Dissecting gonadal cell lineage specification and sex determination during human development using a single-cell transcriptomics approach

Fuchou Tang (tangfuchou@pku.edu.cn)
Biomedical Pioneering Innovation Center, Peking University https://orcid.org/0000-0002-8625-7717
Rui Wang
Peking University
Xixi Liu
Biomedical Pioneering Innovation Center, Peking University
Li Li
Peking University
Ming Yang
Peking University Third Hospital
Jun Yong
Peking University
Xiaoying Fan
Peking University
Yuqiong Hu
Beijing Advanced Innovation Center for Genomics (ICG), College of Life Science, Peking University
Fan Zhai
Department of Obstetrics and Gynecology Third Hospital, Peking University
Lu WEN
BIOPIC, Peking University https://orcid.org/0000-0002-1773-1876
Liying Yan
Peking University Third Hospital https://orcid.org/0000-0001-9572-9440
Jie Qiao
Peking University Third Hospital https://orcid.org/0000-0003-2126-1376

Biological Sciences - Article

Keywords: gonadal cells, sex determination, human development, transcriptomics

Posted Date: November 11th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-88057/v1

Abstract

Gonadal somatic cells are the main players in gonad development and are important for sex determination and germ cell development. Here, using a time-series scRNA-seq strategy, we analyzed the fetal germ cells (FGCs) and gonadal somatic cells in human embryos and fetuses. Clustering analysis of testes and ovaries revealed several novel cell subsets, including *POU5F1*$^{+}$*SPARC*$^{+}$ FGCs and *KRT19*$^{+}$ somatic cells. Furthermore, our data indicated that *DLK1*$^{+}$ cells may be the progenitors of steroidogenic cell lineages in both sexes and that *TAC1*$^{+}$ cells may be the progenitors of granulosa cells in females. Intriguingly, the testosterone synthesis function transitioned from fetal Sertoli cells to adult Leydig cells in a step-wise manner. Moreover, interactions between gonadal somatic cells were systematically explored and verified in our study. In detail, we observed that Sertoli cells interacted with Leydig cells through *DHH-PTCH1* and *PDGFA-PDGFRA/ PDGFRB* ligand-receptor gene pairs. More importantly, we identified cell type-specific developmental defects of both FGCs and gonadal somatic cells in a Turner syndrome embryo (45, XO). Our work provides a blueprint of the complex yet highly ordered development and interactions of human FGCs and gonadal microenvironment cells.

Introduction

During the development of the human gonad, gonadal somatic cells first initiate sex-specific differentiation with the specification of two major cell lineages: the supporting cell lineage (Sertoli cells in the testes and granulosa cells in the ovaries) and the steroidogenic cell lineage (Leydig cells in the testes and theca cells in the ovaries). The somatic cell environment plays vital roles in germ cell development, hormone production, and the establishment of the reproductive tracts[1].

In humans, abnormal specification of gonadal cell lineages and dysfunctions of gonadal somatic cells can result in disorders of sexual development (DSDs)[2]. Turner syndrome is a DSD caused by loss or structural alteration of X chromosome in females. Generally, it occurs in approximately 1 in 2,500 newborn girls worldwide and approximately half of the individuals with Turner syndrome have monosomy X (45, XO)[1].

Insights gained from mouse models have largely advanced our understanding of gonad development, sex-specific differentiation and DSDs[1-8]. However, these topics remain poorly explored in humans, especially the development of gonadal somatic cells. Recently, we and other groups have described the developmental transcriptomic landscapes of human fetal germ cells (FGCs) and spermatogenesis in studies focusing mainly on the development of germ cells instead of gonadal somatic cells[9-17]. To further expand our understanding of the specification of gonadal somatic cells and sex differentiation in humans, we performed large-scale single-cell RNA sequencing (scRNA-seq) of gonadal cells in both sexes in a time series spanning gonad development from 6W to as late as 23W of gestation. In total, we obtained 56,399 gonadal cells from 16 embryos, and we found that male FGCs underwent proliferation and expansion much earlier than female FGCs. In addition to the known somatic cell lineages, we

identified a new *KRT19*$^{+}$ gonadal somatic cell type that highly expressed *ALDH1A2* in both male and female embryos and fetuses. Furthermore, our data implied that the steroidogenic cell lineage was derived from *DLK1*$^{+}$ precursor cells in both sexes and that theca cells originated from *TAC1*$^{+}$ precursor cells that existed only at a relatively early embryonic stage. Interactions between Sertoli cells and Leydig cells were also comprehensively dissected. More importantly, we systematically analyzed the transcriptomic profiles of both primordial germ cells (PGCs) and gonadal somatic cells in a 7W XO human embryo. Together, our results provide an in-depth characterization of the transcriptomic landscape of human gonads and key insights into the transcriptomic features of germ cells and gonadal somatic cells in a human embryo with monosomy X.

Classification Of Gonadal Cell Types And Identification Of Embryo Sexuality

In this study, we used scRNA-seq to analyze all cells in the testes and ovaries from human embryos and fetuses at different developmental stages. This allowed us to identify all major fetal germ cell (FGC) and somatic cell subsets, and to further trace the developmental trajectories of these subsets. After filtering out low-quality cells, we retained a total of 56,399 gonadal cells from 16 embryos and fetuses ranging from 6W to 23W of age for subsequent analyses. Based on UMAP clustering, we classified the major cell populations present in the developing gonad, including FGCs (5,866 cells), gonadal somatic cells (45,660 cells), blood cells (1,267 cells), immune cells (3,043 cells) and endothelial cells (563 cells) (Fig. 2a, Supplementary Fig. 1a, 1c and Table 1). There was minimal batch influence based on the clustering results for two technical replicates for a 15W embryo, and on average, 1,694 genes were detected in each individual cell (Supplementary Fig.1b, 2d).

At early developmental stages, such as 6W to 10W, we could not distinguish female and male gonads by morphology. Therefore, we calculated the ratio of reads mapped to the Y chromosome and the expression of genes located on sex chromosomes (*XIST* on the X chromosome and *RPS4Y1* on the Y chromosome) and then judged the sex of each embryo and fetus. The results determined by this method were highly consistent with those determined by morphology for late-stage fetuses, verifying the accuracy of our strategy (Fig.2b and Supplementary Fig.1e). Female embryos and fetuses strongly expressed *XIST* but barely expressed *RPS4Y1*, and few reads mapped to the Y chromosome (Fig. 2b and Supplementary Fig. 1e). However, these genes were abnormally expressed in a 7W embryo; this embryo expressed neither *XIST* nor *RPS4Y1*, and few reads were mapped to the Y chromosome. These results implied that this embryo exhibited monosomy X (45, XO). The gonad of this XO embryo did not show a significant difference with normal embryos in morphology. To further verify this hypothesis, we performed SNP calling in the scRNA-seq dataset (Supplementary Fig. 1g-h). We found that the allele frequency of the X chromosome inactivation escape gene (*RBM3*) in female embryos was approximately 50%, but the allele frequency in male embryos and the 7W XO embryo was 100% (Supplementary Fig. 1f). We also counted the number of SNPs on the Y chromosome and found that all female embryos and the 7W XO embryo had no SNPs on the Y chromosome, as expected (Supplementary Fig. 1g). To evaluate the technical

influence, we investigated the frequency of SNPs on chromosome 1 of the 7W XO embryo and found that most of the SNP loci were heterozygous, as expected (Supplementary Fig.1h). Therefore, the 7W embryo was confirmed as an XO embryo.

Transcriptome Profile Comparison Of Pgcs And Gonadal Somatic Cells Among Xo, Xx And Xy Embryos

To characterize the phenotype of the monosomy X female embryo, we further clustered cells of XO and normal XX, XY embryos at the same developmental stage (7W). PGCs from normal female and male embryos clustered together whereas PGCs from the XO embryo clustered separately from them (Fig. 1c and Supplementary Fig.2a-b). In addition, the gonadal somatic cells from the female and male embryos clustered separately, but the transcriptomes of gonadal somatic cells from the normal embryos (XX and XY) were relatively similar to each other (Fig. 1c). Next, we identified differentially expressed genes (DEGs) in gonadal somatic cells between these embryos (Fig.1d and Supplementary Fig.2a-b, d-e). The normal XX female embryo highly expressed genes enriched for Gene Ontology (GO) terms such as regulation of the estrogen signaling pathway (*FOS* and *JUN*, p-value=$1.2x10^{-5}$) and positive regulation of cell migration and differentiation (p-value = $2.5x10^{-2}$). In contrast, the XO embryo highly expressed genes enriched for GO terms such as engulfment of apoptotic cells (p-value = $3.1x10^{-3}$), the WNT signaling pathway (RAC1, p-value = $4.4x10^{-3}$), and regulation of cell morphogenesis (p-value = $8.8x10^{-3}$). We also compared the expression profiles of PGCs between XO and normal XX & XY embryos (Supplementary Fig. 2c, d, Fig. 3 and Table 2). Similar to gonadal somatic cells, PGCs in the XO embryo clearly enriched genes in the WNT signaling pathway (p-value = $5.3x10^{-3}$) and those related to engulfment of apoptotic cells (p-value = $1.4x10^{-2}$). On the other hand, different from the gonadal somatic cells, PGCs in the XO embryo depleted expression of genes associated with cell cycle (p-value = $4.3x10^{-7}$) and p53 signaling pathway (p-value = $1.7x10^{-2}$). Specifically, PGCs in XO embryo specifically downregulated expression of *HAT1* and *PSMD12*, genes that participate in replication-dependent chromatin assembly and cell cycle progression, whereas they upregulated expression of *UTF1*, *PDCD2* and *FRAT2*, genes that play critical roles in the WNT signaling pathway, regulation of cell proliferation and embryo differentiation (Supplementary Fig. 2d, e). Notably, these five genes did not show expression differences in somatic cells among XO embryos and normal XX embryos, which indicated that monosomy X has a cell-type-specific defects in PGCs and gonadal microenvironment cells. There were also cell-type shared changes of gene expression between PGCs and gonadal somatic cells, such as effects on *PDIA3*, *RAC1*, *ATP5D* and *IGFBP2* (Supplementary Fig. 3a-b).

Next, we classified the somatic cell populations present in all 7W embryos using UMAP clustering (Fig. 1e and Supplementary Fig. 2f-g). Three cell clusters were identified in the XO embryo and two of them also existed in both normal XX and XY embryos ($DLK1^+$ and $KRT19^+$ in both XX and XY embryos, and $TAC1^+$ in the XX embryo). According to the clustering results, we speculated that PCP^+ cells were a subgroup of supporting lineage in both sexes (Fig. 1e). Notably, that cell clusters identified in all these three 7W embryos highly expressed the somatic progenitor cell marker *LHX9* (Fig.1d). The cell cluster that highly

expressed *TAC1* was shared only by XX and XO embryos, which may imply that this cell type is unique to female embryos at an early developmental stage. Specifically, at this embryonic stage, XY male gonadal cells had differentiated into the supporting cell lineage, Sertoli cells, which indicates that the specification of the supporting cell lineage in male embryos is much earlier than that in female embryos (Fig. 1e and Supplementary Fig. 2g-h, Fig. 3a). In addition, both XX and XY embryos contained a subset of cells that highly expressed *PCP4*, a gene that functions as a modulator of calmodulin activity (Fig. 1e and Supplementary Fig. 2g-h, Fig. 3a). Through immunostaining, all cell populations identified in early-stage testes were verified (Supplementary Fig. 2f). To summarize, XO embryo exhibited abnormal gene expression profiles in both germ cells and gonadal somatic cells, and lacked $PCP4^+$ cell populations compared to normal XX embryos (Supplementary Fig. 3b).

Identification Of Novel Fgc Subtypes

Next, we analyzed the transcriptional characteristics of the FGCs. By performing UMAP unsupervised clustering, we clearly identified 8 clusters of FGCs (Supplementary Table 3). Based on the expression of well-known FGC markers, these clusters were classified as mitotic (proliferative & quiescent), mitotic arrest and $SPARC^+$ FGCs in male gonads as well as mitotic, meiotic prophase, retinoic acid (RA)-responsive, oogenesis, and Mix in female gonads. The Mix cluster of FGCs in female ovaries simultaneously expressed somatic cell markers and FGC markers, which are potentially germ cells contaminated by debris of neighboring somatic cells. So we omitted them in the subsequent analyses (Fig. 1a-b, Supplementary Fig. 4a and Table 4). The transcriptomic patterns of early- and late-stage FGCs differed dramatically in females, but showed milder differences in males (Fig. 1a-b). Most of these clusters have been identified in our previous studies, verifying the accuracy of our analyses. Due to the increased number of cells analyzed, we identified a new FGC subtypes, $SPARC^+$ cells in male gonads. $SPARC^+$ cells were mainly found in the first-trimester embryos and fetuses (Fig. 2a-b). Through immunofluorescence staining of 6W and 20W human fetal testes, we verified the existence of the newly identified $SPARC^+POU5F1^+$ FGCs. To further investigate the expression features of these FGC subtypes, we performed DEG analysis (Supplementary Fig. 4c). Since SPARC plays important roles in cell migration, we speculate that this group of cells may retain the ability to migrate.

We then assessed the ratios of the defined FGC subtypes across different developmental stages in the testes and ovaries (Fig. 2c and Supplementary Fig. 4b). In 6W-7W gonads, both male and female FGCs consisted mainly of mitotic FGCs (Supplementary Fig. 4b). From 13 week onwards, RA-responsive, meiotic prophase and oogenesis FGCs began to appear in the ovaries. By 22W, more than half of the FGCs in females were in meiosis, and the rest were mainly RA-responsive and oogenetic FGCs. In 8W to 10W testes, more than 25% of FGCs were in the active proliferation phase, and this result was consistent with the significant increase in the total number of FGCs in testes from 8W onwards (Fig. 2c). Mitotic arrest FGCs appeared at 10W, and they became a major FGC cell type at 23w, accounting for 68.1% of the FGCs.

Proliferation Of Human Fgcs In Males And Females

Since we sequenced all cells in gonads without FACS enrichment, we could explore the dynamic changes in the percentages of FGCs during gonad development. The results showed that male FGCs underwent proliferation and expansion as early as 8W and increased rapidly in number until 15W of gestation, while the proliferation of female FGCs began at 16W and persisted until as late as 22W of gestation (Fig. 2c). We further performed immunofluorescence staining of FGC markers in the early, mid- and late stages of human fetal testes to verify the changes in the percentages of FGCs (Fig. 2e and Supplementary Fig. 4d). The evidences suggested that FGCs have a sex-specific proliferation pattern in the human gonad.

Transcriptomic Features Of Prespermatogonia

Spermatogenesis is a continuous and highly coordinated process of sperm production [8,12,13,16]. In humans, the initiation of spermatogenesis is believed to occur during puberty. Due to the scarcity of biological materials and the limitations of technologies, there is no clear definition for the gonocyte-to-spermatogonium transition (GST) and the precise timing of the GST remains unclear in human testes[18].

In males, after arrival at the gonadal anlage, PGCs usually give rise to prespermatogonia (gonocytes), which form a homogenous population of single, round cells in embryonic testes. During the first trimester, despite some morphological differences, gonocytes continue to express markers of early PGCs such as *OCT4 (POU5F1), NANOG* and *DDX4*. During the second trimester, most but not all gonocytes gradually lose their mitotic activity along with early PGC marker expression. At this time, a new type of mitotic arrest germ cell begins to appear, called the transition (T)-spermatogonium[19,20].

To characterize the progression of GST, we combined expression data for fetal, neonatal and adult germ cells together and clustered them by principal component analysis (PCA) (Fig. 2f and Supplementary Fig. 5a-d)[13]. According to the PCA results and gene expression features, germ cells in fetal stages were grouped into mitotic FGC, and mitotic arrest FGCs, with the former being further divided into proliferative and quiescent subtypes as we previously shown (Fig. 2f, 2i and Supplementary Fig. 6a, c-d). (Supplementary Fig. 6c-d). Mitotic arrest FGC included mainly cells that came from testes in the second trimester, accounting for 85% of all FGCs, while proliferative mitotic FGC consisted mainly of cells from the first-trimester (Fig. 2g). According to PCA, mitotic arrest FGC represented as a transition state from gonocytes to spermatogonia. The BMP signaling pathway and HOX family genes may play roles in the GST process, since *ID1* to *ID4* as well as *HOXC9*, *HOXB7* and *HOXC9* were highly expressed from the mitotic arrest FGC stage to postnatal stages (Fig. 2i and Supplementary Fig. 6a). Through immunostaining of ID1 and the late-stage FGC marker DDX4, we verified BMP signaling pathway target gene ID1 was expressed by mitotic arrest FGC cells in human 14W testis (Supplementary Fig. 6b).

Male Gonadal Somatic Cells Can Be Classified Into Three Transcriptionally Distinct Cell Populations

To investigate the transcriptomic features of male gonadal somatic cells, we first classified these cells by UMAP clustering, and three distinct cell populations were identified (Fig. 3a, Supplementary Fig. 7a and Table 5). In addition to the already known groups of male gonadal somatic cells, Leydig and Sertoli cells, we identified a new group of cells that highly expressed *KRT19*, and a small fraction of these *KRT19*$^{+}$ cells also expressed the Sertoli cell marker *SOX9* at an early stage (Fig. 3b and Supplementary Fig. 7a). Notably, KRT19^{+} cells were mainly detected in gonads of embryos from 6W to 8W and were minimally detected after 9W (Fig. 3b). Through immunostaining, we verified that many cells in the gonads at 8W were *KRT19*$^{+}$ and a small fraction of these cells were also *SOX9*$^{+}$; this occurred only in early-stage testes and was consistent with our scRNA-seq data (Fig. 3b-d and Supplementary Fig. 7a, 6b). Starting from 9W, only a small fraction of cells that located around the Sertoli cells expressed *KRT19* (Supplementary Fig. 7b). To enrich specific somatic cells for further study, we first identified DEGs for these three gonadal somatic cell types, selected cell surface markers for them and verified the specificity of the markers by immunostaining (Supplementary Fig. 7c, e, f). To analyze the structural changes of somatic and germ cells in the testes, we performed immunostaining of early-and late-stage testes with FGC and Leydig cell markers (Supplementary Fig. 7d). We observed that in the 6W testes, germ cells and Leydig cells were scattered. In the late stage, Leydig cells formed a tubular structure that surrounded the germ cells (Supplementary Fig. 7d).

Development Of Fetal Leydig Cells And Sertoli Cells

Our data covered male embryos from 6 to 23W; thus, we were able to explore the developmental trajectories of Leydig cells and Sertoli cells in detail (Supplementary Fig. 8). At early developmental stages, Leydig cells highly expressed BMP and WNT signaling pathway genes, such as *ID2* and *ID3* in the BMP pathway, and *CTNNB1* and *SFRP1* in the WNT pathway (Supplementary Fig. 8a-c). At late developmental stages, Leydig cells highly expressed genes in the FOS family and JUN family, such as *FOS*, *FOSB*, *JUN* and *JUNB*. Interestingly, the expression patterns of these marker genes in Sertoli cells were opposite to those in Leydig cells. Sertoli cells expressed FOS family genes mainly in the early stages of development rather than in the late stages (Supplementary Fig. 8d-f).

Interaction Between Leydig Cells And Sertoli Cells

In mammals, sex-specific differentiation is initiated by gonadal somatic cells with specification of the supporting cell lineages (Sertoli cells and granulosa cells). Then the supporting cell lineages facilitate the differentiation of the steroidogenic cell lineages (Leydig cells and theca cells). To elucidate how supporting cell lineages facilitate the differentiation of steroidogenic cell lineages, we further explored the interactions between Sertoli cells and Leydig cells in the testes (Fig. 3e-f)[1].

According to previous studies, DHH/PTCH1 signaling might not only function in Leydig cell differentiation but also play roles in signaling between Sertoli and peritubular myoid cells in mice[21]. Here, we found that the *DHH/PTCH1* gene pair interacts between Sertoli cells and Leydig cells in the gonads of

human fetuses, rather than between Sertoli cells and peritubular muscle cells, as occurs in mice (Fig. 3e). We further verified this by immunofluorescence staining (Fig. 3f).

The platelet-derived growth factor (PDGF) family genes participates primarily in the migration, proliferation, and differentiation of cells in various organ systems, and it has also been reported to be an important downstream target of Sry during testicular organogenesis and Leydig cell differentiation in mice. In our study, we found that Sertoli cells and Leydig cells can interact with each other through PDGF family genes (Fig. 3e)[22]. Sertoli cells expressed the ligand *PDGFA* and Leydig cells expressed the receptors *PDGFRA* and *PDGFRB* (Fig. 3e). We also verified this interaction through immunofluorescence staining (Fig. 3f).

In addition, we found that Leydig cells may still interact with Sertoli cells through *DHH/PTCH1* and *PDGFA/PDGFRA-PDGFRB* at the neonatal, pubertal and adult stages in the testes. The cell type-specific markers we identified in fetal testicular data were also universally applicable to neonatal, pubertal and adult testicular cells (Fig. 3g and Supplementary Fig. 9b).

Differences Between Antenatal And Postnatal Gonadal Somatic Cells

Androgens are critical participants in the development and homeostasis of male reproductive functions. It is well known that Leydig cells in testes are the primary sources of physically active androgens. In mammals, testosterone, the most effective androgen is synthesized from cholesterol through several reactions that mediated by steroidogenic enzymes. In mammals, two distinct populations of Leydig cells, fetal Leydig cells (fLCs) and adult Leydig cells (aLCs), develop sequentially in the fetal and adult testes, respectively, and exhibit distinct characteristics[4]. The developmental relationship between fLCs and aLCs is not fully understood. To explore this issue, we clustered male gonadal somatic cells from fetal (6W-23W, 8 stages), neonatal (D2-D7, 2 stages), pubertal (1-14 years old, 5 stages) and adult (17-42 years old, 5 stages) stages by UMAP clustering(Fig 3g and Supplementary Fig. 5a-e, Fig. 9a)[10,11,13]. The results showed that both Leydig and Sertoli cells exhibit continuous developmental trajectories from the first-trimester to the second-trimester, and then through neonatal, pubertal and adult periods. Globally, the neonatal gonadal somatic cells were more similar to those at the fetal stages (Fig. 3g), while gonadal cells originated from individuals older than 1 year old clustered together, which may indicate that the transcriptomes of gonadal somatic cells change greatly during the first year after birth. Next, we identified DEGs among Leydig cells and Sertoli cells at different stages (Supplementary Fig. 9c-e). Only adult Leydig cells highly expressed genes such as *LGALS3*, *DPT*, *CFD* (Supplementary Fig. 9d). Both Sertoli cells and Leydig cells at the neonatal stage highly expressed *HSPA1A* and *HSPA1B* (Supplementary Fig. 9d). Interestingly, *SEFRPINF1*, a WNT signaling pathway inhibitor was highly expressed by Sertoli cells at all stages analyzed, but was highly expressed in only a small fraction of neonatal Leydig cells and the majority of the aLCs (Supplementary Fig. 9c, e). Another WNT signaling pathway inhibitor was also highly expressed in aLCs. In addition, the expression of *SYNE2*, a gene involved in meiosis and the cell cycle, gradually increased along the developmental path of Sertoli cells; this gene may play a role in meiosis initiation in male germ cells (Supplementary Fig.9c, e).

aLCs can produce testosterone to maintain male reproductive function. However, it remains unknown whether fLCs can produce testosterone. Generally, Sertoli cells have been accepted as nonsteroidogenic cells[4,23]. However, we found that at the fetal stage, Sertoli cells rather than Leydig cells expressed *HSD17B3*, an enzyme that mediates the final step of testosterone synthesis (Supplementary Fig. 10a-c)[23]. Shima *et al* also observed similar patterns in mice during the fetal period and they concluded that androstenedione produced by fLCs is transferred to fetal Sertoli cells and then converted to testosterone in mice[23]. In addition to neonatal Sertoli cells expressing *HSD17B3*, there was also a small fraction of neonatal Leydig cells expressing *HSD17B3*, consistent with the findings of a study conducted on postnatal fetal rats (P8)[24]. Moreover, the ratio of Leydig cells expressing *HSD17B3* gradually increased from the neonatal stages; however, unexpectedly, the ratio of Sertoli cells that expressed *HSD17B3* remained high after birth (Supplementary Fig. 9b-c). Cholesterol sidechain cleavage P450 (*CYP11A1*), an intermediated product during testosterone synthesis, also exhibited an expression pattern similar to that of *HSD17B3*[4]. Taken together, the data suggest that even after birth Sertoli cells still contributes to testosterone synthesis, which is gradually taken over by Leydig cells.

Both testosterone and RA signaling are important for spermatogenesis, and dysregulation of either RA or androgens can cause testicular dysfunction and result in male infertility. RA is tightly regulated by a group of RA-synthesizing (ALDH1A1, ALDH1A2 and ALDH1A3) and RA-metabolizing enzymes (CYP26A1, CYP26B1 and CRY26C1) (Fig. 3h)[25-28]. Here, we found that the three RA synthesis enzymes were expressed by different types of gonadal somatic cells in male fetuses (Fig. 3i and Supplementary Fig. 7d). *ALDH1A1* was expressed by fetal and neonatal Sertoli cells as well as aLCs. *ALDH1A2* was expressed mainly by KRT19$^+$ gonadal somatic cells and a small fraction of Leydig cells in the first trimester (Fig. 3i and Supplementary Fig. 2h, 8d). *ALDH1A3* was expressed primarily in Leydig cells during the second trimester, neonatal period and adulthood. In addition, the circulating retinol binding protein II (*CRBP2*) was expressed by all types of male gonadal somatic cells in mainly the first trimester. Once the fetal supporting cell lineage (Sertoli cells in the testes and granulosa cells in the ovaries) was determined, FGCs in the ovaries began to undergo meiosis because of exposure to RA, while FGCs in the testes entered mitotic arrest because of the expression of the RA-degrading cytochrome P450 enzyme CYP26B1 in Sertoli cells. Consistent with previous studies, we found that a small fraction of fetal and neonatal Sertoli cells expressed *CYP26B1*.

Female Gonadal Somatic Cells Can Be Classified Into Seven Transcriptionally Distinct Cell Populations

Having characterized somatic cells in male gonads, we next focused on the gonadal somatic cells of female embryos and fetuses. We identified seven somatic cell types in fetal ovaries (Fig. 4a-c and Supplementary Table 6). Globally, early-stage female gonadal somatic cells showed expression patterns that were distinct from those of cells that derived from fetuses after 13W (Fig. 4b). In detail, DLK1$^+$, DLK1$^+$TOP2A$^+$ and TAC1$^+$ groups were detected only in the gonads of 7W embryos, and granulosa cells (FOXL2$^+$ cells) appeared after 13W (Fig.4b). In addition to granulosa cells, we also found KRT19$^+$ cells in

the gonads of embryos and fetuses at all developmental stages we analyzed. Similar to KRT19$^+$ cells in male gonads, KRT19$^+$ cells in female gonads expressed the RA-synthesizing enzyme *ALDH1A2* (Fig. 4a, c). However, *ALDH1A2* was also expressed by granulosa cells at later stages. We also examined the other two RA synthesizing enzymes. The results showed that *ALDH1A1* was expressed by gonadal somatic cells since 13W, but more granulosa cells expressed this RA synthesizing enzyme. In addition, the same as somatic cells in fetal testes, *CRABP2* was also highly expressed by 7W female gonadal somatic cells. Starting from 13W, there was a group of cells (Mid cells) located between KRT19$^+$ and granulosa cells (FOXL2$^+$ cells). To further explored the relationship between these three cell types, we performed RNA velocity analysis on them. The results showed that Mid cells has potential to differentiate into KRT19$^+$ cells and granulosa cells (FOXL2$^+$ cells) (Fig. 4f and Supplementary Fig. 10d).

Dlk1 Cells Are The Progenitors Of Steroidogenic Cell Lineage

Once the FGCs enter meiosis, the proximal somatic cells will later differentiate into two subsets of granulosa cells: mural granulosa cells and cumulus cells. To explore the correspondence between fetal and adult gonadal somatic cell types in females, we mapped our data to single-cell transcriptome data of adult ovarian (Fig. 4d and Supplementary Fig. 10e)[14]. As expected, the transcriptome of FOXL2$^+$ cells in fetal ovaries most closely resembled that of adult granulosa cells (Fig. 4d-e). Specifically, TAC1$^+$ cells in 7W female gonads were most similar to adult progenitor granulosa cells (proGCs), which indicated that *TAC1$^+$* cells may be progenitor cells of granulosa cells. It has been reported that theca cells arise from at least two progenitor populations: most theca cells are derived from *WT1$^+$* somatic progenitors in the gonadal primordium and a small fraction of theca progenitors migrate from the mesonephros to the ovaries[1]. In agreement with previous studies, we found that DLK1$^+$ cells and some KRT19$^+$ cells were clustered together with adult theca cells in ovaries, and the corresponding steroidogenic cells in males, Leydig cells also highly expressed *DLK1*, a paternally imprinted gene; these findings may indicate that DLK1$^+$ cells are steroidogenic lineage progenitor cells (Fig. 4c, e and Supplementary Fig. 7f). However, some differences were observed: in the male steroidogenic cell lineage, *DLK1* continued to be expressed in Leydig cells, while in the female lineage, *DLK1* was expressed only at 7W. In addition, both DLK1$^+$ and some KRT19$^+$ cells had the potential to differentiate into theca cells (Fig. 3a,3c and Supplementary Fig. 7f). Notably, KRT19$^+$ cells seemed to have the potential to differentiate into theca and granulosa cell types.

Discussion

We previously performed scRNA-seq analysis on 2,486 (319 with the Tang protocol and 2,167 with a modified STRT-Seq protocol) human FGCs and their microenvironment somatic cell for 44 embryos between 4 and 26 weeks of gestation[9]. We comprehensively and extensively analyzed the transcriptomic features of FGCs and clearly identified four and two types of FGCs in females and males, respectively. In females, these types were mitotic-phase, RA signaling-response-phase, meiotic-prophase and oogenesis-

phase FGCs, while in males, the mitotic-phase and mitotic arrest-phase FGCs were identified. However, insufficient numbers of gonadal somatic cells were analyzed; thus, the developmental characteristics of gonadal somatic cells were not fully explored in our previous studies.

In this study, we created a comprehensive map of cell populations in the human ovaries and testes, including both FGCs and somatic cells in their niche. To characterize the differentiation and maturation of FGCs and gonadal somatic cells, we generated developmental expression profiles of 56,399 gonadal cells for 16 human embryos and fetuses between 6 and 23W of gestation. Specifically, we identified 7 groups of FGCs, including newly identified FGCs that simultaneously expressed *POU5F1* and *SPARC*. In addition to immune cells, endothelial cells and blood cells, we identified three and seven types of somatic cells in male gonads (testes) and female gonads (ovaries), respectively. In both testes and ovaries, we identified a new KRT19^{+} cell population that highly expressed the RA synthesis enzyme *ALDH1A2* at early developmental stages (Fig. 4f). We have found and verified surface protein markers that can be used to isolate specific somatic cell types in future studies, such as CDH2 for Sertoli cells and PDGFRA for Leydig cells. In addition, we also combined the postnatal 55,404 gonadal single cells (20,676 ovarian single cells and 34,728 testicular single cells) from four published datasets, and comprehensively and systematically explored the gene expression pattern of gonadal cells[10,11,13,14]. Especially for the testis, we covered single cell transcriptomic data of male individuals from 4W to 25W before birth, and from 2 days to 42 years of ages after birth, and conducted in-depth research on the expression patterns of testicular somatic and germ cells[9,11,13].

Second, we assessed the developmental origins of granulosa cells in females and of steroidogenic lineages in both sexes (Fig. 4f). Our data indicated that DLK1^{+} cells in the gonads of 7W of male and female embryos may further differentiate into steroidogenic cell lineages: Leydig cells in the testes and theca cells in the ovaries. However, *DLK1* expression continued in Leydig cells of the testes in subsequent developmental stages, while it disappeared in theca-like cells of the ovaries. A human fetus does not undergo sex differentiation until approximately 7W. Our data indicated that in 7W embryos, the expression feature and cell types of both germ cells and gonadal somatic cells are very similar if not identical in male and female embryos. Next, by mapping our ovarian somatic cell data to published adult ovarian scRNA-seq data, we found that the expression profiles of TAC1^{+} cells and proGCs are very similar; thus, we speculate that TAC1^{+} cells in 7W female gonads may further differentiate into granulosa cells in later developmental stages. In addition, through RNA velocity analysis, we found that Mid cells that appearing after 13W have the potential to differentiate to KRT19^{+} cells and granular cells (FOXL2^{+} cells).

Although the molecular characteristics, cellular origins and developmental and functional links of fLCs and aLCs have been extensively studied in rodents, we are just beginning to unravel these details in humans due to the scarcity of samples and technical limitations. In mammals, there are two main types of Leydig cells, fLCs and aLCs. Leydig cells are the main sources of androgens, while Sertoli cells are generally accepted as nonsteroidogenic cells. However, in this study, we found that Sertoli cells, rather than Leydig cells express *HSD17B3*, an enzyme that mediates the final step of testosterone synthesis, in

the fetal (18.5% in the first trimester and 34.5% in the second trimester) and neonatal periods (30%). Leydig cells did not express *HSD17B3* until neonatal stages (6.8%) and the ratio of cells that expressed *HSD17B3* gradually increased in the later stages. This finding indicates that the androstenedione produced by fLCs is transferred to fetal Sertoli cells and then converted to testosterone. Then, during the neonatal stages, testosterone production function is gradually taken over by Leydig cells. At the adult stage, Leydig cells can produce testosterone by themselves since they can express *HSD17B3* and other steroidogenic enzymes simultaneously. Additionally, we elucidated the cell-cell interactions between Leydig cells and Sertoli cells. Firstly, we found that the Hh ligand genes *DHH* and its downstream receptor gene *PTCH1* were expressed in Sertoli cells and Leydig cells, respectively. However, the *DHH-PTCH1* ligand-receptor regulates mainly the interactions between Sertoli and peritubular myoid cells in mice[21]. Secondly, we found that Sertoli cells and Leydig cells interacted with each other through *PDGFB-PDGFRA/PDGFRB* ligand-receptor gene pairs. Notably, these interactions were also observed in neonatal and adult testes.

To the best of our knowledge, this study is the first to systematically and comprehensively compare the gene expression patterns and cell type compositions among monosomy X (45, XO), normal female (46, XX) and male (46, XY) embryos at the same developmental stage (7W). Globally, monosomy X caused abnormalities not only in the transcriptome of FGCs, but also in gonadal somatic cells. More importantly, monosomy X was associated with cell type-specific developmental defects in FGCs and gonadal somatic cells. Furthermore, it also caused the loss of a *PCP4*-positive cell population in XO gonad compared with normal XX gonads (Supplementary Fig. 3).

Altogether, our data identified and characterized new types of FGCs and gonadal somatic cells. We speculated that the $DLK1^+$ cell population was a progenitor cell population of the steroidogenic cell lineage and that the $TAC1^+$ cell population was a progenitor cell population of granulosa cells. In addition, we illustrated the transcriptomic changes in monosomy X embryos. Our data also have potential applications for the isolation of specific gonadal somatic cell populations *in vivo*. It deepens our view of both germ cell and gonadal somatic cell development in the testes and ovaries. Finally, our work provides insights and valuable data resources for further elucidation of the molecular mechanisms of infertility.

Declarations

Acknowledgments

This work was supported by National Natural Science of China (No.81521002 and No.31625018) and National Key R&D Program of China (2017YFA0102702 and 2018YFA0107601). Part of analysis was performed on the High Performance Computing Platform of the Center for Life Science. We thank all the patients for tissue donation in this study. We thank Chunyan Shan from the National Center for Protein Sciences at Peking University in Beijing, China for help with immunofluorescent staining training.

Author contributions

F.T. and Q.J. conceived the project. With the help of X.F., Y.H. and L.W., L.L. and X.L. performed single cell RNA sequencing and whole-genome sequencing experiments. M.Y., J.Y. F.Z. and L.Y. performed the patient's enrollment as well as clinical embryo stages and sexes determinations. With the guidance of R.W., X.L. performed the H&E and immunohistochemical staining. R.W. conducted the bioinformatics analyses. R.W. and F.T. wrote the manuscript. R.W., F.T., X.L. and L.L participated in the revision and polishing of the manuscript with the help from all other authors.

Competing interests

The authors declare they have no competing interests.

References

1. Rotgers, E., Jorgensen, A. & Yao, H. H. At the Crossroads of Fate-Somatic Cell Lineage Specification in the Fetal Gonad. *Endocr Rev* **39**, 739–759, doi:10.1210/er.2018-00010 (2018).
2. Sarraj, M. A. & Drummond, A. E. Mammalian foetal ovarian development: consequences for health and disease. *Reproduction* **143**, 151–163, doi:10.1530/REP-11-0247 (2012).
3. Bowles, J. & Koopman, P. Retinoic acid, meiosis and germ cell fate in mammals. *Development* **134**, 3401–3411, doi:10.1242/dev.001107 (2007).
4. Shima, Y. Development of fetal and adult Leydig cells. *Reprod Med Biol* **18**, 323–330, doi:10.1002/rmb2.12287 (2019).
5. Su, D. M. *et al.* Influence of fetal Leydig cells on the development of adult Leydig cell population in rats. *The Journal of reproduction and development* **64**, 223–231, doi:10.1262/jrd.2017-102 (2018).
6. Wen, Q. *et al.* Wt1 dictates the fate of fetal and adult Leydig cells during development in the mouse testis. *Am J Physiol Endocrinol Metab* **307**, E1131-1143, doi:10.1152/ajpendo.00425.2014 (2014).
7. Stevant, I. *et al.* Deciphering Cell Lineage Specification during Male Sex Determination with Single-Cell RNA Sequencing. *Cell reports* **22**, 1589–1599, doi:10.1016/j.celrep.2018.01.043 (2018).
8. Chen, Y. *et al.* Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. *Cell research* **28**, 879–896, doi:10.1038/s41422-018-0074-y (2018).
9. Li, L. *et al.* Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell stem cell* **20**, 891–892, doi:10.1016/j.stem.2017.05.009 (2017).
10. Guo, J. *et al.* The adult human testis transcriptional cell atlas. *Cell research* **28**, 1141–1157, doi:10.1038/s41422-018-0099-2 (2018).
11. Guo, J. *et al.* The Dynamic Transcriptional Cell Atlas of Testis Development during Human Puberty. *Cell stem cell* **26**, 262–276 e264, doi:10.1016/j.stem.2019.12.005 (2020).
12. Hermann, B. P. *et al.* The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. *Cell reports* **25**, 1650–1667 e1658, doi:10.1016/j.celrep.2018.10.026 (2018).

13. Sohni, A. *et al.* The Neonatal and Adult Human Testis Defined at the Single-Cell Level. *Cell reports* **26**, 1501–1517 e1504, doi:10.1016/j.celrep.2019.01.045 (2019).
14. Fan, X. *et al.* Single-cell reconstruction of follicular remodeling in the human adult ovary. *Nature communications* **10**, 3164, doi:10.1038/s41467-019-11036-9 (2019).
15. Wagner, M. *et al.* Single-cell analysis of human ovarian cortex identifies distinct cell populations but no oogonial stem cells. *Nature communications* **11**, 1147, doi:10.1038/s41467-020-14936-3 (2020).
16. Wang, M. *et al.* Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. *Cell stem cell* **23**, 599–614 e594, doi:10.1016/j.stem.2018.08.007 (2018).
17. Zhang, Y. *et al.* Transcriptome Landscape of Human Folliculogenesis Reveals Oocyte and Granulosa Cell Interactions. *Molecular cell* **72**, 1021–1034 e1024, doi:10.1016/j.molcel.2018.10.029 (2018).
18. Pui, H. P. & Saga, Y. Gonocytes-to-spermatogonia transition initiates prior to birth in murine testes and it requires FGF signaling. *Mech Dev* **144**, 125–139, doi:10.1016/j.mod.2017.03.002 (2017).
19. Abe, S. Differentiation of spermatogenic cells from vertebrates in vitro. *Int Rev Cytol* **109**, 159–209, doi:10.1016/s0074-7696(08)61722-2 (1987).
20. Dym, M. & Fawcett, D. W. Further observations on the numbers of spermatogonia, spermatocytes, and spermatids connected by intercellular bridges in the mammalian testis. *Biology of reproduction* **4**, 195–215, doi:10.1093/biolreprod/4.2.195 (1971).
21. Yao, H. H., Whoriskey, W. & Capel, B. Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Genes & development* **16**, 1433–1440, doi:10.1101/gad.981202 (2002).
22. Brennan, J., Tilmann, C. & Capel, B. Pdgfr-alpha mediates testis cord organization and fetal Leydig cell development in the XY gonad. *Genes & development* **17**, 800–810, doi:10.1101/gad.1052503 (2003).
23. Shima, Y. *et al.* Contribution of Leydig and Sertoli cells to testosterone production in mouse fetal testes. *Molecular endocrinology* **27**, 63–73, doi:10.1210/me.2012-1256 (2013).
24. Weisser, J., Landreh, L., Soder, O. & Svechnikov, K. Steroidogenesis and steroidogenic gene expression in postnatal fetal rat Leydig cells. *Mol Cell Endocrinol* **341**, 18–24, doi:10.1016/j.mce.2011.03.008 (2011).
25. Childs, A. J., Cowan, G., Kinnell, H. L., Anderson, R. A. & Saunders, P. T. Retinoic Acid signalling and the control of meiotic entry in the human fetal gonad. *PLoS One* **6**, e20249, doi:10.1371/journal.pone.0020249 (2011).
26. Jauregui, E. J., Mitchell, D., Topping, T., Hogarth, C. A. & Griswold, M. D. Retinoic acid receptor signaling is necessary in steroidogenic cells for normal spermatogenesis and epididymal function. *Development* **145**, doi:ARTN dev160465
27. 1242/dev.160465 (2018).
28. Minkina, A. *et al.* Retinoic acid signaling is dispensable for somatic development and function in the mammalian ovary. *Developmental biology* **424**, 208–220, doi:10.1016/j.ydbio.2017.02.015 (2017).

29. Vernet, N. *et al.* Retinoic acid metabolism and signaling pathways in the adult and developing mouse testis. *Endocrinology* **147**, 96–110, doi:10.1210/en.2005-0953 (2006).

Tables

Due to technical limitations, the tables are provided in the Supplementary Files section.

Figures

Figure 0

Figure 0

Figure 1

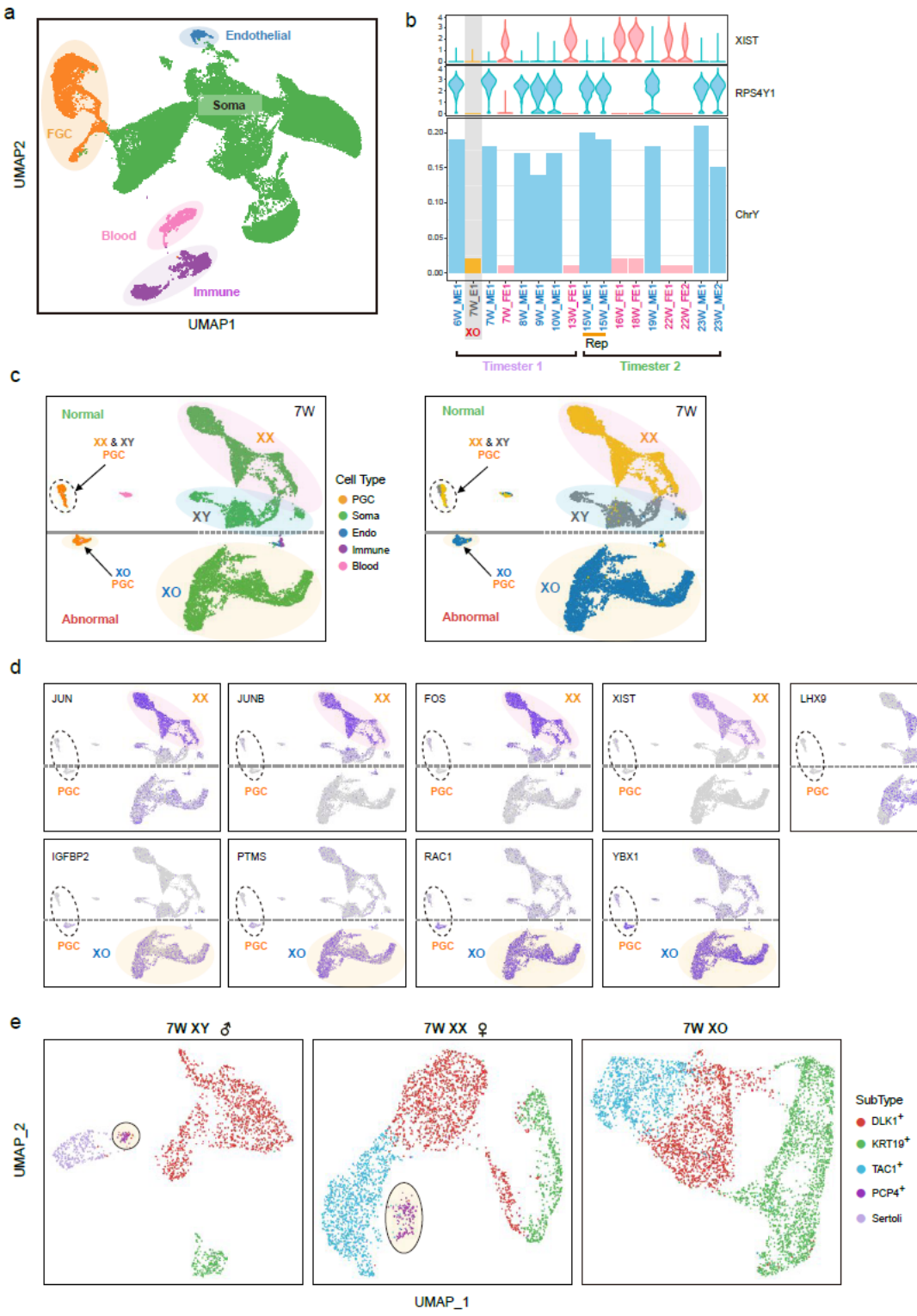

Figure 1

Legend not available in this version.

Figure 1

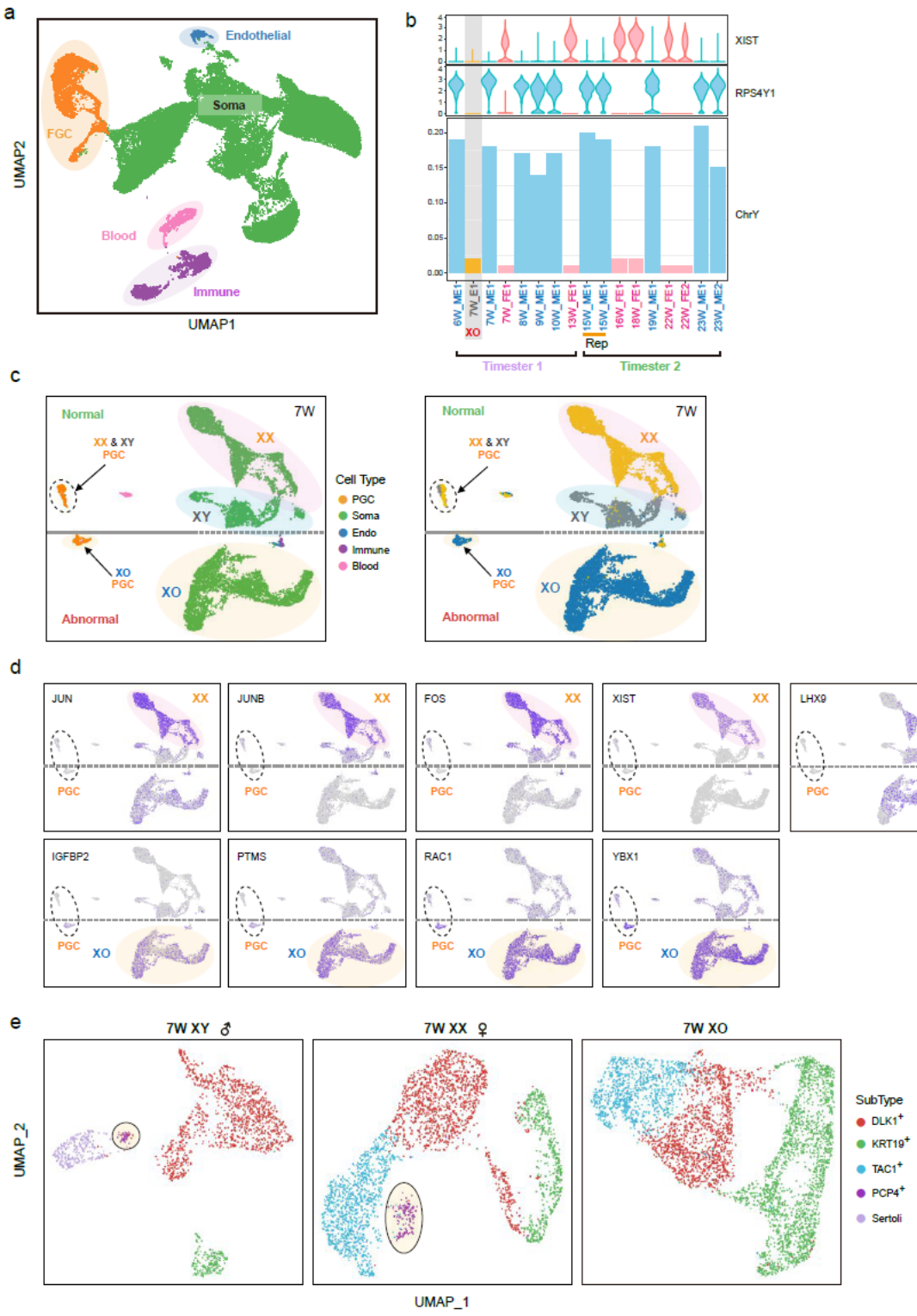

Figure 1

Legend not available in this version.

Figure 2

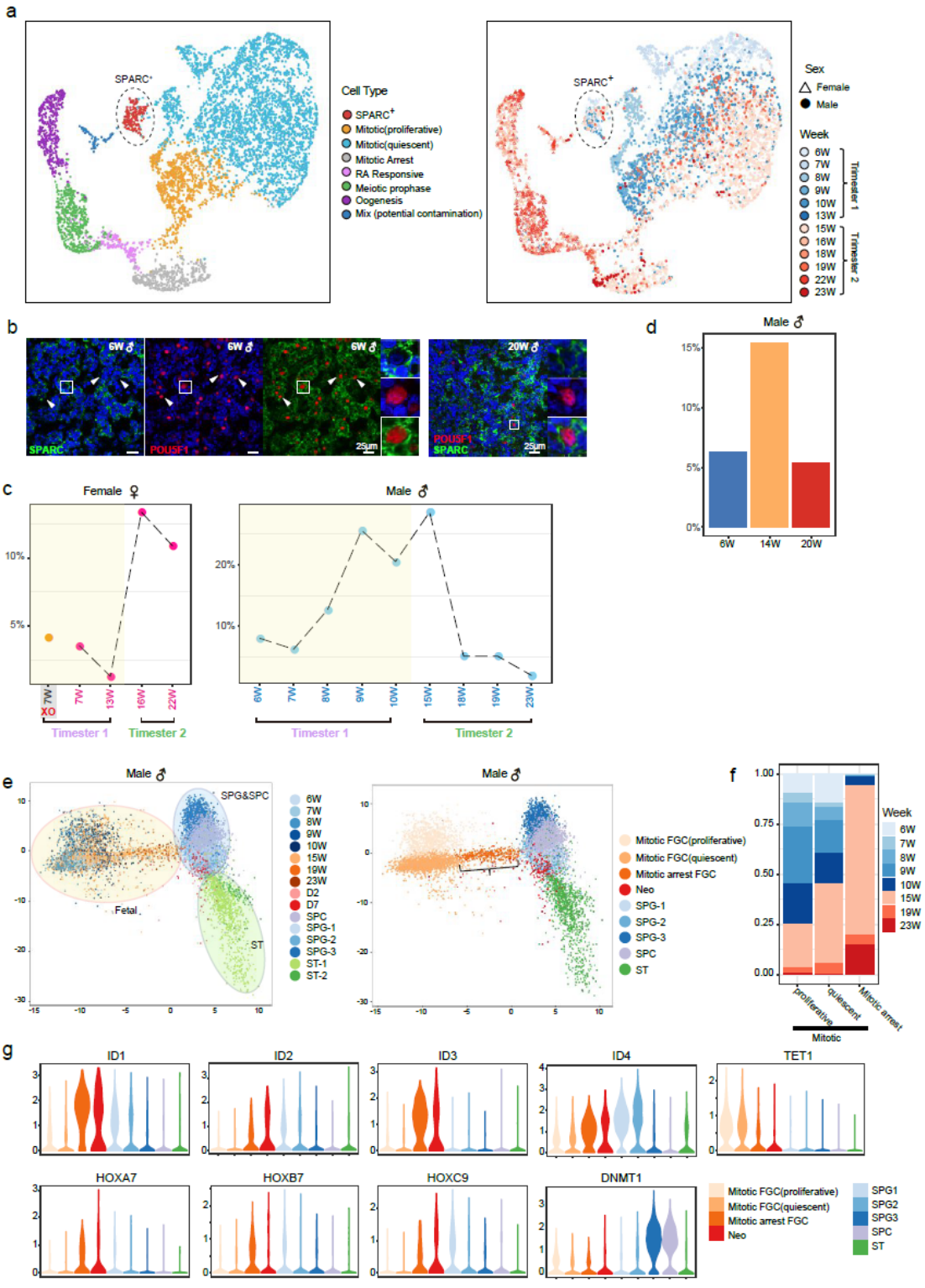

Figure 2

Legend not available in this version.

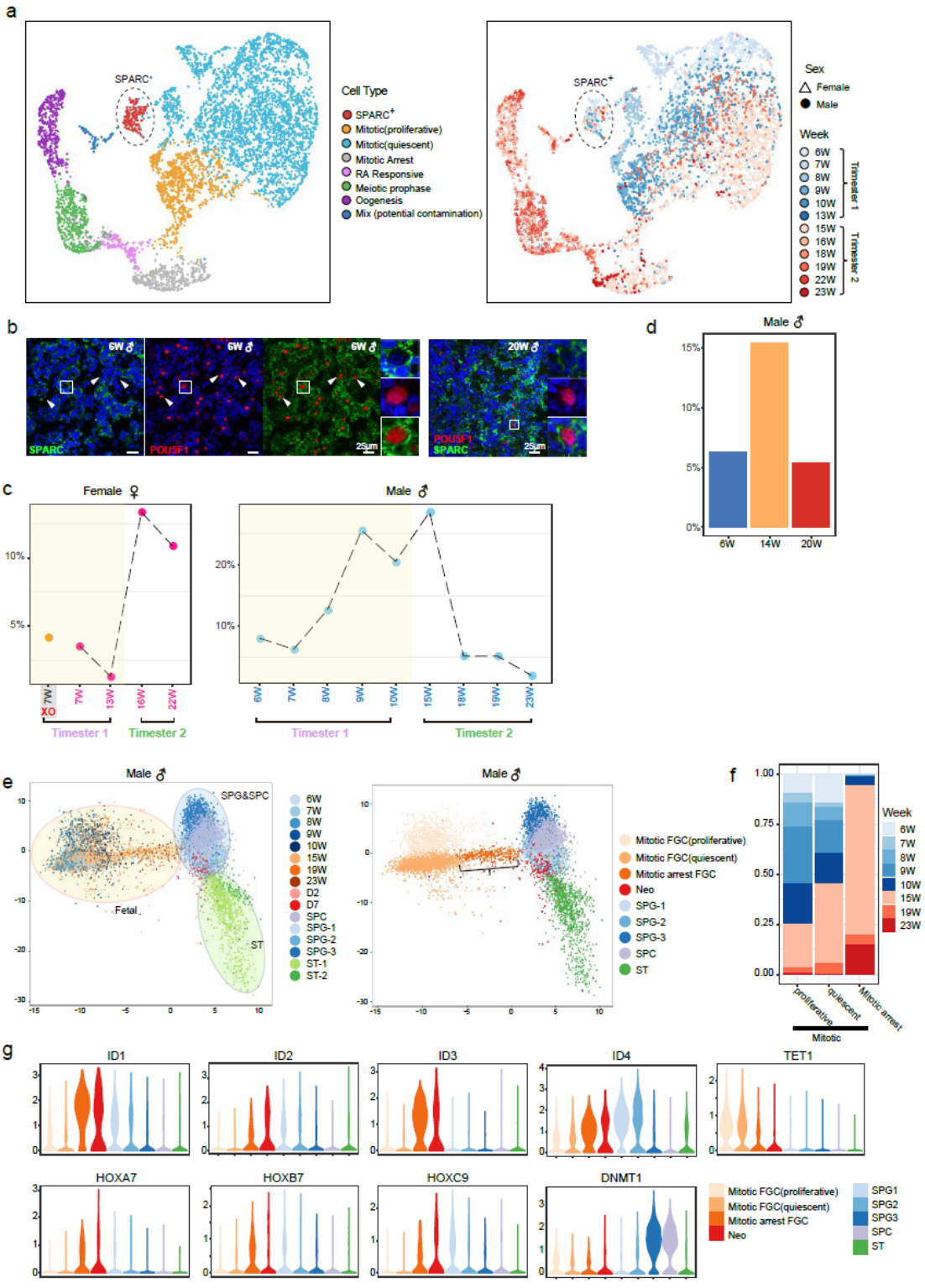

Figure 2

Legend not available in this version.

Figure 2

Figure 3

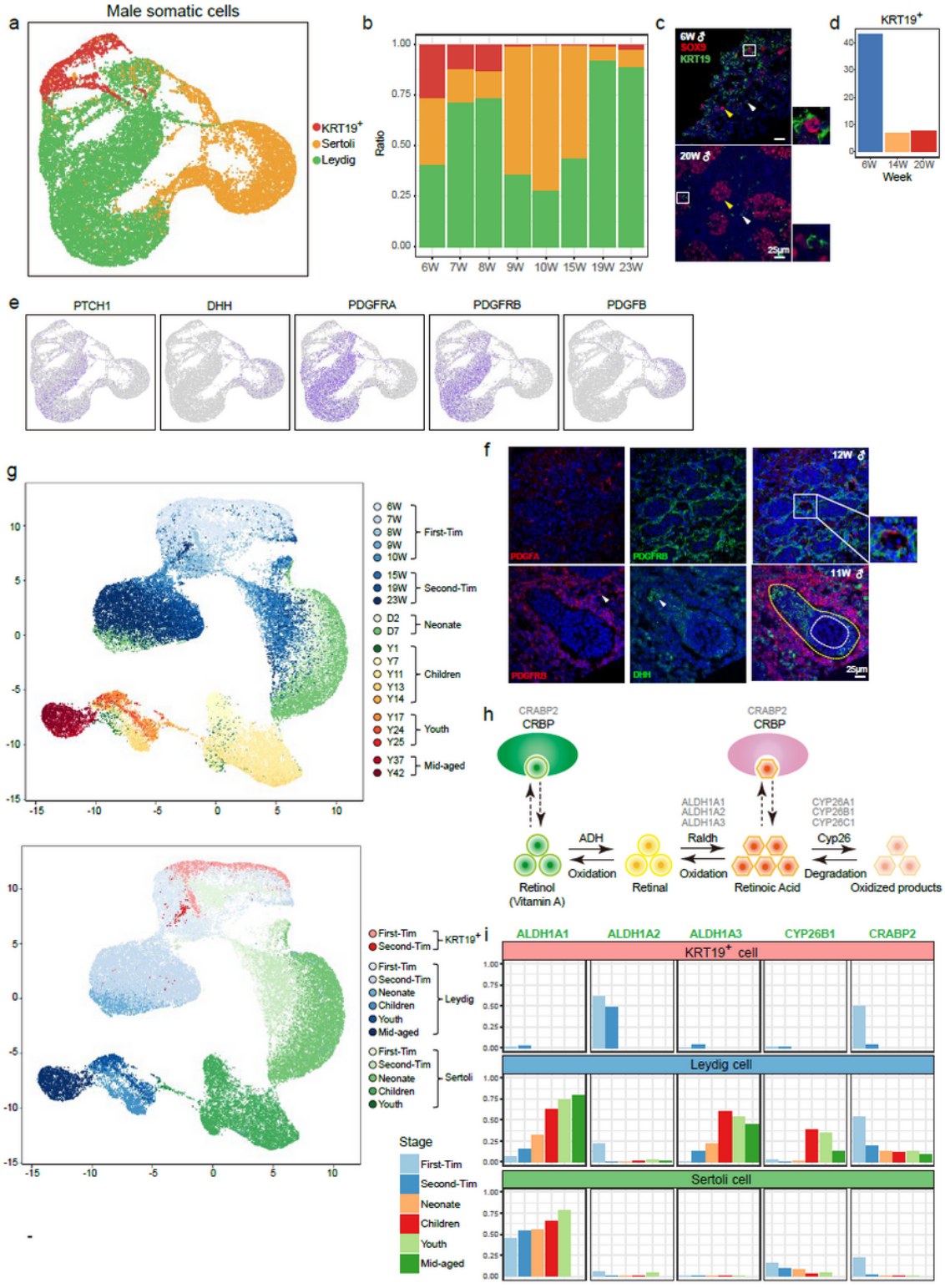

Figure 3

Legend not available in this version.

Figure 3

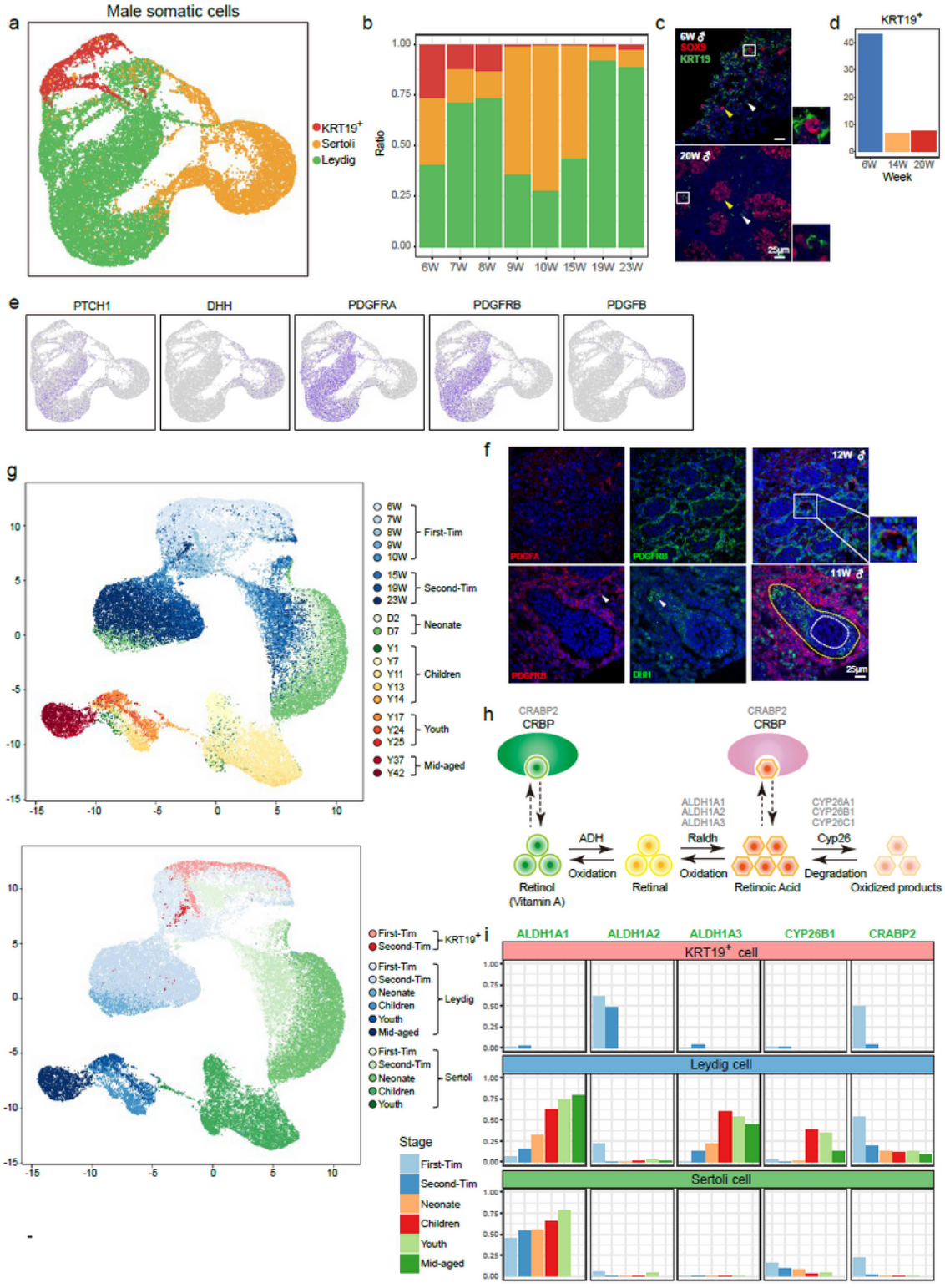

Figure 3

Legend not available in this version.

Figure 4

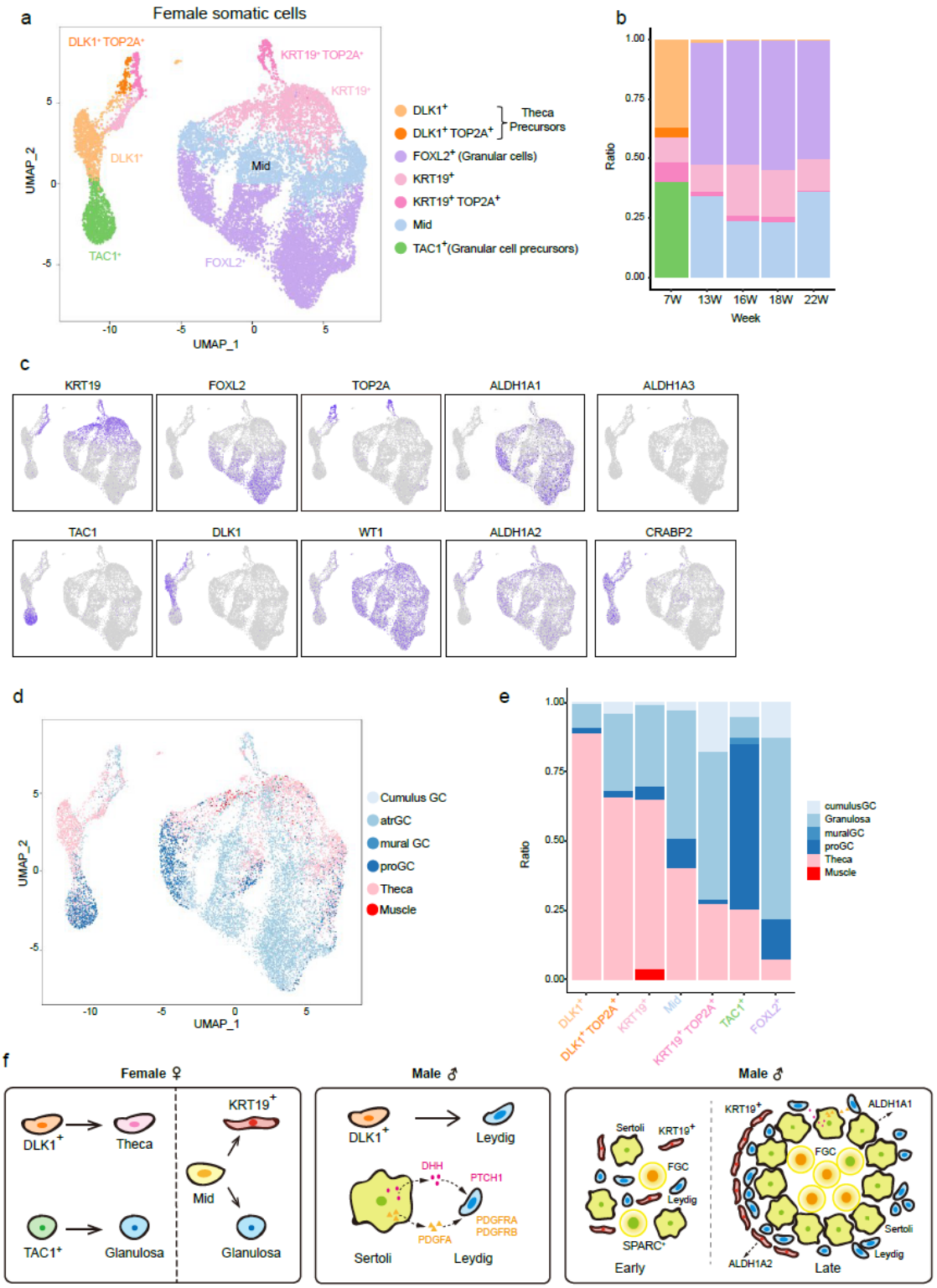

Figure 4

Legend not available in this version.

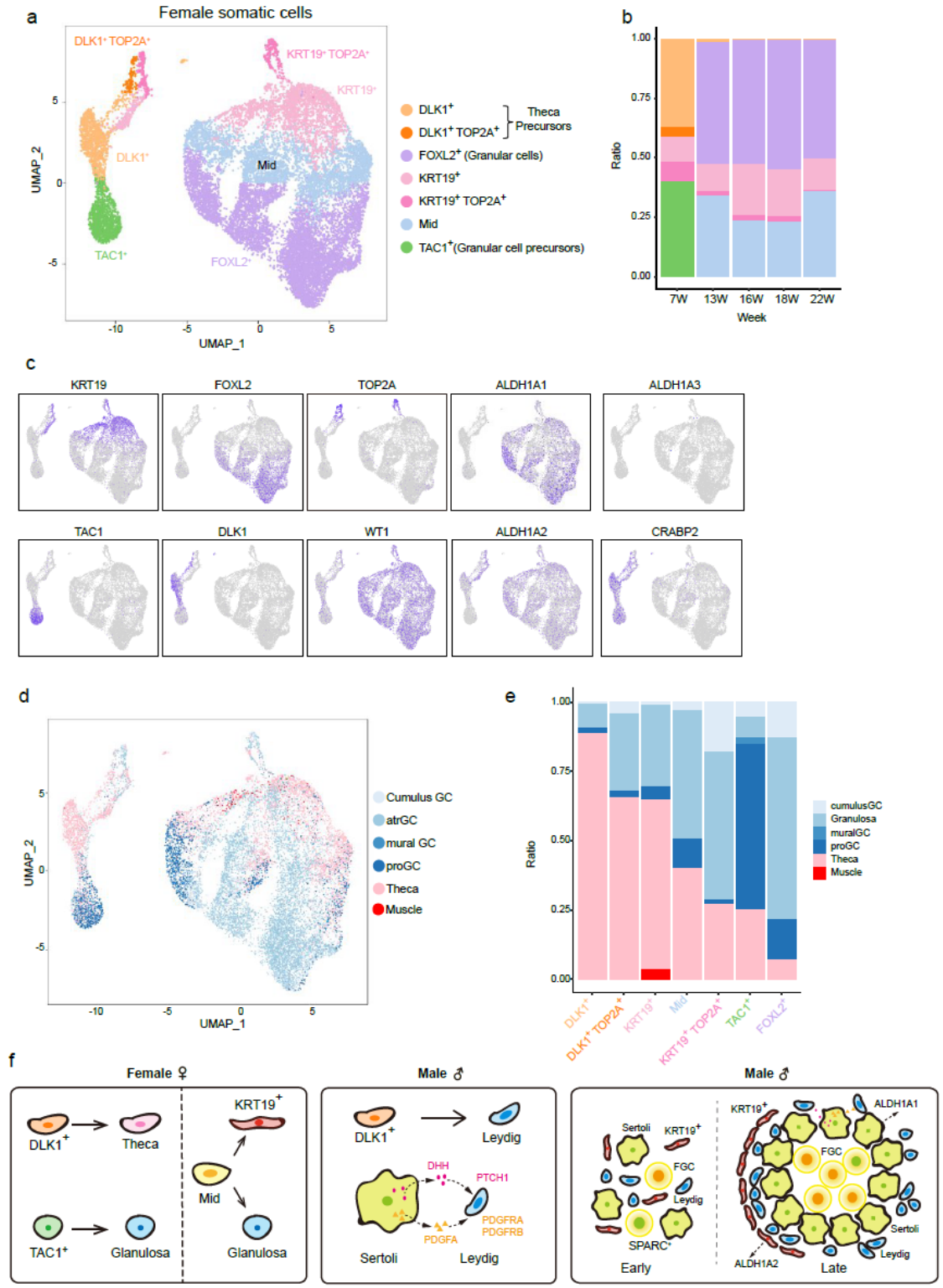

Figure 4

Legend not available in this version.

Figure 4

Figure 5

Figure 5

Figure 6

Figure 6

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- WRHumanGonadMethodV5.docx
- WRHumanGonadMethodV5.docx
- Table1GlobalCellTypeIdentificationUMAP.xlsx
- Table1GlobalCellTypeIdentificationUMAP.xlsx
- Table2HE7WPGCDEGs.xlsx
- Table2HE7WPGCDEGs.xlsx
- Table3PGCUMAP.xlsx
- Table3PGCUMAP.xlsx
- Table4PGCsubtypeDEGs.xlsx
- Table4PGCsubtypeDEGs.xlsx
- Table5MaleSomaticCellUMAP.xlsx
- Table5MaleSomaticCellUMAP.xlsx
- Table6FemaleSomaticCellUMAP.xlsx
- Table6FemaleSomaticCellUMAP.xlsx
- figures1v5.pdf
- figures1v5.pdf
- figures2v8.pdf
- figures2v8.pdf
- figures3v5.pdf

- figures3v5.pdf
- supplement11.pdf
- supplement12.pdf
- supplement13.pdf
- supplement14.pdf
- supplement15.pdf
- supplement16.pdf
- supplement17.pdf