Wu et al. *Zool. Res.* 2023, 44(2): 269−272
https://doi.org/10.24272/j.issn.2095-8137.2022.312

ZR | Zoological Research

Letter to the editor Open Access

Single-cell RNA sequencing reveals gonadal dynamic changes during sex differentiation in hermaphroditic protogynous orange-spotted grouper (*Epinephelus coioides*)

DEAR EDITOR,

Sex differentiation is a complex process that requires many factors to regulate gonadal proliferation, differentiation, development, and organization. In teleosts, the molecular mechanisms of sex differentiation are diverse and unclear, especially in hermaphrodites. In the present study, single-cell RNA sequencing (scRNA-seq) was used to identify 15 cell types, including germ, follicle, and theca cells, in the gonads during sex differentiation in the hermaphroditic orange-spotted grouper (*Epinephelus coioides*). Two pre-follicle cell types were defined, and the differentiation trajectories of follicle cells were outlined. Notably, male-related genes were highly expressed in both pre-follicle and follicle cells, including *amh*, *sox9*, and *dmrt3* in pre-follicle cells and *dmrt1* in follicle cells. Oocytes exhibited two distinct states, with high expression of oocyte development-related genes in one state and spermatogenesis-related genes in the other state. Our findings provide novel insights into cell type and lineage tracing in the gonads during sex differentiation in a hermaphroditic species.

The reproductive strategies of teleosts are complex, including both gonochoristic and hermaphroditic species. Hermaphroditic fish can reverse sex during their life cycle, with the capability of producing mature male and female gametes within a single individual. Female-first maturation is defined as protogynous and male-first maturation is defined as protandry (Nagahama et al., 2021). The orange-spotted grouper (*E. coioides*), an important maricultural species in Asia, is a protogynous hermaphrodite and considered a good model for studying sex differentiation and sex reversal (Liu & De Mitcheson, 2009). Most orange-spotted groupers will develop into females after sex differentiation, with only a small proportion developing into males. The origin of germ cells and the reasons for this preference toward female-first maturation remain unknown and have long been of interest to scientists. Many genes are involved in sex differentiation (Han et al., 2019; Huang et al., 2019; Sun et al., 2017), but the mechanism underlying sex differentiation in groupers remains unclear. Therefore, to better understand how sex differentiation is regulated at the cellular level, we performed scRNA-seq and revealed gene expression patterns in sex differentiation.

We collected gonads from orange-spotted groupers at three developmental stages: undifferentiated gonads (80 days post-hatching (dph)), differentiating gonads (120 dph), and differentiated gonads (180 dph). Histological examination was conducted to identify the time point of sex differentiation. In the 80 dph grouper gonads, blood vessels had emerged, a hallmark of paired gonads (Figure 1A), and primordial germ cell (PGC)-like cells were found due to their unique morphology (Porras-Gómez et al., 2021). In the 120 dph grouper gonads, several germ cells were detected, and the ovarian cavity was formed (Figure 1A). In the 180 dph grouper gonads, primary growth stage oocytes were observed, indicating completion of sex differentiation (Figure 1A). Subsequently, we prepared single-cell suspensions from mixed and freshly dissected gonads at 80 dph (n=30), 120 dph (n=20), and 180 dph (n=5), each with two biological replicates (Figure 1B). In the sequencing data, 83 160 cells were profiled with approximately 37 000 reads per cell, and the median number of genes detected per cell was 1 112 (Supplementary Table S1 and Figure S1).

To further define the major cell types of the six samples, canonical correlation analysis (CCA, Stuart et al., 2019) was used to minimize batch effects, and all cells were combined from the three developmental stages. The aggregated cells were clustered into 25 clusters using uniform manifold approximation and projection (UMAP, Figure 1C; Supplementary Figure S2A), and cell distribution was changed in the three developmental stages (Supplementary Figure S2B). In our experiment, all clusters were divided into two classes, i.e., parenchymal and stromal cells. Parenchymal cells were further divided into follicle, theca, and germ cells based on the following marker genes: follicle cells (*gsdf*, *amh*, and *fshr*), theca cells (*cyp11a1*, *cyp17a1*, and *star*), gonocytes (*ddx4*, *nanos2*, *piwil1*, and *piwil2*), and oocytes (*h1foo*, *zar1*, and *lhx8*). As stromal cells accounted for over 90% of the total detected cells, they were further divided into four populations, i.e., vasculature, epithelial, immune, and uncharacterized stromal cells. The following marker genes were used to

Received: 02 November 2022; Accepted: 05 January 2023; Online: 07 January 2023

Foundation items: This work was supported by the National Key Research and Development Program (2018YFD0900203), National Natural Science Foundation of China (32273132, 32202928), China Agriculture Research System of MOF and MARA (CARS-47), Project of Hainan Academician Team Innovation Center (YSPTZX202122), Fundamental Research Funds for the Central Universities, Sun Yat-sen University (22qntd2614), and Science and Technology Program of Guangdong Provincial (2020A1515011477)

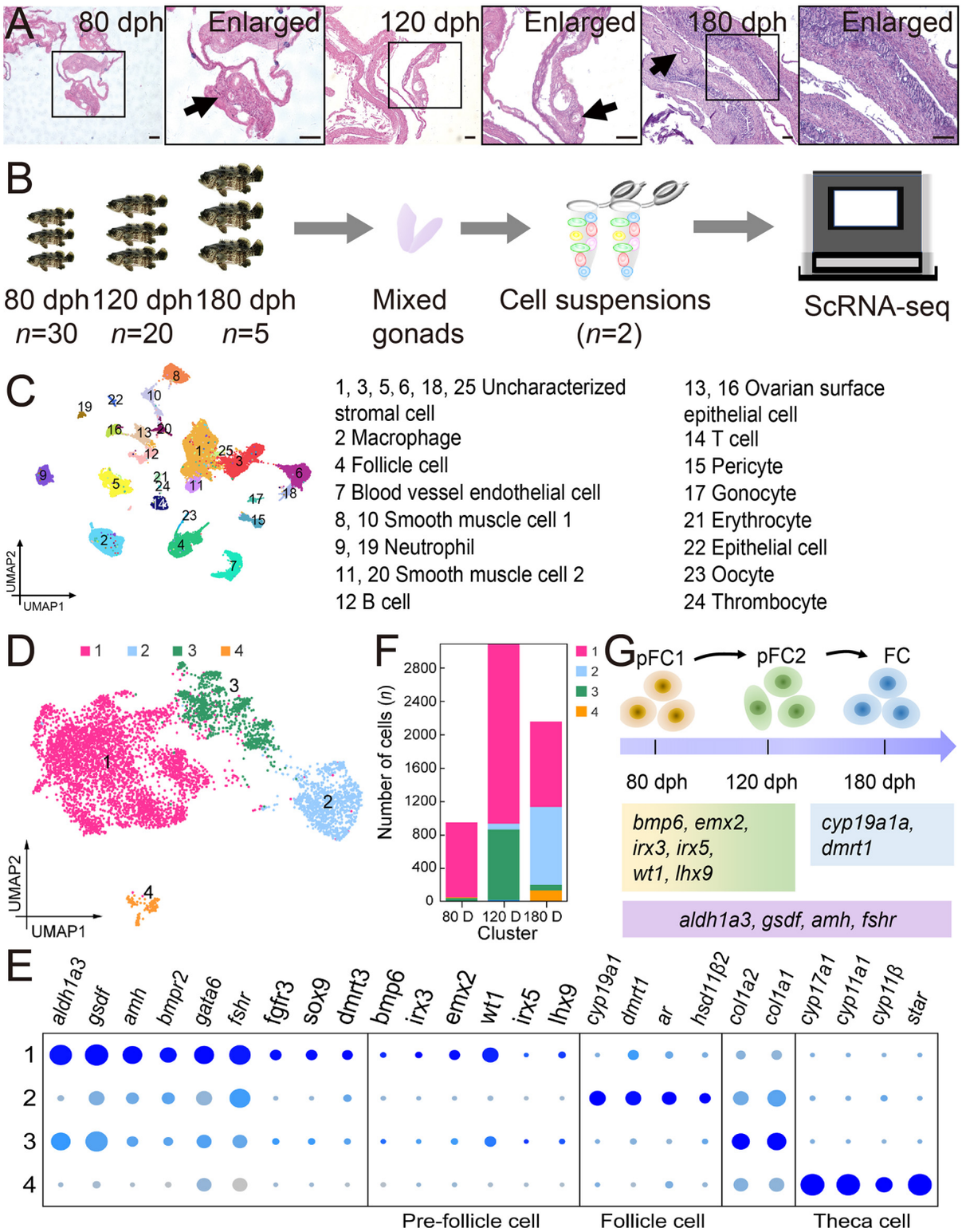


Figure 1 ScRNA-seq reveals dynamic changes during sex differentiation in orange-spotted grouper gonads

A: Gonadal histological morphology of orange-spotted grouper by hematoxylin-eosin (H&E) staining at 80 dph, 120 dph, and 180 dph. Enlarged figure shows higher magnification view of boxed area in previous figure. Black arrows represent blood vessels. Dph, day post-hatching. Scale bar: 50 μm. B: Overview of whole experiment. Fish were dissected from the cloacal aperture, with gonads and connective tissue then removed. Gonads were dissociated into single-cell suspensions for sequencing (*n*=2) and analysis. C: Visualization of 15 cell types among six samples in UMAP plot. Colors mark 25 distinct clusters identified by unsupervised clustering and characterized with differential gene expression and gene set enrichment analyses. Each dot represents a cell. D: Cluster 4 subcluster UMAP plot, with cells color-coded by computationally determined cell subtypes. E: Gene expression plots of selected genes. Cells expressing indicated genes are in blue, and relative intensity indicates relative expression levels. F: Histogram of distribution of cell clusters in three developmental stages. G: Schematic representation of potential differentiation of cluster 4 subclusters from 80 dph to 180 dph. FC, follicle cell; pFC, pre-follicle cell.

classify stromal cell subgroups: (1) vasculature cells: blood vessel endothelial cells (*vmf* and *fli1*), smooth muscle cells (*mylk*, *myh11*, *des*, and *smtnl1*), pericytes cells (*notch3*); (2) epithelial cells: ovarian surface epithelium (*stm* and *gata3*), epithelial cells (*fabp* and *sox10*); (3) immune cells: macrophages (*csf1r2* and *mpeg1*), neutrophils (*lce* and *cebpe*), B cells (*iglc5* and *iglc6*), T cells (*gfi1* and *il2rb*), erythrocytes (*hba* and *hbad*), thrombocytes (*thbs1* and *mpl*); and (4) uncharacterized stromal cells (*col1a1a* and *dcn*) (Supplementary Figure S3 and Supplementary Materials).

Cluster 17 cells were identified as gonocytes based on the expression of germ cell marker genes *ddx4*, *piwil1*, and *piwil2* and germ stem cell marker gene *nanos2* (Supplementary Figure S3). To improve the reliability of the dataset, *Ddx4* and *Piwil1* antibodies were used for immunohistochemical (IHC) analysis during sex differentiation (Supplementary Figure S4A–L), and proliferating cell nuclear antigen (PCNA) was used to detect the proliferative potential of the gonads (Supplementary Figure S4M–R). Detailed information on the antibodies used is provided in Supplementary Table S2. Results showed that *Ddx4* signals were located in all germ cells, especially in the PGC-like cells in 80 dph gonads (Supplementary Figure S4A, D). *Piwil1* signals were also detected in germ cells, especially in oogonia and oocytes in 180 dph gonads (Supplementary Figure S4I, L). PCNA signals were detected in the nuclei of PGC-like cells and oogonia, indicating that these cells had more proliferative activity (Supplementary Figure S4P, R). In addition, three encoded Tudor domain-containing (TDRD) protein genes (*tdrd6*, *tdrd7b*, and *tdrd9*) showed specific expression in the gonocytes, and may therefore be potential markers of gonocytes (Supplementary Figure S5). These results should facilitate future characterization and profiling of gonocytes.

We also analyzed the expression trends of all genes during sex differentiation when comparing gonocytes in the three developmental stages. We identified 2 555 genes that showed consistent increases across the three developmental stages (profile 7, Supplementary Figure S6A), including many reproductive and developmental genes, such as *ddx4*, *piwil1*, *piwil2*, *tdrd6*, *tdrd7*, and *tdrd9* (Supplementary Figure S6B). Profile 7 genes were mainly enriched in DNA replication and mTOR signaling pathway (Supplementary Figure S6C).

Cluster 23 cells were identified as oocytes based on the expression of many oocyte development-related genes (Supplementary Figure S3). However, trajectory analysis revealed that a small portion of cells in cluster 23 expressed many spermatogonia proliferation and differentiation-related genes, such as *zbtb16*, etv5, *sox9*, and *tdrd5* (Supplementary Figure S7; Supplementary Materials). These cells exhibited a distinct developmental state, suggesting that oogonia may possess differential potencies.

Cluster 4 cells were initially identified as follicle cells due to the high expression of follicle cell markers. To explore cluster 4 subtypes during sex differentiation, we performed re-clustering analysis and revealed the existence of four subclusters (sc4.1, sc4.2, sc4.3, and sc4.4; Figure 1D). Sc4.1, sc4.2, and sc4.3 showed high expression of many follicle cell markers, including *amh*, *fshr*, and *gsdf*. However, both sc4.1 and sc4.3 expressed pre-follicle cell-specific genes (*bmp6*, *emx2*, *irx3*, *irx5*, *wt1*, and *lhx9*), as reported in zebrafish (Liu et al., 2022). Sc4.3 showed a similar gene expression pattern to sc4.1, as well as enrichment of collagen-related genes, such as *col1a1* and *col1a2* (Figure 1E). Therefore, sc4.1 and sc4.3 cells were identified as pre-follicle cells (pFC1 and pFC2, respectively), while sc4.2 cells were identified as follicle cells with high *cyp19a1a* expression in the early ovary, consistent with previous research (Dranow et al., 2016; Nakamura et al., 2009). Notably, both pre-follicle and follicle cells showed high expression of male-related genes, including *amh*, *dmrt3*, and *sox9* in pre-follicle cells and *dmrt1* in follicle cells. In addition, *cyp11a1*, *cyp17a1*, and *star* were all expressed in sc4.4 of cluster 4, and thus sc4.4 cells were identified as theca cells (TC, Figure 1E; Supplementary Materials).

Our results showed that pFC1 was the dominant cell type at 80 dph, pFC1 and pFC2 were the dominant cell types at 120 dph, and all four subclusters were present at 180 dph (Figure 1F). Depending on the emergence points of the three subclusters, we predicted the developmental trajectories of the follicle cells. The earliest cell type (pFC1) may be a precursor of follicle cells, while pFC2 may be an intermediate state between pFC1 and follicle cells (Figure 1G).

In addition to its expression in follicle cells, *cyp19a1* was also expressed in uncharacterized stromal cells (cluster 3). Cluster 3 was divided into two subclusters (sc3.1 and sc3.2) based on different marker genes and time of occurrence (Supplementary Figure S8A, B). Sc3.1 specifically expressed *cldn7a*, *tnfsf15*, *tacstd2*, *ccxl4*, *f2rl1*, and *fndc5*, whereas sc3.2 strongly expressed *cyp19a1a*, *fshr*, *gsdf*, and *foxl2* (Supplementary Figure S8C, D; Supplementary Materials), indicating that stromal cells in sc3.2 may also participate in hormone production. These results are consistent with previous research on zebrafish (Liu et al., 2022). We also outlined the developmental trajectories of cluster 3 during sex differentiation. First, a small number of progenitor cells from 80 dph preferentially developed into sc3.1 at 120 dph. Subsequently, number of sc3.1 cells decreased and number of sc3.2 cells increased sharply at 180 dph (Supplementary Figure S8E). The reason for this change may be the transition from sc3.1 to sc3.2 from 120 dph to 180 dph.

In this study, we performed the first scRNA-seq analysis of the orange-spotted grouper, providing a precise map of the gonads during sex differentiation. From the scRNA-seq data, we identified 15 cell types in the early developing gonads, as well as various novel cell markers. Furthermore, we defined two pre-follicle cell types and one follicle cell type and confirmed the order in which they appear during sex differentiation. In addition, transcriptomic differences in gonocytes were revealed across the three developmental stages. This study provides an important dataset and new avenues for research in sex differentiation. This study also reveals a previously unanticipated complexity of the developing gonad and provides a comprehensive resource for future systematic analysis of sex differentiation.

DATA AVAILABILITY

Datasets generated in this study are available in the GSA at the National Genomics Data Center (CRA004056 and CRA007271), NCBI Gene Expression Omnibus (GSE218835), and Science Data Bank (DOI: 10.57760/sciencedb.06672).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

X.C.L., Z.N.M., and X.W. conceived and supervised the study. X.W. and Y.Y. analyzed the data. C.Y.Z., T.W., and Y.H.T., performed the experiments. X.W., C.Y.Z., and Y.Y. wrote the manuscript. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

We are grateful to Guangzhou Genedenovo Biotechnology Company, Ltd. for assistance in sequencing and bioinformatics analysis.

Xi Wu[1], Chao-Yue Zhong[1], Yang Yang[1], Tong Wang[1], Yu-Hao Tao[1], Zi-Ning Meng[1,2], Xiao-Chun Liu[1,2,*]

[1] *State Key Laboratory of Biocontrol*, *Guangdong Province Key Laboratory for Improved Variety Reproduction of Aquatic Economic Animals*, *Institute of Aquatic Economic Animals*, *School of Life Sciences*, *Sun Yat-sen University*, *Guangzhou*, *Guangdong* 510275, *China*

[2] *Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai)*, *Zhuhai*, *Guangdong* 519000, *China*

*Corresponding author, E-mail: lsslxc@mail.sysu.edu.cn

REFERENCES

Dranow DB, Hu K, Bird AM, et al. 2016. Bmp15 is an oocyte-produced signal required for maintenance of the adult female sexual phenotype in zebrafish. *PLoS Genetics*, **12**(9): e1006323.

Han YL, Zhao M, Wang L, et al. 2019. Overexpression of anti-müllerian hormone gene *in vivo* affects gonad sex differentiation in undifferentiated orange-spotted groupers (*Epinephelus coioides*). *Frontiers in Endocrinology*, **10**: 210.

Huang MW, Chen JX, Liu Y, et al. 2019. New insights into the role of follicle-stimulating hormone in sex differentiation of the protogynous orange-spotted grouper. *Epinephelus coioides. Frontiers in Endocrinology*, **10**: 304.

Liu M, De Mitcheson YS. 2009. Gonad development during sexual differentiation in hatchery-produced orange-spotted grouper (*Epinephelus coioides*) and humpback grouper (*Cromileptes altivelis*) (Pisces: Serranidae, Epinephelinae). *Aquaculture*, **287**(1–2): 191–202.

Liu YL, Kossack ME, McFaul ME, et al. 2022. Single-cell transcriptome reveals insights into the development and function of the zebrafish ovary. *eLife*, **11**: e76014.

Nagahama Y, Chakraborty T, Paul-Prasanth B, et al. 2021. Sex determination, gonadal sex differentiation, and plasticity in vertebrate species. *Physiological Reviews*, **101**(3): 1237−1308.

Nakamura S, Kurokawa H, Asakawa S, et al. 2009. Two distinct types of theca cells in the medaka gonad: germ cell-dependent maintenance of *cyp19a1*-expressing theca cells. *Developmental Dynamics*, **238**(10): 2652−2657.

Porras-Gómez TJ, Villagrán-SantaCruz M, Moreno-Mendoza N. 2021. Biology of primordial germ cells in vertebrates with emphasis in urodeles amphibians. *Molecular Reproduction and Development*, **88**(12): 773−792.

Stuart T, Butler A, Hoffman P, et al. 2019. Comprehensive integration of single-cell data. *Cell*, **177**(7): 1888−1902.e21.

Sun ZH, Wang Y, Lu WJ, et al. 2017. Divergent expression patterns and function implications of four nanos genes in a hermaphroditic fish. *Epinephelus coioides. International Journal of Molecular Sciences*, **18**(4): 685.

ZR Zoological Research

Supplementary Materials

Single-cell RNA sequencing reveals gonadal dynamic changes during sex differentiation in hermaphroditic protogynous orange-spotted grouper (*Epinephelus coioides*)

Xi Wu [1], Chao-Yue Zhong [1], Yang Yang [1], Tong Wang [1], Yu-Hao Tao [1], Zi-Ning Meng [1, 2], Xiao-Chun Liu [1,2,*]

1. State Key Laboratory of Biocontrol, Guangdong Province Key Laboratory for Improved Variety Reproduction of Aquatic Economic Animals, Institute of Aquatic Economic Animals, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong 510275, China

2. Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, Guangdong 519000, China

[*]Corresponding author, E-mail: lsslxc@mail.sysu.edu.cn

Supplementary Materials and Methods

Animals

Orange-spotted groupers were raised in the Guangdong Daya Bay Fishery Development Center (Huizhou 516081, Guangdong, China). The fish were kept in indoor pools under controlled water temperature (22.7–27.8 °C). Fish gonads were sampled at different periods, i.e., 80 days post-hatching (dph), 120 dph, and 180 dph, to confirm developmental stages based on histological morphology and to prepare single-cell suspensions. All fish were anesthetized with MS222 and sacrificed. All animal experiments were conducted following the guidelines and approval of the Animal Research and Ethics Committees of Sun Yat-sen University (Approval No. SYSU-IACUC-2022-B0123).

Hematoxylin-eosin (H&E) staining

Fresh gonads were fixed in Bouin's solution for 24 h. Afterwards, the tissues were orderly dehydrated using gradient alcohol and embedded in paraffin. The tissues were further cut into 5 μm thick slices. Lastly, the nucleus and cytoplasm were stained using H&E, respectively, and the sections were observed by light microscopy (Nikon IQ50, Japan) to classify gonadal stages.

Preparation of gonadal single-cell suspension

A mixture of 30 gonad tissue samples at 80 dph was dissociated for single-cell transcriptomics with two replicates. With development, 20 gonad tissue samples at 120 dph and five gonad tissue samples at 180 dph were mixed and dissociated for single-cell RNA sequencing (scRNA-seq). Briefly, gonad samples were cut into pieces and twice cleaned using 1×Hank's Balanced Salt Solution (Hyclone, USA). The samples were then centrifuged at 1 000 ×*g* for 3 min and incubated with 500 U/mL collagenase IV (diluted with dulbecco's phosphate-buffered saline buffer (PBS) without calcium and magnesium ions) at 37 °C for 2 h. The samples were centrifuged at 1 000 ×*g* for 3 min, followed by filtration through a 40 μm strainer (Corning, USA). Cell numbers in the suspensions were calculated using a blood counting chamber, with the cells then stained using 0.4% trypan blue to estimate cell viability. The suspensions (cell viability >85%) were then used for scRNA-seq.

Library preparation and scRNA-seq

ScRNA-seq libraries were constructed using Chromium Next GEM Single-Cell 3'Reagent Kit v3.1 (10X Genomics, USA). The prepared suspensions were used to generate Single-Cell gel beads-in-emulsions (GEMs) by loading onto a 10X Genomics GemCode Single-Cell instrument. Upon dissolution of Single-Cell 3' gel beads in GEMs, primers containing (i) Illumina® R1 sequence (Read 1 sequencing primer), (ii) 16 nt 10X barcode, (iii) 10 nt unique molecular identifier (UMI), and (iv) poly-dT primer sequence were released and mixed with the cell lysate and Master Mix. The GEMs were incubated to produce barcoded full-length cDNA from polyadenylated mRNA. The cDNA was then amplified by polymerase chain reaction (PCR) to generate sufficient mass for library construction. R1 (Read 1 primer sequence) was added to the molecules during GEM incubation. P5, P7, sample index, and R2 (Read 2 primer sequence) were added during library construction via end repair, A-tailing, adaptor

ligation, and PCR. The final libraries contained the P5 and P7 primers used for Illumina bridge amplification.

The Single-Cell 3' protocol produced Illumina-ready sequencing libraries. The Single-Cell 3' library comprised standard Illumina paired-end constructs, which began and ended with P5 and P7. The Single-Cell 3' 16 bp 10X barcode and 10 bp UMI were encoded in Read 1, while Read 2 was used to sequence the cDNA fragment. Sample index sequences were incorporated as the i7 index read. Read 1 and Read 2 were standard Illumina® sequencing primer sites used in paired-end sequencing. All raw scRNA-seq data were stored in the Genome Sequence Archive (https://ngdc.cncb.ac.cn/gsa/) under accession numbers CRA004056 and CRA007271.

Data quality control and gene expression quantify

Raw BCL files were converted to FASTQ files using Cell Ranger (v5.0). After mapping to the orange-spotted grouper genome (unpublished data), reads that uniquely intersected at least 50% of an exon were considered for UMI counting. Valid barcodes were identified using the EmptyDrops method (Lun et al., 2019). Seurat (v3.1.1) was used to perform quality control (QC), analysis, and exploration of scRNA-seq data. Cells with ≥17 000 UMIs, ≥10% mitochondrial genes, ≤230 detected genes, or ≥2 200 detected genes were filtered out. After removing unwanted cells from the dataset, global-scaling normalization ("LogNormalize") was used to normalize gene expression measurements for each cell based on total expression, which were then multiplied by a scale factor (10 000 by default) and log-transformed using gene expression level=log(1+(UMI A÷UMI Total)×10 000).

Canonical correlation analysis (CCA) was used to minimize batch effects (Stuart et al., 2019). The scaled expression matrix was used for dimensionality reduction by principal component analysis (PCA). Significant principal components (PCs) were identified by strong enrichment of genes with low *P*-values, with the cells then clustered using Seurat depending on PCA scores (Chung and Storey, 2015). To cluster cells, the modularity optimization technique selective mapping (SLM) algorithm was applied to iteratively group cells together and optimize standard modularity function (Blondel et al., 2008). In this experiment, we selected a resolution of 0.2 as the clustering parameter to identify 25 clusters. UMAP was then performed to visualize the data in two-dimensional space.

Differential expression analysis

To identify enriched genes in a specific cluster, the median expression patterns of all clusters were calculated, and the expression value of each gene in the cluster was compared to the other cells using the Wilcoxon rank-sum test (Camp et al., 2017).

Significantly up-regulated genes were identified using several criteria, including 1) $P \leq 0.01$; 2) >1.28-fold higher than expression in other cluster; and 3) genes expressed in over 25% of cells in the target cluster.

Pathway enrichment analysis identified significantly enriched metabolic or signal transduction pathways in differentially expressed genes (DEGs) compared to the whole genome background. *P*-values were false discovery rate (FDR) corrected, taking FDR≤0.05 as a threshold. Pathways satisfying this condition were defined as significantly enriched pathways in DEGs.

Single-cell trajectory analysis

Single-cell trajectories were analyzed using Monocle (v2.6.4) with a matrix of cells and gene expression levels. Monocle reduced the space to two dimensions and ordered the cells (sigma=0.001, lambda=NULL, param.gamma=10, tol=0.001). Once the cells were ordered, the trajectory was visualized in the reduced dimensional space.

Immunohistochemical analysis

Gonads from orange-spotted groupers (150 dph) were sampled and embedded in optimal cutting temperature compound (Sakura, Japan). Fast-cooled tissue blocks were sectioned into serial slides (5 μm thick). Immunohistochemical analysis was performed using a previously modified method (Wu et al., 2020). Briefly, frozen gonad sections were rehydrated in PBS for 30 min at room temperature. After antigen retrieval, the sections were treated with hydrogen peroxide for 30 min, blocked with 5% bovine serum albumin (BSA, Sigma, Germany) in PBS, then incubated with primary antibodies of Vasa, Piwi, and PCNA (1:100 diluted with 2% BSA in PBS) overnight at 4 °C. After washing several times with PBS, the sections were blocked with 5% goat serum in PBS for 1 h, then incubated at room temperature for 1 h with horseradish peroxidase (HRP)-conjugated anti-mouse/rabbit IgG secondary antibodies (Boster, China) at 1:2 000 dilution with 2% goat serum in PBS. Sources of antibodies are listed in Supplementary Table S2. Finally, the sections were observed under a light microscope (Leica, TCS-SP5, Germany).

REFERENCES

Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*: *Theory and Experiment*, **2008**: P10008.

Camp JG, Sekine K, Gerber T, Loeffler-Wirth H, Binder H, Gac M, et al. 2017. Multilineage communication regulates human liver bud development from pluripotency. *Nature*, **546**(7659): 533–538.

Chung NC, Storey JD. 2015. Statistical significance of variables driving systematic variation in high-dimensional data. *Bioinformatics*, **31**(4): 545–554.

Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, Marioni JC. 2019. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biology*, **20**(1): 63.

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck III WM, et al. 2019. Comprehensive integration of single-cell data. *Cell*, **177**(7): 1888–1902.e21.

Wu X, Yang Y, Zhong CY, Guo Y, Li SS, Lin HR, et al. 2020a. Transcriptome profiling of laser-captured germ cells and functional characterization of *zbtb40* during 17alpha-methyltestosterone-induced spermatogenesis in orange-spotted grouper (*Epinephelus coioides*). *BMC Genomics*, **21**(1): 73.

Wu X, Yang Y, Zhong CY, Guo Y, Wei TY, Li SS, et al. 2020b. Integration of ATAC-seq and RNA-seq unravels chromatin accessibility during sex reversal in orange-spotted grouper (*Epinephelus coioides*). *International Journal of Molecular Sciences*, **21**(8): 2800.

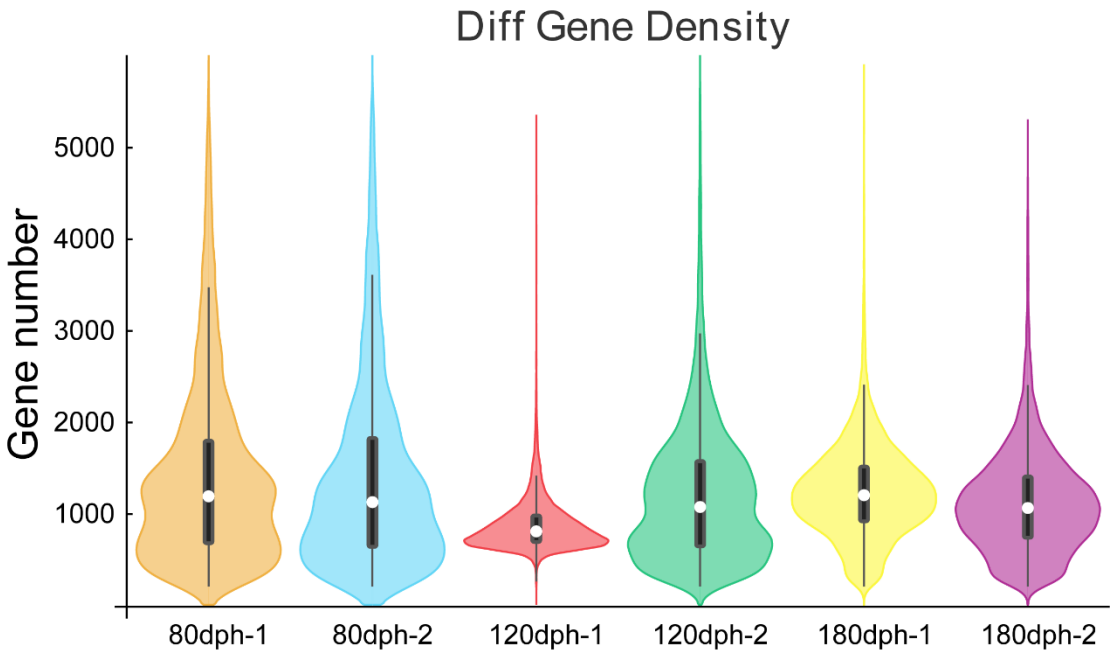
Supplementary Table S1 Overview quality of scRNA-seq data

Sample	Number of reads	Estimated number of cells	Mean reads per cell	Median genes per cell	Total genes detected	Reads mapped confidently to genome
80 dph-1	400,926,869	8,960	44,746	1,169	23,006	91.20%
80 dph-2	390,048,989	6,456	60,416	1,123	22,619	92.50%
120 dph-1	414,906,530	26,019	15,946	875	23,756	91.20%
120 dph-2	426,797,777	15,210	28,060	1,153	23,826	91.90%
180 dph-1	470,318,209	12,892	36,481	1,246	24,323	91.20%
180 dph-2	458,294,574	13,623	33,641	1,108	23,868	91.40%

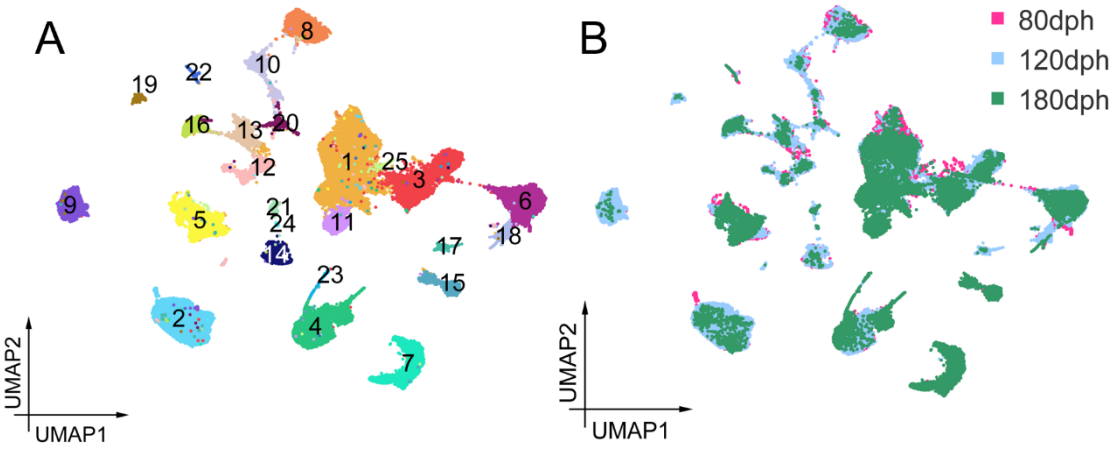
Supplementary Table S2 Key resources used in this paper

REAGENT or RESOURCE	Source	Identifier
Antibodies		
Rabbit anti-Vasa	Xu et al., 2005	N/A
Rabbit anti-Piwi	this paper	N/A
Rabbit anti-Pcna	Proteintech	10205-2-AP
Goat anti-mouse IgG secondary antibody, HRP	Boster	BA1050
Goat anti-rabbit IgG secondary antibody, HRP	Boster	BA1054
Chemicals, Peptides and Recombinant Proteins		
MS222	Sigma	W246719-1KG-K
Hematoxylin	Sigma	H3136-25G
Eosin	Sigma	230251-25G
Paraffin	Lecia	39601006
1×Hank's Balanced Salt Solution	Hyclone	SH30588.01
Collagenase IV	Sigma	C1889-50MG
Anti-Digoxigenin-AP Fab fragments	Roche	11093274910
Anti-Digoxigenin-POD, Fab fragments	Roche	11207733910
DAPI	Sigma	D9542

Optimal cutting temperature compound	Sakura	4583
Hydrogen peroxide	Thermofish	C20246
BSA	Sigma	A3311
Goat serum	Sigma	G9023
Critical Commercial Assays		
DIG RNA Labeling Kit	Roche	18587020
TSA fluorescein system	Roche	NEL701A001KT
Deposited Data		
Single cell RNA-seq data	GSA NCBI Science Data Bank	CRA004056, CRA007271 GSE218835 DOI: 10.57760/sciencedb.06672
Software and Algorithms		
STAR	Alexander Dobin et al., 2013	https://github.com/alexdobin/STAR
Seurat	Stuart et. Al. 2019	https://satijalab.org/seurat/v3.0/integratio n.html R Package Harmony Korsunsky et al., 2019 https://github.com/immunogenomics/harm ony

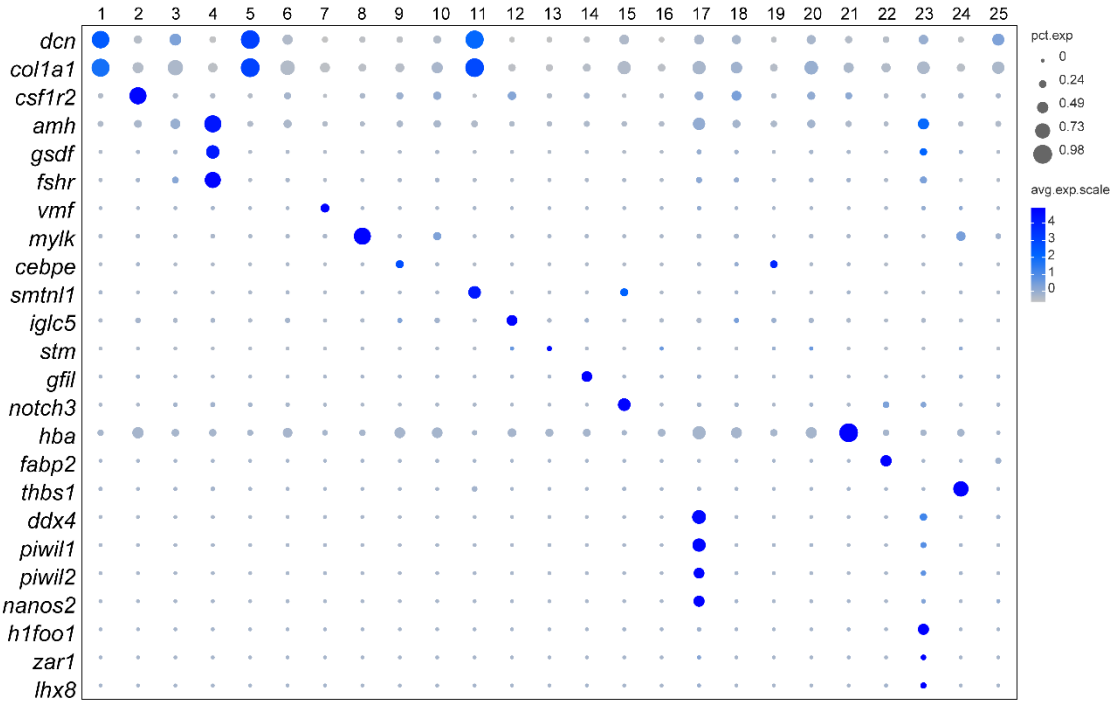


Supplementary Figure S1 Different gene densities of single-cell transcriptomic data in six samples



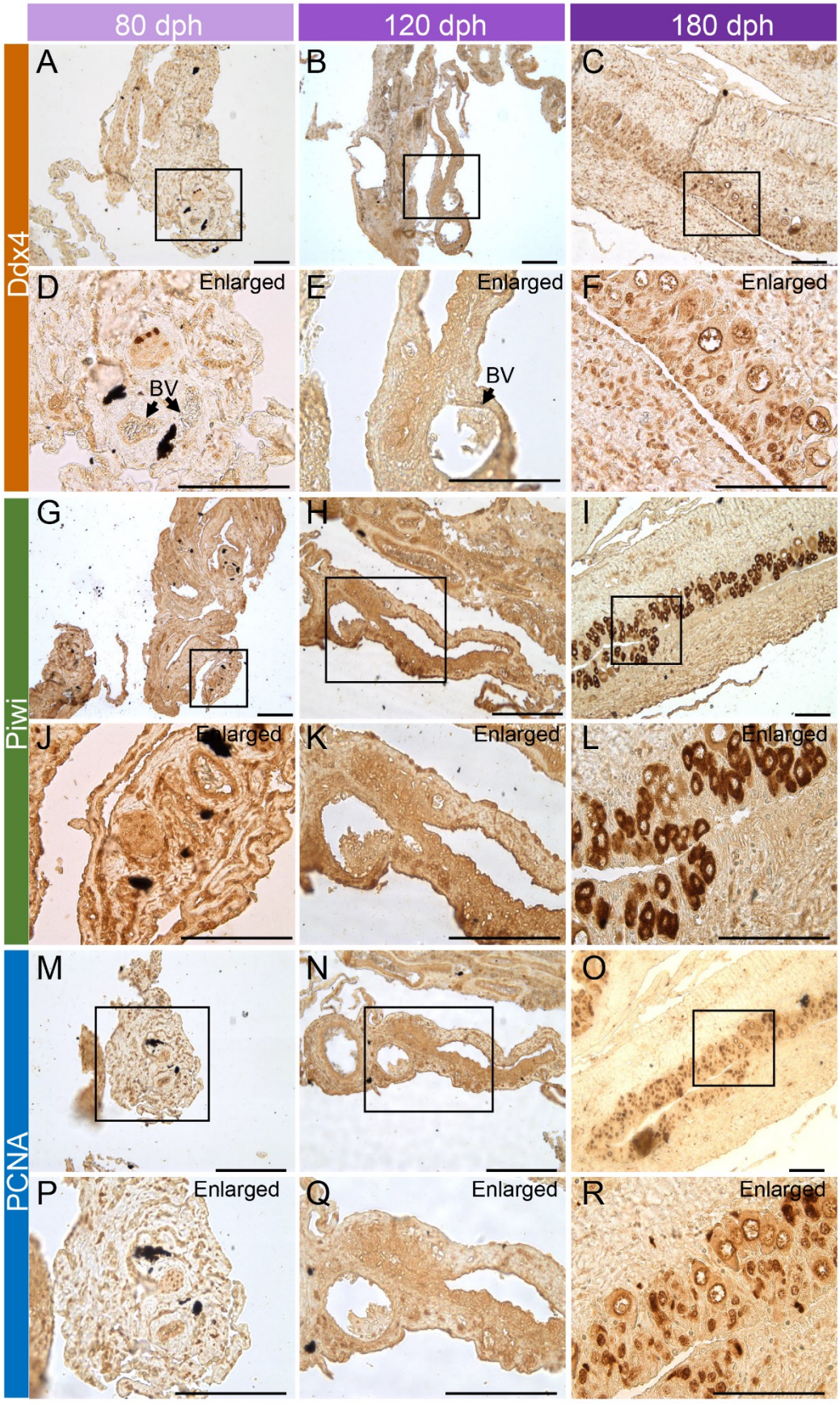
Supplementary Figure S2 UMAP plots of cell clusters in scRNA-seq dataset

A: UMAP plot showing 25 cell clusters in scRNA-seq dataset. **B:** UMAP plot showing changes in 25 cell clusters during three developmental stages.



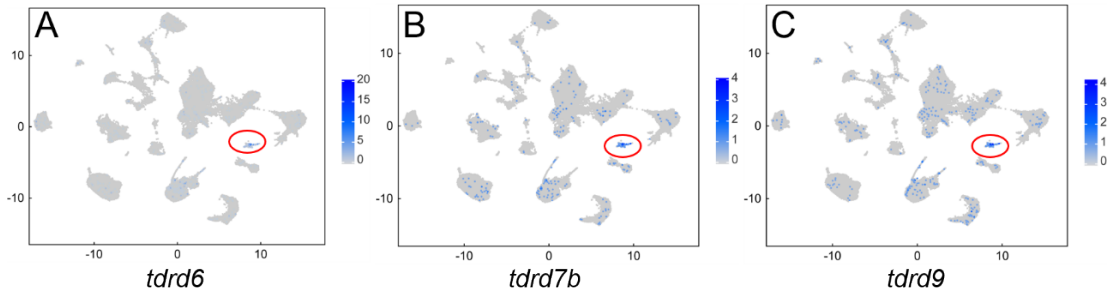
Supplementary Figure S3 Dot plot showing corresponding expression of cluster marker genes

Color indicates mean expression and dot size represents percentage of cells within the cluster expressing the marker.

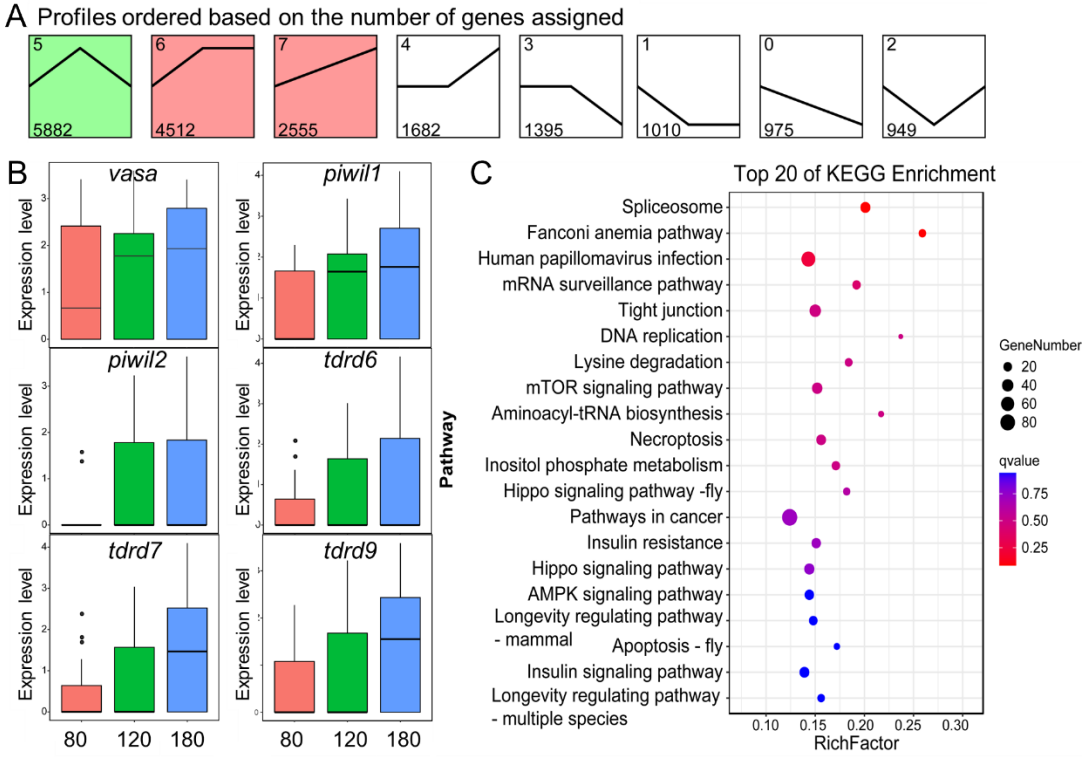


Supplementary Figure S4 IHC validation of locations of *Ddx4*, *Piwi*, and PCNA in gonads at three developmental stages (80 dph, 120 dph, and 180 dph)

D-F, **J-L**, and **P-R** are magnifications of box areas in **A-C**, **G-I**, and **M-O,** respectively. BV, blood vessel; Dph, day post-hatching. Scale bar = 50 μm.

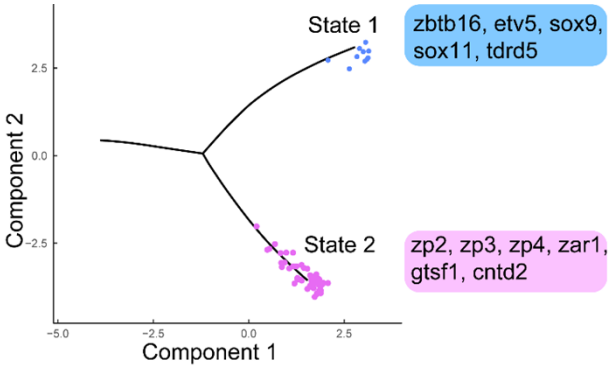


Supplementary Figure S5 Distribution of selected genes in all clusters

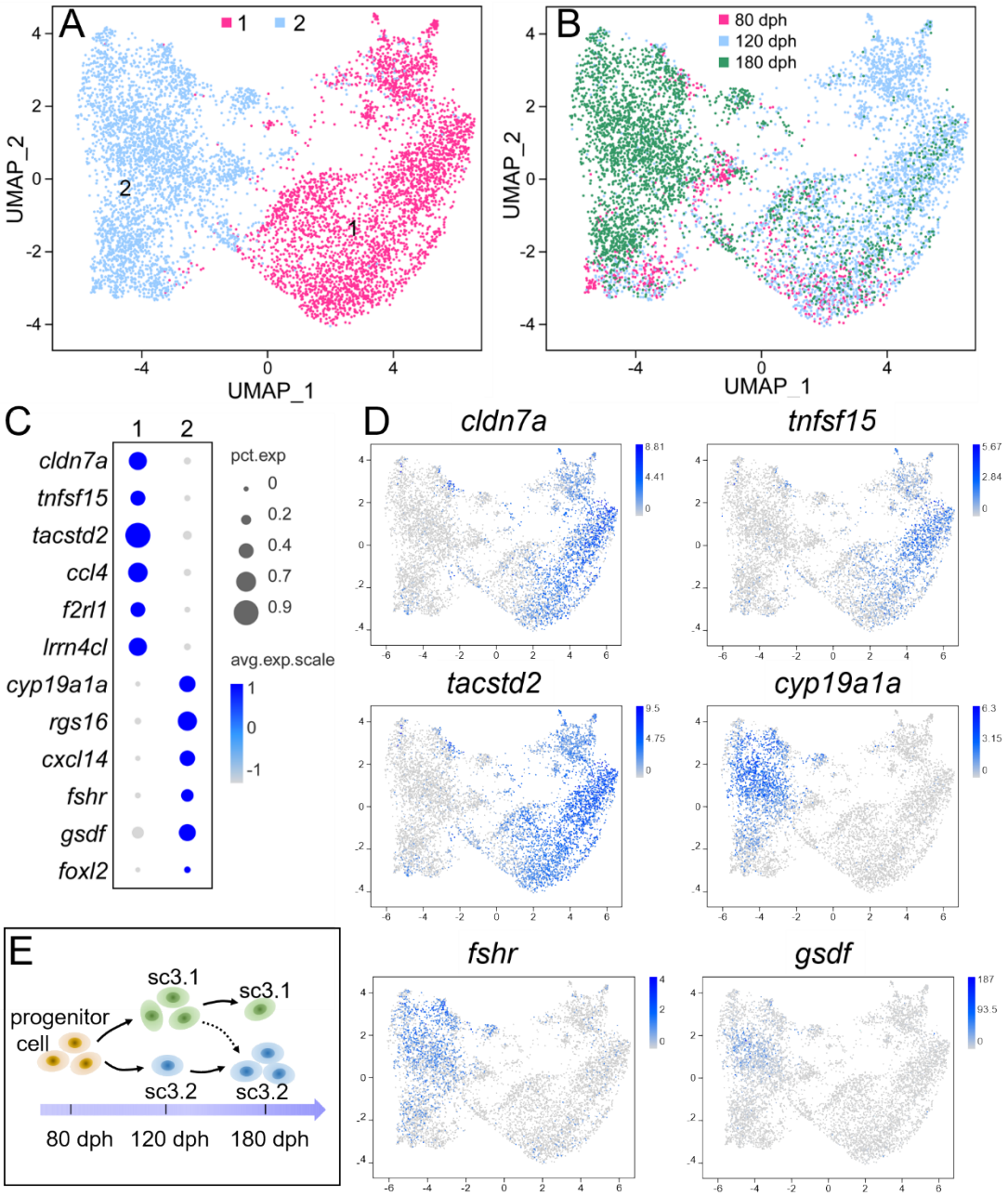


Supplementary Figure S6 Differential analysis of gonocytes at three developmental stages

A: Profiles of genes with differential expression trends. Profiles were ordered based on number of genes assigned. **B:** Expression changes in six selected genes in PGCs at 80 dph, 120dph, and 180 dph. **C:** Top 20 KEGG enrichment pathways in 2 555 genes in profile 7.



Supplementary Figure S7 Differential state analysis of oocytes in 180 dph gonads



Supplementary Figure S8 Subcluster analysis of cluster 3 showing two cell subtypes

A: Cluster 3 subcluster UMAP plot. **B:** UMAP plot of cell distribution at three developmental stages (80 dph, 120 dph, and 180 dph). **C:** Dot-plot showing relative expression of selected genes in subclusters. **D:** Gene expression UMAP plots of selected genes. **E:** Schematic representation of potential differentiation of cluster 3 subclusters from 80 dph to 180 dph. Sc3.1, subcluster 1 in cluster 3; sc3.2, subcluster 2 in cluster 3. Dph, days post-hatching.