nanos2* or *rec8a*, are only expressed in distinct subclusters. C. Gene expression UMAP plots of select genes. 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). D-H. Triple single-molecule fluorescent in situ hybridization for *nanos2* (green), *foxl2l* (magenta) and *rec8a* (blue) in 40 days old zebrafish whole mount ovary. Asterix 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, 10µm; in H, 10µm.

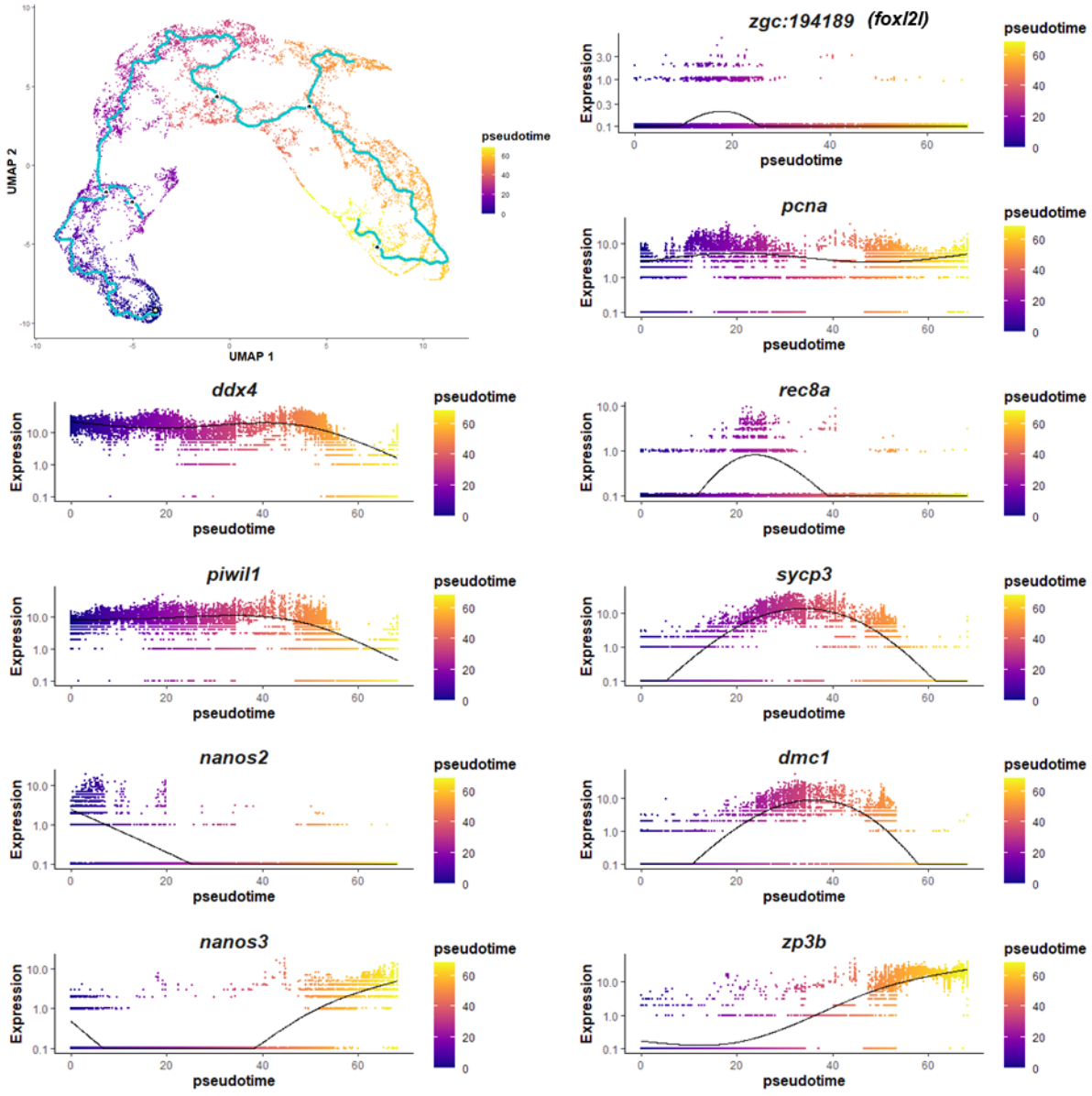


Figure 6. Trajectory analysis of the sorted GC library reproduced a similar trajectory in cluster analysis. Germ cell sub-cluster trajectory UMAP plot and pseudotime gene expression plots.

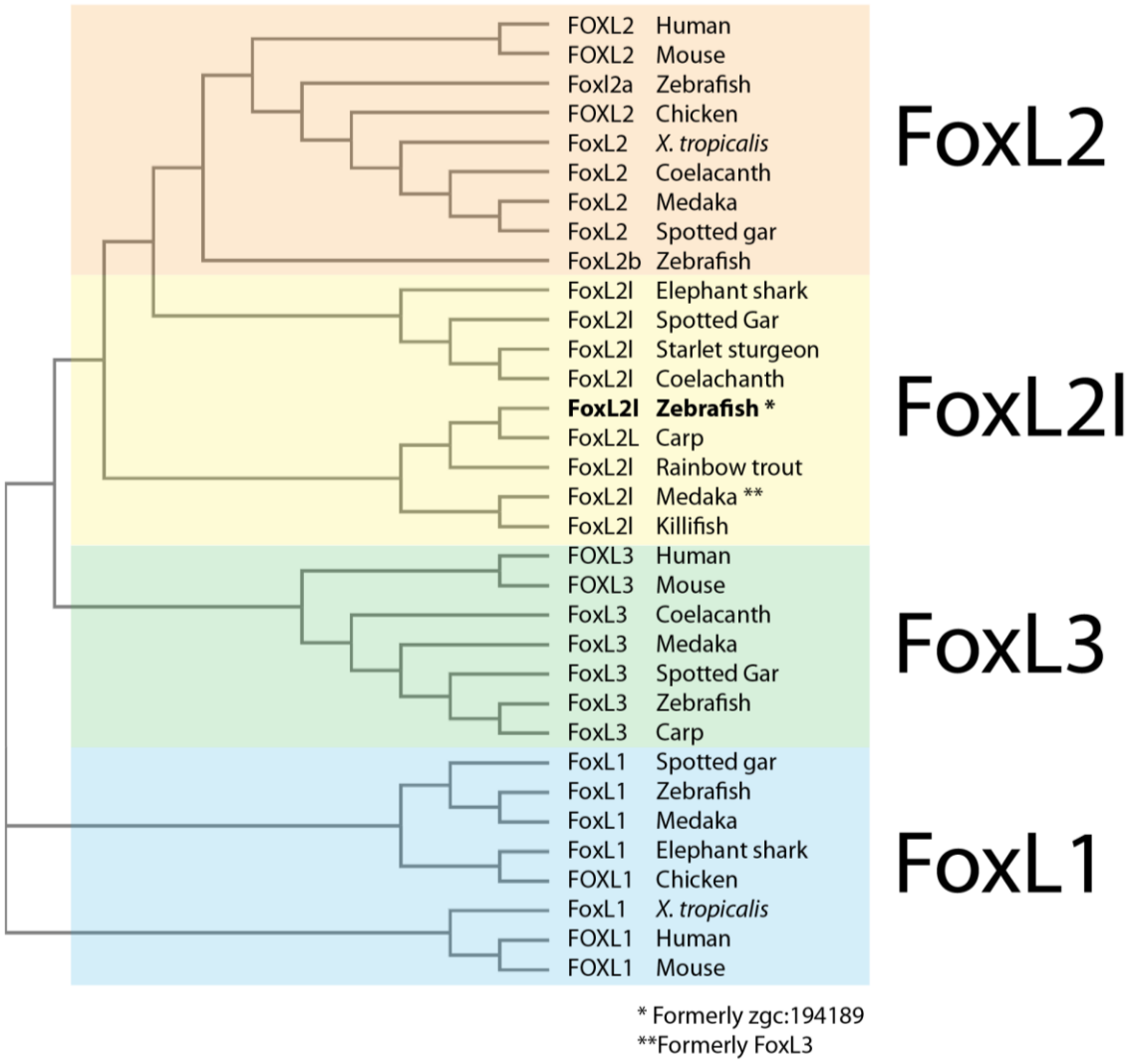


Figure 7. Phylogenetic analysis of Foxl2, Foxl2l, Foxl3 and Foxl1 proteins. *denotes zebrafish Foxl2l (formerly annotated as zgc:194189) and **denotes medaka Foxl2l (formerly called Foxl3).

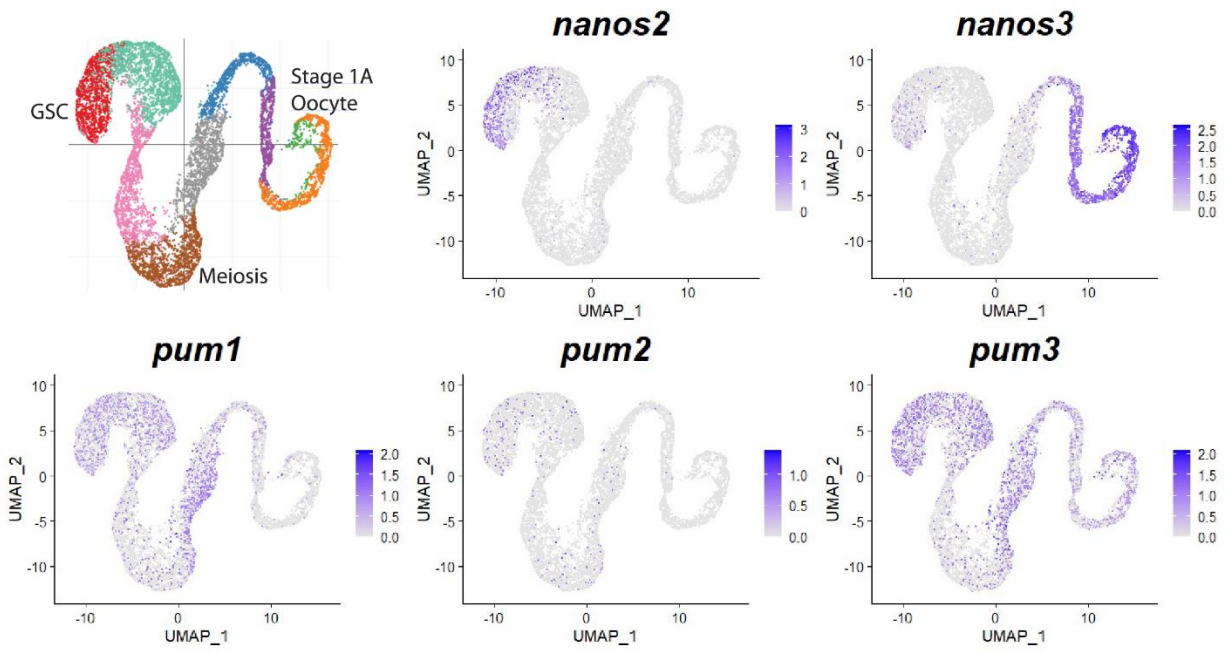


Figure 8. Expression of *nanos* and *pumilio* orthologs in zebrafish germ cells. Gene expression UMAP plots of select genes. Cells expressing the indicated gene are colored blue, and the relative intensity indicates relative expression levels (intensity scale for each plot is on the right). The UMAP in the upper left corner shows cells color-coded by computationally determined cell subtypes.

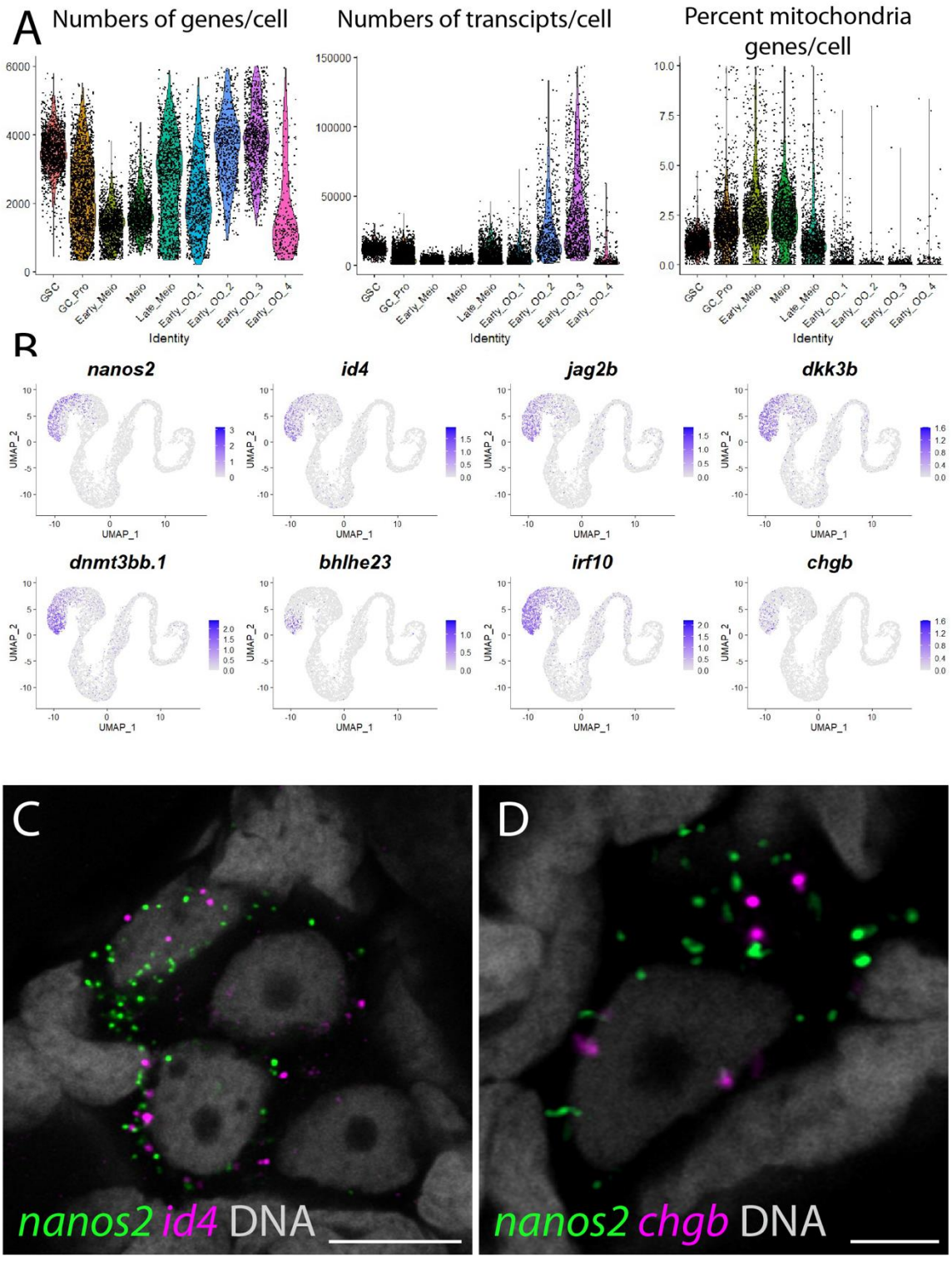


Figure 9. Germ cell library statistics and novel zebrafish GSC markers

A. Violin plot of the number of genes per cell, number of transcripts per cell, and percent of mitochondria genes in each germ cell sub-clusters. **B**. Gene expression UMAP plots of select GSC-enriched genes. Cells expressing the indicated gene are colored blue, and the relative intensity indicates relative expression levels (intensity scale for each plot is on the right). **C, D.** mFISH on whole-mount 40 dpf ovaries reveals the location confirms expression of *id4* (red in panel **C**) and *chgb* (red in panel **D**) in *nanos2* expressing germline stem cells (green). In all panels, DNA is gray. All 5µm scale bars.

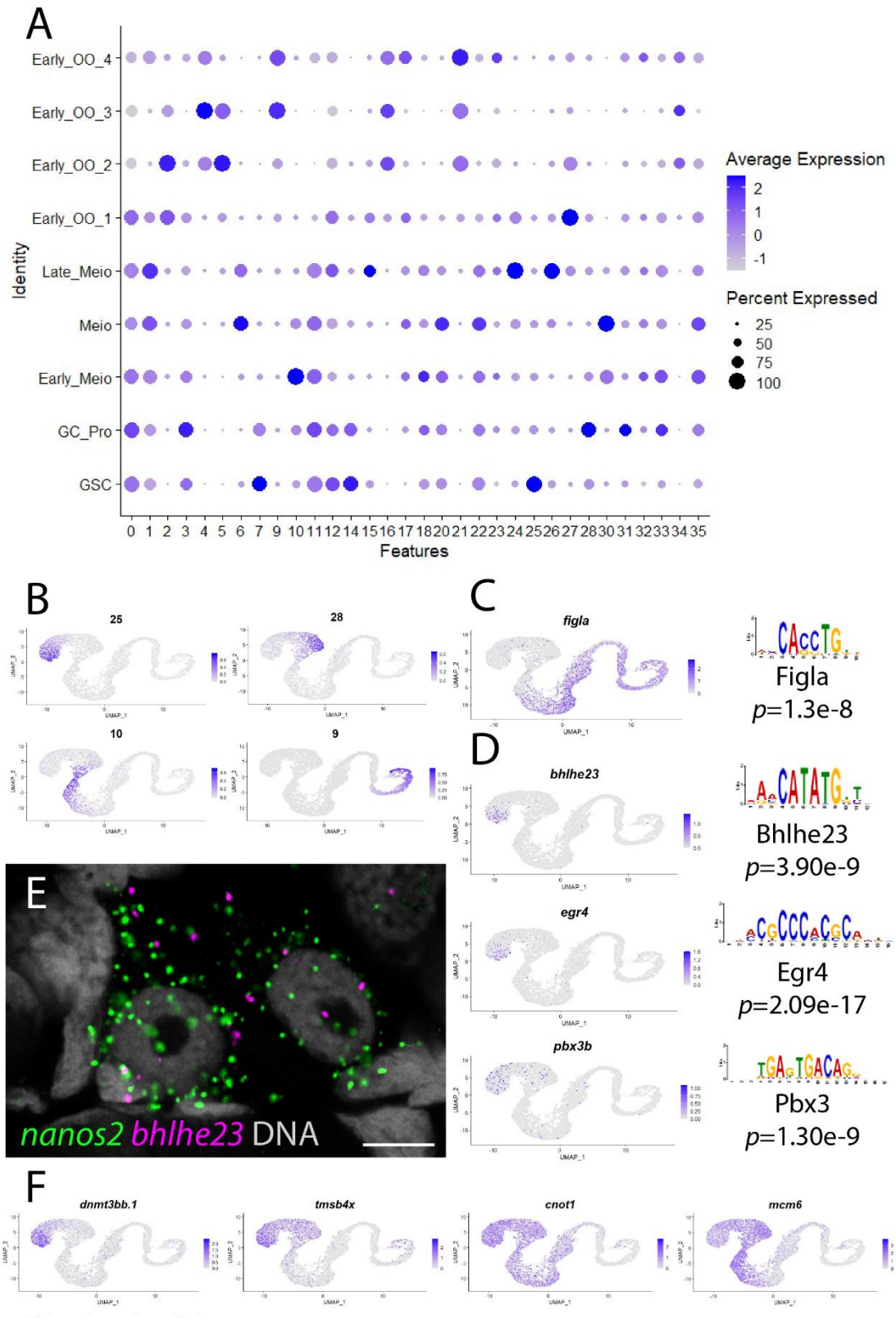

A
Early_OO_4
Early_OO_3
Early_OO_2
Early_OO_1
Late_Meio
Meio
Early_Meio
GC_Pro
GSC
Identity
0 1 2 3 4 5 6 7 9 10 11 12 14 15 16 17 18 20 21 22 23 24 25 26 27 28 30 31 32 33 34 35
Features
Average Expression
2
1
0
-1
Percent Expressed
25
50
75
100
B
25
28
10
9
C
figla
Figla
p=1.3e-8
D
bhlhe23
Bhlhe23
p=3.90e-9
egr4
Egr4
p=2.09e-17
pbx3b
Pbx3
p=1.30e-9
E
nanos2 bhlhe23 DNA
F
dnmt3bb.1
tmsb4x
cnot1
mcm6

Figure 10. Gene module analysis and motif enrichment identifies putative GSC-specific transcription factors. **A.** Dot plot of expression level and percent expressed of identified gene modules corresponding to germ cell clusters. **B.** Average expression of genes in selected gene modules plotted in germ cell subcluster. **C-D** Gene expression plots and its corresponding binding motifs of selected transcription factors that were identified from motif enrichment of module 9 (**C**) and module 25(**D**) **E.** mFISH on whole-mount 40 dpf ovaries reveals the location confirms expression of *bhlhe23* (red) in nanos2-expressing GSC (green). DNA is grey. 5µm scale bar. **F.** Putative targets of *bhlhe23* identified through gene motif enrichment.

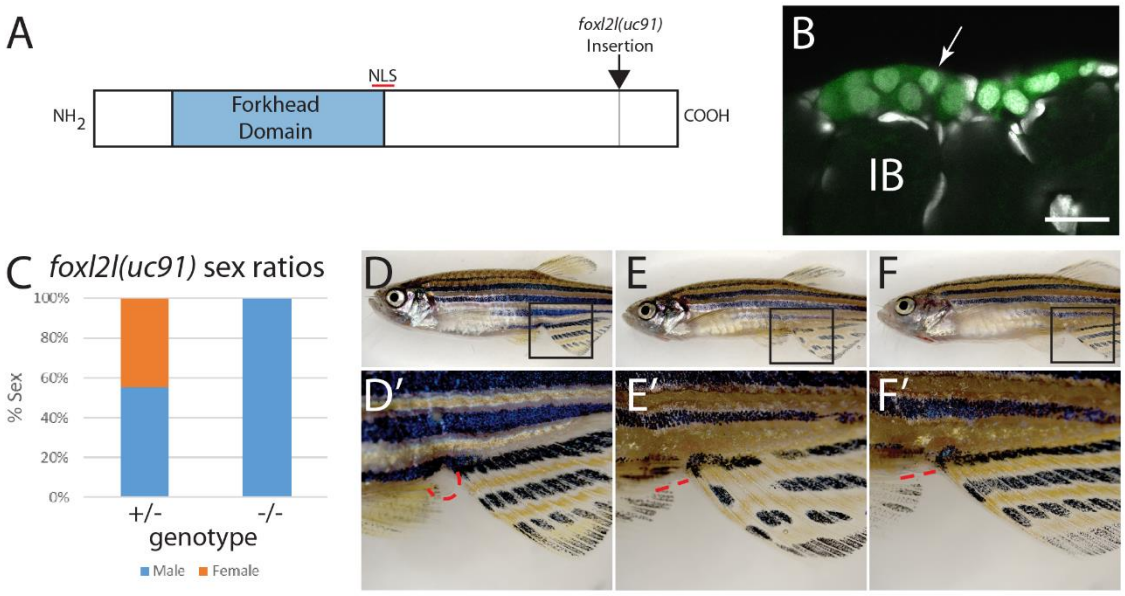


Figure 11. Mutational analysis of *foxl2l*. A. Schematic diagram of the Foxl2l protein showing the DNA binding Forkhead homology domain (blue), the location of the nuclear localization signal (NLS), and the *viral-2A-egfp* insertion site in the *foxl2l*(uc91) allele. B. GFP expression in germ cells from a *foxl2l*(uc91) knock-in allele heterozygote recapitulates endogenous *foxl2l* expression (compare to Figure 5E-F). C. Sex ratios of *foxl2l*(uc91) heterozygotes and homozygotes. D-F Representative light micrographs of fish examined in C. 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'). E. 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 IB oocyte.

2.8 References

Bahrami, N., & Childs, S. J. (2018). Pericyte Biology in Zebrafish. *Adv Exp Med Biol*, *1109*, 33–51.

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME suite. *Nucleic Acids Research*, *43*(W1), 39–49.

Bauer, M. P., Bridgham, J. T., Langenau, D. M., Johnson, A. L., & Goetz, F. W. (2000). Conservation of steroidogenic acute regulatory (StAR) protein structure and expression in vertebrates. *Molecular and Cellular Endocrinology*, *168*(1–2), 119–125.

Beer, R. L., & Draper, B. W. (2013a). Nanos3 maintains germline stem cells and expression of the conserved germline stem cell gene nanos2 in the zebrafish ovary. *Developmental Biology*, *374*(2), 308–318. https://doi.org/10.1016/j.ydbio.2012.12.003

Beer, R. L., & Draper, B. W. (2013b). Nanos3 maintains germline stem cells and expression of the conserved germline stem cell gene nanos2 in the zebrafish ovary. *Developmental Biology*, *374*(2), 308–318. https://doi.org/10.1016/j.ydbio.2012.12.003

Birk, O. S., Casiano, D. E., Wassif, C. A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J. A., Parker, K. L., Porter, F. D., & Westphal, H. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. *Nature*, *403*(6772), 909–913. https://doi.org/10.1038/35002622

Blokhina, Y. P., Nguyen, A. D., Draper, B. W., & Burgess, S. M. (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*(1), e1007730. https://doi.org/10.1371/journal.pgen.1007730

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M. J., & Rossant, J. (2006). Retinoid signaling determines germ cell fate in mice. *Science*, *312*, 596–600.

Brown, L. A., Rodaway, A. R., Schilling, T. F., Jowett, T., Ingham, P. W., Patient, R. K., & Sharrocks, A. D. (2000). Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Fli-1 in wild-type and mutant zebrafish embryos. *Mech Dev*, *90*, 237–252.

Brunet, J.-P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences*, *101*(12), 4164–4169. https://doi.org/10.1073/pnas.0308531101

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, *566*(7745), 496–502. https://doi.org/10.1038/s41586-019-0969-x

Cao, Z., Mao, X., & Luo, L. (2019). Germline Stem Cells Drive Ovary Regeneration in Zebrafish. *Cell Rep*, *26*, 1709-1717 1703.

Carmona, S. J., Teichmann, S. A., Ferreira, L., Macaulay, I. C., Stubbington, M. J., Cvejic, A., & Gfeller, D. (2017). Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. *Genome Res*, *27*, 451–461.

Chen, D., & McKearin, D. (2003). Dpp signaling silences bam transcription directly to establish asymmetric divisions of germline stem cells. *Current Biology*, *13*(20), 1786–1791.

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. https://doi.org/10.1016/j.ygcen.2013.08.016

Crespo, D., Assis, L. H. C., Kant, H. J. G., Waard, S., Safian, D., Lemos, M. S., Bogerd, J., & Schulz, R. W. (2019). Endocrine and local signaling interact to regulate spermatogenesis in zebrafish: Follicle-stimulating hormone, retinoic acid and androgens. *Development*, *146*.

Cui, S., Ross, A., Stallings, N., Parker, K. L., Capel, B., & Quaggin, S. E. (2004). Disrupted gonadogenesis and male-to-female sex reversal in Pod1 knockout mice. *Development*, *131*, 4095–4105.

Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M. R., Armean, I. M., Bennett, R., Bhai, J., Billis, K., Boddu, S., Cummins, C., Davidson, C., Dodiya, K. J., Gall, A., Girón, C. G., Gil, L., Grego, T., Haggerty, L., Haskell, E., … Flicek, P. (2019). Ensembl 2019. *Nucleic Acids Research*, *47*(D1), D745–D751. https://doi.org/10.1093/nar/gky1113

De Keuckelaere, E., Hulpiau, P., Saeys, Y., Berx, G., & Roy, F. (2018). Nanos genes and their role in development and beyond. *Cell Mol Life Sci*, *75*, 1929–1946.

Devlin, R. H., & Nagahama, Y. (2002). Sex determination and sex differentiation in fish: An overview of genetic, physiological, and environmental influences. *Aquaculture*, *208*, 191–364.

Doitsidou, M., Reichman-Fried, M., Stebler, J., Köprunner, M., Dörries, J., Meyer, D., Esguerra, C. V., Leung, T., & Raz, E. (2002). Guidance of Primordial Germ Cell Migration by the Chemokine SDF-1. *Cell*, *111*(5), 647–659. https://doi.org/10.1016/S0092-8674(02)01135-2

Dranow, D. B., Hu, K., Bird, A. M., Lawry, S. T., Adams, M. T., Sanchez, A., Amatruda, J. F., & Draper, B. W. (2016). Bmp15 Is an Oocyte-Produced Signal Required for Maintenance of the Adult Female Sexual Phenotype in Zebrafish. *PLOS Genetics*, *12*(9), e1006323. https://doi.org/10.1371/journal.pgen.1006323

Draper, B. W. (2012). Identification of Oocyte Progenitor Cells in the Zebrafish Ovary. In K. A. Mace & K. M. Braun (Eds.), *Progenitor Cells: Methods and Protocols* (pp. 157–165). Humana Press. https://doi.org/10.1007/978-1-61779-980-8_12

Draper, B. W., McCallum, C. M., & Moens, C. B. (2007a). Nanos1 Is Required To Maintain Oocyte Production in Adult Zebrafish. *Developmental Biology*, *305*(2), 589–598. https://doi.org/10.1016/j.ydbio.2007.03.007

Draper, B. W., McCallum, C. M., & Moens, C. B. (2007b). Nanos1 is required to maintain oocyte production in adult zebrafish. *Developmental Biology*, *305*(2), 589–598. https://doi.org/10.1016/j.ydbio.2007.03.007

Elkouby, Y. M., & Mullins, M. C. (2017). Methods for the analysis of early oogenesis in Zebrafish. *Developmental Biology*, *430*(2), 310–324. https://doi.org/10.1016/j.ydbio.2016.12.014

Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science (New York, N.Y.)*, *360*(6392), eaar3131. https://doi.org/10.1126/science.aar3131

Feng, J., Mantesso, A., & Sharpe, P. T. (2010). Perivascular cells as mesenchymal stem cells. *Expert Opin Biol Ther*, *10*, 1441–1451.

Forbes, A., & Lehmann, R. (1998a). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. *Development*, *125*(4), 679–690.

Forbes, A., & Lehmann, R. (1998b). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. *Development*, *125*, 679–690.

Gaengel, K., Genové, G., Armulik, A., & Betsholtz, C. (2009). Endothelial-mural cell signaling in vascular development and angiogenesis. *Arterioscler Thromb Vasc Biol*, *29*, 630–638.

Gautier, A., Sohm, F., Joly, J.-S., Le Gac, F., & Lareyre, J.-J. (2011). The Proximal Promoter Region of the Zebrafish gsdf Gene Is Sufficient to Mimic the Spatio-Temporal Expression Pattern of the Endogenous

Gene in Sertoli and Granulosa Cells. *Biology of Reproduction*, *85*(6), 1240–1251. https://doi.org/10.1095/biolreprod.111.091892

Gilchrist, R. B., Lane, M., & Thompson, J. G. (2008). Oocyte-secreted factors: Regulators of cumulus cell function and oocyte quality. *Hum Reprod Update*, *14*, 159–177.

Gore, A. V., Athans, B., Iben, J. R., Johnson, K., Russanova, V., Castranova, D., Pham, V. N., Butler, M. G., Williams-Simons, L., & Nichols, J. T. (2016). Epigenetic regulation of hematopoiesis by DNA methylation. *Elife*, *5*, 11813.

Guiguen, Y., Fostier, A., Piferrer, F., & Chang, C. F. (2010). Ovarian aromatase and estrogens: A pivotal role for gonadal sex differentiation and sex change in fish. *Gen Comp Endocrinol*, *165*, 352–366.

Haddon, C., Jiang, Y. J., Smithers, L., & Lewis, J. (1998). Delta-Notch signalling and the patterning of sensory cell differentiation in the zebrafish ear: Evidence from the mind bomb mutant. *Development*, *125*, 4637–4644.

Hsu, R. J., Lin, C. Y., Hoi, H. S., Zheng, S. K., Lin, C. C., & Tsai, H. J. (2010). Novel intronic microRNA represses zebrafish myf5 promoter activity through silencing dickkopf-3 gene. *Nucleic Acids Res*, *38*, 4384–4393.

Hu, Y. C., Okumura, L. M., & Page, D. C. (2013). Gata4 is required for formation of the genital ridge in mice. *PLoS Genet*, *9*, 1003629.

Kidder, G. M., & Vanderhyden, B. C. (2010). Bidirectional communication between oocytes and follicle cells: Ensuring oocyte developmental competence. *Can J Physiol Pharmacol*, *88*, 399–413.

Kikuchi, K., Holdway, J. E., Major, R. J., Blum, N., Dahn, R. D., Begemann, G., & Poss, K. D. (2011). Retinoic Acid Production by Endocardium and Epicardium Is an Injury Response Essential for Zebrafish Heart Regeneration. *Developmental Cell*, *20*(3), 397–404. https://doi.org/10.1016/j.devcel.2011.01.010

Kim, B., Kim, Y., Cooke, P. S., Rüther, U., & Jorgensen, J. S. (2011). The fused toes locus is essential for somatic-germ cell interactions that foster germ cell maturation in developing gonads in mice. *Biol Reprod*, *84*, 1024–1032.

Kimble, J., & Crittenden, S. L. (2007). Controls of germline stem cells, entry into meiosis, and the sperm/oocyte decision in Caenorhabditis elegans. *Annu. Rev. Cell Dev. Biol.*, *23*, 405–433.

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. *Journal of Cell Biology*, *149*(4), 875–888. https://doi.org/10.1083/jcb.149.4.875

Köprunner, M., Thisse, C., Thisse, B., & Raz, E. (2001). A zebrafish nanos-related gene is essential for the development of primordial germ cells. *Genes Dev*, *15*, 2877–2885.

Kossack, M. E., & Draper, B. W. (2019). Genetic regulation of sex determination and maintenance in zebrafish (Danio rerio. *Curr Top Dev Biol*, *134*, 119–149.

Koubova, J., Menke, D. B., Zhou, Q., Capel, B., Griswold, M. D., & Page, D. C. (2006). Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proc Natl Acad Sci U S A*, *103*, 2474–2479.

Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D., & Jaenisch, R. (1993). WT-1 is required for early kidney development. *Cell*, *74*, 679–691.

Kusz, K. M., Tomczyk, L., Sajek, M., Spik, A., Latos-Bielenska, A., Jedrzejczak, P., Pawelczyk, L., & Jaruzelska, J. (2009). The highly conserved NANOS2 protein: Testis-specific expression and significance for the human male reproduction. *Molecular Human Reproduction*. https://doi.org/10.1093/molehr/gap003

Kwok, H. F., So, W. K., 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. *Biol Reprod*, *72*, 1370–1381.

Lawson, N. D., & Weinstein, B. M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. *Developmental Biology*, *248*(2), 307–318.

Lee, J., Kanatsu-Shinohara, M., Inoue, K., Ogonuki, N., Miki, H., Toyokuni, S., Kimura, T., Nakano, T., Ogura, A., & Shinohara, T. (2007). Akt mediates self-renewal division of mouse spermatogonial stem cells. *Development (Cambridge, England)*, *134*(10), 1853–1859. https://doi.org/10.1242/dev.003004

Leerberg, D. M., Sano, K., & Draper, B. W. (2017). Fibroblast growth factor signaling is required for early somatic gonad development in zebrafish. *PLOS Genetics*, *13*(9), e1006993. https://doi.org/10.1371/journal.pgen.1006993

Leu, D. H., & Draper, B. W. (2010). The ziwi promoter drives germline-specific gene expression in zebrafish. *Developmental Dynamics*, *239*(10), 2714–2721. https://doi.org/10.1002/dvdy.22404

Li, T. Y., Colley, D., Barr, K. J., Yee, S. P., & Kidder, G. M. (2007). Rescue of oogenesis in Cx37-null mutant mice by oocyte-specific replacement with Cx43. *J Cell Sci*, *120*, 4117–4125.

Lieschke, G. J., Oates, A. C., Crowhurst, M. O., Ward, A. C., & Layton, J. E. (2001). Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. *Blood*, *98*, 3087–3096.

Luo, X., Ikeda, Y., & Parker, K. L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell*, *77*, 481–490.

Marlow, F. L., & Mullins, M. C. (2008). Bucky ball functions in Balbiani body assembly and animal–vegetal polarity in the oocyte and follicle cell layer in zebrafish. *Developmental Biology*, *321*(1), 40–50. https://doi.org/10.1016/j.ydbio.2008.05.557

McLeay, R. C., & Bailey, T. L. (2010). Motif Enrichment Analysis: A unified framework and an evaluation on ChIP data. *BMC Bioinformatics*, *11*(1), 1–11.

McMillan, D. B. (2007). *Fish histology: Female reproductive systems*. Springer Science & Business Media.

Meng, X., Lindahl, M., Hyvönen, M. E., Parvinen, M., Rooij, D. G. de, Hess, M. W., Raatikainen-Ahokas, A., Sainio, K., Rauvala, H., Lakso, M., Pichel, J. G., Westphal, H., Saarma, M., & Sariola, H. (2000). Regulation of Cell Fate Decision of Undifferentiated Spermatogonia by GDNF. *Science*, *287*(5457), 1489–1493. https://doi.org/10.1126/science.287.5457.1489

Mikalsen, S. O., S, Í. K., & Tausen, M. (2021). Connexins during 500 Million Years-From Cyclostomes to Mammals. *Int J Mol Sci*, *22*.

Miller, W. L., & Auchus, R. J. (2011). The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr Rev*, *32*, 81–151.

Mindnich, R., Deluca, D., & Adamski, J. (2004). Identification and characterization of 17 beta-hydroxysteroid dehydrogenases in the zebrafish, Danio rerio. *Mol Cell Endocrinol*, *215*, 19–30.

Mindnich, R., Haller, F., Halbach, F., Moeller, G., 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. *J Mol Endocrinol*, *35*, 305–316.

Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I., & Aizawa, S. (1997). Defects of urogenital development in mice lacking Emx2. *Development*, *124*, 1653–1664.

Morvan-Dubois, G., Le Guellec, D., Garrone, R., Zylberberg, L., & Bonnaud, L. (2003). Phylogenetic analysis of vertebrate fibrillar collagen locates the position of zebrafish alpha3(I) and suggests an evolutionary link between collagen alpha chains and hox clusters. *J Mol Evol*, *57*, 501–514.

Nakamura, S. (2010). *Identification of Germline Stem Cells*. *1561*(2004). https://doi.org/10.1126/science.1185473

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. *Dev Dyn*, *238*, 2652–2657.

Nishimura, T., Sato, T., Yamamoto, Y., Watakabe, I., Ohkawa, Y., Suyama, M., Kobayashi, S., & Tanaka, M. (2015). Foxl3 is a germ cell–intrinsic factor involved in sperm-egg fate decision in medaka. *Science*, *349*(6245), 328–331. https://doi.org/10.1126/science.aaa2657

Oatley, J. M., & Brinster, R. L. (2012). The germline stem cell niche unit in mammalian testes. *Physiological Reviews*, *92*(2), 577–595. https://doi.org/10.1152/physrev.00025.2011

Oatley, M. J., Kaucher, A. V., Racicot, K. E., & Oatley, J. M. (2011). Inhibitor of DNA Binding 4 Is Expressed Selectively by Single Spermatogonia in the Male Germline and Regulates the Self-Renewal of Spermatogonial Stem Cells in Mice. *Biology of Reproduction*, *85*(2), 347–356. https://doi.org/10.1095/biolreprod.111.091330

Onichtchouk, D., Aduroja, K., Belting, H.-G., Gnügge, 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 gonad differentiation in zebrafish. *Developmental Dynamics*, *228*(3), 393–404. https://doi.org/10.1002/dvdy.10392

Parajes, S., Griffin, A., Taylor, A. E., Rose, I. T., Miguel-Escalada, I., Hadzhiev, Y., Arlt, 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.

Park, K.-E., Kaucher, A. V., Powell, A., Waqas, M. S., Sandmaier, S. E. S., Oatley, M. J., Park, C.-H., Tibary, A., Donovan, D. M., Blomberg, L. A., Lillico, S. G., Whitelaw, C. B. A., Mileham, A., Telugu, B. P., & Oatley, J. M. (2017). Generation of germline ablated male pigs by CRISPR/Cas9 editing of the NANOS2 gene. *Scientific Reports*, *7*(December 2016), 40176. https://doi.org/10.1038/srep40176

Pepling, M. E., Cuevas, M., & Spradling, A. C. (1999). Germline cysts: A conserved phase of germ cell development? *Trends Cell Biol*, *9*, 257–262.

Prasasya, R. D., & Mayo, K. E. (2018). Notch Signaling Regulates Differentiation and Steroidogenesis in Female Mouse Ovarian Granulosa Cells. *Endocrinology*, *159*, 184–198.

Qin, M., Zhang, Z., Song, W., Wong, Q. W.-L., Chen, W., Shirgaonkar, N., & Ge, W. (2018). Roles of Figla/figla in Juvenile Ovary Development and Follicle Formation During Zebrafish Gonadogenesis. *Endocrinology*, *159*(11), 3699–3722. https://doi.org/10.1210/en.2018-00648

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*(W1), 191–198.

Rodriguez-Mari, A., Canestro, C., BreMiller, R. A., Catchen, J. M., Yan, Y. L., & Postlethwait, J. H. (2013). Retinoic acid metabolic genes, meiosis, and gonadal sex differentiation in zebrafish. *PLoS One*, *8*, 73951.

Rodriguez-Mari, A., Yan, Y. L., Bremiller, R. A., Wilson, C., Canestro, C., & Postlethwait, J. H. (2005). Characterization and expression pattern of zebrafish Anti-Mullerian hormone (Amh) relative to sox9a, sox9b, and cyp19a1a, during gonad development. *Gene Expr Patterns*, *5*, 655–667.

Ruzicka, L., Howe, D. G., Ramachandran, S., Toro, S., Van Slyke, C. E., Bradford, Y. M., Eagle, A., Fashena, D., Frazer, K., & Kalita, P. (2019). The Zebrafish Information Network: New support for non-coding genes, richer Gene Ontology annotations and the Alliance of Genome Resources. *Nucleic Acids Res*, *47*, 867–873.

RyanKJ. (1979). Granulosa-thecal cell interaction in ovarian steroidogenesis. *J. Steroid Biochem*, *11*, 799–800.

Sada, A., Suzuki, A., Suzuki, H., & Saga, Y. (2009). The RNA-binding protein NANOS2 is required to maintain murine spermatogonial stem cells. *Science*, *325*(5946), 1394–1398.

Selman, K., Wallace, R. A., Sarka, A., & Qi, X. (1993). Stages of oocyte development in the zebrafish, Brachydanio rerio. *Journal of Morphology*, *218*(2), 203–224. https://doi.org/10.1002/jmor.1052180209

Shen, Y. C., Shami, A. N., Moritz, L., Larose, H., Manske, G. L., Ma, Q., Zheng, X., Sukhwani, M., Czerwinski, M., & Sultan, C. (2021). TCF21(+) mesenchymal cells contribute to testis somatic cell development, homeostasis, and regeneration in mice. *Nat Commun*, *12*, 3876.

Siebert, S., Farrell, J. A., Cazet, J. F., Abeykoon, Y., Primack, A. S., Schnitzler, C. E., & Juliano, C. E. (2019). Stem cell differentiation trajectories in Hydra resolved at single-cell resolution. *Science*, *365*(6451). https://doi.org/10.1126/science.aav9314

Siegfried, K. R., & Draper, B. W. (2020). Chapter 16—The Reproductive System. In S. C. Cartner, J. S. Eisen, S. C. Farmer, K. J. Guillemin, M. L. Kent, & G. E. Sanders (Eds.), *The Zebrafish in Biomedical Research* (pp. 151–164). Academic Press.

Siu, M. K., & Cheng, C. Y. (2012). The blood-follicle barrier (BFB) in disease and in ovarian function. *Adv Exp Med Biol*, *763*, 186–192. https://doi.org/10.1007/978-1-4614-4711-5_9.

Söllner, C., Burghammer, M., Busch-Nentwich, E., Berger, J., Schwarz, H., Riekel, C., & Nicolson, T. (2003). Control of Crystal Size and Lattice Formation by Starmaker in Otolith Biomineralization. *Science*, *302*(5643), 282–286. https://doi.org/10.1126/science.1088443

Song, X., Wong, M. D., Kawase, E., Xi, R., Ding, B. C., McCarthy, J. J., & Xie, T. (2004a). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. *Development*, *131*(6), 1353–1364.

Song, X., Wong, M. D., Kawase, E., Xi, R., Ding, B. C., McCarthy, J. J., & Xie, T. (2004b). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. *Development*, *131*, 1353–1364.

Song, X., Zhu, C.-H., Doan, C., & Xie, T. (2002). Germline stem cells anchored by adherens junctions in the Drosophila ovary niches. *Science*, *296*(5574), 1855–1857.

Soyal, S. M., Amleh, A., & Dean, J. (2000). FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development (Cambridge, England)*, *127*(21), 4645–4654.

Su, Y.-Q., Sugiura, K., & Eppig, J. J. (2009). Mouse Oocyte Control of Granulosa Cell Development and Function: Paracrine Regulation of Cumulus Cell Metabolism. *Seminars in Reproductive Medicine*, *27*(1), 32–42. https://doi.org/10.1055/s-0028-1108008

Subramaniam, K., & Seydoux, G. (1999). Nos-1 and nos-2, two genes related to Drosophila nanos, regulate primordial germ cell development and survival in Caenorhabditis elegans. *Development (Cambridge, England)*, *126*(21), 4861–4871.

Suzuki, A., Saba, R., Miyoshi, K., Morita, Y., & Saga, Y. (2012). Interaction between NANOS2 and the CCR4-NOT Deadenylation Complex Is Essential for Male Germ Cell Development in Mouse. *PLoS ONE*, *7*(3), e33558. https://doi.org/10.1371/journal.pone.0033558

Suzuki, A., Saga, Y., Suzuki, A., & Saga, Y. (2008). *Nanos2 suppresses meiosis and promotes male germ cell differentiation service Nanos2 suppresses meiosis and promotes male germ cell differentiation*. 430–435. https://doi.org/10.1101/gad.1612708

Takahashi, H. (1977). Juvenile Hermaphroditism in the zebrafish, Brachydanio rerio. *Bull Fac Fish Hokkaido Univ*, *28*, 57–65.

Tang, Q., Iyer, S., Lobbardi, R., Moore, J. C., Chen, H., Lareau, C., Hebert, C., Shaw, M. L., Neftel, C., & Suva, M. L. (2017). Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. *J Exp Med*, *214*, 2875–2887.

Tourtellotte, W. G., Nagarajan, R., Auyeung, A., Mueller, C., & Milbrandt, J. (1999). Infertility associated with incomplete spermatogenic arrest and oligozoospermia in Egr4-deficient mice. *Development*, *126*(22), 5061–5071. https://doi.org/10.1242/dev.126.22.5061

Villanueva, R. A. M., & Chen, Z. J. (2019).

Wagner, M., Yoshihara, M., Douagi, I., Damdimopoulos, A., Panula, S., Petropoulos, S., Lu, H., Pettersson, K., Palm, K., & Katayama, S. (2020). Single-cell analysis of human ovarian cortex identifies distinct cell populations but no oogonial stem cells. *Nat Commun*, *11*, 1147.

Wang, Y., Pan, L., Moens, C. B., & Appel, B. (2014). Notch3 establishes brain vascular integrity by regulating pericyte number. *Development*, *Jan;141(2):307-17*.

Wang, Z., & Lin, H. (2004). Nanos maintains germline stem cell self-renewal by preventing differentiation. *Science (New York, N.Y.)*, *303*(5666), 2016–2019. https://doi.org/10.1126/science.1093983

Weidmann, C. A., Qiu, C., Arvola, R. M., Lou, T.-F., Killingsworth, J., Campbell, Z. T., Hall, T. M. T., & Goldstrohm, A. C. (2016). Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio. *ELife*, *5*, e17096.

Westerfield, M. (n.d.). *THE ZEBRAFISH BOOK, 5th Edition; A guide for the laboratory use of zebrafish (Danio rerio)*. University of Oregon Press.

Whitesell, T. R., Kennedy, R. M., Carter, A. D., Rollins, E.-L., Georgijevic, S., Santoro, M. M., & Childs, S. J. (2014). An α-Smooth Muscle Actin (acta2/αsma) Zebrafish Transgenic Line Marking Vascular Mural Cells and Visceral Smooth Muscle Cells. *PLOS ONE*, *9*(3), e90590. https://doi.org/10.1371/journal.pone.0090590

Wierson, W. A., Welker, J. M., Almeida, M. P., Mann, C. M., Webster, D. A., Torrie, M. E., Weiss, T. J., Kambakam, S., Vollbrecht, M. K., & Lan, M. (2020). Efficient targeted integration directed by short homology in zebrafish and mammalian cells. *Elife*, *9*.

Winterhager, E., & Kidder, G. M. (2015). Gap junction connexins in female reproductive organs: Implications for women's reproductive health. *Hum Reprod Update*, *21*, 340–352.

Xie, J., Wang, W. Q., Liu, T. X., Deng, M., & Ning, G. (2008). Spatio-temporal expression of chromogranin A during zebrafish embryogenesis. *J Endocrinol*, *198*, 451–458.

Xie, T. (2008). Germline stem cell niches. *StemBook*, 1–18. https://doi.org/10.3824/stembook.1.23.1

Xie, T., & Spradling, A. C. (2000). A Niche Maintaining Germ Line Stem Cells in the Drosophila Ovary. *Science*, *290*(5490), 328–330. https://doi.org/10.1126/science.290.5490.328

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*, 134. https://doi.org/10.1186/1471-2105-13-134

Yoshida, K., Kondoh, G., Matsuda, Y., Habu, T., Nishimune, Y., & Morita, T. (1998). The Mouse RecA-like Gene Dmc1 Is Required for Homologous Chromosome Synapsis during Meiosis. *Molecular Cell*, *1*(5), 707–718. https://doi.org/10.1016/S1097-2765(00)80070-2

Zakrzewska, A., Cui, C., Stockhammer, O. W., Benard, E. L., Spaink, H. P., & Meijer, A. H. (2010). Macrophage-specific gene functions in Spi1-directed innate immunity. *Blood*, *116*(3), e1–e11. https://doi.org/10.1182/blood-2010-01-262873

Chapter 3 scRNA-seq Transcriptome of Somatic Cells in the Zebrafish Ovary

3.1 Preface

This chapter included works from the following contributors:

Yulong Liu, Michelle E. Kossack, Matthew E. McFaul, Stefan Siebert, Samuel Horst, Nayeli Arroyo, Celina E. Juliano and Bruce W. Draper

3.2 Abstract

The zebrafish ovary is comprised of multiple somatic cell types that function to support the renewal and development of germ cells and to produce estrogen (E2) which coordinates female sex differentiation and maintenance. While the cell types are not well defined in zebrafish, there are three main gonadal somatic cell populations: follicle, theca, and stromal. The follicle cells and the theca cells form the follicle layers surrounding the oocytes, which is crucial for proper oocyte development. The inner follicle cell layer produces and transports signals and nutrients to the oocytes via gap junctions, and the outer steroidogenic theca layer produces the precursor for E2, which is then converted to E2 by the follicle cells. The stromal cells provide structure to the ovary. Here we performed a comprehensive analysis of the gonadal cell populations in the developing zebrafish ovary through single-cell RNA sequencing. We identified and validated multiple somatic cell subpopulations, including multiple follicle cell stages and two cell subpopulations that are candidate GSC niche cells.

3.3 Introduction

The zebrafish ovary is composed of an array of somatic gonad cell types. There are three primary female gonadal somatic cell populations: the follicle cells, theca cells, and stromal cells. The follicle cells and the theca cells form follicle cell layers that surround individual oocytes and are crucial for proper oocyte development. The follicle cells localize to the inner cell layer and are in direct contact with the oocyte. Follicle cells produce signals and nutrients, some of which are transported into the oocyte cytoplasm via gap junctions (Su et al., 2009). In mammals, follicle cells are also the major source of ovarian estradiol (E2) production. Surrounding the follicle cells is a layer of theca cells which are steroidogenic and produce the major precursor for E2 synthesis. Stromal cells refer to all other somatic cell types present in the ovary, including connective tissues, blood vessels and immune cells. The role of stromal cells in the ovary is not well characterized in vertebrates. Although these general cell populations have been observed in the zebrafish ovary, the full diversity of cell types present and the functional roles they play during ovarian development is not well understood.

In this study, we characterized the complex cell environment in the zebrafish ovary using single-cell RNA sequencing, allowing us to identify the somatic cells, including follicle, theca, and stromal cells. Using sub-cluster analysis we define sub-populations within these broad cell-type classifications and validate these subpopulations by determining where they reside in the ovary using smFISH. Our data provide strong support that orthologs of genes involved in mammalian ovary development and function likely play parallel roles in zebrafish. This reference data set, together with the analytical pipeline developed here, will serve as a resource to greatly enhance future studies of ovarian development and function in unprecedented detail.

3.4 Materials and Methods

See Chapter 2 Materials and Methods.

3.5 Results

Identification of three follicle cell subpopulations

Follicle cells are epithelial somatic gonad cells that encapsulate developing oocytes once they arrest at the diplotene phase 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 gene, such as *forkhead-box protein L2* (*foxl2a* and *foxl2b*), *anti-Mullerian hormone* (*amh*), *follicle stimulating hormone receptor* (*fshr*), *aromatase/cytochrome P450 19a1a* (*cyp19a1a*), and *notch receptor 3* (*notch3*) (Figure 1B and Figure 2) (B. Crespo et al., 2013; Dranow et al., 2016; Kwok et al., 2005; Prasasya & Mayo, 2018; Rodriguez-Mari et al., 2013). By contrast, unlike mammalian granulosa cells that surround the oocytes with multiple cell layers, zebrafish follicle cells surround the oocytes with only a single cell layer. Oocyte and follicle cell development is coordinated through bidirectional oocyte-follicle cell-cell interactions (Kidder & Vanderhyden, 2010). Soon after an oocyte has formed, it is surrounded by follicle cells and becomes dependent on these cells for compounds it cannot produce, such as pyruvate, cholesterol and select amino acids, which in some cases are transported to the oocyte via connexin-mediated gap junctions (Su et al., 2009). Mammalian granulosa cells are divided into two major cell types, the steroidogenic mural cells that form an outlet layer and produce estradiol, and cumulus cells that are in direct contact with the oocyte and provide nutrients through gap junctions (Gilchrist et al., 2008). By contrast, the single layered teleost follicle cells are likely the source of

estradiol production and select nutrients required for further oocyte development (Devlin & Nagahama, 2002). One major difference between mammals and teleost fish is that in mammals, new oocytes are only formed during embryogenesis, while in many teleost fish, including zebrafish, new oocytes are formed continuously from a population of GSCs, which are then surrounded by a layer a follicle cells produced by a presumed population of pre-follicle cells (Beer & Draper, 2013b; Z. Cao et al., 2019). Though follicle cells have been identified in zebrafish based on their expression of select genes, such as *gsdf*, little is known about stage-specific transcriptional changes that may occur in coordination with oocyte developmental progression. In addition, pre-follicle cells have not been identified in any teleost.

To further characterize early follicle cells in zebrafish, we performed sub-cluster analysis on the *gsdf*+ cells (Figure 1). As previously mentioned, the follicle cells in the data set likely represent pre-follicle cells and those derived from Stage IB or early Stage II follicles. Consistent with this, *cyp19a1a*, which is expressed exclusively in follicle cells surrounding mid-stage II and older follicles, is only detected in a very small subset of follicle cells in the dataset (Figure 1B and Figure 2). While the majority of genes expressed in the follicle cells are expressed ubiquitously throughout the cluster regardless of follicle stage (e.g. *gsdf*, *foxl2a/b* and *amh*) Figure 1A,B and Figure 2), our sub-cluster analysis distinguished six cell sub-clusters (0-5 in Fig. 3A) based on differences in gene expression (Figure 1A). However, upon more detailed inspection five of these six sub-clusters likely represent three developmentally relevant cell populations (e.g. one sub-cluster represents a mitotic subpopulation of otherwise developmentally similar cells; Figure 1A), while sub-cluster 5 may be an artifact of cell doublets and contains too few cells to warrant further consideration.

Stage IB follicle cells (Sub-cluster 1): Cells within sub-cluster 1 have enriched expression of the heparan sulfate proteoglycan-encoding gene, *glypican 1a* (*gpc1a*), *RERG/RAS-like b* (*rerglb*) and *Htra serine protease 1b* (*htra1b*) (Figures 1A-C). Using *gpc1a* as a representative sub-cluster 1 gene, we performed smFISH and determined that *gpc1a*+ cells are found only surrounding stage IB oocytes (Figure 1D) (Selman et al., 1993), and thus represent the primary follicle cells that first enclose oocytes to form the functional follicle. Based on the lack of expression of genes required for cell division, such as *proliferating cell nuclear antigen* (*pcna*) and *dihydrofolate reductase* (*dhfr*), these primary follicle cells are not highly proliferative (Figures 1B,C).

Stage II follicle cells (Sub-clusters 0, 2 and 4): Sub-cluster 0 contains the majority of follicle cells in our data set, which are characterized by the enriched expression of several genes, including gap junctional protein alpha 11 (*gja11*; formerly *cx43.5*; Figures 1A-C). To determine the location of this cell population within the ovary, we performed smFISH for *gja11*. These results show that follicle cells expressing *gja11* are restricted to Stage II follicles (Figure 1E). Gene expression analysis for sub-cluster 2 cells suggests that these cells are transitioning from Stage IA to Stage II follicle cells, as they express genes, such as *sidekick cell adhesion molecule 1a* (*sdk1a*), that are expressed in both populations (Figure 1B). Gene expression analysis suggests that cells within sub-cluster 4 are a mitotic subpopulation of sub-cluster 0 cells, as these cells have nearly identical gene expression as sub-cluster 0 cells but also express high levels of genes necessary for mitosis, such as *pcna*; Figures 1B,C). Assuming follicle cells maintain a constant size, to surround the surface of the oocyte, which increases 1,400-fold during oogenesis, follicle cells must increase in number from ~25 cells per early stage IB oocyte to ~1.5 million cells per mature oocyte (Stage IV). Thus, regulation of follicle cell proliferation is essential to maintain oocyte development. In mammals there is evidence that granulosa cell

proliferation is regulated by pituitary-produced follicle stimulating hormone (Fsh), and consistent with this, we found the gene encoding the Fsh receptor, called *fshr*, is expressed in zebrafish follicle cells including those in sub-cluster 4 (Figure 2). Finally, we detect only a small percentage of cells expressing *cyp19a1a* in sub-cluster 0, arguing that cells in this cluster are predominantly derived from early Stage II follicles, which do not express *cyp19a1a* (Figure 1B and Figure 2) (Dranow et al., 2016).

Pre-follicle cells (Sub-cluster 3): The sub-cluster 3 cells are part of the follicle cell lineage as they express *gsdf* (Figures 1A-C), yet these cells also express many genes that are unique to this subpopulation, suggesting that these cells are distinct from the Stage IB and II follicle cells identified above. Interestingly, these cells express orthologs of the core genes associated with early undifferentiated somatic gonad cell in mammals, including the *transcription factors LIM homeobox 9* (*lhx9*), *Wilms tumor 1a* (*wt1a*), *empty spiracles homeobox 2* (*emx2*), *nuclear receptor 5a1b* (*nr5a1b*; formerly *sf1b*; (Figure 1B, C and Figure 3C,D) (Birk et al., 2000; Kreidberg et al., 1993; Luo et al., 1994; Miyamoto et al., 1997). Notably, sub-cluster 3 cells also uniquely express *iroquois homeobox 3a* and *5a* (*irx3a* and *irx5a*) whose orthologs in mammals are required for pre-granulosa cell development in the embryonic mouse ovary (Figure 1B, C and Figure 3C,D) (Kim et al., 2011). Interestingly, sub-cluster 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 suggests that these cells are not developmentally equivalent to the somatic gonad precursors present in the bipotential gonad. Finally, sub-cluster 3 cells also express lower levels of genes associated with follicle cell differentiation and function, such as *foxl2a/b*, *amh*, *fshr* and *notch3* (Figure 1B, C and Figure 2). Together these results strongly suggest that sub-cluster 3 cells are pre-follicle cells.

Many of the genes with enriched expression in sub-cluster 3 did not appear uniformly expressed in all cells within this sub-cluster. For example, *lhx9* have higher expression in cells distal to the sub-cluster 1 in the UMAP while *gsdf* and *bmp6* are higher in cells proximal to sub-cluster 1 (Fig. 3A, C). To explore this further, we performed sub-cluster analysis using only those cells derived from follicle cell sub-clusters 1 and 3 (Figure 3A,B). This analysis showed that sub-cluster 3 cells can be further partitioned into two distinct sub-clusters which we have designated sub-clusters 3.1 and 3.2 (Figure 3B). While many genes, such as *emx2*, *irx3a* and *irx5a* are expressed uniformly in sub-clusters 3.1 and 3.2, this analysis confirmed that *lhx9* and *wt1a* had enriched expression in sub-cluster 3.1 while *gsdf* and *bmp6* appeared enriched in sub-cluster 3.2 (Figure 3C,D). Another intriguing difference between these two cell populations is that the *bmp6*-expressing cells have enriched expression of other genes that encode cell signaling components, such as *fibroblast growth factor receptors 3* and *4* (*fgfr3* and *fgfr4*) and Fgf-responsive genes *etv4* and *spry4*, while the *lhx9*-expressing cells have enriched expression of genes that encode cell signaling attenuators, such as the Bmp signaling antagonists *noggin1* and *3* (*nog1* and *nog3*), the Wnt signaling antagonist *dikkopf 1b* (*dkk1b*), and the retinoic acid degrading enzyme encoded by *cytochrome P450 26a1* (*cyp26a1*). Retinoic acid is known to promote entry into meiosis in mammals (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.

To determine where these cells reside in the ovary, we performed smFISH, 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 tracts of cells that are located on the surface of the ovary (arrows in Figures 1F,F',F"). We next asked if these cells were associated with

pre-follicle stage germ cells, which would include pre-meiotic and/or early meiotic germ cells. We performed triple smFISH for *lhx9*, GSC-expressed *nanos2* and early meiosis-expressed *dmc1*. Remarkably, we found that these early-stage germ cells 100% colocalized with *lhx9*+ pre-follicle cells (n=31 clusters of *nanos2* or *dmc1* in 4 biological replicates) (Figures 1F,F',F"). We next determined the location of *bmp6* expressing cells together with *nanos2* and *dmc1*. Similar to *lhx9*, *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 (Figures 1G,G'). Based on the lateral location we hypothesized that this region will eventually form the GSC niche (Beer & Draper, 2013b; Draper et al., 2007b).

In mammals, retinoic acid signaling is required to induce germ cell entry into meiosis (Bowles et al., 2006; Koubova et al., 2006). Previous studies show 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 4A) (Rodriguez-Mari et al., 2013). Given that the retinoic acid degrading enzyme *Cyp26a1* is produced by the *lhx9*-expressing pre-follicle cells (Figure 3C,D and Figure 4B), and the retinoic acid receptor-encoding genes *rxrba* and *rxrbb* are expressed in premeiotic germ cells (Figure 4C), we speculate that the close association of premeiotic germ cells with these cells is important for regulating germline stem cell maintenance vs. differentiation.

Identification of connexin-encoding genes that likely function in oocyte:follicle cell gap junction formation

In mammals, granulosa cells are connected to oocytes through gap junctions formed by the extracellular interactions of oocyte and granulosa cell-expressed connexin proteins, which serve as a conduit for the transfer of essential nutrients that oocytes stop producing soon after reaching dictyate arrest (Su et al., 2009). Oocyte-to-follicle cell signaling instructs follicle cells to upregulate the enzymatic pathways necessary to produce these nutrients. For example, oocyte-produced Bone morphogenetic protein 15 (Bmp15) and Growth differentiation factor 9 (Gdf9) instruct follicle cells to increase expression of the glycolytic enzymes necessary to produce pyruvate (Su et al., 2009). Like mammals, the genes encoding Bmp15 and Gdf9 are expressed in zebrafish oocytes, and mutational analysis has shown that *bmp15* mutants have defects in maintaining oocyte development, suggesting that this pathway plays a conserved role in regulating follicle cell-oocyte communication across vertebrates (Dranow et al., 2016). To gain more insight into this, we asked if connexin protein-encoding genes are expressed in oocyte and follicle cells, as this would suggest that like mammalian granulosa cells, zebrafish follicle cells are a source for oocyte-required nutrients.

Gap junctions are formed by connexin proteins, which are four-pass integral membrane proteins. Six connexin proteins oligomerize to form a hemi-channel complex called a connexon, which then form end-to-end complexes with connexons on adjacent cells to create the functional intercellular channel (Winterhager & Kidder, 2015). In mammals, the oocyte-expressed connexin encoded by Gap junction protein A4 (Gja4; formerly Cx37) likely forms gap junction complexes with granulosa cell-expressed Gja4 (Li et al., 2007). In addition, follicle cell-expressed Gja1 (formerly CX43) forms gap junction complexes between adjacent follicle cells (Li et al., 2007). The zebrafish genome

contains two genes encoding ohnologs of Gja1, called *gja1a* and *gja1b*, but our data indicates only *gja1b* is expressed at significant levels in follicle cells (Figure 5). By contrast, neither zebrafish oocytes or follicle cells express the ortholog of gja4 (formerly cx39.4). Instead, we found that oocytes express the ohnologs *gjc4a* and *gjc4b* (formerly *cx44.2* and *cx43.4*), orthologs of which are found in most vertebrates except mammals (Mikalsen et al., 2021). While follicle cells express significant levels of *gja11* (formerly *cx34.5*), orthologs of which appear to have been lost at the base of the lobe-finned lineage that includes mammals (Mikalsen et al., 2021).Thus, though Gja1 orthologs may play conserved roles in forming gap junctions between neighboring follicle cells in both mammals and fish, the connexins that are likely used to link follicle cells to oocytes do not appear to be conserved (Figure 5C). Further analysis will be required to determine which connexin(s) is required for gap junction formation between zebrafish oocytes and follicle cells.

Identification of five stromal cell subpopulations

Ovarian stromal cells are broadly defined as all ovarian cell types other than germ, follicle, or theca cells, and are the least characterized cell types in the ovary. Known stromal cell types include the interstitial cells that form collagen-producing connective tissue (fibroblast) that provides structural integrity to the ovary, the surface epithelial cells that surround the ovary, the vasculature and associated perivasculature cells, and the resident blood cell types. 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* (Figure 1A) and reasoned that these were likely the stromal cell population. This is further supported by the expression

of *decorin* (*dcn*) in these cells, which encodes a proteoglycan component of the extracellular matrix and is a gene expressed in ovarian stromal cells in mammals (Figure 6A) (Wagner et al., 2020). To further characterize this cell population, we performed sub-cluster analysis and identified five probable cell subpopulations, which are described below (Figure 6A). Of the somatic ovarian cells, the stromal cell cluster is the most transcriptionally diverse. This is also supported by the Gene Ontology (GO) analysis. Compared to the follicle cells, the stromal cell sub-clusters have minimal shared top GO terms between them (Figure 7).

Interstitial cells (Cluster 0 and 4): Stromal cell sub-cluster 0 is the largest sub-cluster of the stromal cell population and likely represents the collagen-producing fibroblast-like connective tissue, or interstitial cells, that provide structural integrity to the ovary (Figure 6B). In addition to the expression of the collagen, *col1a1a* (Figure 1A), notable genes expressed in these cells include the *chemokine (C-X-C motif) ligand 12a* (*cxcl12a*, formerly known as *sdf1a*), the basement membrane component *nidogen 1b* (*nid1b*) and the cell signaling molecule bmp4 (Figure 6A,B). Cells within sub-clusters 0 and 4 have nearly identical gene expression profiles, with the exception that sub-cluster 4 cells also have strong expression of genes associated with cell cycle progression, such as *pcna* (Figure 6A,B). Given that the ovary continues to increase in size as the fish grows, sub-cluster 4 cells may serve as a source for new interstitial cells similar to follicle cell sub-cluster 4 cells described above. Previous studies have shown that *cxcl12a* is essential for attracting migrating primordial germ cells to the embryonic gonad (Doitsidou et al., 2002). However, we did not detect expression of the *cxcl12a* receptor, *cxcr4a*, in any cells of the 40 dpf ovary data set, so the relevant receptor or functional significance of *cxcl12a* expression at this stage is not apparent. Regardless, this gene provides a convenient marker for determining the location of this cell population in the ovary. Using fluorescent

in situ hybridization, 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 sub-clusters 4 and 0 are proliferative and non-proliferative interstitial cells, respectively.

Ovarian Surface Epithelium (sub-cluster 5): Stromal cell sub-cluster 5 has enriched expression of *fgf24*, a gene that is expressed in the outer epithelial layer of the early bi-potential gonad (10-12 dpf), and that is required for early gonad development in zebrafish (Leerberg et al., 2017). In addition to *fgf24*, these cells also express signaling ligand-encoding genes *fgf18b* and *wnt7aa*, *transcription factors paired box 2b* (*pax2b*) and *empty spiracles homeobox 2* (*emx2*), and cell adhesion proteins encoded by *cadherin 6* (*cdh6*) and *claudin 3d* (*cldn3d*; formerly called *cldnc*) and *starmaker* (*stm*; Figure 6A,B and Figure 8). Given the expression of *fgf24* in surface epithelial cells in the early gonad (Leerberg et al., 2017) and the role of claudin proteins in forming tight junctions between epithelial cells, we propose that sub-cluster 5 cells correspond to the ovarian surface epithelium. Previous studies of *stm* have shown it has an important role in inner ear development, specifically in the formation of the mineralized otoliths (Söllner et al., 2003). However, its function in the ovary is not known. Using smFISH, we found *stm* is specifically expressed at the medial and lateral edges of the 40dpf ovary, where there is an enrichment of *nanos2*+ GSCs and *dmc1*+ meiotic GCs (Figure 6D-D'), a region that likely gives rise to the germinal zone of the adult ovary (Beer & Draper, 2013b; Draper et al., 2007b). To examine this in more detail we performed smFISH for *stm* together with hematoxylin and eosin (H&E) histology on 3-month-old adult females. We found *stm* expressing cells formed the membrane that lines the dorsal surface of the ovarian cavity. Interestingly, this epithelium is connected to the medial and lateral edges of the ovary, a site that precisely correlates with the location of early stage germ cells (Figure 8B, C) (Takahashi, 1977). Histological sections confirmed that premeiotic and early meiotic

germ cells were found precisely where the ovarian epithelium attaches to the ovary at the lateral and medial sides (Figure 8D,E and 8F,G, respectively). In the adult ovary, early-stage germ cells, including GSC and early meiotic germ cells have previously been shown to localize to a discrete zone on the lateral and medial surfaces of the ovary termed the germinal zone. These data suggest that the location of the germinal zone correlates to where the ovarian epithelium that forms the ovarian cavity is attached to the ovary. All together these data suggest that *stm*+ stromal cells form the ovarian cavity epithelium and may also play a role in germline stem cell niche function.

Perivascular Mural cells: Vascular smooth muscle (sub-cluster 1) and pericytes (sub-cluster 2): sub-clusters 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 establishment of the blood-brain barrier, these cells are implicated in formation of a blood-follicle barrier (Siu & Cheng, 2012). Pericytes are generally solitary cells associated with small diameter blood vessels, (arterioles, capillaries, and venules) while vascular smooth muscle cells (vSMC) 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 ß* (*pdgfrb*) and *notch3*, two genes whose expression are enriched in our stromal cell sub-cluster 2 (Figure 6A and Figure 9) (Y. Wang et al., 2014). In addition, we find that sub-cluster 2 cells also have enriched expression of *proteolipid protein 1b* (*plp1b*; Figure 6A,B), *T-box transcription factor 2* (*tbx2*), *melanoma cell adhesion molecule B* (*mcamb*), and *regulator of G protein signaling 5a and 5b* (*rgs5a* and *rgs5b*), orthologs of which are expressed in human ovarian pericytes (Figure 9A) (Wagner et al., 2020). vSMC are characterized by expression of *α-smooth muscle actin* (*acta2*) and *transgelin* (*tagln*, formally known as *SM22α*) (Bahrami & Childs, 2018). We found that both *acta2* and *tagln* are enriched in stromal cell sub-cluster 1 (Figure 6A,B and Figure 9B). In addition,

we found that sub-cluster 1 also has enriched expression of *myosin light chain kinase b* (*mylkb*), *tropomyosin 2* (*tpm2*) and *cysteine and glycine-rich protein 1b* (*csrp1b*), further supporting the conclusion that these cells are vascular smooth muscle (Figure 9B).

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 smFISH for *plp1b* using ovaries isolated from *Tg(fli1:egfp)* which express eGFP in vascular endothelial cells (Lawson & Weinstein, 2002). Consistent with *plp1b* labeling pericytes, *plp1b*+ cells were generally solitary, but always adjacent to blood vessels (Figure 6E-E'). To determine the localization of vSMC in the zebrafish ovary, we imaged ovaries from *Tg(kdrl:dsRed)*; *Tg(acta2:eGFP)* double transgenic animals. *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 6F-F'). This evidence supports the conclusion that cells in SC cluster 1 are vascular smooth muscle cells. Together our data argue strongly that sub-cluster 1 and 2 represent pericytes and vSMC, respectively. Interestingly, it has been proposed that pericytes may also function as mesenchymal stem cells capable of repairing stroma to maintain organ integrity (Feng et al., 2010).

Stromal progenitor cells (sub-cluster 3): sub-cluster 3 cells, in addition to expressing high levels of the general stromal genes *col1a1a* and *dcn* (Figure 6A), also express many genes that are developmentally relevant (Figure 6A,B and Figure 10B,C). For example, *transcription factor 21* (*tcf21*) is expressed specifically in stromal cell sub-cluster 3 cells, and in mice Tcf21-expressing cells were identified through lineage labeling as multipotent mesenchymal/stromal progenitor cells that can

produce all somatic cell types in the mouse ovary and testis during development, and in the adult testis could contribute to somatic cell turnover during aging or cell replacement following injury (Cui et al., 2004; Shen et al., 2021). Interestingly, sub-cluster 3 cells and mouse Tcf21-expressing mesenchymal progenitor cells share common expression of the transcription factor encoded by *odd-skipped related 1* (*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 10B, C). We therefore propose that these cells are stromal progenitor cells.

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 & Nagahama, 2002). In mammals, the production of E2 requires two cell types: theca and granulosa cells (RyanKJ, 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 follicle cells for conversion to E2 by two additional enzymes, most notably the Cyp19a1 aromatase (Miller & 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 lead to E2 production. For example, while orthologs of the enzymes necessary for E2 production in mammals have been identified, the specific cell type(s) that produce E2 in zebrafish have not been determined, nor have theca cells been definitively identified.

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. The expression of both *star* and *cyp11a1* have been detected in zebrafish ovaries (Bauer et al., 2000). Unlike other vertebrates, teleost fish 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 scRNA-seq data set demonstrates that *cyp11a1* and *cyp11a2* are expressed in a distinct population of somatic cells, thus identifying these as theca cells (Figure 1 and Figure 12). Additionally, we found that *cyp11a2* is expressed in theca cells at a higher level than *cyp11a1* and therefore possibly encodes the primary cholesterol side chain-cleaving enzyme required for E2 synthesis in the ovary (Figure 12). By contrast, the majority of *cyp11a1* expression in the dataset is localized to oocytes (Stage IA; Figure 12), consistent with previous results showing that it is a maternally loaded transcript present in early embryos (Parajes et al., 2013). Finally, expression of the three remaining genes encoding enzyme orthologs required for androstenedione production by theca cells, star, *cyp17a1* and hydroxy-delta-5-steroid dehydrogenase, *3 beta- and steroid delta-isomerase 1* (*hsd3b1*), are also detected in theca cells (Figure 12). As the location of theca cells in the zebrafish ovary has not been previously determined, we performed smFISH using *cyp17a1* as a marker. As expected, we found that *cyp17a1*-expressing cells localize to the interstitial spaces between developing oocytes (Figure 11B,C). Thus, we have identified 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 suggests that theca cells are not a homogeneous cell population (Figure 11A). Following sub-clustering, we found several genes that had non-uniform expression within theca cells in zebrafish, leading to the identification of four theca cell sub-populations (Figure 11A). Examples of genes expressed non-uniformly in theca cells include star (sub-cluster 1), *cyp19a1a* and *nidogen 1b* (*nid1b*), which encodes a basement membrane-associated protein (sub-cluster 3), *v-ets avian erythroblastosis virus E26 oncogene homolog 2* (*ets2*; sub-clusters 0 and 1), and *glutamate ammonia ligase A* (*glula*), which encodes an enzyme involved in glutamine synthesis (sub-cluster 0; Figure 11A). Non-uniformity 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.

Identification of the probable pathway for estrogen production in zebrafish ovary

There are two possible biosynthetic pathways for the conversion of androstenedione to E2 (Figure 13A). In the first pathway, testosterone is the intermediate while in the second it is estrone (E1). Pathway choice is dependent on the expression of the enzymes 17ß-hydroxysteroid dehydrogenase types 1 and 3 (Hsd17b1 and Hsd17b3). While these enzymes catalyze similar reactions, in zebrafish Hsd17b3 prefers androstenedione as a substrate while Hsd17b1 prefers estrone (Figure 13A, Figure 14) (Mindnich et al., 2004, 2005). By contrast, current evidence argues that the aromatase Cyp19a1a does not have preference for androstenedione vs. testosterone (Figure 13A) (Guiguen et al., 2010). 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 fish (Devlin & Nagahama, 2002), the pathway by which E2 is produced in the zebrafish ovary remains to be verified. Thus, knowing which genes are expressed, and in what cells is key to understanding how E2 production is regulated in the zebrafish ovary.

We previously reported that *cyp19a1a* was likely expressed in theca cells in the early ovary and then was upregulated in follicle cells surrounding oocytes that had progressed past mid-stage II, a stage that is not present in 40dpf ovaries (Dranow et al., 2016). Our scRNA-seq data verifies that *cyp19a1a* is expressed in theca cells (Figure 13B) but not in early follicle cells (Figure 13B, see also Figure 1B and Figure 5). Unexpectedly, however, we found significant expression of *cyp19a1a* in interstitial cells of the stroma. We therefore propose that three distinct ovarian cell populations contribute to E2 production in zebrafish (Figure 13B,C). Finally, we could not detect expression of *hsd17b3* in any cells in the data set, but instead found that *hsd17b1* was expressed at high levels specifically in the follicle cells, indicating that the second E2 pathway is the major pathway for E2 synthesis in the zebrafish ovary. This data also argues that Hsd17b1, not Cyp19a1a, catalyzes the final enzymatic step in E2 synthesis, and consistent with mammals indicates that follicle cells are the major source of E2 production in the zebrafish ovary. Thus, future studies using gene expression as a proxy for determining E2 production in zebrafish should include *hsd17b1* in addition to *cyp19a1a*. These results demonstrate how analysis of gene expression at single-cell resolution can profoundly affect our understanding of steroid synthesis in the zebrafish ovary.

3.6 Discussion

Among the somatic cell populations, we identified three populations that warrant further analysis. First, we found a subpopulation of follicle cells, marked by expression of *lhx9*, that we propose are pre-follicle cells. Unlike mammals, zebrafish need to maintain a population of pre-follicle cells to serve as a source of new follicle cells to enclose oocytes that are continually produced in the adult ovary. However, pre-follicle cells had not been identified. We also found that these cells co-localization with pre-meiotic and early meiotic germ cells, and therefore may play a role in regulating the balance between germline stem cell maintenance and differentiation. Alternatively, close association between pre-follicle cells and early meiotic oocytes may be important for formation of the oocyte-follicle cell complex.

The second interesting cell type we identified is the stromal cell subpopulation 5, which expresses *stm* and that appear to form the ovarian epithelium (OE). Teleost ovaries can be categorized as two general types- gymnovarian and cystovarian. Gymnovarian ovaries ovulate mature eggs directly into the coelomic cavity whereas cystovarian ovaries ovulate eggs into an ovarian cavity that forms between the dorsal surface of the ovary and an overlying OE. The OE connects to the reproductive ducts at the posterior end of the ovary, through which eggs are spawned during mating (Kossack & Draper, 2019). The presence or absence of an OE is therefore a defining feature of ovarian type, yet previously this cell type had only been identified based on histological criteria. Zebrafish ovaries form an ovarian cavity and thus are cystovarian-type (Takahashi, 1977) and our analysis has identified stromal cell subpopulation 5 as representing the probable ovarian epithelial cells. Interestingly, we found that the OE attaches to the lateral and medial sides of the ovary at the precise point where we have previously proposed the germline stem cell niche, referred to as the germinal zone, develops (Beer & Draper, 2013b; Draper et al., 2007b). It is therefore an exciting possibility that this asymmetry

on the surface of the ovary may play an important role, either directly or indirectly, in the formation and/or function of the ovarian germline stem cell niche in zebrafish.

A second subpopulation of stromal cells that warrant further investigation are those contained within subpopulation 4 that express *tcf21*. In the mouse ovary and testis, Tcf21-expressing cells have been shown, through lineage trace analysis, to be capable of producing all known somatic cell types of the ovary and testis during development or in the adult following wounding, and therefore may have stem cell-like properties (Shen et al., 2021). We do not yet know if these cells are a stable population that persist in the adult ovary. It will also be important to determine when these cells are first detected during gonad development and if they are a precursor population of multiple somatic gonadal cell types, as they appear to be in the mouse.

One of the most exciting aspects of this data set is that it can be used to quickly and accurately determine if a gene is expressed in the ovary and in what cell type(s). As an example, we determine the expression patterns of the genes that encode the enzymes necessary for sex hormone synthesis. As expected, all the enzymes required for the conversion of cholesterol to the androgen androstenedione are expressed in theca cells. We also confirmed that the *cyp19a1a*, which encodes aromatase, is not expressed in early-stage follicle cells, as it is in mammals, but is instead expressed in a subset of theca cells in the juvenile ovary. Unexpectedly, we also detected low levels of *cyp19a1a* in the interstitial stromal cells, though we have not confirmed this by RNA in situ hybridization. The most significant result from this analysis was the discovery that expression of *hsd17b3* was not detected in any somatic ovarian cell type, while *hsd17b1* appears to be expressed abundantly in follicle cells. Given that Hsd17b1 has a strong preference for using estrone as a substrate, we conclude that the pathway for E2

synthesis is as follows: androstenedione is converted to estrone by Cyp19a1a and then Hsd17b1 converts estrone to E2. Thus, testosterone is not produced by the ovary and Hsd17b1, not Cyp19a1a, is the final enzyme in the pathway. Our conclusions are therefore in stark contrast to those of previous studies that proposed E2 was produced from a testosterone intermediate. These conclusions were based on RT-PCR studies showing *hsd17b3* expression in the ovary. Previous studies showing that *hsd17b3* is expressed in ovaries used RT-PCR, and base on this evidence concluded that the zebrafish ovary produces testosterone which is the converted to E2 by Cyp19a1a. However, mutational analysis will be necessary to confirm this is the correct mechanism of E2 synthesis.

3.7 Figures

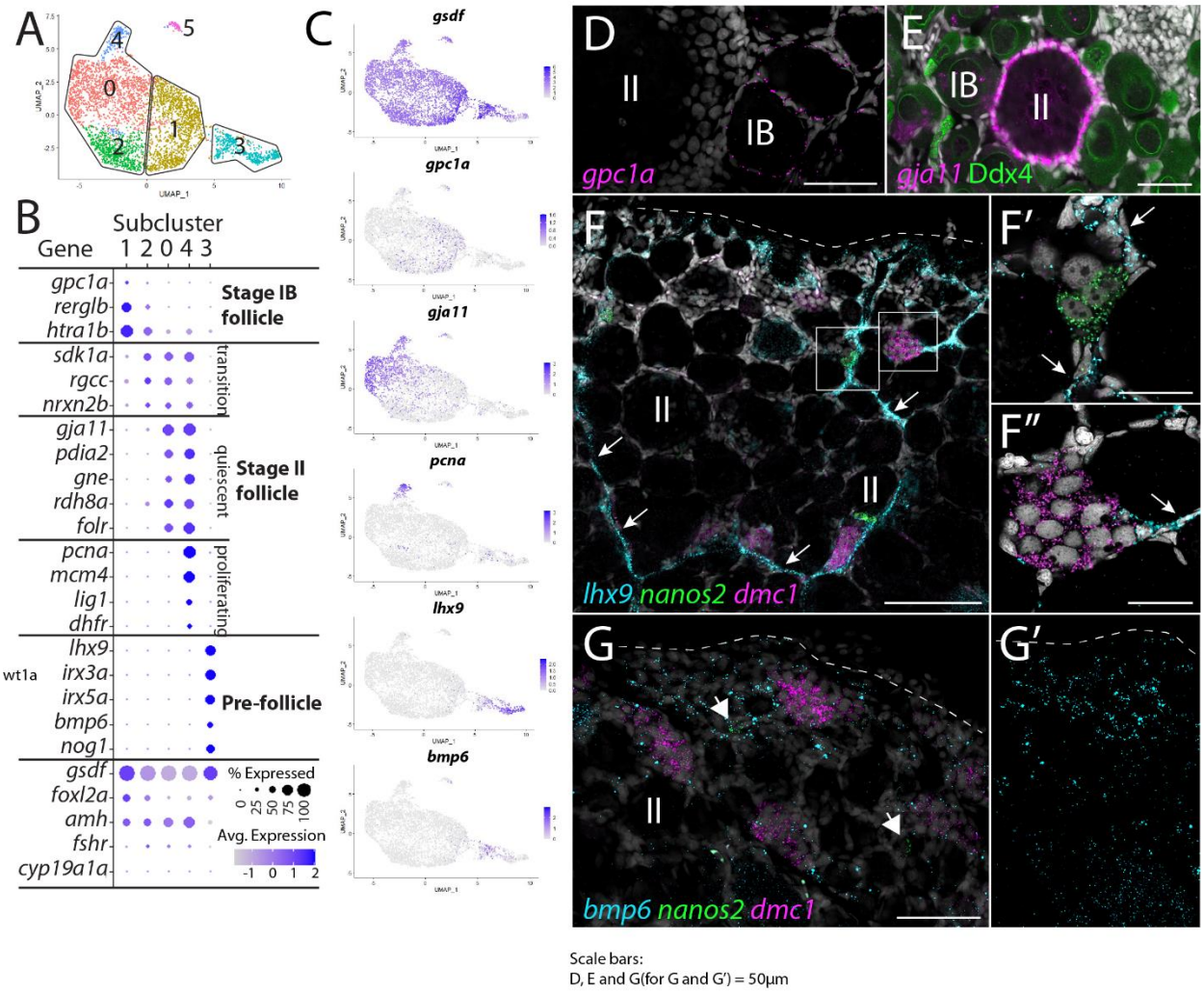


Figure 1. Follicle cell sub-cluster analysis reveals three main cell subtypes. **A.** Follicle cell sub-cluster UMAP plot, with cells color-coded by computationally determined cell subtypes. The three main subtypes are outlined. **B.** Dot-plot showing the relative expression of select genes in the follicle cell sub-clusters. 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 genes. 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). **D-G.**

smFISH on whole-mount 40 dpf ovaries reveals the location of cell subtypes. In all panels, DNA is gray. **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 immunofluorescences (green) labels all germ cells. **F.** Triple smFISH shows *lhx9* expressing cells (blue) form tracts on the surface of the ovary (arrows) that colocalize with *nanos2* (green) and *dmc1* (pink) expressing germline stem cells and early meiotic cells, respectively. Lateral edge of ovary is indicated with dashed line. **F'** and **F"** Higher magnification views of regions boxed in **F**. **G.** Triple smFISH shows *bmp6* expressing cells (blue) are concentrated near the lateral edge of the ovary, a region that contains *nanos2* (green) and *dmc1* (pink) expressing germline stem cells and early meiotic cells, respectively. **G'**. *bmp6* channel only. Scale bars in D, E and G9 for G and G'), 50 µm; F, 100µm; F' and F", 20µm. IB, stage IB oocyte; II, stage II oocyte.

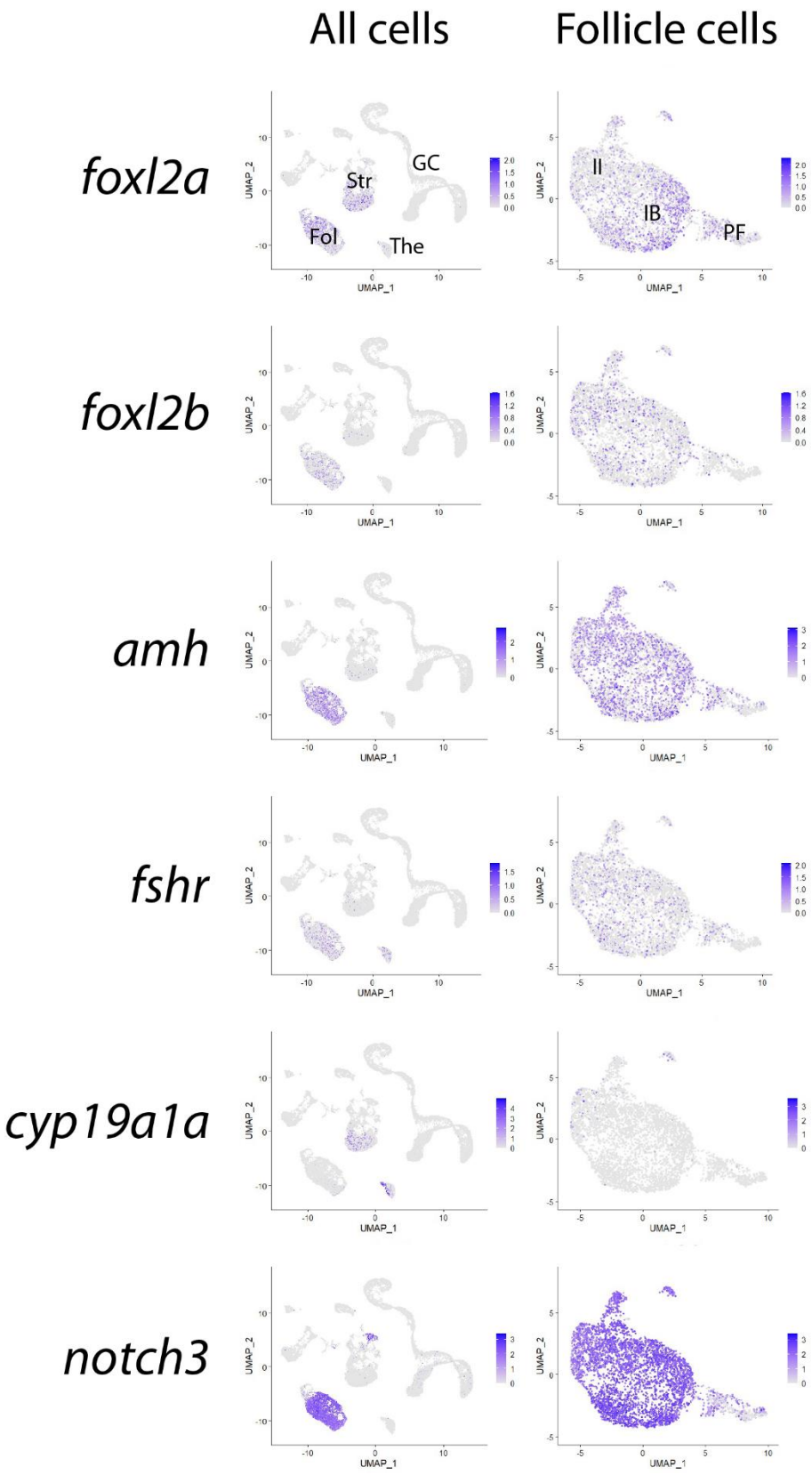

All cells
Follicle cells
foxl2a
foxl2b
amh
fshr
cyp19a1a
notch3
Str
GC
Fol
The
II
IB
PF
UMAP_1
UMAP_2

Figure 2. Gene expression UMAP plots of select follicle cell-enriched genes. Cells expressing the indicated gene in all cells (left column) or in follicle cell subcluster (right column) are colored blue, and the relative intensity indicates relative expression levels (intensity scale for each plot is on the right). Fol, follicle cells; Str, stromal cells; The, theca cells; GC, germ cells; IB, stage IB follicle cells; II, stage II follicle cells, PF, pre-follicle cells.

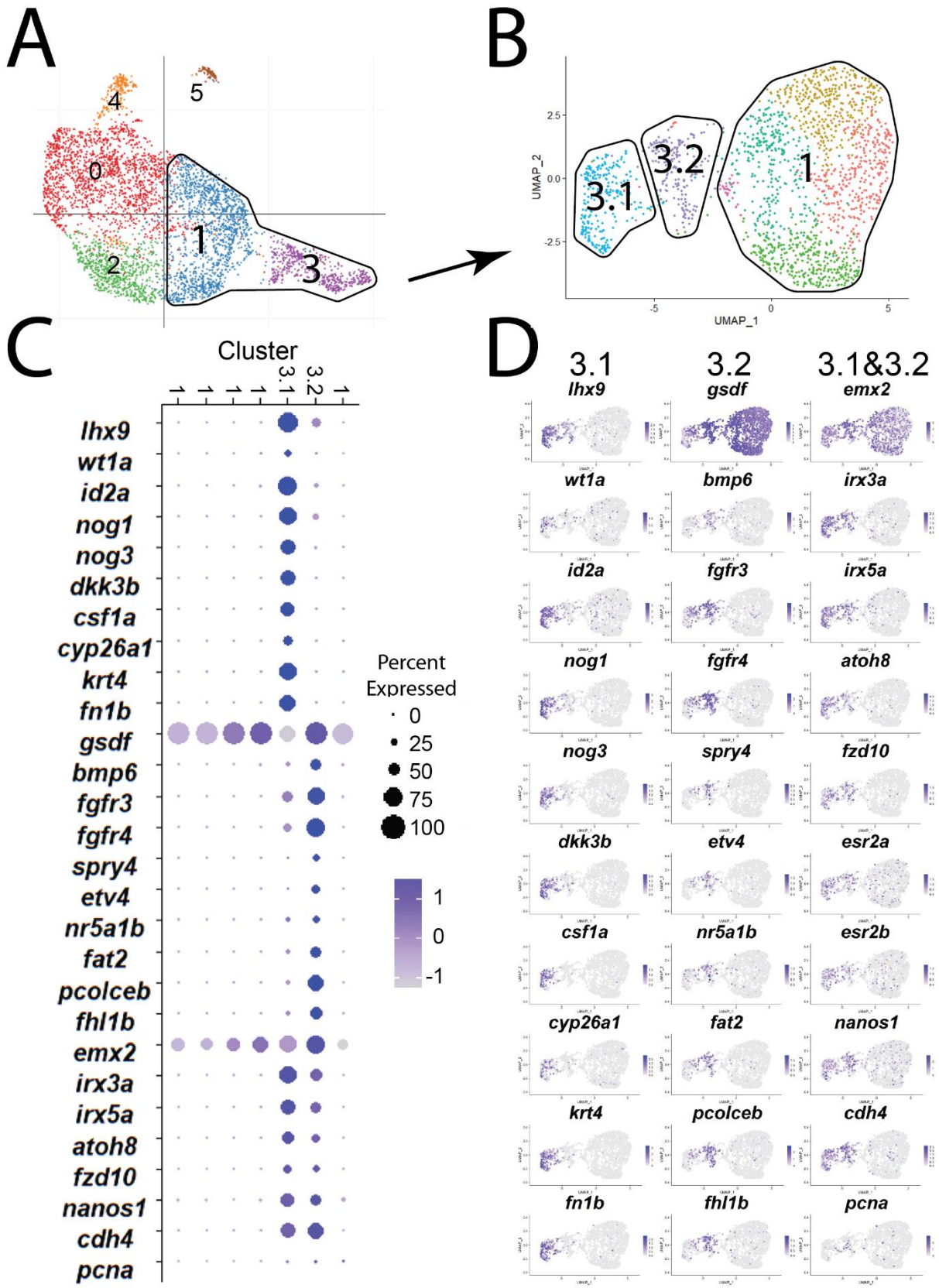

A
4
5
0
1
2
3
B
3.1
3.2
1
UMAP_2
UMAP_1
C
Cluster
1
1
1
1
3.1
3.2
1
lhx9
wt1a
id2a
nog1
nog3
dkk3b
csf1a
cyp26a1
krt4
fn1b
gsdf
bmp6
fgfr3
fgfr4
spry4
etv4
nr5a1b
fat2
pcolceb
fhl1b
emx2
irx3a
irx5a
atoh8
fzd10
nanos1
cdh4
pcna
Percent Expressed
0
25
50
75
100
1
0
-1
D
3.1
3.2
3.1&3.2
lhx9
gsdf
emx2
wt1a
bmp6
irx3a
id2a
fgfr3
irx5a
nog1
fgfr4
atoh8
nog3
spry4
fzd10
dkk3b
etv4
esr2a
csf1a
nr5a1b
esr2b
cyp26a1
fat2
nanos1
krt4
pcolceb
cdh4
fn1b
fhl1b
pcna

Figure 3. Subcluster analysis of pre-follicle cells (*lhx9+).* **A**. Follicle cell sub-cluster UMAP plot, with cells color-coded by computationally determined cell subtypes. The cells outlined, subclusters 1 and 3, were further subclustered, generating the UMAP plot shown in **B**. **B**. UMAP plot of the *lhx9* subcluster. For simplicity, these subclusters have been group into three subclusters (3.1, 3.2 and 1), referencing their origin in the UMAP plot shown in **B**. **C**. Dot-plot showing the relative expression of select genes in the pre-follicle cell sub-clusters. **D**. Gene expression UMAP plots of genes listed in **B**.

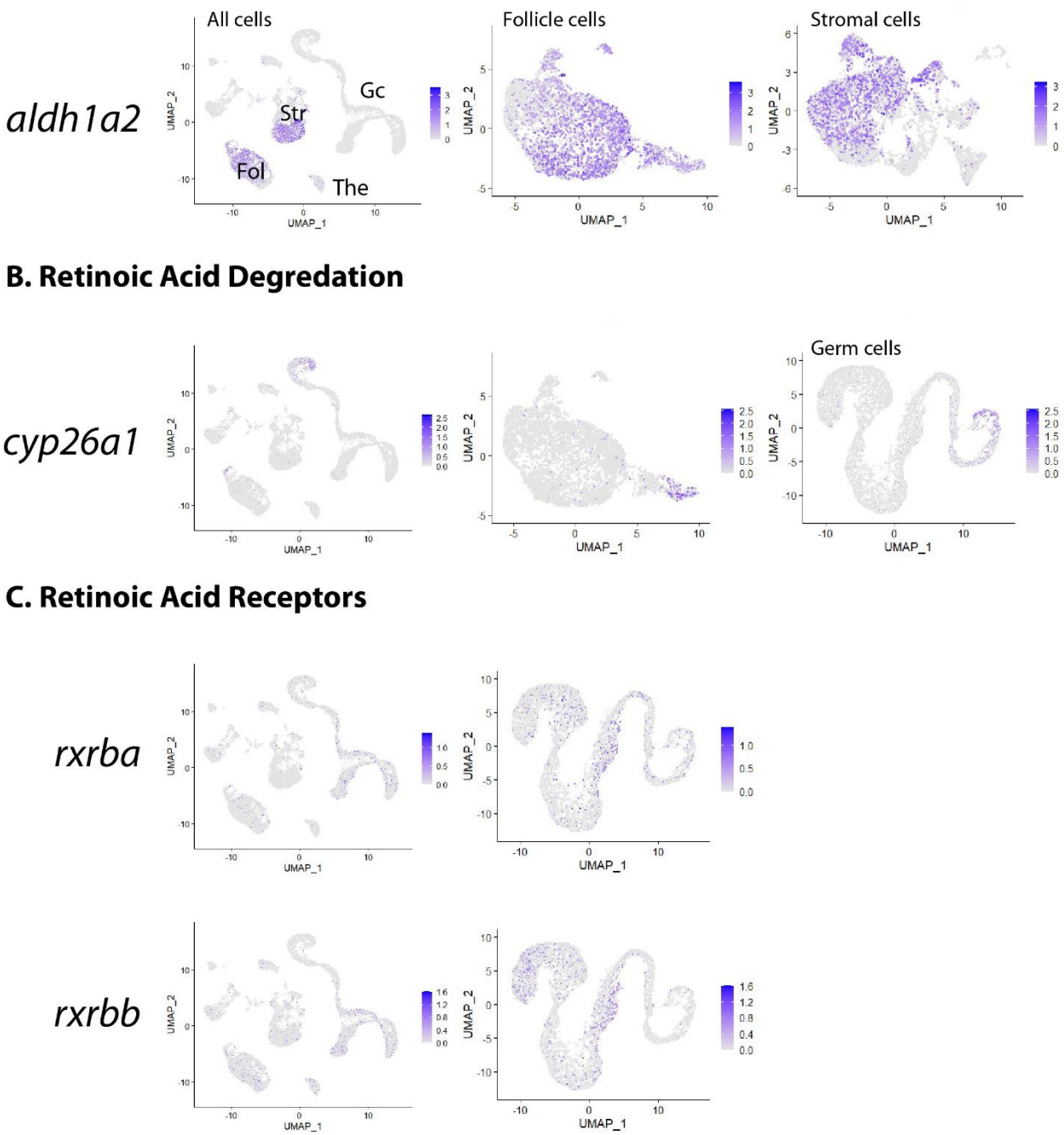
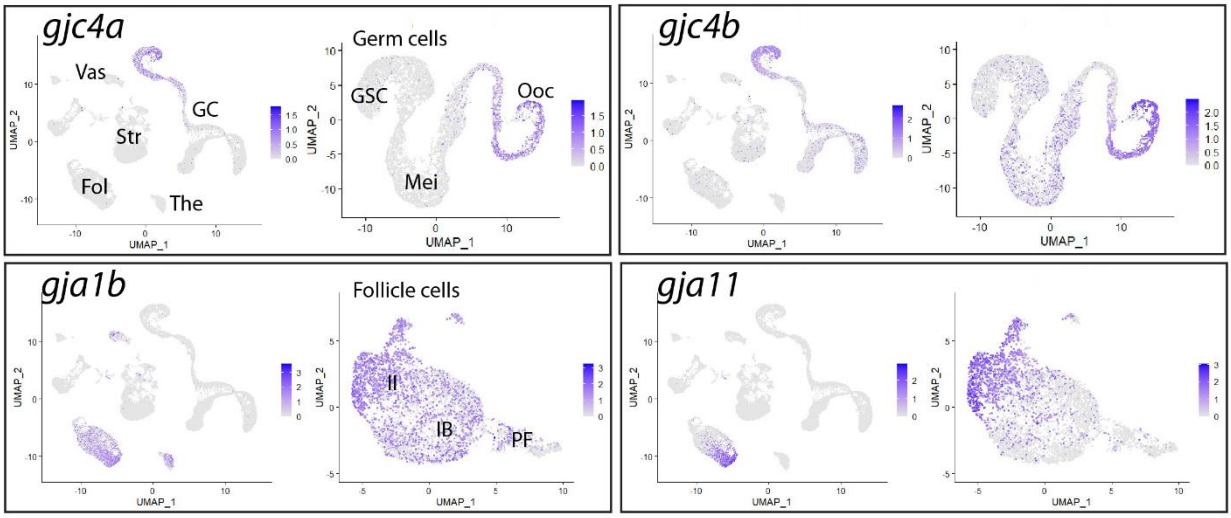


Figure 4. Gene expression UMAP plots for genes that function in the retinoic acid signaling pathway. Fol, follicle cells; Str, stromal cells; The, theca cells; GC, germ cells.

A



B

Zebrafish		Mouse	
cx	*gj*	*Cx*	*Gj*
cx43	*gja1b*	*Cx43*	*Gja1*
cx43.4	*gjc4b*		
cx44.2	*gjc4a*		
cx34.5	*gja11*		
cx39.4	*gja4*	*Cx37*	*Gja4*

C

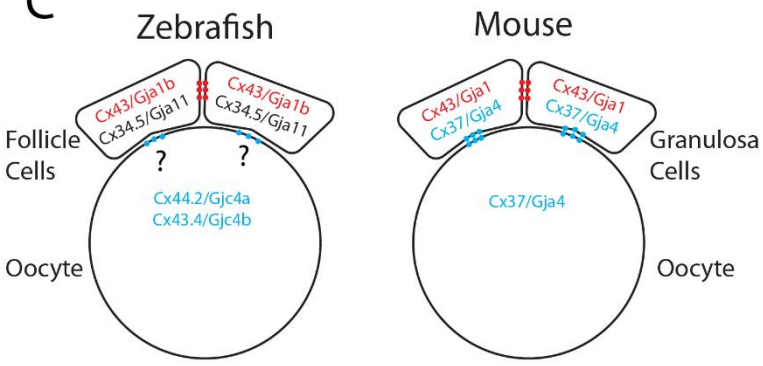


Figure 5. Expression of connexon encoding genes in the zebrafish ovary. A. Gene expression UMAP plots for connexin-encoding genes. B. Gene homology between zebrafish and mouse connexin genes, using the old *cx* and current *gj* nomenclature. C. Model showing gene expression in both zebrafish and mouse connexin proteins in follicle cells (trapezoids) and oocytes (circles) together with probable localization of specific connexons. Fol, follicle cells; Str, stromal cells; The, theca cells; Gc, germ cells.

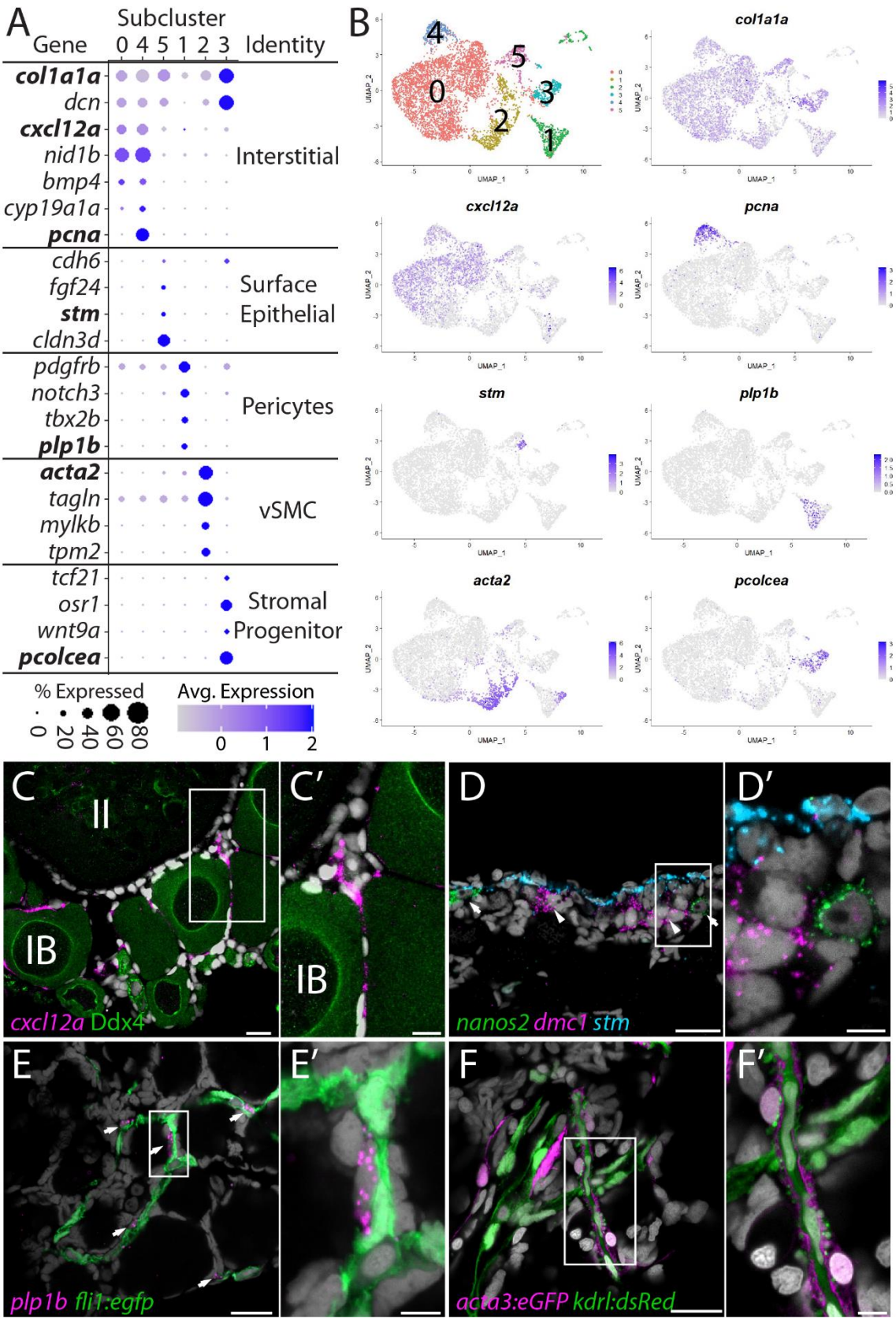

A
Gene
Subcluster
0 4 5 1 2 3
Identity
col1a1a
dcn
cxcl12a
nid1b
bmp4
cyp19a1a
pcna
Interstitial
cdh6
fgf24
stm
cldn3d
Surface Epithelial
pdgfrb
notch3
tbx2b
plp1b
Pericytes
acta2
tagln
mylkb
tpm2
vSMC
tcf21
osr1
wnt9a
pcolcea
Stromal Progenitor
% Expressed
0 20 40 60 80
Avg. Expression
0 1 2
B
col1a1a
cxcl12a
pcna
stm
plp1b
acta2
pcolcea
UMAP_1
UMAP_2
C
II
IB
cxcl12a Ddx4
C′
IB
D
nanos2 dmc1 stm
D′
E
plp1b fli1:egfp
E′
F
acta3:eGFP kdrl:dsRed
F′

Figure 6. Stromal cell sub-cluster analysis reveals five main cell subtypes. **A.** Dot-plot showing the relative expression of select genes in the stromal cell sub-clusters. Some genes, like *col1a1a*, are expressed in all stromal cells, while others, such as *stm*, are only expressed in a specific subcluster. UMAP plots of genes in bold are shown in **B. 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 indicates relative expression levels (intensity scale for each plot is on the right). **C-G.** smFISH on whole-mount 40 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 immunofluorescences (green) labels all germ cells. **C'.** higher magnification of region boxed in **C. E.** *stm* (blue) expressing cells localize to the lateral margin of the ovary and colocalize with *nanos2* (green) and *dmc1* (pink) germline stem cells and early meiotic cells, respectively. **E'.** Higher magnification of region boxed in **E. F.** *plp1b*-expression pericytes co-localize with *fli1:eg*fp expressing blood vessels (green). **F'.** higher magnification of region boxed in **F.** *acta3:egfp* expression vascular smooth muscle cells (red) surround *kdrl:dsRed* expression blood vessels (green). **G'.** higher magnification of region boxed in **G.** Scale bar in C, D and E 20µm; C', 10µm, D', E' and F', 5 µm. IB, stage IB oocyte; II, stage II oocyte.

A. Follicle cell subcluster GO terms

Subcluster	Id	source	term_id	term_name	term_size	p_value cluster.0	p_value cluster.1	p_value cluster.2	p_value cluster.3	p_value cluster.4	p_value cluster.5
0 q-Stage II FC	1	GO:BP	GO:0006412	translation	519	3.0e-43	NA	2.7e-04	NA	NA	NA
	2	GO:BP	GO:0043043	peptide biosynthetic process	526	5.6e-43	NA	3.0e-04	NA	NA	NA
	3	GO:BP	GO:0043604	amide biosynthetic process	609	4.4e-40	NA	6.9e-04	NA	NA	NA
	4	GO:BP	GO:0006518	peptide metabolic process	625	1.4e-39	NA	1.1e-03	NA	NA	NA
	5	GO:BP	GO:0043603	cellular amide metabolic process	806	1.2e-34	NA	7.1e-03	NA	NA	NA
1 Stage I FC	6	GO:BP	GO:0019882	antigen processing and presentation	36	NA	1.1e-13	NA	NA	NA	NA
	7	GO:BP	GO:0007423	sensory organ development	673	NA	1.8e-04	NA	NA	3.3e-02	NA
	8	GO:BP	GO:0009653	anatomical structure morphogenesis	2275	NA	2.3e-03	NA	NA	NA	2.4e-03
	9	GO:BP	GO:0048839	inner ear development	178	NA	2.3e-03	NA	NA	NA	NA
	10	GO:BP	GO:0043583	ear development	178	NA	2.3e-03	NA	NA	NA	NA
2 Transitioning	11	GO:BP	GO:0042773	ATP synthesis coupled electron transport	56	NA	NA	3.9e-04	NA	NA	NA
	12	GO:BP	GO:0046034	ATP metabolic process	149	NA	NA	6.5e-04	NA	9.0e-03	NA
	13	GO:BP	GO:1902600	proton transmembrane transport	169	NA	NA	8.9e-04	NA	NA	NA
	14	GO:BP	GO:0022904	respiratory electron transport chain	69	NA	NA	9.1e-04	NA	NA	NA
	15	GO:BP	GO:0006119	oxidative phosphorylation	69	NA	NA	9.1e-04	NA	1.5e-02	NA
3 pre-FC	16	GO:BP	GO:0048856	anatomical structure development	4984	4.0e-02	2.6e-03	NA	6.4e-08	NA	2.8e-03
	17	GO:BP	GO:0032502	developmental process	4900	NA	7.4e-03	NA	1.1e-07	NA	7.0e-03
	18	GO:BP	GO:0009888	tissue development	1594	NA	5.2e-03	NA	4.0e-06	NA	1.9e-02
	19	GO:BP	GO:0007275	multicellular organism development	4280	5.4e-03	1.4e-03	NA	2.0e-05	NA	1.7e-03
	20	GO:BP	GO:0061448	connective tissue development	192	NA	1.4e-02	NA	1.2e-04	NA	NA
4 p-Stage II FC	21	GO:BP	GO:0006260	DNA replication	142	NA	NA	NA	NA	9.3e-35	NA
	22	GO:BP	GO:0006259	DNA metabolic process	585	NA	NA	NA	NA	3.9e-27	NA
	23	GO:BP	GO:0006261	DNA-dependent DNA replication	100	NA	NA	NA	NA	4.4e-25	NA
	24	GO:BP	GO:0006281	DNA repair	363	NA	NA	NA	NA	2.0e-17	NA
	25	GO:BP	GO:0034641	cellular nitrogen compound metabolic process	5168	9.9e-12	NA	NA	NA	2.5e-16	NA
5	26	GO:BP	GO:0030198	extracellular matrix organization	161	NA	NA	NA	NA	NA	6.9e-06
	27	GO:BP	GO:0043062	extracellular structure organization	161	NA	NA	NA	NA	NA	6.9e-06
	28	GO:BP	GO:0048731	system development	3749	NA	1.3e-03	NA	1.9e-04	NA	1.0e-04
	29	GO:BP	GO:0001501	skeletal system development	404	NA	NA	NA	2.6e-03	NA	7.0e-04
	30	GO:BP	GO:0048589	developmental growth	401	NA	NA	NA	NA	NA	7.1e-03

B. Stromal cell subcluster GO terms

Subcluster	id	term_id	term_name	p_value cluster.0	p_value cluster.1	p_value cluster.2	p_value cluster.3	p_value cluster.4	p_value cluster.5
0 q-Interstitial	1	GO:0051085	chaperone cofactor-dependent protein refolding	1.1e-02	NA	NA	NA	NA	NA
	2	GO:0051084	'de novo' posttranslational protein folding	1.1e-02	NA	NA	NA	NA	NA
	3	GO:0006458	'de novo' protein folding	1.3e-02	NA	NA	NA	NA	NA
	4	GO:0001945	lymph vessel development	1.4e-02	NA	NA	NA	NA	NA
	5	GO:0061077	chaperone-mediated protein folding	2.8e-02	NA	NA	NA	NA	NA
1 pericytes	6	GO:0006412	translation	NA	7.0e-20	NA	NA	NA	NA
	7	GO:0043043	peptide biosynthetic process	NA	9.3e-20	NA	NA	NA	NA
	8	GO:0043604	amide biosynthetic process	NA	2.2e-18	NA	NA	NA	NA
	9	GO:0006518	peptide metabolic process	NA	3.9e-18	NA	NA	NA	NA
	10	GO:0043603	cellular amide metabolic process	NA	9.2e-16	NA	NA	NA	NA
2 VSM	11	GO:0050920	regulation of chemotaxis	NA	NA	4.1e-03	NA	NA	NA
	12	GO:0031175	neuron projection development	NA	NA	2.9e-02	NA	NA	NA
	13	GO:0010959	regulation of metal ion transport	NA	NA	4.2e-02	NA	NA	NA
3 Stromal progenitors	14	GO:0001501	skeletal system development	NA	NA	NA	2.6e-02	NA	NA
	15	GO:0048385	regulation of retinoic acid receptor signaling pathway	NA	NA	NA	2.7e-02	NA	NA
	16	GO:0030198	extracellular matrix organization	NA	NA	NA	3.4e-02	NA	NA
	17	GO:0043062	extracellular structure organization	NA	NA	NA	3.4e-02	NA	NA
4 p-Interstitial	18	GO:0006259	DNA metabolic process	NA	NA	NA	NA	1.1e-27	NA
	19	GO:0006260	DNA replication	NA	NA	NA	NA	1.5e-26	NA
	20	GO:0006261	DNA-dependent DNA replication	NA	NA	NA	NA	8.1e-20	NA
	21	GO:0051276	chromosome organization	NA	NA	NA	NA	9.7e-20	NA
	22	GO:0006281	DNA repair	NA	NA	NA	NA	2.5e-18	NA
5 Epithelial	23	GO:0071696	ectodermal placode development	NA	NA	NA	NA	NA	6.9e-03
	24	GO:0060114	vestibular receptor cell differentiation	NA	NA	NA	NA	NA	1.4e-02
	25	GO:0060118	vestibular receptor cell development	NA	NA	NA	NA	NA	1.4e-02

Figure 7. GO terms associated with the follicle and stromal cell sub-clusters. A. Top 5 unique GO terms associated with individual follicle sub-clusters. B. Top unique GO terms associated

with the stromal sub-clusters. In both A and B, blue indicates a high and yellow indicates low significance.

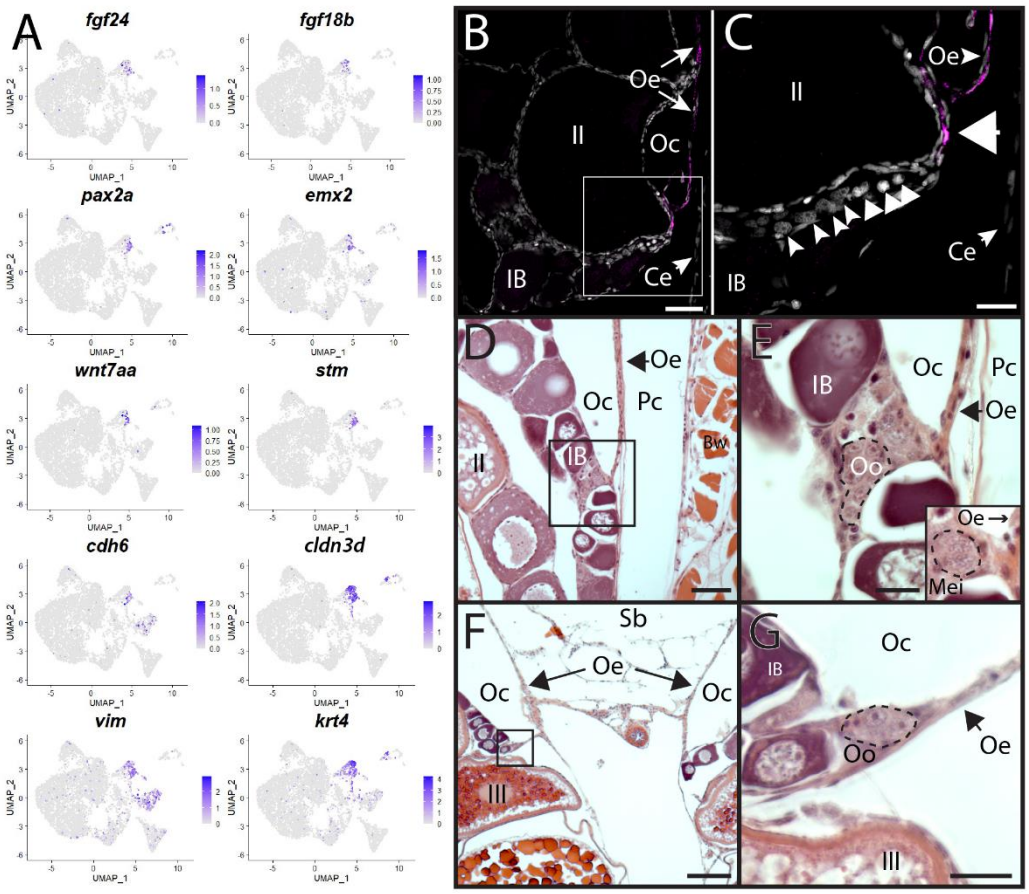


Figure 8. Stromal cell subcluster 3: Ovarian epithelium. **A.** Gene expression UMAP plots for select genes whose expression is enriched in stromal cell subcluster 3. ligand-encoding genes. **B.** smFISH on transverse sections from a 3-month-old ovary showing that *stm* (red) is expressed in the epithelium that lines the ovarian cavity. DNA is grey. **C.** Higher magnification of region boxed in **B** showing that early-stage germ cells localize to the region subjacent to where the ovarian cavity epithelium is attached to the lateral side of the ovary (arrowhead). **D-G.** Histological transverse sections from a 3-month-old ovary showing correlation between where the ovarian cavity epithelium attaches to the ovary at the lateral (**D** and **E**) and medial (**F** and **G**) sides, and the presence of pre-meiotic germ cells, characterized by large, dark staining nucleoli, and early meiotic germ cells, characterized by condensed chromosomes (inset in **E**). **E** and **G** are

higher magnification views of regions boxed in **D** and **F**, respectively. Oe, ovarian epithelium; Oc, ovarian cavity; Pc, peritoneal cavity; Ce, celomic epithelium; Sb, swim bladder; Oo, premeiotic oogonia; Mei, early meiotic germ cell; IB, stage IB oocyte; III, stage III oocyte.

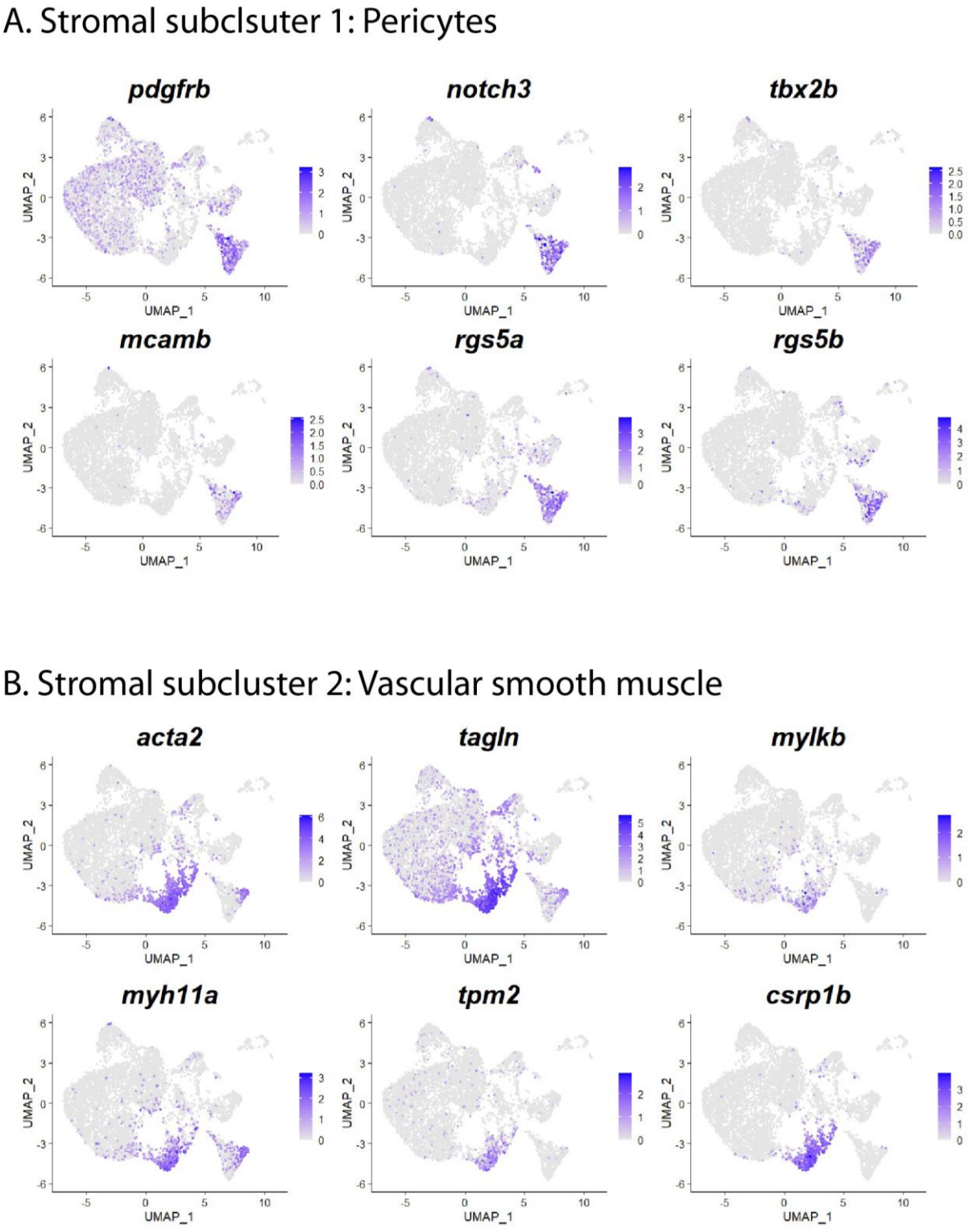


Figure 9. Stromal cell subclusters 1 and 2: pericytes and vascular smooth muscle cells. **A.** Gene expression UMAP plots for select genes whose expression is enriched in stromal cell subcluster

1 and are markers of pericytes. **B.** Gene expression UMAP plots for select genes whose expression is enriched in stromal cell subcluster 2 and are markers of vascular smooth muscle cells.

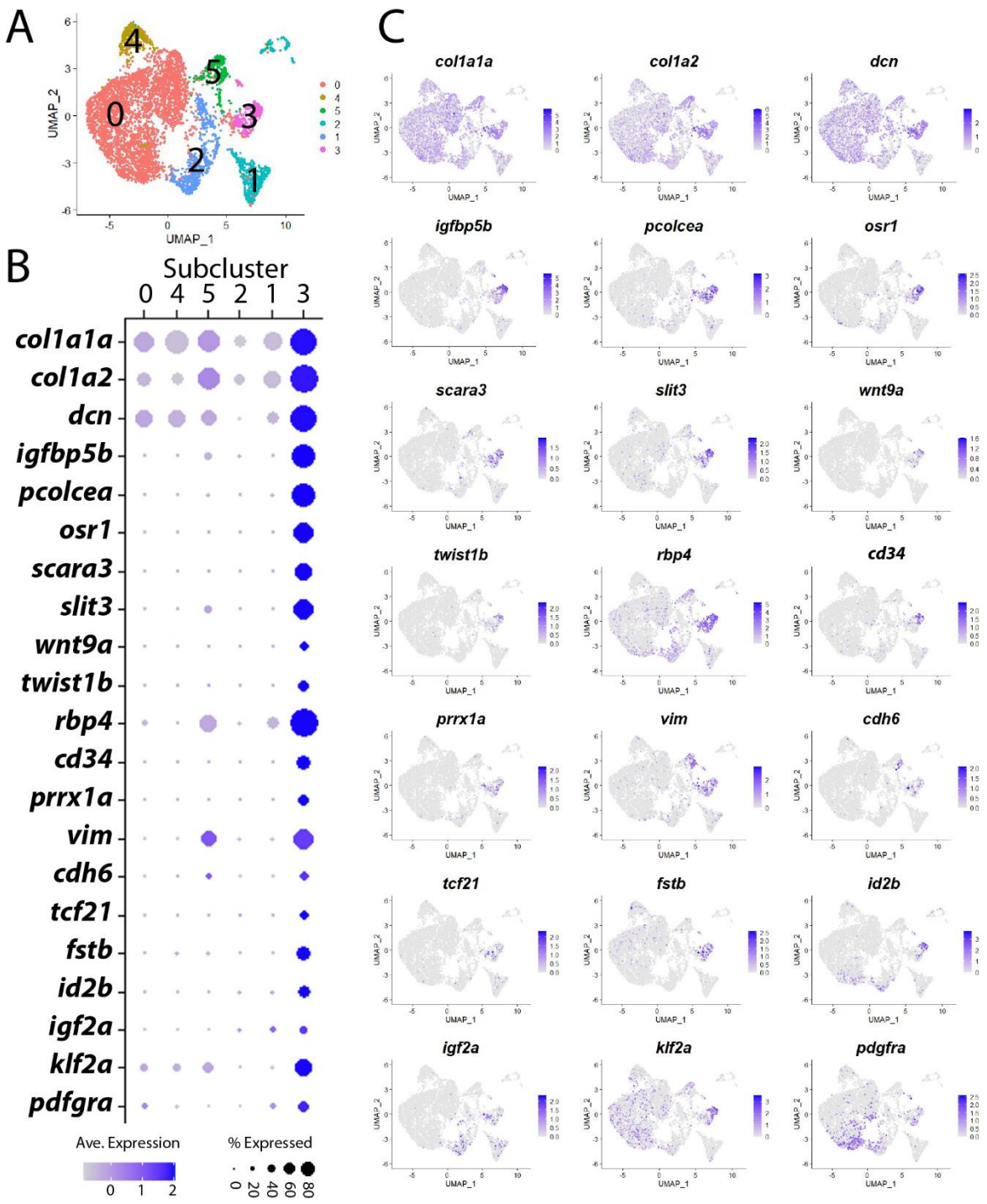


Figure 10. Stromal cell subcluster 3: stromal progenitor cells. **A**. Stomal cell sub-cluster UMAP plot, with cells color-coded by computationally determined cell subtypes. **B**. Dot-plot showing

the relative expression of select genes in stromal cell subclusters that have enriched expression is sub-cluster 3. **C**. UMAP gene expression plot for genes listed in **B**.

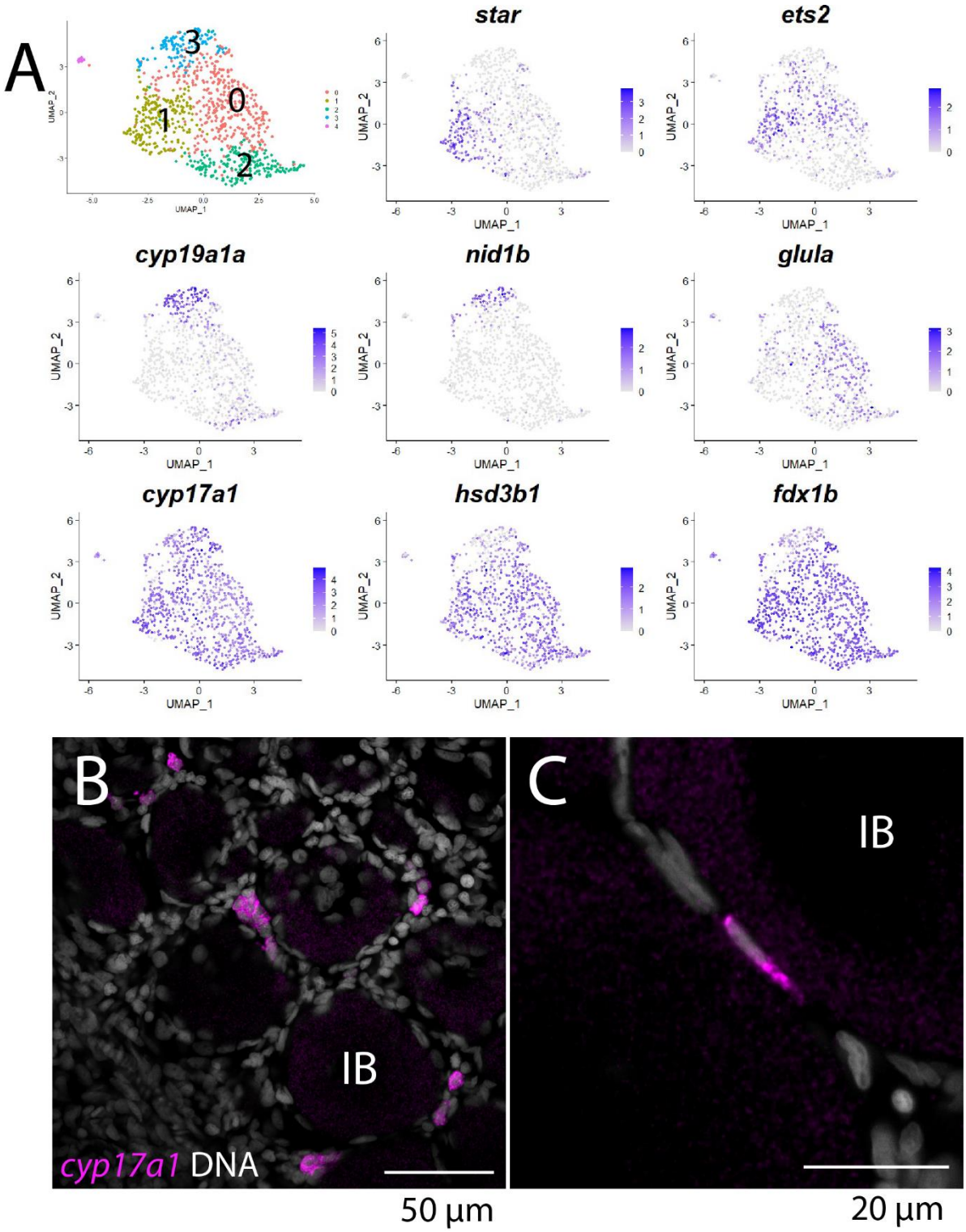


Figure 11. Theca cell sub-cluster analysis. **A.** Gene expression UMAP plots of select genes. Top left panel shows cells color-coded by computationally determined cell subtype. Cells expressing

the indicated gene are colored purple, and the relative intensity indicates relative expression levels (intensity scale for each plot is on the right). **B.** smFISH on whole-mount 40 dpf ovaries reveals the location of *cyp17a1* expressing theca cells. DNA is gray. **C.** Higher magnification of a *cyp17a1* expressing theca cells. IB, stage IB oocyte. Scale bars in B, 50µm and C, 20 µm.

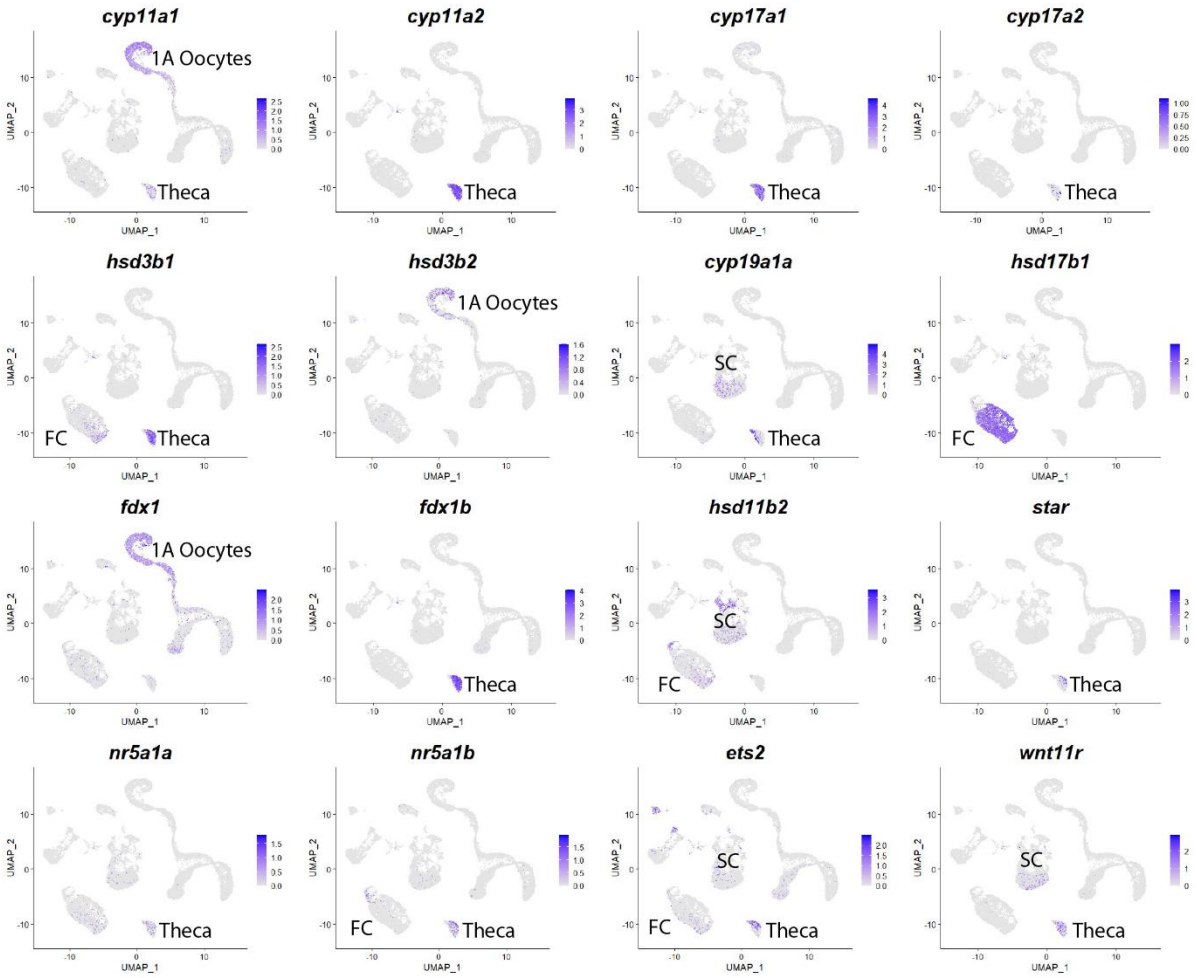


Figure 12. Expression of genes involved in theca cell development and steroid synthesis. UMAP gene expression plot for orthologs of genes known to be involved in sex steroid synthesis or theca cell development, mapped onto all cells. Fc, follicle cell; Sc, stromal cell.

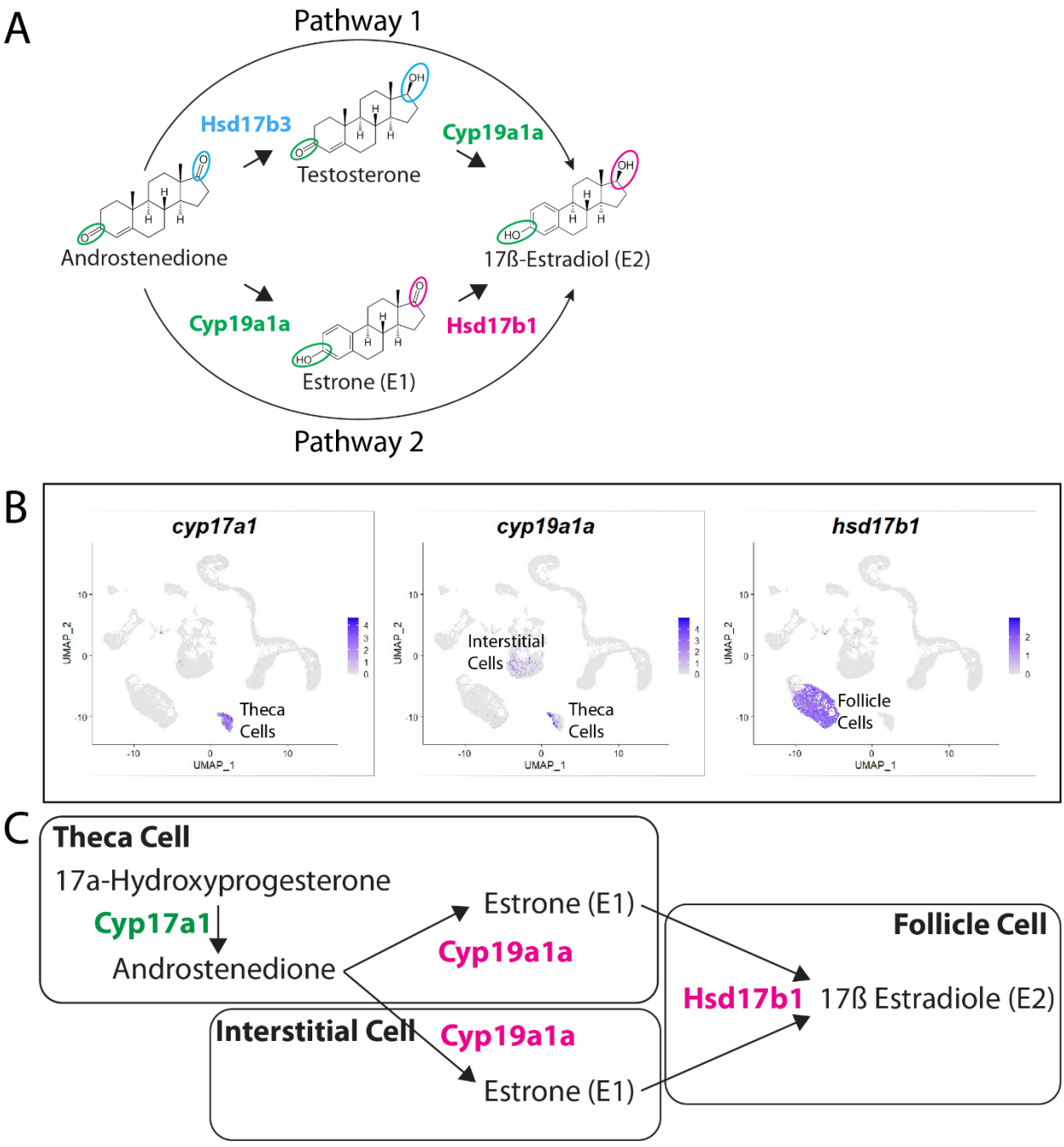


Figure 13. Pathway for 17ß-estrodiol (E2) synthesis in the zebrafish ovary. **A.** Two possible pathways for E2 synthesis starting with androstenedione. Colored circles indicate the region of the molecules being modified, and the color corresponds to the enzyme that catalyzes the modification. **B.** 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 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 17a-Hydroyprogesterone intermediate precursor, together with the cell types where each of the reactions occur.

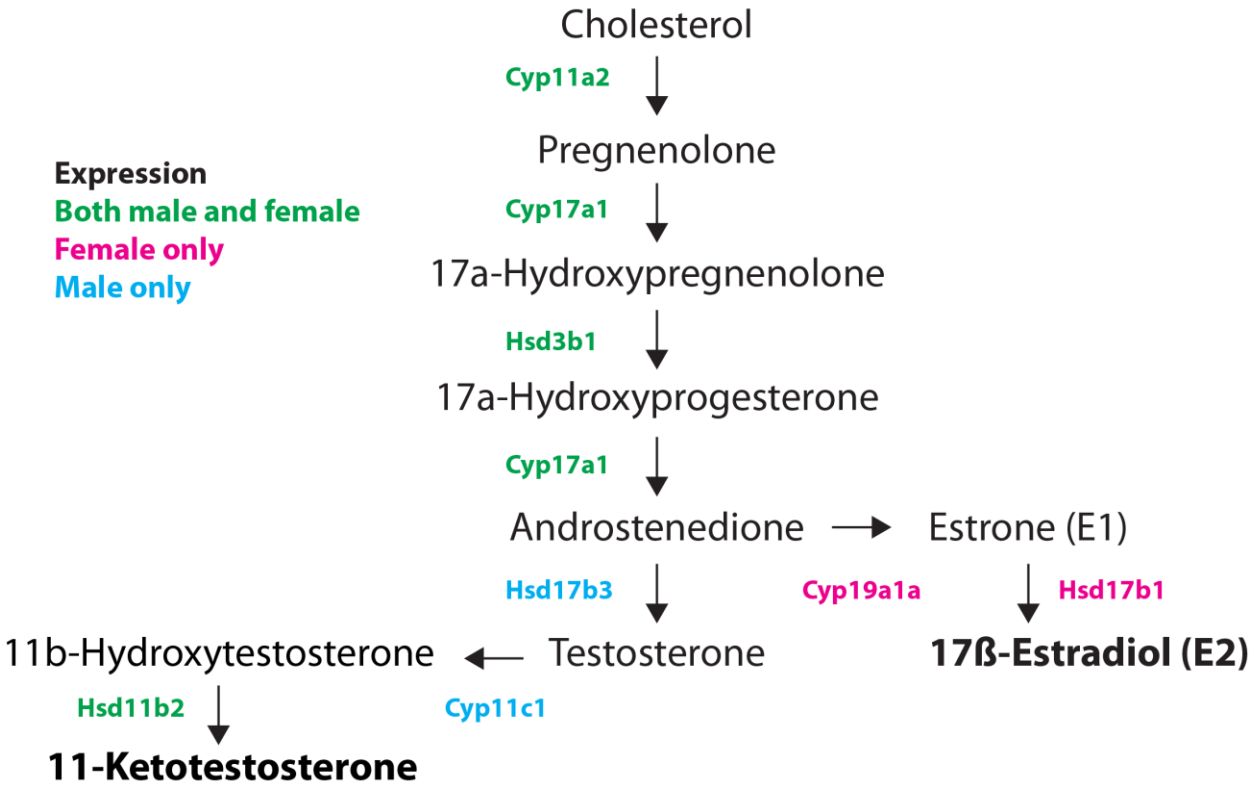


Figure 14. Pathway for sex hormone synthesis in the juvenile zebrafish. Diagram show a likely pathway for the synthesis of the two main sex steroids in zebra fish 17ß-estradiol (E2; female) and 11-ketotestosterone (male), starting from cholesterol. Enzymes that canalize each step are shown, with color indicating sex specific expression. For example, Cyp17a1 is expressed in both males and females, while Cyp19a1a and Hsd17b1 are female specific and Hsd17b3 and Cyp11c1 are male-specific.

3.8 References

Bahrami, N., & Childs, S. J. (2018). Pericyte Biology in Zebrafish. *Adv Exp Med Biol*, *1109*, 33–51.

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME suite. *Nucleic Acids Research*, *43*(W1), 39–49.

Bauer, M. P., Bridgham, J. T., Langenau, D. M., Johnson, A. L., & Goetz, F. W. (2000). Conservation of steroidogenic acute regulatory (StAR) protein structure and expression in vertebrates. *Molecular and Cellular Endocrinology*, *168*(1–2), 119–125.

Beer, R. L., & Draper, B. W. (2013a). Nanos3 maintains germline stem cells and expression of the conserved germline stem cell gene nanos2 in the zebrafish ovary. *Developmental Biology*, *374*(2), 308–318. https://doi.org/10.1016/j.ydbio.2012.12.003

Beer, R. L., & Draper, B. W. (2013b). Nanos3 maintains germline stem cells and expression of the conserved germline stem cell gene nanos2 in the zebrafish ovary. *Developmental Biology*, *374*(2), 308–318. https://doi.org/10.1016/j.ydbio.2012.12.003

Birk, O. S., Casiano, D. E., Wassif, C. A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J. A., Parker, K. L., Porter, F. D., & Westphal, H. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. *Nature*, *403*(6772), 909–913. https://doi.org/10.1038/35002622

Blokhina, Y. P., Nguyen, A. D., Draper, B. W., & Burgess, S. M. (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*(1), e1007730. https://doi.org/10.1371/journal.pgen.1007730

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M. J., & Rossant, J. (2006). Retinoid signaling determines germ cell fate in mice. *Science*, *312*, 596–600.

Brown, L. A., Rodaway, A. R., Schilling, T. F., Jowett, T., Ingham, P. W., Patient, R. K., & Sharrocks, A. D. (2000). Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Fli-1 in wild-type and mutant zebrafish embryos. *Mech Dev*, *90*, 237–252.

Brunet, J.-P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences*, *101*(12), 4164–4169. https://doi.org/10.1073/pnas.0308531101

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, *566*(7745), 496–502. https://doi.org/10.1038/s41586-019-0969-x

Cao, Z., Mao, X., & Luo, L. (2019). Germline Stem Cells Drive Ovary Regeneration in Zebrafish. *Cell Rep*, *26*, 1709-1717 1703.

Carmona, S. J., Teichmann, S. A., Ferreira, L., Macaulay, I. C., Stubbington, M. J., Cvejic, A., & Gfeller, D. (2017). Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. *Genome Res*, *27*, 451–461.

Chen, D., & McKearin, D. (2003). Dpp signaling silences bam transcription directly to establish asymmetric divisions of germline stem cells. *Current Biology*, *13*(20), 1786–1791.

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. https://doi.org/10.1016/j.ygcen.2013.08.016

Crespo, D., Assis, L. H. C., Kant, H. J. G., Waard, S., Safian, D., Lemos, M. S., Bogerd, J., & Schulz, R. W. (2019). Endocrine and local signaling interact to regulate spermatogenesis in zebrafish: Follicle-stimulating hormone, retinoic acid and androgens. *Development*, *146*.

Cui, S., Ross, A., Stallings, N., Parker, K. L., Capel, B., & Quaggin, S. E. (2004). Disrupted gonadogenesis and male-to-female sex reversal in Pod1 knockout mice. *Development*, *131*, 4095–4105.

Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M. R., Armean, I. M., Bennett, R., Bhai, J., Billis, K., Boddu, S., Cummins, C., Davidson, C., Dodiya, K. J., Gall, A., Girón, C. G., Gil, L., Grego, T., Haggerty, L., Haskell, E., … Flicek, P. (2019). Ensembl 2019. *Nucleic Acids Research*, *47*(D1), D745–D751. https://doi.org/10.1093/nar/gky1113

De Keuckelaere, E., Hulpiau, P., Saeys, Y., Berx, G., & Roy, F. (2018). Nanos genes and their role in development and beyond. *Cell Mol Life Sci*, *75*, 1929–1946.

Devlin, R. H., & Nagahama, Y. (2002). Sex determination and sex differentiation in fish: An overview of genetic, physiological, and environmental influences. *Aquaculture*, *208*, 191–364.

Doitsidou, M., Reichman-Fried, M., Stebler, J., Köprunner, M., Dörries, J., Meyer, D., Esguerra, C. V., Leung, T., & Raz, E. (2002). Guidance of Primordial Germ Cell Migration by the Chemokine SDF-1. *Cell*, *111*(5), 647–659. https://doi.org/10.1016/S0092-8674(02)01135-2

Dranow, D. B., Hu, K., Bird, A. M., Lawry, S. T., Adams, M. T., Sanchez, A., Amatruda, J. F., & Draper, B. W. (2016). Bmp15 Is an Oocyte-Produced Signal Required for Maintenance of the Adult Female Sexual Phenotype in Zebrafish. *PLOS Genetics*, *12*(9), e1006323. https://doi.org/10.1371/journal.pgen.1006323

Draper, B. W. (2012). Identification of Oocyte Progenitor Cells in the Zebrafish Ovary. In K. A. Mace & K. M. Braun (Eds.), *Progenitor Cells: Methods and Protocols* (pp. 157–165). Humana Press. https://doi.org/10.1007/978-1-61779-980-8_12

Draper, B. W., McCallum, C. M., & Moens, C. B. (2007a). Nanos1 Is Required To Maintain Oocyte Production in Adult Zebrafish. *Developmental Biology*, *305*(2), 589–598. https://doi.org/10.1016/j.ydbio.2007.03.007

Draper, B. W., McCallum, C. M., & Moens, C. B. (2007b). Nanos1 is required to maintain oocyte production in adult zebrafish. *Developmental Biology*, *305*(2), 589–598. https://doi.org/10.1016/j.ydbio.2007.03.007

Elkouby, Y. M., & Mullins, M. C. (2017). Methods for the analysis of early oogenesis in Zebrafish. *Developmental Biology*, *430*(2), 310–324. https://doi.org/10.1016/j.ydbio.2016.12.014

Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science (New York, N.Y.)*, *360*(6392), eaar3131. https://doi.org/10.1126/science.aar3131

Feng, J., Mantesso, A., & Sharpe, P. T. (2010). Perivascular cells as mesenchymal stem cells. *Expert Opin Biol Ther*, *10*, 1441–1451.

Forbes, A., & Lehmann, R. (1998a). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. *Development*, *125*(4), 679–690.

Forbes, A., & Lehmann, R. (1998b). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. *Development*, *125*, 679–690.

Gaengel, K., Genové, G., Armulik, A., & Betsholtz, C. (2009). Endothelial-mural cell signaling in vascular development and angiogenesis. *Arterioscler Thromb Vasc Biol*, *29*, 630–638.

Gautier, A., Sohm, F., Joly, J.-S., Le Gac, F., & Lareyre, J.-J. (2011). The Proximal Promoter Region of the Zebrafish gsdf Gene Is Sufficient to Mimic the Spatio-Temporal Expression Pattern of the Endogenous Gene in Sertoli and Granulosa Cells. *Biology of Reproduction*, *85*(6), 1240–1251. https://doi.org/10.1095/biolreprod.111.091892

Gilchrist, R. B., Lane, M., & Thompson, J. G. (2008). Oocyte-secreted factors: Regulators of cumulus cell function and oocyte quality. *Hum Reprod Update*, *14*, 159–177.

Gore, A. V., Athans, B., Iben, J. R., Johnson, K., Russanova, V., Castranova, D., Pham, V. N., Butler, M. G., Williams-Simons, L., & Nichols, J. T. (2016). Epigenetic regulation of hematopoiesis by DNA methylation. *Elife*, *5*, 11813.

Guiguen, Y., Fostier, A., Piferrer, F., & Chang, C. F. (2010). Ovarian aromatase and estrogens: A pivotal role for gonadal sex differentiation and sex change in fish. *Gen Comp Endocrinol*, *165*, 352–366.

Haddon, C., Jiang, Y. J., Smithers, L., & Lewis, J. (1998). Delta-Notch signalling and the patterning of sensory cell differentiation in the zebrafish ear: Evidence from the mind bomb mutant. *Development*, *125*, 4637–4644.

Hsu, R. J., Lin, C. Y., Hoi, H. S., Zheng, S. K., Lin, C. C., & Tsai, H. J. (2010). Novel intronic microRNA represses zebrafish myf5 promoter activity through silencing dickkopf-3 gene. *Nucleic Acids Res*, *38*, 4384–4393.

Hu, Y. C., Okumura, L. M., & Page, D. C. (2013). Gata4 is required for formation of the genital ridge in mice. *PLoS Genet*, *9*, 1003629.

Kidder, G. M., & Vanderhyden, B. C. (2010). Bidirectional communication between oocytes and follicle cells: Ensuring oocyte developmental competence. *Can J Physiol Pharmacol*, *88*, 399–413.

Kikuchi, K., Holdway, J. E., Major, R. J., Blum, N., Dahn, R. D., Begemann, G., & Poss, K. D. (2011). Retinoic Acid Production by Endocardium and Epicardium Is an Injury Response Essential for Zebrafish Heart Regeneration. *Developmental Cell*, *20*(3), 397–404. https://doi.org/10.1016/j.devcel.2011.01.010

Kim, B., Kim, Y., Cooke, P. S., Rüther, U., & Jorgensen, J. S. (2011). The fused toes locus is essential for somatic-germ cell interactions that foster germ cell maturation in developing gonads in mice. *Biol Reprod*, *84*, 1024–1032.

Kimble, J., & Crittenden, S. L. (2007). Controls of germline stem cells, entry into meiosis, and the sperm/oocyte decision in Caenorhabditis elegans. *Annu. Rev. Cell Dev. Biol.*, *23*, 405–433.

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. *Journal of Cell Biology*, *149*(4), 875–888. https://doi.org/10.1083/jcb.149.4.875

Köprunner, M., Thisse, C., Thisse, B., & Raz, E. (2001). A zebrafish nanos-related gene is essential for the development of primordial germ cells. *Genes Dev*, *15*, 2877–2885.

Kossack, M. E., & Draper, B. W. (2019). Genetic regulation of sex determination and maintenance in zebrafish (Danio rerio. *Curr Top Dev Biol*, *134*, 119–149.

Koubova, J., Menke, D. B., Zhou, Q., Capel, B., Griswold, M. D., & Page, D. C. (2006). Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proc Natl Acad Sci U S A*, *103*, 2474–2479.

Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D., & Jaenisch, R. (1993). WT-1 is required for early kidney development. *Cell*, *74*, 679–691.

Kusz, K. M., Tomczyk, L., Sajek, M., Spik, A., Latos-Bielenska, A., Jedrzejczak, P., Pawelczyk, L., & Jaruzelska, J. (2009). The highly conserved NANOS2 protein: Testis-specific expression and

significance for the human male reproduction. *Molecular Human Reproduction*. https://doi.org/10.1093/molehr/gap003

Kwok, H. F., So, W. K., 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. *Biol Reprod*, *72*, 1370–1381.

Lawson, N. D., & Weinstein, B. M. (2002). In vivo imaging of embryonic vascular development using transgenic zebrafish. *Developmental Biology*, *248*(2), 307–318.

Lee, J., Kanatsu-Shinohara, M., Inoue, K., Ogonuki, N., Miki, H., Toyokuni, S., Kimura, T., Nakano, T., Ogura, A., & Shinohara, T. (2007). Akt mediates self-renewal division of mouse spermatogonial stem cells. *Development (Cambridge, England)*, *134*(10), 1853–1859. https://doi.org/10.1242/dev.003004

Leerberg, D. M., Sano, K., & Draper, B. W. (2017). Fibroblast growth factor signaling is required for early somatic gonad development in zebrafish. *PLOS Genetics*, *13*(9), e1006993. https://doi.org/10.1371/journal.pgen.1006993

Leu, D. H., & Draper, B. W. (2010). The ziwi promoter drives germline-specific gene expression in zebrafish. *Developmental Dynamics*, *239*(10), 2714–2721. https://doi.org/10.1002/dvdy.22404

Li, T. Y., Colley, D., Barr, K. J., Yee, S. P., & Kidder, G. M. (2007). Rescue of oogenesis in Cx37-null mutant mice by oocyte-specific replacement with Cx43. *J Cell Sci*, *120*, 4117–4125.

Lieschke, G. J., Oates, A. C., Crowhurst, M. O., Ward, A. C., & Layton, J. E. (2001). Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. *Blood*, *98*, 3087–3096.

Luo, X., Ikeda, Y., & Parker, K. L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell*, *77*, 481–490.

Marlow, F. L., & Mullins, M. C. (2008). Bucky ball functions in Balbiani body assembly and animal–vegetal polarity in the oocyte and follicle cell layer in zebrafish. *Developmental Biology*, *321*(1), 40–50. https://doi.org/10.1016/j.ydbio.2008.05.557

McLeay, R. C., & Bailey, T. L. (2010). Motif Enrichment Analysis: A unified framework and an evaluation on ChIP data. *BMC Bioinformatics*, *11*(1), 1–11.

McMillan, D. B. (2007). *Fish histology: Female reproductive systems*. Springer Science & Business Media.

Meng, X., Lindahl, M., Hyvönen, M. E., Parvinen, M., Rooij, D. G. de, Hess, M. W., Raatikainen-Ahokas, A., Sainio, K., Rauvala, H., Lakso, M., Pichel, J. G., Westphal, H., Saarma, M., & Sariola, H. (2000). Regulation of Cell Fate Decision of Undifferentiated Spermatogonia by GDNF. *Science*, *287*(5457), 1489–1493. https://doi.org/10.1126/science.287.5457.1489

Mikalsen, S. O., S, Í. K., & Tausen, M. (2021). Connexins during 500 Million Years-From Cyclostomes to Mammals. *Int J Mol Sci*, *22*.

Miller, W. L., & Auchus, R. J. (2011). The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr Rev*, *32*, 81–151.

Mindnich, R., Deluca, D., & Adamski, J. (2004). Identification and characterization of 17 beta-hydroxysteroid dehydrogenases in the zebrafish, Danio rerio. *Mol Cell Endocrinol*, *215*, 19–30.

Mindnich, R., Haller, F., Halbach, F., Moeller, G., 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. *J Mol Endocrinol*, *35*, 305–316.

Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I., & Aizawa, S. (1997). Defects of urogenital development in mice lacking Emx2. *Development*, *124*, 1653–1664.

Morvan-Dubois, G., Le Guellec, D., Garrone, R., Zylberberg, L., & Bonnaud, L. (2003). Phylogenetic analysis of vertebrate fibrillar collagen locates the position of zebrafish alpha3(I) and suggests an evolutionary link between collagen alpha chains and hox clusters. *J Mol Evol*, *57*, 501–514.

Nakamura, S. (2010). *Identification of Germline Stem Cells*. *1561*(2004). https://doi.org/10.1126/science.1185473

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. *Dev Dyn*, *238*, 2652–2657.

Nishimura, T., Sato, T., Yamamoto, Y., Watakabe, I., Ohkawa, Y., Suyama, M., Kobayashi, S., & Tanaka, M. (2015). Foxl3 is a germ cell–intrinsic factor involved in sperm-egg fate decision in medaka. *Science*, *349*(6245), 328–331. https://doi.org/10.1126/science.aaa2657

Oatley, J. M., & Brinster, R. L. (2012). The germline stem cell niche unit in mammalian testes. *Physiological Reviews*, *92*(2), 577–595. https://doi.org/10.1152/physrev.00025.2011

Oatley, M. J., Kaucher, A. V., Racicot, K. E., & Oatley, J. M. (2011). Inhibitor of DNA Binding 4 Is Expressed Selectively by Single Spermatogonia in the Male Germline and Regulates the Self-Renewal of Spermatogonial Stem Cells in Mice. *Biology of Reproduction*, *85*(2), 347–356. https://doi.org/10.1095/biolreprod.111.091330

Onichtchouk, D., Aduroja, K., Belting, H.-G., Gnügge, 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 gonad differentiation in zebrafish. *Developmental Dynamics*, *228*(3), 393–404. https://doi.org/10.1002/dvdy.10392

Parajes, S., Griffin, A., Taylor, A. E., Rose, I. T., Miguel-Escalada, I., Hadzhiev, Y., Arlt, 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.

Park, K.-E., Kaucher, A. V., Powell, A., Waqas, M. S., Sandmaier, S. E. S., Oatley, M. J., Park, C.-H., Tibary, A., Donovan, D. M., Blomberg, L. A., Lillico, S. G., Whitelaw, C. B. A., Mileham, A., Telugu, B. P., & Oatley, J. M. (2017). Generation of germline ablated male pigs by CRISPR/Cas9 editing of the NANOS2 gene. *Scientific Reports*, *7*(December 2016), 40176. https://doi.org/10.1038/srep40176

Pepling, M. E., Cuevas, M., & Spradling, A. C. (1999). Germline cysts: A conserved phase of germ cell development? *Trends Cell Biol*, *9*, 257–262.

Prasasya, R. D., & Mayo, K. E. (2018). Notch Signaling Regulates Differentiation and Steroidogenesis in Female Mouse Ovarian Granulosa Cells. *Endocrinology*, *159*, 184–198.

Qin, M., Zhang, Z., Song, W., Wong, Q. W.-L., Chen, W., Shirgaonkar, N., & Ge, W. (2018). Roles of Figla/figla in Juvenile Ovary Development and Follicle Formation During Zebrafish Gonadogenesis. *Endocrinology*, *159*(11), 3699–3722. https://doi.org/10.1210/en.2018-00648

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*(W1), 191–198.

Rodriguez-Mari, A., Canestro, C., BreMiller, R. A., Catchen, J. M., Yan, Y. L., & Postlethwait, J. H. (2013). Retinoic acid metabolic genes, meiosis, and gonadal sex differentiation in zebrafish. *PLoS One*, *8*, 73951.

Rodriguez-Mari, A., Yan, Y. L., Bremiller, R. A., Wilson, C., Canestro, C., & Postlethwait, J. H. (2005). Characterization and expression pattern of zebrafish Anti-Mullerian hormone (Amh) relative to sox9a, sox9b, and cyp19a1a, during gonad development. *Gene Expr Patterns*, *5*, 655–667.

Ruzicka, L., Howe, D. G., Ramachandran, S., Toro, S., Van Slyke, C. E., Bradford, Y. M., Eagle, A., Fashena, D., Frazer, K., & Kalita, P. (2019). The Zebrafish Information Network: New support for non-coding genes, richer Gene Ontology annotations and the Alliance of Genome Resources. *Nucleic Acids Res*, *47*, 867–873.

RyanKJ. (1979). Granulosa-thecal cell interaction in ovarian steroidogenesis. *J. Steroid Biochem*, *11*, 799–800.

Sada, A., Suzuki, A., Suzuki, H., & Saga, Y. (2009). The RNA-binding protein NANOS2 is required to maintain murine spermatogonial stem cells. *Science*, *325*(5946), 1394–1398.

Selman, K., Wallace, R. A., Sarka, A., & Qi, X. (1993). Stages of oocyte development in the zebrafish, Brachydanio rerio. *Journal of Morphology*, *218*(2), 203–224. https://doi.org/10.1002/jmor.1052180209

Shen, Y. C., Shami, A. N., Moritz, L., Larose, H., Manske, G. L., Ma, Q., Zheng, X., Sukhwani, M., Czerwinski, M., & Sultan, C. (2021). TCF21(+) mesenchymal cells contribute to testis somatic cell development, homeostasis, and regeneration in mice. *Nat Commun*, *12*, 3876.

Siebert, S., Farrell, J. A., Cazet, J. F., Abeykoon, Y., Primack, A. S., Schnitzler, C. E., & Juliano, C. E. (2019). Stem cell differentiation trajectories in Hydra resolved at single-cell resolution. *Science*, *365*(6451). https://doi.org/10.1126/science.aav9314

Siegfried, K. R., & Draper, B. W. (2020). Chapter 16—The Reproductive System. In S. C. Cartner, J. S. Eisen, S. C. Farmer, K. J. Guillemin, M. L. Kent, & G. E. Sanders (Eds.), *The Zebrafish in Biomedical Research* (pp. 151–164). Academic Press.

Siu, M. K., & Cheng, C. Y. (2012). The blood-follicle barrier (BFB) in disease and in ovarian function. *Adv Exp Med Biol*, *763*, 186–192. https://doi.org/10.1007/978-1-4614-4711-5_9.

Söllner, C., Burghammer, M., Busch-Nentwich, E., Berger, J., Schwarz, H., Riekel, C., & Nicolson, T. (2003). Control of Crystal Size and Lattice Formation by Starmaker in Otolith Biomineralization. *Science*, *302*(5643), 282–286. https://doi.org/10.1126/science.1088443

Song, X., Wong, M. D., Kawase, E., Xi, R., Ding, B. C., McCarthy, J. J., & Xie, T. (2004a). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. *Development*, *131*(6), 1353–1364.

Song, X., Wong, M. D., Kawase, E., Xi, R., Ding, B. C., McCarthy, J. J., & Xie, T. (2004b). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. *Development*, *131*, 1353–1364.

Song, X., Zhu, C.-H., Doan, C., & Xie, T. (2002). Germline stem cells anchored by adherens junctions in the Drosophila ovary niches. *Science*, *296*(5574), 1855–1857.

Soyal, S. M., Amleh, A., & Dean, J. (2000). FIGalpha, a germ cell-specific transcription factor required for ovarian follicle formation. *Development (Cambridge, England)*, *127*(21), 4645–4654.

Su, Y.-Q., Sugiura, K., & Eppig, J. J. (2009). Mouse Oocyte Control of Granulosa Cell Development and Function: Paracrine Regulation of Cumulus Cell Metabolism. *Seminars in Reproductive Medicine*, *27*(1), 32–42. https://doi.org/10.1055/s-0028-1108008

Subramaniam, K., & Seydoux, G. (1999). Nos-1 and nos-2, two genes related to Drosophila nanos, regulate primordial germ cell development and survival in Caenorhabditis elegans. *Development (Cambridge, England)*, *126*(21), 4861–4871.

Suzuki, A., Saba, R., Miyoshi, K., Morita, Y., & Saga, Y. (2012). Interaction between NANOS2 and the CCR4-NOT Deadenylation Complex Is Essential for Male Germ Cell Development in Mouse. *PLoS ONE*, *7*(3), e33558. https://doi.org/10.1371/journal.pone.0033558

Suzuki, A., Saga, Y., Suzuki, A., & Saga, Y. (2008). *Nanos2 suppresses meiosis and promotes male germ cell differentiation service Nanos2 suppresses meiosis and promotes male germ cell differentiation*. 430–435. https://doi.org/10.1101/gad.1612708

Takahashi, H. (1977). Juvenile Hermaphroditism in the zebrafish, Brachydanio rerio. *Bull Fac Fish Hokkaido Univ*, *28*, 57–65.

Tang, Q., Iyer, S., Lobbardi, R., Moore, J. C., Chen, H., Lareau, C., Hebert, C., Shaw, M. L., Neftel, C., & Suva, M. L. (2017). Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing. *J Exp Med*, *214*, 2875–2887.

Tourtellotte, W. G., Nagarajan, R., Auyeung, A., Mueller, C., & Milbrandt, J. (1999). Infertility associated with incomplete spermatogenic arrest and oligozoospermia in Egr4-deficient mice. *Development*, *126*(22), 5061–5071. https://doi.org/10.1242/dev.126.22.5061

Villanueva, R. A. M., & Chen, Z. J. (2019).

Wagner, M., Yoshihara, M., Douagi, I., Damdimopoulos, A., Panula, S., Petropoulos, S., Lu, H., Pettersson, K., Palm, K., & Katayama, S. (2020). Single-cell analysis of human ovarian cortex identifies distinct cell populations but no oogonial stem cells. *Nat Commun*, *11*, 1147.

Wang, Y., Pan, L., Moens, C. B., & Appel, B. (2014). Notch3 establishes brain vascular integrity by regulating pericyte number. *Development*, *Jan;141(2):307-17*.

Wang, Z., & Lin, H. (2004). Nanos maintains germline stem cell self-renewal by preventing differentiation. *Science (New York, N.Y.)*, *303*(5666), 2016–2019. https://doi.org/10.1126/science.1093983

Weidmann, C. A., Qiu, C., Arvola, R. M., Lou, T.-F., Killingsworth, J., Campbell, Z. T., Hall, T. M. T., & Goldstrohm, A. C. (2016). Drosophila Nanos acts as a molecular clamp that modulates the RNA-binding and repression activities of Pumilio. *ELife*, *5*, e17096.

Westerfield, M. (n.d.). *THE ZEBRAFISH BOOK, 5th Edition; A guide for the laboratory use of zebrafish (Danio rerio)*. University of Oregon Press.

Whitesell, T. R., Kennedy, R. M., Carter, A. D., Rollins, E.-L., Georgijevic, S., Santoro, M. M., & Childs, S. J. (2014). An α-Smooth Muscle Actin (acta2/αsma) Zebrafish Transgenic Line Marking Vascular Mural Cells and Visceral Smooth Muscle Cells. *PLOS ONE*, *9*(3), e90590. https://doi.org/10.1371/journal.pone.0090590

Wierson, W. A., Welker, J. M., Almeida, M. P., Mann, C. M., Webster, D. A., Torrie, M. E., Weiss, T. J., Kambakam, S., Vollbrecht, M. K., & Lan, M. (2020). Efficient targeted integration directed by short homology in zebrafish and mammalian cells. *Elife*, *9*.

Winterhager, E., & Kidder, G. M. (2015). Gap junction connexins in female reproductive organs: Implications for women's reproductive health. *Hum Reprod Update*, *21*, 340–352.

Xie, J., Wang, W. Q., Liu, T. X., Deng, M., & Ning, G. (2008). Spatio-temporal expression of chromogranin A during zebrafish embryogenesis. *J Endocrinol*, *198*, 451–458.

Xie, T. (2008). Germline stem cell niches. *StemBook*, 1–18. https://doi.org/10.3824/stembook.1.23.1

Xie, T., & Spradling, A. C. (2000). A Niche Maintaining Germ Line Stem Cells in the Drosophila Ovary. *Science*, *290*(5490), 328–330. https://doi.org/10.1126/science.290.5490.328

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*, 134. https://doi.org/10.1186/1471-2105-13-134

Yoshida, K., Kondoh, G., Matsuda, Y., Habu, T., Nishimune, Y., & Morita, T. (1998). The Mouse RecA-like Gene Dmc1 Is Required for Homologous Chromosome Synapsis during Meiosis. *Molecular Cell*, *1*(5), 707–718. https://doi.org/10.1016/S1097-2765(00)80070-2

Zakrzewska, A., Cui, C., Stockhammer, O. W., Benard, E. L., Spaink, H. P., & Meijer, A. H. (2010). Macrophage-specific gene functions in Spi1-directed innate immunity. *Blood*, *116*(3), e1–e11. https://doi.org/10.1182/blood-2010-01-262873

Chapter 4 Closing remarks and future directions

The dissertation work presented here provided a comprehensive examination of the cellular and expression diversity of the juvenile zebrafish ovary. In chapter 2, we identified the expression progression of the germ cells from GSC to early-stage oocytes and found multiple GSC-specific markers and putative transcriptional regulators, such as *bhlhe23*. Both the expression domain and the gene enrichment analysis supported that *bhlhe23* may have a regulatory function for GSC. In the future, knockout studies of *bhlhe23* should be performed to determine if *bhlhe23* is essential for GSC development and identify its function.

We also identified and validated the first reported marker for the GC progenitor population in zebrafish ovary, *foxl2l*, and showed that it is crucial for female development through genetic perturbation. Although we observed the pronounced phenotype of foxl2l mutant, the down-stream signals that foxl2l regulates is still unknown. Future studies will need to determine the mechanism by which foxl2 regulates GC progenitors.

In chapter 3, we identified all major somatic populations of the juvenile zebrafish ovary and discovered the diverse subpopulations of the follicle and stromal cells. Two populations were particularly interesting. We found the *stm*-expressing stromal cells specifically localized and attached at the germinal zone, where the putative GSC niche is located, and most of the early mitotic germ cells can be found. We also found that all early germ cells ranging from GSCs to early meiotic GCs co-localized to the *lhx9*-expressing pre-follicle cells and form track-like features that crisscross the ovary. The characteristics of these two populations of cells indicated they may play a role in early germ cell regulation and function as a GSC niche. It would be

practically interesting to dissect how these cells affect GSC development. An inducible cell ablation system could be used to conditionally eliminate those cells to determine if they are necessary to maintain the GSC and then identify the signaling pathway that those cells use to regulate the GSC. Based on how the GSC niche signaling works in other organisms, a ligand(s) is likely expressed in those cells to signal to the GSCs and maintain their stem cell identity.

In conclusion, this dissertation revealed novel cell types and signaling in the zebrafish ovary. The single-cell RNA sequencing data generated in this study will provide a valuable resource for the community to used for further discoveries.