Spring 2024 – Systems Biology of Reproduction
Discussion Outline – Testis Systems Biology
Michael K. Skinner – Biol 475/575
CUE 418, 10:35-11:50 am, Tuesday & Thursday
February 22, 2024
Week 7

Testis Systems Biology

Primary Papers:

1. Guo, et al. (2021) Cell Stem Cell 28,764-778
2. Endo, et al. (2015) PNAS E2347-2356
3. Guo, et al. (2018) Cell Research 28:1141-1157

Discussion

Student 5: Reference 1 above

- What was the technology used?
- What experimental design was used?
- What insights were obtained on testis somatic cell and germ cell origins?

Student 6: Reference 2 above

- What was the experimental design and culture system used?
- What spermatogenic process occurred in vitro?
- How could this technology be applied?

Student 7: Reference 3 above

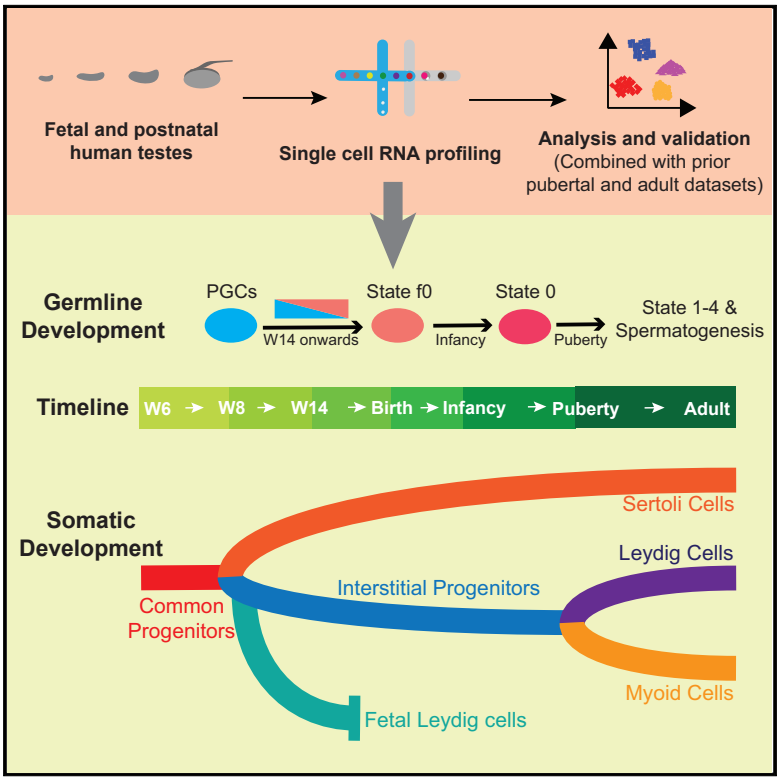
- What is the experimental and systems approach?
- What single cell expression and epigenetic relationships exist?
- What insights are provided on testis cell biology?

Cell Stem Cell

Resource

Single-cell analysis of the developing human testis reveals somatic niche cell specification and fetal germline stem cell establishment

Graphical Abstract



Highlights

- A transcriptional single-cell atlas of the fetal and postnatal human testes
- Somatic niche cell types derive from a common progenitor ~7 weeks after fertilization
- PGCs transition directly into fetal state 0-like cells (state f0) starting at week 14
- Fetal somatic niche cell specification precedes the PGC-to-state f0 transition

Authors

Jingtao Guo, Enrique Sosa, Tsotne Chitiashvili, ..., Jan-Bernd Stukenborg, Amander T. Clark, Bradley R. Cairns

Correspondence

jingtao.guo@hci.utah.edu (J.G.), clarka@ucla.edu (A.T.C.), brad.cairns@hci.utah.edu (B.R.C.)

In Brief

Guo et al. provide a transcriptional cell atlas of the fetal and postnatal human testes. Remarkably, starting from ~14 weeks postfertilization, fetal primordial germ cells transition to a cell state highly similar to postnatal spermatogonial stem cells. Furthermore, somatic niche specification precedes this transition, which is consistent with guiding fetal germline development.

Guo et al., 2021, Cell Stem Cell *28*, 764–778
April 1, 2021
https://doi.org/10.1016/j.stem.2020.12.004

CellPress

CellPress

Cell Stem Cell

Resource

Single-cell analysis of the developing human testis reveals somatic niche cell specification and fetal germline stem cell establishment

Jingtao Guo,[1,2,*] Enrique Sosa,[3,7] Tsotne Chitiashvili,[3,4,7] Xichen Nie,[1] Ernesto Javier Rojas,[3] Elizabeth Oliver,[5] DonorConnect,[6] Kathrin Plath,[4] James M. Hotaling,[2] Jan-Bernd Stukenborg,[5] Amander T. Clark,[3,*] and Bradley R. Cairns[1,8,*]

[1]Howard Hughes Medical Institute, Department of Oncological Sciences and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT 84112, USA
[2]Division of Urology, Department of Surgery, University of Utah School of Medicine, Salt Lake City, UT 84112, USA
[3]Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, Los Angeles, CA 90095, USA
[4]Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, CA 90095, USA
[5]NORDFERTIL Research Laboratory Stockholm, Childhood Cancer Research Unit, Bioclinicum J9:30, Department of Women's and Children's Health, Karolinska Institutet and Karolinska University Hospital, Solna 17164, Sweden
[6]DonorConnect, Murray, UT 84107, USA
[7]These authors contributed equally
[8]Lead contact
*Correspondence: jingtao.guo@hci.utah.edu (J.G.), clarka@ucla.edu (A.T.C.), brad.cairns@hci.utah.edu (B.R.C.)
https://doi.org/10.1016/j.stem.2020.12.004

SUMMARY

Human testis development in prenatal life involves complex changes in germline and somatic cell identity. To better understand, we profiled and analyzed ~32,500 single-cell transcriptomes of testicular cells from embryonic, fetal, and infant stages. Our data show that at 6–7 weeks postfertilization, as the testicular cords are established, the Sertoli and interstitial cells originate from a common heterogeneous progenitor pool, which then resolves into fetal Sertoli cells (expressing tube-forming genes) or interstitial cells (including Leydig-lineage cells expressing steroidogenesis genes). Almost 10 weeks later, beginning at 14–16 weeks postfertilization, the male primordial germ cells exit mitosis, downregulate pluripotent transcription factors, and transition into cells that strongly resemble the state 0 spermatogonia originally defined in the infant and adult testes. Therefore, we called these fetal spermatogonia "state f0." Overall, we reveal multiple insights into the coordinated and temporal development of the embryonic, fetal, and postnatal male germline together with the somatic niche.

INTRODUCTION

As the germline stem cells of the adult testis, spermatogonial stem cells (SSCs) must properly balance self-renewal and differentiation to maintain lifelong spermatogenesis and fertility (Kanatsu-Shinohara and Shinohara, 2013). Adult SSCs are the culmination of a complex developmental process that begins in the embryo and continues through distinct fetal, juvenile, pubertal, and adult stages. The human germline is specified through the formation of primordial germ cells (PGCs), which occurs in the peri-implantation human embryo around the time of gastrulation (Chen et al., 2019; Tang et al., 2016; Witchi, 1948). Here, studies in the cynomolgus macaque and the porcine embryo (Kobayashi et al., 2017; Sasaki et al., 2016), as well as through the differentiation of human embryonic stem cells, suggest that primate and human PGCs originate during amnion specification and also from the posterior end of the nascent primitive streak (Chen et al., 2019; Zheng et al., 2019). Following specification, PGCs migrate through the hindgut, dorsal mesentery, and ultimately into the genital ridges at ~4–5 weeks postfertilization (Witchi, 1948). At ~6 weeks postfertilization, the genital ridges differentiate into either the male or female gonads, with sex-determining region on the Y chromosome (SRY) being essential for testicular development in males (Hanley et al., 2000; Mamsen et al., 2017; Yang et al., 2018). One of the earliest morphological changes in the male gonad at 6 weeks is the formation of nascent "cord-like" structures comprising PGCs and Sertoli-lineage cells surrounded by fetal Leydig and interstitial cells. In humans, this basic niche structure persists through the fetal and postnatal stages, as the formation of an organized seminiferous tubule does not occur until the pubertal stages in humans (Guo et al., 2020; Paniagua and Nistal, 1984).

Within the developing fetal testicular niche, recent genomics profiling and immunofluorescence (IF) imaging approaches have revealed that male germline cells undergo major developmental changes (Gkountela et al., 2013, 2015; Guo et al., 2015;

Li et al., 2017; Tang et al., 2015). Notably, the germline transitions from pluripotent-like PGCs migrating to and into the developing gonad to pluripotent-like and mitotically active PGCs in the gonad (called fetal germ cells [FGCs] or gonocytes), followed by the transition to "mitotically arrested" germ cells that repress the pluripotency-like program at/after weeks 14–18 (Li et al., 2017). Here, a key unanswered question during this stage of germline development involving the relationship between the mitotically arrested germ cells that arise during weeks 14–18 and the postnatal SSCs is as follows: are prenatal germ cells nearly identical to postnatal SSCs or are there major additional developmental stages that occur during prenatal stages? Notably, our prior work on the adult testis identified five distinct spermatogonial states (called states 0–4) accompanying human spermatogonial differentiation, with state 0 identified as the most naive and undifferentiated state (Guo et al., 2017, 2018, 2020), a result supported by single-cell RNA sequencing (scRNA-seq) profiling from other groups (Hermann et al., 2018; Li et al., 2017; Shami et al., 2020; Sohni et al., 2019; Wang et al., 2018). Consistent with this notion, state 0 is the predominant SSC state present in the infant testis, and state 0 SSCs express hundreds of state-specific markers, including *PIWIL4*, *TSPAN33, MSL3*, and *EGR4* (Guo et al., 2018). The key markers identified in state 0 SSCs are also expressed in the undifferentiated spermatogonial states identified by others in recent studies, such as the SSC1-B (Sohni et al., 2019) or SPG-1 adult spermatogonia population (Shami et al., 2020), as well as in spermatogonia profiled from human neonates (Sohni et al., 2019) and in undifferentiated spermatogonia from macaques (Shami et al., 2020). Here, we explore whether the previously identified mitotically arrested prenatal germ cells transcriptionally resemble state 0 postnatal spermatogonia, or instead represent a unique precursor that undergoes additional prenatal changes before birth.

The testis niche plays an important role in guiding the survival and differentiation of the male germline. In the adult testis, somatic niche cells, including Sertoli, Leydig, and myoid cells, provide physical and hormonal support for the successful execution of spermatogenesis from SSCs (Guo et al., 2018). The development of the functional adult testis and its organized tubule-like structure is completed at puberty, during which the final specification and maturation of all somatic niche cells takes place. Our prior work, which used scRNA-seq to study human testis development during puberty, revealed a common progenitor for Leydig and myoid cells that exists before puberty in humans, which is analogous to the somatic progenitor observed in fetal mice (Guo et al., 2020). However, during prenatal life, several key issues remain elusive, such as how the human testicular niche cell lineages are initially specified, whether they have a common progenitor, how the nascent gonad initially forms cords, and how niche cells differentiate further during subsequent fetal developmental stages to arrive at their postnatal states.

To address these questions, we profiled a total of ~32,500 unsorted single testicular cells from embryonic, fetal, and postnatal samples through the 10x Genomics Chromium platform. This unbiased profiling allowed us to examine the specification process in the somatic cell niche and the development of both the germline and niche cells; this enabled a detailed comparison of the cell types and developmental processes in infant, pubertal, and adult testis.

RESULTS

Single-cell transcriptomes of human embryonic, fetal, and postnatal testes

We obtained human testis tissues from 3 embryonic stages (6, 7, and 8 weeks postfertilization), 3 fetal stages (12, 15, and 16 weeks postfertilization), and 1 young infant stage (5 months postbirth) for comparisons to prior datasets from older infants, juveniles, and adults. To systematically investigate both germ cell and somatic cell development across embryonic and fetal stages, we prepared single-cell suspensions from these testicular tissues and performed scRNA-seq using the 10x Genomics platform. For embryonic and fetal samples, we profiled ~5,000 single cells per sample; for the young infant sample, we performed 2 replicates, and profiled ~2,500 single cells. From a total of ~32,500 cells, 30,045 passed standard quality control dataset filters and were retained for downstream analysis (see Method details). We obtained ~80,000–120,000 reads/cell, which enabled the analysis of ~1,800–2,500 genes/cell.

To analyze the dataset, we first performed UMAP (uniform manifold approximation and projection dimension reduction analysis) on the combined datasets using the Seurat package (Figures 1A and S1A: Butler et al., 2018). Interestingly, we observed a trend in which cells from 6 and 7 weeks cluster closely, and likewise, cells from 8, 12, 15, and 16 weeks cluster closely (Figures 1A and S1A), while also displaying temporal changes in particular cell types (Figures S1B and S1C). Further clustering analyses yielded 17 major clusters or cell types (Figure 1B) that were subsequently annotated using known gene markers (Figures 1C and S2). Clusters 1–4 are testicular niche cells from 6- and 7-week embryos, which uniquely express *NR2F2* and *TCF21*. Clusters 5–9 correspond to somatic cells from the interstitial and Leydig lineage from ≥8-week samples, which express *DLK1*. Clusters 10–11 are Sertoli lineage cells from ≥8-week samples, which express *AMH* and *SOX9*. Cluster 12 includes germ cells from all of the samples, which express known germ cell markers (e.g., *TFAP2C*, *DAZL*) with a subset expressing markers of pluripotency (e.g., *POU5F1*, *NANOG*). Clusters 13–17 correspond to endothelial cells (cluster 13, *PECAM1*$^+$), macrophages (cluster 14, *CD4*$^+$), smooth muscle cells (cluster 15, *RGS5*$^+$), red blood cells (cluster 16, *HBA1*$^+$), and fetal kidney cells (cluster 17, *CYSTM1*$^+$), respectively. We also provide examples of the many additional markers that were used to define these cell types (Figure S2).

Emergence of state 0 SSCs as PGCs exit mitosis and repress pluripotency

Development of the male germline was examined by parsing out and analyzing the germ cells separately from the somatic cells of the prenatal and postnatal (5 months) testes (cluster 12 from Figure 1B). To place the embryonic, fetal, and postnatal germ cells in a more complete developmental timeline and enable comparisons, we combined these data with data from infant germ cells (1 year old) and adult spermatogonial states (states 0–4) from our prior published work (Guo et al., 2018), which was also profiled on the 10x Genomics platform. A combination of dimension reduction (via t-distributed stochastic neighbor embedding [t-SNE]) and pseudotime analysis revealed seven defined clusters and a single pseudo-developmental trajectory that ordered

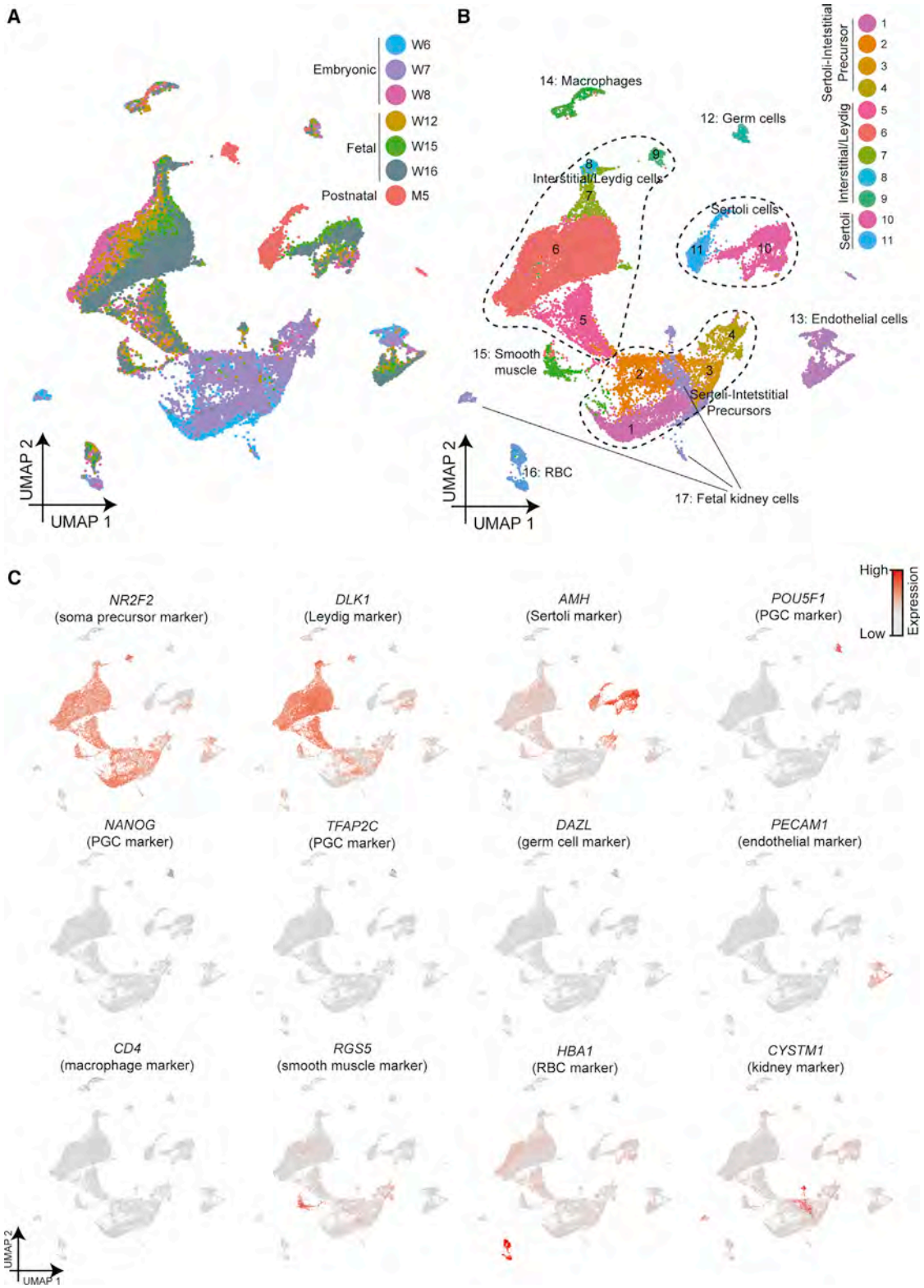


Figure 1. Single-cell transcriptome profiling and analysis of the human fetal and postnatal testis

(A) Dimension reduction presentation (via UMAP) of combined single-cell transcriptome data from embryonic, fetal, and infant human testes (n = 30,045). Each dot represents a single cell and is colored according to its age/donor of origin.

(legend continued on next page)

and linked germ cells from the different stages (Figure 2A). Following the order of pseudotime, we observed that the first cluster of germ cells was largely composed of cells from 6 to 12 weeks, as well as a portion of germ cells from week 15 (Figures 2A and S3A). This cluster was called the "embryonic-fetal group." Their transcriptional identity is consistent with that of PGCs, including the expression of *TFAP2C*, *KIT*, *NANOG*, *POUF51*, *SOX17*, and others (Figure 2B), which is consistent with prior scRNA-seq results (Li et al., 2017). The next developmental stage along pseudotime consists of cells from 15- and 16-week fetal samples that group together with cells from the 5-month- and 1-year-old postnatal samples, and was thus called the "fetal-infant group" (Figures 2A and S3B). Interestingly, cells from the fetal-infant group lacked expression of the PGC markers mentioned above, and instead initiated the expression of multiple key state 0-specific markers (*PIWIL4*, *EGR4*, *MSL3*, *TSPAN33*, others), which were previously defined in the adult, infant, and neonatal testis. The subsequent clusters correspond to states 0–4 spermatogonia from adults, which display the sequential expression of markers associated with the subsequent developmental states: quiescent/undifferentiated (state 1; *GFRA1*$^+$), proliferative (states 2–3; *MKI67*$^+$, *TOP2A*$^+$), and differentiating (state 4; *SYCP3*$^+$) (Figures 2A, 2B, and S3C), which is consistent with our previous work (Guo et al., 2017, 2018). This pseudotime order was further supported by orthogonal Monocle-based pseudotime analysis (Figures S3D and S3E). A more systematic analysis via heatmap and clustering yielded 2,448 dynamic genes and provided a format to explore and display the identity, Gene Ontology (GO) terms, and magnitude of genes that show dynamic expression along this germ cell differentiation timeline (Figure 2C; Table S1). The embryo-fetal group (PGCs) displayed a high expression of genes (cluster 1) associated with signaling and gonad and stem cell development, which were then abruptly repressed between weeks 15 and 16, coinciding with the transition to the subsequent fetal-infant group. Here, we also observe the upregulation of many transcription- and homeobox-related genes (cluster 2) in the fetal-infant group, and the clear upregulation of markers of state 0 spermatogonia. Interestingly, the transition from the fetal-infant group to state 0 spermatogonia is characterized by a deepening and reinforcement of the state 0 gene expression signature, rather than a large number of new genes displaying upregulation. For example, differential gene expression analysis comparing fetal germ cells to adult state 0 spermatogonia identified only 2 genes (*ID3* and *GAGE12H*; 2-fold, $p < 0.05$) that display fetal-specific expression (Figure S4G). Consistent with prenatal-postnatal similarity, we observe germ cells from both younger and older infants located in the fetal-infant and adult state 0 clusters. Our results revealed that the spermatogonia present in young and older infants (called state 0) are highly similar to the fetal germline cells that emerge directly after PGCs exit the pluripotent-like state. Given this similarity, we call these fetal (f) cells state f0.

To validate our scRNA-seq profiles at the protein level, we performed IF staining for key markers. The proportion of NANOG$^+$ (PGC marker) and MKI67$^+$ (proliferation marker) decreased from 5 to 19 weeks (Figures 2D and S3G), supporting the notion that the exit from the pluripotent-like state and entry into G0 are temporally linked. We note that for NANOG, the loss of RNA signal based on transcription profiling appears more abrupt than the loss of protein, suggesting heterogeneity in the rates of protein loss. Regarding the acquisition of state 0 markers, no PIWIL4 positivity was detected in the 8- and 10-week samples; however, from week 14 onward, PIWIL4$^+$ cells were clearly detected, specifically in DDX4$^+$ germ cells (Figures 2E, 2F, and S3H). Thus, for the key pluripotency, proliferation, and state 0 markers tested, our IF staining results validate our scRNA-seq results.

Network expression dynamics during embryonic, fetal, and postnatal germ cell development

To define candidate key genes and networks linked to germline developmental stages and transitions, we conducted network analysis. Using weighted correlation network analysis (WGCNA) (Langfelder and Horvath, 2008), we identified gene-gene interactions that display dynamic expression patterns during PGC differentiation to state f0 spermatogonia. Here, for the PGC upregulated network ("PGC network;" Figures S4A and S4D), we identified 2,126 genes and 122,360 interactions, and present the top 11 hub genes (and their interactions). As expected, several genes with known expression in PGCs were present, including *POU5F1*, *NANOG*, *NANOS3*, *SOX15*, and *TFAP2C* (Gkountela et al., 2015; Guo et al., 2015; Tang et al., 2015), confirming the robustness of our analysis. In addition, this analysis revealed *PHLDA3*, *PDPN*, *ITM2C*, *RNPEP*, *THY1*, and *ETV4* as prominent markers in mitotic PGCs, providing candidates for future analysis. For example, *PDPN*, *ITM2C*, and *THY1* encode cell surface proteins, and PDPN has successfully been used to isolate PGCs differentiated from human pluripotent stem cells (Sasaki et al., 2016). Regarding networks that accompany the differentiation of PGCs into state f0 spermatogonia, a large fraction of the identified genes show relatively broad expression within all subsequent spermatogonia stages, and thus we call this network the "spermatogonia network" (Figures S4B and S4E). We identified 771 genes and 31,557 interactions, and present the top 10 hub genes. Here, roles for *EGR4*, *DDX4*, *TCF3*, and *MORC1* in mammalian germ cells are well known. Interestingly, our analysis also indicates several additional factors (e.g., *RHOXF1*, *STK31*, *CSRP2*, *ASZ1*, *SIX1*, *THRA*) worthy of further exploration. For example, *RHOXF1* mutations in humans confer male infertility (Borgmann et al., 2016), and *MORC1* and *ASZ1* both play important roles in protecting the germline genome by repressing transposable element activity (Ma et al., 2009; Pastor et al., 2014), raising the possibility that they may coordinate with the PIWIL4 factor described below. We also examined the networks that were exclusively expressed in state

(B) Dimension-reduction presentation of combined single-cell transcriptome data from (A), labeled with corresponding cell categories and colored according to its cell type identity.
(C) Expression patterns of selected markers projected on the UMAP plot (A). For each cell cluster, 1 cell marker is shown in the main figure, accompanied by a gallery of additional markers in Figure S2.
See also Figures S1 and S2.

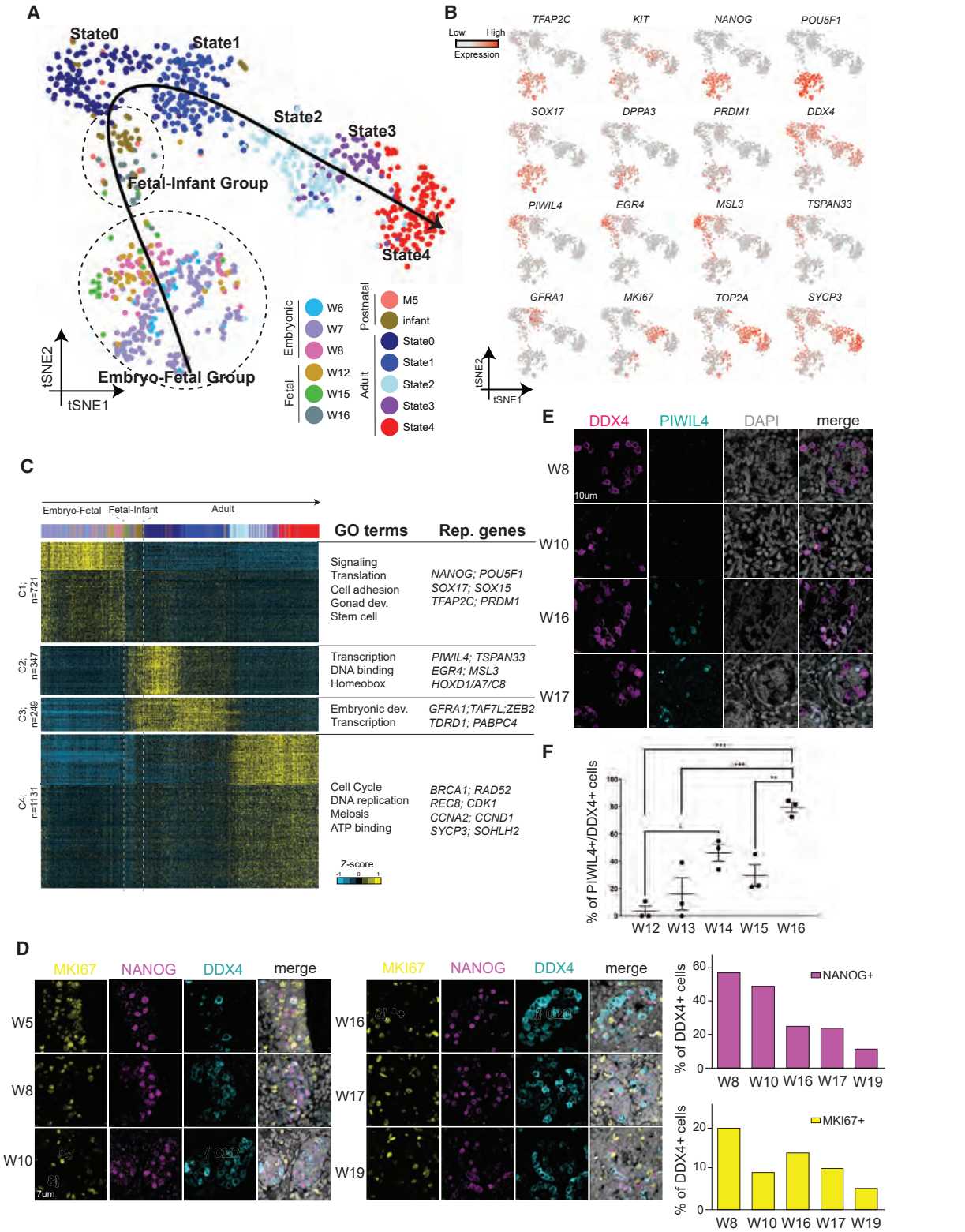


Figure 2. Gene expression dynamics during the development of human PGCs to adult spermatogonia

(A) Focused analysis (t-SNE and pseudotime) of the profiled germ cells (cluster 12 from Figure 1B) combined with infant germ cells and adult spermatogonia states (from Guo et al., 2018) revealed a single pseudo-developmental trajectory for germ cell development from embryo to adult. Cells are colored based on the ages of the donors.

(legend continued on next page)

0 SSCs ("state 0 network"; Figures S4C and S4F). We identified 190 genes and 8,841 interactions, and present the top 9 hub genes. Among them, *EGR4*, *CAMK2B*, *MSL3*, *PLPPR5*, *APBB1*, and *PIWIL4* were already identified in prior work (Guo et al., 2018; Sohni et al., 2019), whereas here, *NRG2*, *RGS14*, and *DUSP5* emerge as additional factors. Thus, our analysis confirms the roles of many known factors and provides a list of key candidate genes with less-studied functions in germ cell development, providing multiple avenues for future studies.

Embryonic specification and fetal development of interstitial and Sertoli lineages

Our cell type analyses revealed that the human embryonic and fetal testis stages consist primarily of somatic niche cells, including Sertoli cells and interstitial cells (including Leydig cells) (Figures S1B and S1C). Notably, we did not observe cells that resemble fetal myoid cells by examining myoid markers, including *ACTA2* and *MYH11*, which contrasts with observations in mice (Wen et al., 2016). Here, our profiling of early embryonic (weeks 6–7) testes provided the opportunity to examine Sertoli and interstitial/Leydig cell specification. To this end, we parsed out the fetal somatic niche cells that belong to the interstitial/Leydig and Sertoli lineages, along with the early cells of indeterminate cell type (clusters 1–8 and 10 from Figure 1B), and performed further analysis. Interestingly, reclustering and subsequent pseudotime analysis revealed one cell cluster at early pseudotime, which transcriptionally bifurcates into two distinct lineages later in pseudotime (Figure 3A). Notably, the early cluster was composed exclusively of cells from weeks 6–7, whereas cells from week 7 onward align along 2 distinct paths (Figures 3A, 3B, and S5A). Examination of known markers suggested that the 2 developmental paths represent Sertoli (left trajectory) or interstitial/Leydig (right trajectory) lineages, respectively (Figures 3C and 3D), and the existence of a heterogeneous pool of cells at weeks 6–7 from which both of these trajectories originate, raising the possibility of a common somatic progenitor population. Based on our clustering analysis, we then classified the embryonic-fetal interstitial and Sertoli development into seven stages (A–G), beginning with candidate common somatic progenitors (A) that differentiate into embryonic interstitial/Leydig progenitors (B), which undergo active proliferation (expressing high *MKI67*). The mostly quiescent embryonic Sertoli progenitors emerge at around week 7 (F). The embryonic interstitial progenitors (A) appear to differentiate into fetal interstitial progenitors (C and D) and also fetal Leydig cells (E), and embryonic Sertoli progenitors will differentiate into fetal Sertoli cells (G). Thus, our computational analysis suggests a heterogeneous multipotential progenitor for interstitial cells and Sertoli cells at 6–7 weeks, which then differentiates into Sertoli and interstitial (including Leydig) lineages between weeks 7 and 8.

To further define the gene expression programs that accompany male sex determination, we performed gene expression clustering analysis (k-means) to identify the gene groups that display dynamic expression patterns along the pseudotime developmental trajectories (Figure 4A; Table S1). Notably, the candidate progenitors (at weeks 6–7) express multiple notable transcription factors, including *GATA2*, *GATA3*, *NR2F1*, *HOXA*, and *HOXC* factors and others, with enriched GO terms that include signaling and vasculature development. In particular, several genes involved in tube development (e.g., *TBX3*, *ALX1*, *HOXA5*) are specifically expressed in these candidate progenitors, which is consistent with the initiation of tubule formation to create the testis cords at week 6 (Figures 4A and S5B).

This population of cells then bifurcates into distinct transcriptional programs consistent with embryonic Leydig or Sertoli cell progenitors. Along the Sertoli lineage, expressed genes are associated with chromatin assembly, extracellular region, and filament formation. Along the Leydig lineage, cells first express genes related to DNA replication, proliferation, and cell cycle, indicating a phase of Leydig lineage amplification, consistent with a much higher number of cells present on the Leydig lineage trajectory at and after 8 weeks compared to the Sertoli lineage (Figures 3B, 4A, and S5A). This is followed in the Leydig lineage by the upregulation of terms linked to extracellular matrix, cell adhesion and glycoproteins, and components and gene targets associated with both Notch and Hedgehog signaling. Consistent with the known roles of fetal Leydig cells in androgen production in mice (Shima et al., 2013, 2015), fetal Leydig cells placed at the end of pseudotime express high levels of genes related to steroid biosynthesis (e.g., *HSD3B2*: Figure 3D) and secretion. Interestingly, these cells emerge very early, by week 7, and persist for the remainder of the stages examined, suggesting both an early and a persistent role. For the Sertoli lineage, the fetal Sertoli cells express high levels of genes associated with structural functions. To validate the temporal features of steroidogenesis genes, we performed IF staining of CYP17A1, a marker for steroidogenesis highly expressed in fetal Leydig cells (Shima et al., 2013; Figures 4B and S5D). Notably, we found that CYP17A1 is absent in the genital ridge epithelium at 5.5 weeks, whereas robust staining is observed in the interstitial (non-cord) areas in all samples at $\geq$7 weeks, strongly suggesting that Leydig cell specification occurs at around week 7, consistent with our scRNA-seq findings. Furthermore, we observed that at week 8, not all interstitial cells are positive for CYP17A1. Here, we speculate that the fetal CYP17A1^{-} interstitial cells may be the interstitial cell population that gives rise to postnatal Leydig and peritubular cells.

(B) Expression patterns of known PGC and germ cell markers projected onto the tSNE plot from (A).
(C) k-means clustering of genes exhibiting differential expression (n = 2,448) along the germ cell pseudo-developmental trajectory. Each row represents a gene, and each column represents a single cell, with columns/cells placed in the pseudotime order defined in (A). Differential gene expression levels use a *Z* score as defined by the color key; associated GO terms (using DAVID version 6.7) are given on the right of the corresponding gene clusters.
(D) Protein co-immunofluorescence for markers of proliferation (MKI67, yellow), pluripotency (NANOG, magenta), and germ cells (DDX4, cyan) in samples from 5 to 19 weeks, and their corresponding quantification.
(E) Protein co-immunofluorescence for germ cell (DDX4) and state 0 (PIWIL4) markers in samples from 8 to 17 weeks.
(F) Quantification of the proportion of PIWIL4^{+} germ cells (DDX4^{+}) in weeks 12–16 fetal testis samples. At least 100 cells per replicate and 3 replicates were counted. Each replicate was from a unique donor. Data show the means ± SEMs (1-way ANOVA followed by a Tukey's post-test). Adjusted *p = 0.0136, **p = 0.0048, and ***p $\leq$ 0.0008.
See also Figures S3 and S4 and Table S1.

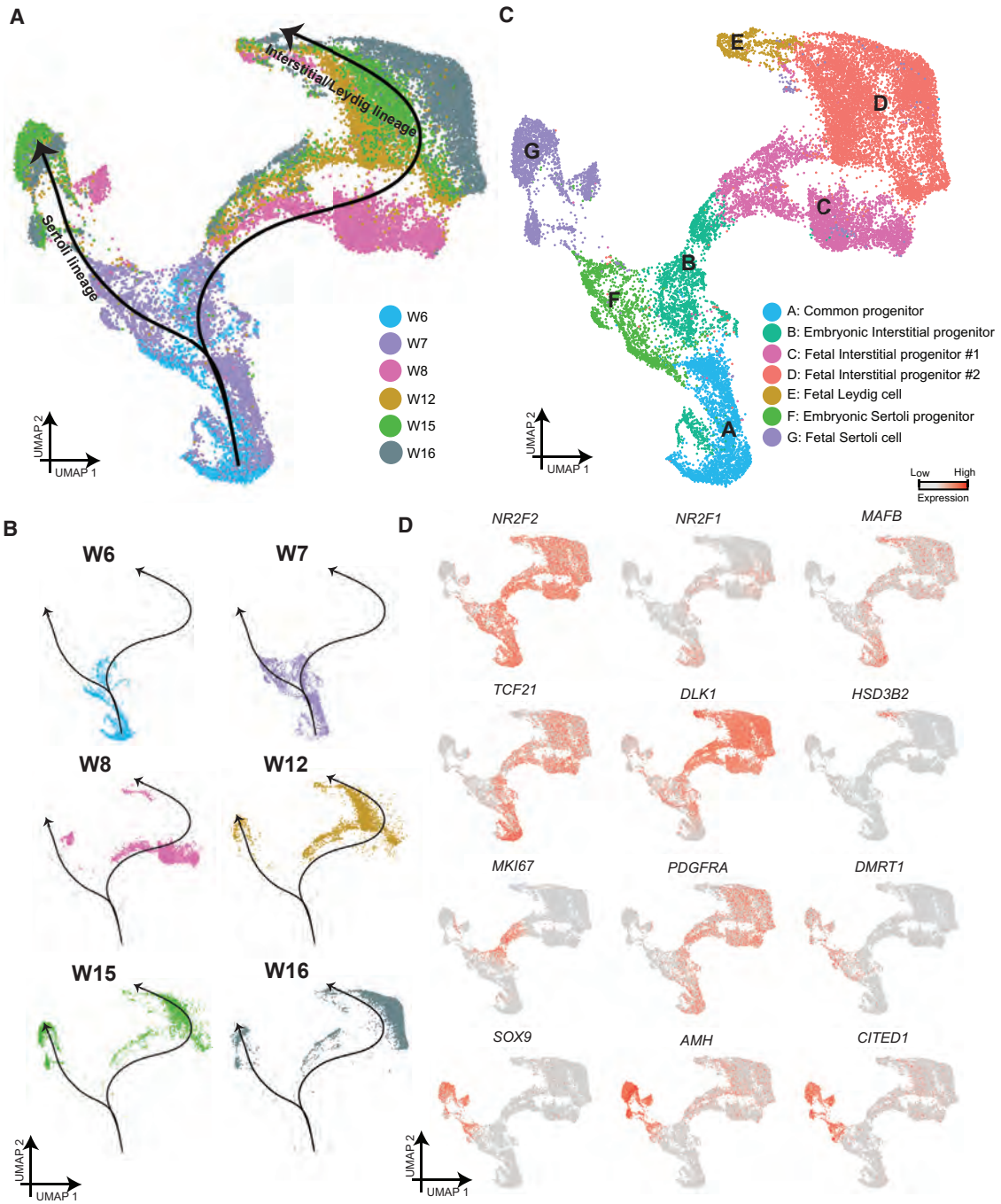


Figure 3. The specification of interstitial and Sertoli lineages

(A) Focused analysis (UMAP and pseudotime) of the testicular niche cells (clusters 1–11 from Figure 1B), with cells colored according to the ages of the donors.

(B) Deconvolution of the plot in (A) according to the ages of the donors.

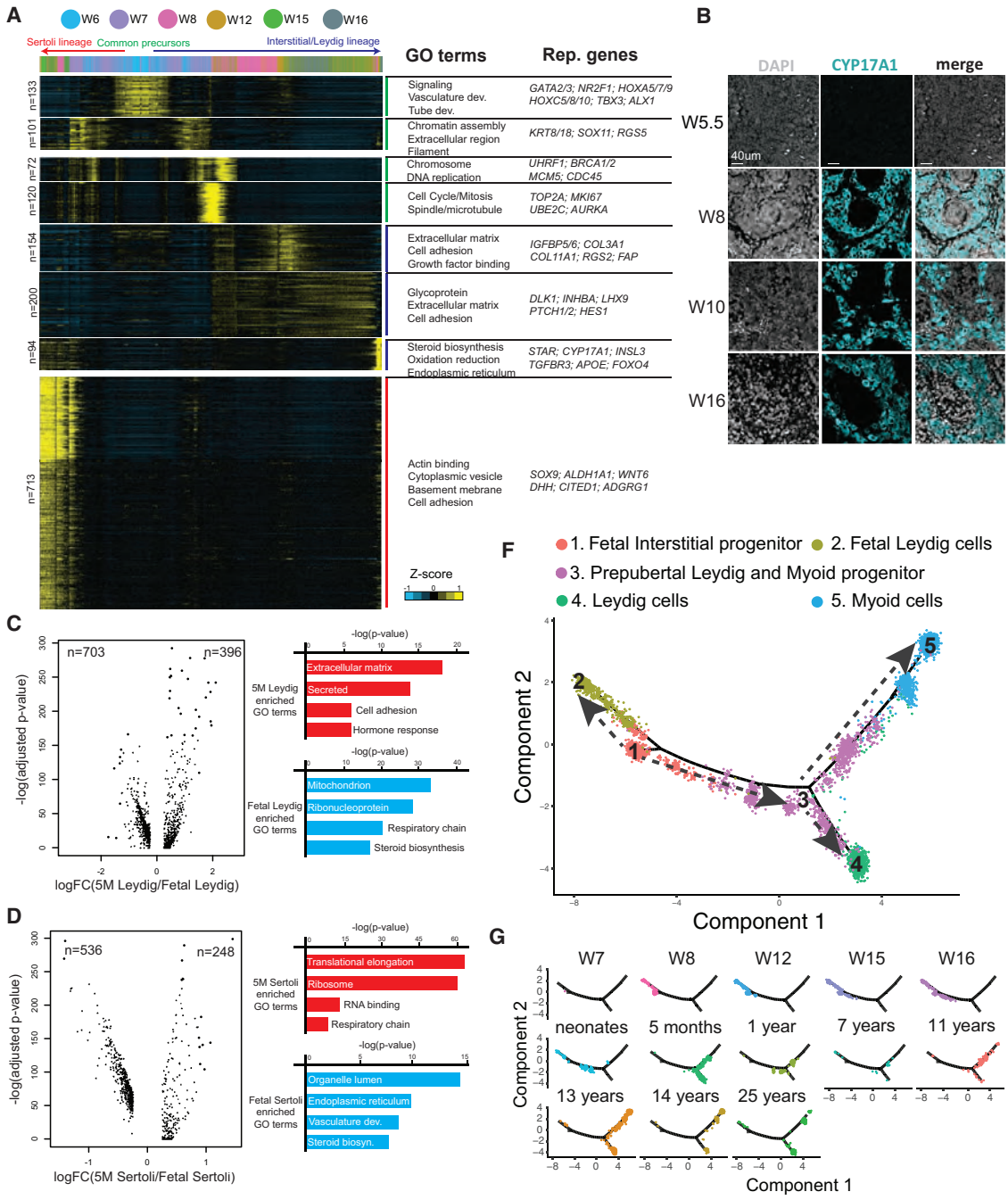
(C) Focused analysis (in A) of the testicular niche cells (clusters 1–11 from Figure 1B), with cells colored according to the ages/donors of origin.

(D) Expression patterns of known progenitor, interstitial/Leydig, and Sertoli markers projected onto the plot from (A).

See also Figure S5.

Relationship between fetal and infant Leydig and Sertoli cells

Our datasets provided an opportunity to compare and contrast fetal versus postnatal human Leydig and Sertoli cells. We found 396 or 703 genes differentially expressed (upregulated or downregulated, respectively) when comparing fetal to infant Leydig cells, respectively (bimodal test; adjusted $p < 0.01$; $|logFC| > 0.25$) (Figure 4C). As Leydig cells transition from fetal to infant, genes associated with the extracellular matrix, secretion, cell adhesion and hormonal response are upregulated, while those



(legend on next page)

with mitochondrial function and steroid biosynthesis (e.g., *CYP17A1*, *HSD3B2*, *STAR*) are downregulated (Figure 4C). Likewise, we found 536 or 248 genes differentially expressed in the infant or fetal Sertoli cells, respectively (Figure 4D). As Sertoli cells transition from fetal to infant, genes associated with translation and respiratory chain are upregulated, and these cells with endoplasmic reticulum and steroid biosynthesis are downregulated (Figure 4D). To confirm, we performed IF staining of CYP17A1 (Shima et al., 2013) and found that its expression is undetected in the postnatal samples (Figure 4E), suggesting that fetal Leydig cells disappear or differentiate after birth in humans, which is consistent with discoveries in mice (Svingen and Koopman, 2013). Our results suggest that human fetal Leydig and Sertoli cells both exhibit expression of steroid biosynthetic genes, whereas this property is downregulated in the postnatal samples tested.

Our prior work based on juvenile human testes showed that Leydig and myoid cells share a common progenitor at prepubertal stages (Guo et al., 2020). To gain a deeper understanding of the relationship between the fetal interstitial progenitors and prepubertal Leydig/myoid progenitors, as well as insight into how the common progenitor for the Leydig and myoid lineage is specified from fetal and postnatal precursor cells, we performed additional analysis. Here, we combined *in silico* the scRNA-seq datasets from fetal interstitial cells (clusters C, D, and E from Figure 3C), neonatal Leydig cells (Sohni et al., 2019), and the postnatal and adult Leydig/myoid cells (Guo et al., 2020). Following cell combination, we performed Monocle pseudotime analysis, which aims to provide the developmental order of the analyzed cells through computational prediction (Figures 4F and 4G). Here, the pseudotime trajectories (depicted by the dashed arrows in Figure 4F) agree nicely with developmental order based on age (Figure 4G), suggesting that fetal interstitial progenitor cells give rise to the postnatal and prepubertal Leydig/myoid progenitor cells. In addition, the analysis suggests that the fetal Leydig cells, which originate from the fetal interstitial progenitors, are absent in the postnatal and infant stages, a result confirmed by our immunostaining data (Figure 4E).

Key factors correlated with embryonic specification of interstitial and Sertoli lineages

Whereas testicular niche cells from 8 to 16 weeks expressed transcription factors characteristic of advanced interstitial or Sertoli cell lineages, the cells from the 6-week gonads lack these late markers, which initially emerge at week 7 (Figures 3A–3C). To better understand the genes expressed during the time of somatic specification, we parsed out the 6- and 7-week cells (from Figure 3A) and performed a more detailed analysis. Here, principal-component analysis (PCA) of the 6- and 7-week cells revealed that a large portion of the cells did not display markers distinctive for either interstitial or Sertoli cells (Figure 5A), suggesting a heterogeneous population in which the Sertoli and Leydig/interstitial precursors are emerging. An orthogonal analysis via Monocle also confirmed similar patterns and properties (Figures S6C–S6E). Based on the gene expression patterns (Figure 5B), we can assign the cells at the bottom as the embryonic interstitial/Leydig lineage (expressing *DLK1* and *TCF21*), and the cells at the top right as the embryonic Sertoli lineage (expressing *SRY*, *DMRT1*, *SOX9*, *AMH*, and others).

Next, we sought to identify candidate key transcription factors that may participate in initial somatic lineage specification (Figure 5B). Interestingly, a set of *GATA* family factors displayed sequential and largely non-overlapping patterns: *GATA3* expression was earliest, at the top and left edge of the PCA plot (mostly 7 week), *GATA2* started to express somewhat later, and *GATA4* was expressed in a later population that was progressing toward the Sertoli lineage. Many other factors also display sequential expression. For example, *NR2F1*, *MAFB*, and *TCF21* show relatively early expression (similar to *GATA2*), while *TCF21* expression persists through the development of the Leydig lineage, but not the Sertoli lineage. Notably, both *ARX* and *NR0B1* are expressed at the bifurcation stage. For the Sertoli lineage, these early markers cease expression at lineage specification, followed by the expression of *SRY* and *DMRT1* as the earliest markers of the lineage, and then followed by *SOX9*.

Finally, we performed extensive IF to validate our genomics findings. We observed GATA3 throughout the genital ridge epithelium at week 5, which became restricted to a subpopulation of interstitial cells at weeks 6–7, and by week 8, GATA3 protein becomes undetectable (Figure 5C). In addition, GATA4 expression is evident both inside and outside the cords from week 6 and onward (Figures 5D and S5B). To evaluate Sertoli lineage specification, we stained for DMRT1 alongside either a germ cell marker (DDX4) or an additional Sertoli cell marker (SOX9) (Figures 5E and 5F). As expected, DMRT1 and SOX9 protein were undetectable in the GATA3/GATA2$^+$ genital ridge epithelium containing DDX4$^+$ PGCs at week 5 (Figure 5E). However, by 8 weeks (after cord formation), DMRT1$^+$ and SOX9$^+$ Sertoli cells are identified (Figure 5F). Taken together, our IF staining results confirm key markers discovered through our genomics approaches and provide additional insights into the physiology of testis cord development in the embryonic and fetal stages.

Figure 4. Gene expression dynamics during specification of interstitial and Sertoli lineages

(A) k-means clustering of genes exhibiting differential expression (n = 1,578) along interstitial/Leydig and Sertoli specification. Each row represents a gene, and each column represents a single cell, with columns/cells placed in the pseudotime order defined in Figure 3A. Differential gene expression levels use a *Z* score, as defined by the color key; associated GO terms (using DAVID version 6.7) are given on the right of the corresponding gene clusters.

(B) Immunostaining of Leydig marker CYP17A1 (cyan) in samples from 5 to 16 weeks.

(C and D) Analysis to reveal differentially expressed genes during Leydig (C) or Sertoli (D) cell differentiation from fetal to infant stages. Violin plot on the left of each panel displays the fold change (x axis) and adjusted p value (y axis). The right part of each panel represents the enriched GO terms and the associated p value.

(E) Immunostaining of Leydig marker CYP17A1 (cyan) in fetal and postnatal testis samples.

(F) Pseudotime trajectory (combined Monocle analysis) of fetal interstitial cells, prepubertal Leydig/myoid cells, and the adult Leydig and myoid cells. Cells are colored according to their predicted locations along pseudotime. Neonatal data were from Sohni et al., 2019; 1-year-old and 25-year-old data were from Guo et al., 2018, and 7- to 14-year-old data were from Guo et al., 2020.

(G) Deconvolution of the Monocle pseudotime plot according to ages/donors of origin.

See also Figure S6 and Table S2.

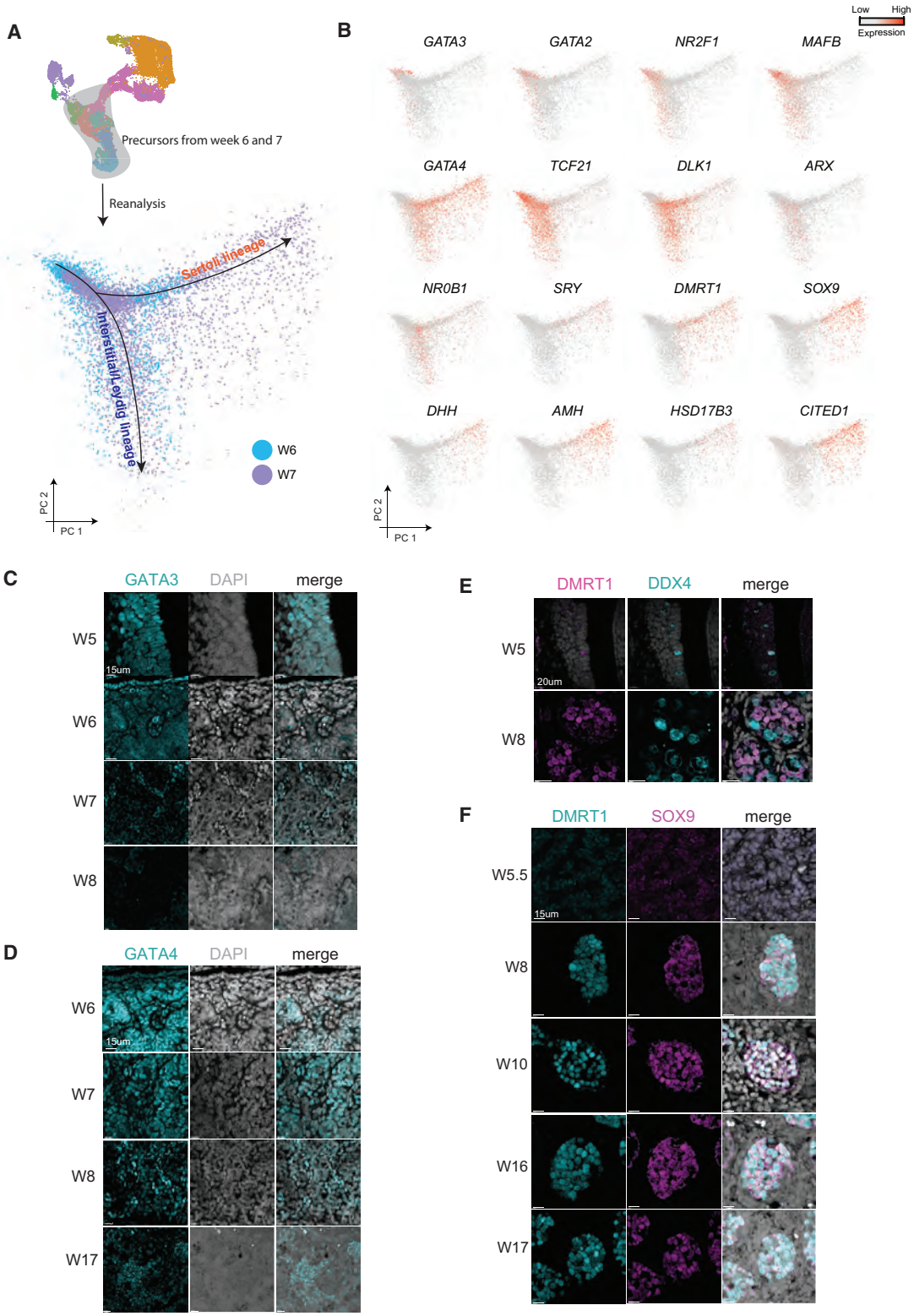


Figure 5. Key transcription factors involving the specification of interstitial and Sertoli cells

(A) Principal-component analysis of testicular niche progenitors from 6- and 7-week cells, revealing the existence of interstitial/Leydig and Sertoli lineage bifurcation.

(legend continued on next page)

DISCUSSION

PGCs are specified in the early embryo, followed by migration to the genital ridge (Chen et al., 2019; Tang et al., 2016; Witchi, 1948). The genital ridge then undergoes exquisite developmental programming to form the somatic cells of the testicular niche that support the survival and differentiation of the male germline during fetal life. Although prior studies from mice provide rich knowledge of the formation and lineage specification in the embryonic testis (reviewed in Svingen and Koopman, 2013), our understanding of human embryonic and fetal testis development has been much less studied, particularly in regard to the specification of the somatic lineages. Here, through the application of single-cell sequencing of unselected testicular cells, together with IF staining, we provide a detailed molecular overview of human fetal testis development, to help delineate the temporal molecular changes involved in human embryonic and fetal testis development and further differentiation.

One critical question we aimed to address is the transition of PGCs into spermatogonia, specifically the transcriptional relationship of differentiating male human PGCs during fetal life to postnatal state 0 SSCs, which have been identified as the most undifferentiated male germline stem cells in human infants and adults (Guo et al., 2018; Sohni et al., 2019), as well as primates (Shami et al., 2020). Combined with prior work (Guo et al., 2017, 2018, 2020; Sohni et al., 2019), our current work provides an evidence-based and detailed model for human germline development that spans embryonic, fetal, infant, pubertal, and adult stages (Figure 6A). During 6–12 weeks postfertilization, as the male somatic cell linages are being specified, human male PGCs express high levels of transcription factors associated with pluripotency (e.g., *POU5F1*, *NANOG*), together with classic well-characterized PGC transcription factors (e.g., *SOX17*, *TFAP2C*) and are proliferative. At 14 weeks, a subpopulation of PGCs initiates repression of the pluripotency-like program, and extinguishes expression of the early PGC genes (Li et al., 2017), while simultaneously turning on the state f0 spermatogonia programs (e.g., *PIWIL4*, *MSL3*, *EGR4*, *TSPAN33*). These state f0 spermatogonia are transcriptionally highly similar to the state 0 spermatogonia, and are found from fetal stages through infants within the seminiferous cords. Interestingly, when we examine the expression patterns of many key PGC or state f0 markers in a prior FGC dataset (Li et al., 2017; Figure S4H), we found that the mitotically arrested FGCs exhibit specific and high expression of state 0 genes (e.g., *PIWIL4*, *EGR4*, *MSL3*, *TSPAN33*) and low expression of PGC genes (e.g., *POU5F1*, *NANOG*, *TFAP2C*, *SOX17*). This observation strongly suggests that the previously defined mitotically arrested FGCs (Li et al., 2017), which also emerge at ~14 weeks postfertilization (Figure S4I), are likely the same cells as the state f0 defined in our study. Here, our prior derivation of infant state 0 cellular identity and their demonstrated similarity to the fetal population in the present study defines a critical linkage: PGCs differentiate and transition into state f0 spermatogonia and reinforce their state 0-like transcriptome as they transition between fetal germ cells and postnatal germ cells. By 5 months, all of the germline cells display a state 0 spermatogonial transcriptome, and cells with a PGC transcriptome are below the limit of detection. Consistent with our observations at 5 months and in infants, state 0 markers are also expressed in human neonatal germ cells (Sohni et al., 2019). We have revealed that state 0-like spermatogonia originate from PGCs at around weeks 14–16 of fetal life and persist through all of the prenatal and postnatal developmental stages, to provide a pool of undifferentiated spermatogonia in adults available for niche-guided transitions to more differentiated spermatogonial states and ultimately gametogenesis (Figure 6A).

Prior work in mouse models has revealed several factors and pathways that play important roles in lineage specification and progression of testicular somatic cells in mice (Liu et al., 2016; Svingen and Koopman, 2013; Yao et al., 2002). Recently, scRNA-seq has proven to be a powerful tool to study embryonic and neonatal mouse testis development (Stévant et al., 2019; Tan et al., 2020). Here, our work demonstrates that several key factors in early somatic lineages (e.g., WT1, NR2F1, SOX9, SRY, DMRT1) are shared between humans and mice. Furthermore, through our systematic examination of prenatal human testes via single-cell profiling and IF staining, we provide many additional candidate factors for future characterization, and reveal multiple human-mouse differences. For example, through IF staining of the genital ridge epithelium, we find no evidence of Sertoli cell or Leydig cell identity before 6 weeks postfertilization. Then, starting at week 6, our unbiased/unselected single cell transcriptome profiling identified rare fetal Leydig- and Sertoli-like cells. We also identified in pseudotime a large, closely related population of cells that is heterogeneously positive for developmental transcription factors, notably *NR2F1*, *GATA3*, and *GATA4* RNA. GATA3 protein analysis demonstrated that GATA3 is uniformly expressed by the genital ridge epithelium at week 5 postfertilization before specification of Sertoli and Leydig cells. Notably, at week 6, when cord formation initiates, GATA3 expression is restricted to a subpopulation of cells in the interstitium. In counterdistinction, GATA4 expression is evident and broad at 6–7 weeks postfertilization, and remains detectable at 17 weeks postfertilization. In the mouse embryo, GATA4 is known to be critical for genital ridge formation, and in the absence of GATA4, the bipotential gonads do not form (Hu et al., 2013). Given that *GATA3* is expressed in the genital ridge epithelium before *GATA4*, we speculate that *GATA3* may have a role in specifying the genital ridge in humans, whereas *GATA4* instead may be involved in maintaining the somatic cell lineages after 6 weeks postfertilization, when *GATA3* expression is reduced. In the mouse, NR5A1 (also called SF1) is another major transcription factor required for specifying the genital ridge

(B) Expression patterns of key factors that show specific patterns during the progenitor differentiation.
(C) Staining of transcription factors GATA3 (cyan) in the 5- and 8-week samples.
(D) Staining of transcription factors GATA4 (cyan) in the 6- and 17-week samples.
(E) Co-staining of Sertoli (DMRT1, magenta) and germ cell (DDX4, cyan) markers in the 5- and 8-week samples.
(F) Co-staining of 2 Sertoli cell markers, DMRT1 and SOX9, in the 5.5- to 17-week samples.
See also Figure S6.

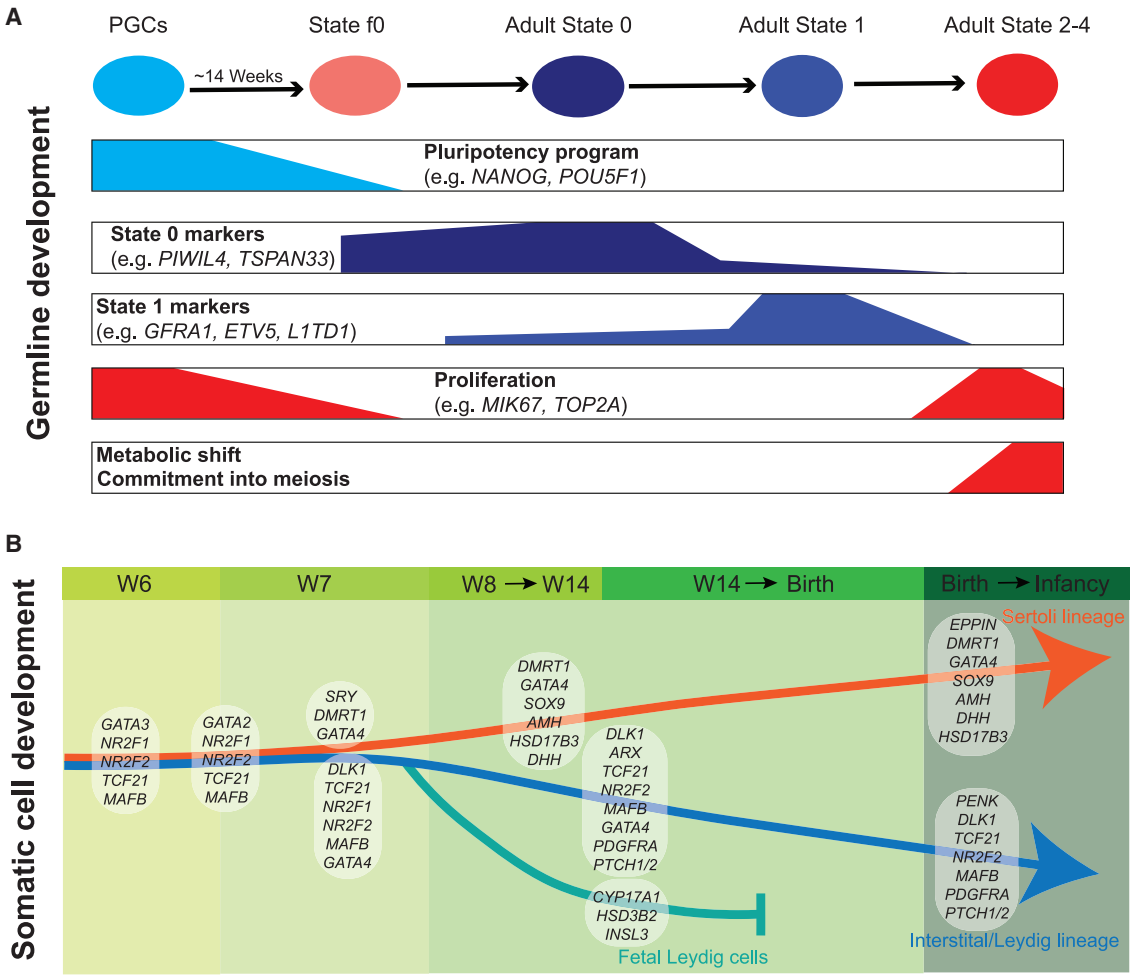


Figure 6. Proposed models for human germline development and somatic niche cell specification during prenatal and postnatal stages

(A) Schematic summarizing the combined gene expression programs and cellular events accompanying human PGC differentiation into adult SSCs.

(B) The timeline and proposed model for human testicular somatic niche cell development at embryonic, fetal, and postnatal stages. Specification of a unique progenitor cell population toward Sertoli and interstitial/Leydig lineages begins at around 7 weeks postfertilization, when the cord formation occurs.

epithelium (Hatano et al., 1996; Luo et al., 1994). However, we do not observe clear expression of *NR5A1* in the *GATA3*$^{+}$ human progenitors, providing a second example in which formation of the genital ridge epithelium in human embryos appears different from the mouse (Figure S5B). Analysis at the week 6–7 time point suggests that Leydig and Sertoli cell specification occurs at or near the same developmental time. Our IF studies at week 7 show both Sertoli cells in cords and Leydig cells outside the cords. This result represents a major difference from the mouse, in which Sertoli cells are specified first, and then Leydig cells are subsequently specified (Svingen and Koopman, 2013). Considering that the size of the fetal human testis is proportionally much larger than that of mice, the human testis progenitors may commit relatively early in development, followed by waves of proliferation, which may partly explain the developmental differences.

In addition to being specified at an equivalent developmental stage, we also discovered that the 6- and 7-week somatic niche progenitors expressed markers consistent with their ability to differentiate into interstitial/Leydig and Sertoli lineages by transiently expressing (in a small subset of cells) key transcription factors, including *ARX*, *NR0B1*, or *SRY*. This identity is further reinforced at 8 weeks, when all cells are distinguishable as interstitial/Leydig or Sertoli lineage cells. Notably, the establishment of the male somatic cell lineages in the embryonic testis occurs almost 2 months before the PGCs begin differentiating into state f0 (at 14–18 weeks). In contrast, in mice, there is only a 2-day delay in the timing of the male niche cell differentiation (at day 12) to the initiation of mouse PGC differentiation into prospermatogonia (at embryonic day 14) (Saitou and Yamaji, 2012; Svingen and Koopman, 2013; Western et al., 2008). The purpose of this 2-month delay in which human PGCs are shielded from initiating differentiation into state f0 spermatogonia in the seminiferous cord niche may be related to the need to increase the number of male germ cells through proliferation, given that these cells are MKI67^{+}, before initiation of state f0 differentiation and male-specific epigenetic reprogramming (Figure 6B).

The testis produces gametes in adult males through continuous niche-guided differentiation of SSCs, and a deep

understanding of this biology is needed to improve male reproductive health. Here, our work provides major insights into defining the timing and strategy of human testis formation and its development before and after birth. Notably, the state f0 germ cells that emerge at ~15 weeks during fetal life display remarkable similarities to the infant and adult state 0 cells, and thus allow us to link and depict the complete developmental progression of PGCs to adult state 0 cells. Furthermore, we provide detailed molecular characterization of a common somatic progenitor pool and its amplification and transition to testicular niche cells, as well as initial insights into testicular cord formation and possible roles in guiding germ cell development. These results should provide a foundation for future hypothesis-driven research, and could also help guide the reconstruction and study of the human early testis *in vitro*.

Limitations of study

Our present work focuses primarily on transcriptomic profiling with the additional protein validation of key markers; however, RNA expression does not always linearly reflect protein abundance. For example, as PGCs transition to state f0 SSCs (at/after 14 weeks), transcript levels for key PGC markers (e.g., *DDX4*, *NANOG*, *MKI67*) fall abruptly, whereas their protein levels reduce gradually, suggesting that complex posttranscriptional mechanisms exist to modulate protein levels. Furthermore, given the variations that may exist among different embryos and the challenges in accurately assessing the embryo ages, future studies with additional samples may refine the temporal aspects of our findings and may also reveal additional details regarding developmental processes and transitions. For example, although we know the transition from PGC to state f0 begins at ~week 14, there could be heterogeneity and individual variation regarding the time at which this conversion is complete. Finally, our identification of a common human fetal somatic cell precursor was based on transcriptional profiling and computational prediction. Here, further studies that use lineage tracing approaches in nonhuman primate models may provide a more definitive test of our model.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Sample transportation and storage
  - Human testis sample preparation for single cell RNA sequencing
  - Single cell RNA-seq performance, library preparation and sequencing
  - Processing of single cell RNA-seq data
  - Immunostaining of testicular tissues
  - Microscopy
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Weighted correlation network analysis

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stem.2020.12.004.

ACKNOWLEDGMENTS

We are grateful to the donors and their families, who made this work possible. We thank Brian Dalley and Opal Allen in the HCI High-Throughput Genomics Shared Resource for sequencing expertise, Chris Stubben and Tim Parnell in the HCI Bioinformatics Share Resource for bioinformatics assistance, and the DonorConnect staff for family consents and postnatal sample handling. Financial support was provided by Howard Hughes Medical Institute (HHMI) to B.R.C. and NCI P30CA042014 to Huntsman Cancer Institute core facilities. We thank the Technology Center for Genomics and Bioinformatics at the UCLA Johnson Comprehensive Cancer Center (JCCC) and the Next Generation Sequencing core at BSCRC for help with genomics approaches, and the Translational Pathology Core Laboratory for help with histology. We thank microscopy cores at the UCLA Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research Center (BSCRC) for help with imaging. Financial support was from the NIH to A.T.C. (R01HD079546). K.P. was supported by the BSCRC at UCLA, the David Geffen School of Medicine at UCLA, the UCLA JCCC, the NIH (R01HD098387 and P01 GM099134), and a Faculty Scholar grant from HHMI. T.C. was supported by a Boehringer Ingelheim PhD fellowship. E.S. was supported by the BSCRC Postdoctoral Fellowship at UCLA. We acknowledge Ian Glass from the University of Washington's Birth Defects laboratory for fetal tissue, supported by NIH human fetal tissue research grant no. 5R24HD000836-53. J.B.S. was supported by the Swedish Childhood Cancer Foundation (PR2019-0123; TJ2020-0026), the Magnus Bergvall Foundation, the Birgitta and Carl-Axel Rydbeck research Grant for Pediatric Research, and the Swedish Research Council (2018-03094).

AUTHOR CONTRIBUTIONS

J.G., A.T.C., and B.R.C. conceived and supervised the project. T.C. collected and processed the embryonic and fetal samples. J.G. collected and processed the postnatal tissues, and conducted all of the computational analyses, with help from X.N. Prenatal sample acquisition was led by A.T.C., K.P., and J.-B.S. Postnatal sample acquisition was led by J.M.H., with input from B.R.C. and J.G. E.S., E.J.R., and E.O. performed all of the IF experiments, which were supervised by A.T.C. and J.B.S. The manuscript was written by J.G. and B.R.C., with input from A.T.C. and the agreement of all of the authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 8, 2020
Revised: October 7, 2020
Accepted: December 8, 2020
Published: January 15, 2021

REFERENCES

Borgmann, J., Tüttelmann, F., Dworniczak, B., Röpke, A., Song, H.-W., Kliesch, S., Wilkinson, M.F., Laurentino, S., and Gromoll, J. (2016). The human RHOX gene cluster: target genes and functional analysis of gene variants in infertile men. Hum. Mol. Genet. *25*, 4898–4910.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Chen, D., Sun, N., Hou, L., Kim, R., Faith, J., Aslanyan, M., Tao, Y., Zheng, Y., Fu, J., Liu, W., et al. (2019). Human Primordial Germ Cells Are Specified from Lineage-Primed Progenitors. Cell Rep. *29*, 4568–4582.e5.

Gkountela, S., Li, Z., Vincent, J.J., Zhang, K.X., Chen, A., Pellegrini, M., and Clark, A.T. (2013). The ontogeny of cKIT+ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation. Nat. Cell Biol. *15*, 113–122.

Gkountela, S., Zhang, K.X., Shafiq, T.A., Liao, W.-W., Hargan-Calvopiña, J., Chen, P.-Y., and Clark, A.T. (2015). DNA Demethylation Dynamics in the Human Prenatal Germline. Cell *161*, 1425–1436.

Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., Yong, J., Hu, Y., Wang, X., Wei, Y., et al. (2015). The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. Cell *161*, 1437–1452.

Guo, J., Grow, E.J., Yi, C., Mlcochova, H., Maher, G.J., Lindskog, C., Murphy, P.J., Wike, C.L., Carrell, D.T., Goriely, A., et al. (2017). Chromatin and Single-Cell RNA-Seq Profiling Reveal Dynamic Signaling and Metabolic Transitions during Human Spermatogonial Stem Cell Development. Cell Stem Cell *21*, 533–546.e6.

Guo, J., Grow, E.J., Mlcochova, H., Maher, G.J., Lindskog, C., Nie, X., Guo, Y., Takei, Y., Yun, J., Cai, L., et al. (2018). The adult human testis transcriptional cell atlas. Cell Res. *28*, 1141–1157.

Guo, J., Nie, X., Giebler, M., Mlcochova, H., Wang, Y., Grow, E.J., Kim, R., Tharmalingam, M., Matilionyte, G., Lindskog, C., et al. (2020). The Dynamic Transcriptional Cell Atlas of Testis Development during Human Puberty. Cell Stem Cell *26*, 262–276.e4.

Hanley, N.A., Hagan, D.M., Clement-Jones, M., Ball, S.G., Strachan, T., Salas-Cortés, L., McElreavey, K., Lindsay, S., Robson, S., Bullen, P., et al. (2000). SRY, SOX9, and DAX1 expression patterns during human sex determination and gonadal development. Mech. Dev. *91*, 403–407.

Hatano, O., Takakusu, A., Nomura, M., and Morohashi, K. (1996). Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1. Genes Cells *1*, 663–671.

Hermann, B.P., Cheng, K., Singh, A., Roa-De La Cruz, L., Mutoji, K.N., Chen, I.-C., Gildersleeve, H., Lehle, J.D., Mayo, M., Westernströer, B., et al. (2018). The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. Cell Rep. *25*, 1650–1667.e8.

Hu, Y.-C., Okumura, L.M., and Page, D.C. (2013). Gata4 is required for formation of the genital ridge in mice. PLoS Genet. *9*, e1003629.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kanatsu-Shinohara, M., and Shinohara, T. (2013). Spermatogonial stem cell self-renewal and development. Annu. Rev. Cell Dev. Biol. *29*, 163–187.

Kobayashi, T., Zhang, H., Tang, W.W.C., Irie, N., Withey, S., Klisch, D., Sybirna, A., Dietmann, S., Contreras, D.A., Webb, R., et al. (2017). Principles of early human development and germ cell program from conserved model systems. Nature *546*, 416–420.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

Li, L., Dong, J., Yan, L., Yong, J., Liu, X., Hu, Y., Fan, X., Wu, X., Guo, H., Wang, X., et al. (2017). Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. Cell Stem Cell *20*, 858–873.e4.

Liu, C., Rodriguez, K., and Yao, H.H.-C. (2016). Mapping lineage progression of somatic progenitor cells in the mouse fetal testis. Development *143*, 3700–3710.

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

Ma, L., Buchold, G.M., Greenbaum, M.P., Roy, A., Burns, K.H., Zhu, H., Han, D.Y., Harris, R.A., Coarfa, C., Gunaratne, P.H., et al. (2009). GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. PLoS Genet. *5*, e1000635.

Mamsen, L.S., Ernst, E.H., Borup, R., Larsen, A., Olesen, R.H., Ernst, E., Anderson, R.A., Kristensen, S.G., and Andersen, C.Y. (2017). Temporal expression pattern of genes during the period of sex differentiation in human embryonic gonads. Sci. Rep. *7*, 15961.

Otasek, D., Morris, J.H., Bouças, J., Pico, A.R., and Demchak, B. (2019). Cytoscape Automation: empowering workflow-based network analysis. Genome Biol. *20*, 185.

Paniagua, R., and Nistal, M. (1984). Morphological and histometric study of human spermatogonia from birth to the onset of puberty. J. Anat. *139*, 535–552.

Pastor, W.A., Stroud, H., Nee, K., Liu, W., Pezic, D., Manakov, S., Lee, S.A., Moissiard, G., Zamudio, N., Bourc'his, D., et al. (2014). MORC1 represses transposable elements in the mouse male germline. Nat. Commun. *5*, 5795.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982.

Saitou, M., and Yamaji, M. (2012). Primordial germ cells in mice. Cold Spring Harb. Perspect. Biol. *4*, a008375.

Sasaki, K., Nakamura, T., Okamoto, I., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Shiraki, N., Takakuwa, T., et al. (2016). The Germ Cell Fate of Cynomolgus Monkeys Is Specified in the Nascent Amnion. Dev. Cell *39*, 169–185.

Shami, A.N., Zheng, X., Munyoki, S.K., Ma, Q., Manske, G.L., Green, C.D., Sukhwani, M., Orwig, K.E., Li, J.Z., and Hammoud, S.S. (2020). Single-Cell RNA Sequencing of Human, Macaque, and Mouse Testes Uncovers Conserved and Divergent Features of Mammalian Spermatogenesis. Dev. Cell *54*, 529–547.e12.

Shima, Y., Miyabayashi, K., Haraguchi, S., Arakawa, T., Otake, H., Baba, T., Matsuzaki, S., Shishido, Y., Akiyama, H., Tachibana, T., et al. (2013). Contribution of Leydig and Sertoli cells to testosterone production in mouse fetal testes. Mol. Endocrinol. *27*, 63–73.

Shima, Y., Matsuzaki, S., Miyabayashi, K., Otake, H., Baba, T., Kato, S., Huhtaniemi, I., and Morohashi, K. (2015). Fetal Leydig Cells Persist as an Androgen-Independent Subpopulation in the Postnatal Testis. Mol. Endocrinol. *29*, 1581–1593.

Sohni, A., Tan, K., Song, H.-W., Burow, D., de Rooij, D.G., Laurent, L., Hsieh, T.-C., Rabah, R., Hammoud, S.S., Vicini, E., and Wilkinson, M.F. (2019). The Neonatal and Adult Human Testis Defined at the Single-Cell Level. Cell Rep. *26*, 1501–1517.e4.

Stévant, I., Kühne, F., Greenfield, A., Chaboissier, M.-C., Dermitzakis, E.T., and Nef, S. (2019). Dissecting Cell Lineage Specification and Sex Fate Determination in Gonadal Somatic Cells Using Single-Cell Transcriptomics. Cell Rep. *26*, 3272–3283.e3.

Svingen, T., and Koopman, P. (2013). Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. Genes Dev. *27*, 2409–2426.

Tan, K., Song, H.-W., and Wilkinson, M.F. (2020). Single-cell RNAseq analysis of testicular germ and somatic cell development during the perinatal period. Development *147*, dev183251.

Tang, W.W.C., Dietmann, S., Irie, N., Leitch, H.G., Floros, V.I., Bradshaw, C.R., Hackett, J.A., Chinnery, P.F., and Surani, M.A. (2015). A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. Cell *161*, 1453–1467.

Tang, W.W.C., Kobayashi, T., Irie, N., Dietmann, S., and Surani, M.A. (2016). Specification and epigenetic programming of the human germ line. Nat. Rev. Genet. *17*, 585–600.

Wang, M., Liu, X., Chang, G., Chen, Y., An, G., Yan, L., Gao, S., Xu, Y., Cui, Y., Dong, J., et al. (2018). Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. Cell Stem Cell *23*, 599–614.e4.

Wen, Q., Wang, Y., Tang, J., Cheng, C.Y., and Liu, Y.-X. (2016). Sertoli Cell Wt1 Regulates Peritubular Myoid Cell and Fetal Leydig Cell Differentiation during Fetal Testis Development. PLoS ONE *11*, e0167920.

Western, P.S., Miles, D.C., van den Bergen, J.A., Burton, M., and Sinclair, A.H. (2008). Dynamic regulation of mitotic arrest in fetal male germ cells. Stem Cells *26*, 339–347.

Witchi, E. (1948). Migration of the germ cells of human embryos from the yolk sac to the primitive gonadal folds. In Contributions to Embryology, Volume 32, no. 209 (Carnegie Institution of Washington), pp. 67–80.

Yang, Y., Workman, S., and Wilson, M. (2018). The molecular pathways underlying early gonadal development. J. Mol. Endocrinol. *62*, R47–R64.

Yao, H.H.-C., Whoriskey, W., and Capel, B. (2002). Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. Genes Dev. *16*, 1433–1440.

Zheng, Y., Xue, X., Shao, Y., Wang, S., Esfahani, S.N., Li, Z., Muncie, J.M., Lakins, J.N., Weaver, V.M., Gumucio, D.L., and Fu, J. (2019). Controlled modelling of human epiblast and amnion development using stem cells. Nature *573*, 421–425.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-PIWIL4, Dilution: 1:200	Thermo Fisher Scientific	Cat#: PA5-3144, RRID: AB_2548922
Mouse monoclonal (CloneB56) anti-MKI67, Dilution: 1:200	BD Biosciences	Cat#: 556003, RRID: AB_396287,
Goat polyclonal anti-DDX4, Dilution: 1:100	R&D Systems,	Cat#: AF2030, RRID: AB_2277369
Rabbit monoclonal (D73G4) anti-NANOG, Dilution: 1:100	Cell Signaling Technology	Cat#:4903, RRID: AB_10559205,
Mouse monoclonal anti-CYP17A1, Dilution: 1:200	Santa Cruz Biotechnology,	Cat#: SC-374244, RRID: AB_10988393
Mouse monoclonal (1A12-1D9) anti-GATA3, Dilution: 1:100	Thermo Fisher Scientific	Cat#: MA1028, RRID: AB_2536713,
Mouse monoclonal (G-4) anti-GATA4, Dilution: 1:100	Santa Cruz Biotechnology,	Cat#: SC-25310, RRID: AB_627667
Mouse monoclonal anti-DMRT1, Dilution: 1:100	Santa Cruz Biotechnology,	Cat#: SC-377167
Rabbit polyclonal anti-SOX9, Dilution: 1:200	Millipore,	Cat#: AB5535, RRID: AB_2239761
AF488 goat-anti mouse IgG2a	Invitrogen	Cat#: A21131, RRID: AB_2535771
AF594 donkey-anti-mouse IgG	Invitrogen	Cat#: A21203, RRID: AB_2535789
AF594 goat-anti-mouse IgG1,	Invitrogen	Cat#: A21125, RRID: AB_2535767
AF594 donkey-anti-rabbit IgG,	Jackson ImmunoResearch	Cat#: 711-585-152, RRID: AB_2340621
AF647 donkey-anti-goat IgG,	Invitrogen	Cat#: A21447, RRID: AB_2535864
Biological samples		
Human testis samples from postnatal donors	DonorConnect	N/A
Human testis samples from embryonic and fetal stages	University of Washington- Birth Defects Research Lab	N/A
Human testis samples from Jan's lab	Karolinska Institutet	N/A
Deposited data		
Single cell RNA-seq for embryonic and fetal human testes	This paper	GEO: GSE143356
Single cell RNA-seq for postnatal testes	This paper	GEO: GSE161617
Software and algorithms		
Seurat (2.3.4)	Butler et al., 2018	https://satijalab.org/seurat/
Monocle (2.10.1)	(Qiu et al., 2017)	http://cole-trapnell-lab.github.io/monocle-release/
GO (David 6.7)	Huang et al., 2009	https://david-d.ncifcrf.gov
Cell Ranger (2.2.0)	NA	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
Cluster 3.0	NA	http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm
WGCNA (1.68)	(Langfelder and Horvath, 2008)	https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/
Cytoscape (3.7.2)	(Otasek et al., 2019)	https://cytoscape.org

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
Single cell RNA-seq for infant and adult human testes	Guo et al., 2018	GEO: GSE120508
Single cell RNA-seq for neonatal human testes	Sohni et al., 2019	GEO: GSE124263

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Bradley R. Cairns (brad.cairns@hci.utah.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All software tools can be found online (see Key resources table). The accession number for all sequencing data reported in this paper is GEO: GSE143356 and GEO:GSE161617.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Prenatal male gonads from 6 to 16 weeks post-fertilization were obtained from three collaborating laboratories at University of Washington Birth Defects Research Laboratory (BDRL), University of Tubingen and Karolinska Institutet. At BRDL, the prenatal gonads were obtained with regulatory oversight from the University of Washington IRB approved Human Subjects protocol, combined with a Certificate of Confidentiality from the Federal Government. The research project was also approved by the research ethics committee of the University of Tübingen. All consented material was donated anonymously and carried no personal identifiers. Human first trimester tissue was collected after elective surgical terminations with maternal written informed consent. The Regional Human Ethics Committee, Stockholm, Sweden, approved the collection (Dnr 2007/1477-31 with complementary permissions 2011/1101-32 and 2013/564-32. The ethical approval to perform the gonadal studies: Dnr 2013/457-31/4). Developmental age was documented by BDRL and University of Tübingen as days post fertilization using a combination of prenatal intakes and Carnegie staging. Developmental age was documented by Karolinska Institutet as days post fertilization by the examination of anatomical landmarks such as nervous system, limb, eye and gonadal development according to the atlas of England. Formalin fixed and paraffin embedded adult testis from biobank samples without underlaying testicular pathologies was obtained at the Department of Pathology at the Karolinska Institutet, and Karolinska University Hospital (ethical approval: Dnr 2014/267-31/4).

Postnatal human testicular sample (5 months old) was obtained through the University of Utah Andrology laboratory and DonorConnect. This sample was removed from deceased individuals who consented to organ donation for transplantation and research.

METHOD DETAILS

Sample transportation and storage

The prenatal samples collected at BDRL used for single cell transcriptome profiling were shipped overnight in HBSS with an ice pack for immediate processing in Los Angeles. From University of Tübingen samples were delivered to UCLA within 24-48 hours after the procedure.

The postnatal whole testis was transported to the research laboratory on ice in saline and processed within 1 hour of removal by surgery. Around 90% of each testis was divided into smaller portions (~500 mg – 1g each) using scissors and directly transferred into cryovials (Corning cat # 403659) in DMEM medium (Life Technologies cat # 11995073) containing 10% DMSO (Sigma-Aldrich cat # D8779), 15% fetal bovine serum (FBS) (GIBCO cat # 10082147) and cryopreserved in Mr. Frosty container (Thermo Fisher Scientific cat #5100-0001) at a controlled slow rate, and stored at −80°C for overnight. Cryovials were transferred to liquid nitrogen for long-term storage.

Human testis sample preparation for single cell RNA sequencing

Prenatal tissues were processed within 24-48 hours after termination. Upon arrival to UCLA tissues were gently washed with PBS and placed in dissociation buffer containing collagenase IV 10mg/ml (Life Technologies #17104-019), Dispase II 250 ug/ml (Life Technologies #17105041), DNase I 1:1000 (Sigma 4716728001), 10% FBS (Life Technologies 10099141) in 1x PBS. After every 5 minutes tissues were gently pipetted with P1000 pipette against the bottom of Eppendorf tube. This process was repeated 3 times for a total of 15 minutes. Afterward, cells were centrifuged for 5 minutes at 500 g and pellet was resuspended in 1x PBS with 0.04% BSA and

strained through 40μm strainer and counted using automated cell counter (Thermo Fisher, Countess II). The cell concentration was adjusted to 800-1200 cells per microliter and immediately used for scRNA-seq. For postnatal tissues, 1 cryovial of tissue was thawed quickly, which was then washed twice with PBS, and subject to digestion as described previously (Guo et al., 2018). Tissues were washed twice in 1 x PBS and minced into small pieces for better digestion outcome. Tissues were then treated with trypsin/ethylenediaminetetraacetic acid (EDTA; Invitrogen cat # 25300054) for 20-25 min and collagenase type IV (Sigma Aldrich cat # C5138-500MG) at 37°C. Single testicular cells were obtained by filtering through 70 μm (Fisher Scientific cat # 08-771-2) and 40 μm (Fisher Scientific cat # 08-771-1) strainers. The cells were pelleted by centrifugation at 600 g for 15 min and washed with PBS twice. Cell number was counted using a hemocytometer, and the cells were then resuspended in PBS + 0.4% BSA (Thermo Fisher Scientific cat # AM2616) at a concentration of ~1,000 cells/uL ready for single-cell sequencing.

Single cell RNA-seq performance, library preparation and sequencing

We targeted to capture ~6,000-7,000 cells. The prenatal sequencing was conducted in UCLA, and the postnatal sequencing was conducted at University of Utah. Briefly, cells were diluted following manufacturer's instructions, and 33.8 μL of total mixed buffer together with cells were loaded into 10x Chromium Controller using the Chromium Single Cell 3′ v3 reagents. The sequencing libraries were prepared following the manufacturer's instructions, using 13 cycles of cDNA amplification, followed by an input of ~100 ng of cDNA for library amplification using 12 cycles. The resulting libraries were then sequenced on a 2 X 150 cycle paired-end run on an Illumina Novaseq 6000 instruments.

Processing of single cell RNA-seq data

Raw data were demultiplexed using mkfastq application (Cell Ranger v2.2.0) to make Fastq files. Fastq files were then run with count application (Cell Ranger v2.2.0) using default settings, which performs alignment (using STAR aligner), filtering and UMI counting. The UMI count tables were used for further analysis.

Immunostaining of testicular tissues

Intact testes were fixed in 4% PFA at room temperature for 2 hours on a platform rocker. Tissues were washed 3 times with PBS for 10 minutes each wash then placed into paraffin blocks (Histogel,Thermo Scientific HG4000012) for sectioning onto slides. Sections were deparaffinized and rehydrated in a Xylene then ethanol series (100%, 95%, 70%, 50%, water) respectively. Antigen retrieval was performed in either Tris-EDTA solution (pH 9.0) or Sodium Citrate Solution (pH 6.0) in a hot water bath (95°C) for 40 minutes. Sections were washed in PBS, 0.2% Tween-20 (PBS-T) 3 times, 5 minutes each then permeabilized in PBS, 0.05% Trition X-100 for 20 minutes. Sections were blocked with blocking solution (10% Normal Donkey Serum (NDS), PBS-T) for 30 minutes at room temperature in a humid chamber. Primary Antibodies were diluted in 2.5% NDS, PBS-T at the appropriate dilutions (see Key resources table) and incubated overnight at 4°C in a humid chamber. After 3 washes in PBS-T (5 minutes each) secondary antibodies were added and allowed to incubate at room temperature for 1 hour in a humid chamber. After 3 washes in PBS-T, DAPI was added to the sections for approximately 5 minutes, then washed 3 times 5 minutes each in PBS-T. Prolong Gold antifade mountant (Invitrogen P10144) was added to the sections. Coverslips were placed onto slides then sealed with nail polish. Slides were allowed to cure overnight, in the dark, at room temperature then subsequently stored at 4°C until ready to image. For sections stained with PIWIL4 antibody, the blocking buffer used was Superblock blocking buffer (Thermo Scientific 37580). In addition, the SignalBoost Immunoreaction Enhancer Kit (Millipore 407207) was used to dilute primary and secondary antibodies for experiments involving PIWIL4 antibody.

Microscopy

A Zeiss LSM 880 with Airyscan controlled by the Zen Black software, equipped with the Plan-Apochromat 20×/0.8 NA and the Plan-Apochromat 63×/1.4 NA M27 oil immersion objective, was used to acquire confocal images. Saved CZI files were converted to Imaris format files (.ims) using the Imaris File converter (Bitplane), then processed using the image analysis software IMARIS 9.3 (Bitplane). An Olympus BX-61 light microscope was used to examine Hematoxylin and Eosin (H&E) stained slides. The ImageJ stitch function uses similar features/structures from a collection of images to make a fused image, therefore each image has some overlap with the previous image taken. Briefly, H&E images were taken with the 20x objective. In ImageJ under the Plugins dropdown box we chose the Stitching plugin and then selected the Grid/Collection Stitching function. In the "Type" box we selected "unknown position" and chose "all files in directory" for the "Order". For the Fusion Method we used Linear Blending. The Regression threshold was set at 0.30. The Max/avg displacement threshold was set at 2.50 and the Absolute displacement threshold was set to 3.50. Stitched images were built using the ImageJ2(NIH) Grid/Collection Stitching plugin.

QUANTIFICATION AND STATISTICAL ANALYSIS

The Seurat program (https://satijalab.org/seurat/, R package, v.2.3.4) was used as a first analytical package. To start with, UMI count tables from both replicates from all four juvenile donors were loaded into R using Read10X function, and Seurat objects were built from each experiment. Each experiment was filtered and normalized with default settings. Specifically, cells were retained only if they contained > 500 expressed genes and had < 25% reads mapped to mitochondrial genome. t-SNE and clustering analysis were first run on each replicate, which resulted in similar t-SNE map. Data matrices from different donors and replicates were then combined with the previously published infant and adult data (Guo et al., 2018). Next, cells were normalized to the total UMI read counts, as

instructed in the tutorial (https://satijalab.org/seurat/). t-SNE and clustering analyses were performed on the combined data using the top 6,000 highly variable genes and 1-30 PCs, which showed the most significant p values.

Detailed pseudotime for different cell types were performed using the Monocle package (v2.10.1) following the default settings. After pseudotime coordinates/order were determined, gene clustering analysis was performed to establish the accuracy of pseudotime ordering. Here, cells (in columns) were ordered by their pseudotime, and genes (in rows) were clustered by k-means clustering using Cluster 3.0. Different k-mean numbers were performed to reach the optimal clustering number. Cell cycle analysis was performed using scran program (https://bioconductor.org/packages/3.7/bioc/vignettes/scran/inst/doc/scran.html, R Package; v1.6.5).

Weighted correlation network analysis

Hub genes in PGC, spermatogonia and State 0 were found by WGCNA (https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/, R package, v1.68). When finding hub genes in PGC and spermatogonia, gene expression data of 40 cells from PGC and State 0 respectively were randomly extracted from the UMI count tables of scRNA-seq data. Genes were filtered by selecting those genes expressed in more than 20 cells since scRNA-seq data had a high drop-out rate and low expression genes may represent noise. Then the counts were normalized by total reads (x*100000/total reads) and then log-transformed (log2(x+1)). Afterward, one-step network construction and module detection were performed. In this step, we chose parameters including signed hybrid network type, Pearson correlation method and the default soft-threshold power β to reach the scale-free network topology. To identify the modules that were significantly correlated with PGC or spermatogonia, bi-weight mid-correlation (robustY = FALSE) was used. The quality of the modules was checked by the strong correlation between module eigengenes and traits of interest as well as the strong correlation between gene module membership and gene-trait correlation. Finally, hub genes inside those modules were selected from the top 40 genes with the highest intramodular connectivity (sum of in-module edge weights). Specifically, in order to find hub genes in State 0 rather than spermatogonia, we added gene expression data of 40 cells from State1 to rule out the genes expressing broadly in States 0-4 and performed the same analysis to determine the modules that were significantly correlated with State 0. Ten hub genes were selected by the same standard. Finally, the networks were visualized by Cytoscape Software 3.7.2.

Periodic retinoic acid–STRA8 signaling intersects with periodic germ-cell competencies to regulate spermatogenesis

Tsutomu Endo[a,1,2], Katherine A. Romer[a,b,1], Ericka L. Anderson[a,c,1], Andrew E. Baltus[a,c], Dirk G. de Rooij[a], and David C. Page[a,c,d,2]

[a]Whitehead Institute, Cambridge, MA 02142; [b]Computational and Systems Biology Program and [c]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; and [d]Howard Hughes Medical Institute, Whitehead Institute, Cambridge, MA 02142

Contributed by David C. Page, March 24, 2015 (sent for review January 18, 2015; reviewed by William W. Wright)

Mammalian spermatogenesis—the transformation of stem cells into millions of haploid spermatozoa—is elaborately organized in time and space. We explored the underlying regulatory mechanisms by genetically and chemically perturbing spermatogenesis in vivo, focusing on spermatogonial differentiation, which begins a series of amplifying divisions, and meiotic initiation, which ends these divisions. We first found that, in mice lacking the retinoic acid (RA) target gene *Stimulated by retinoic acid gene 8* (*Stra8*), undifferentiated spermatogonia accumulated in unusually high numbers as early as 10 d after birth, whereas differentiating spermatogonia were depleted. We thus conclude that *Stra8*, previously shown to be required for meiotic initiation, also promotes (but is not strictly required for) spermatogonial differentiation. Second, we found that injection of RA into wild-type adult males induced, independently, precocious spermatogonial differentiation and precocious meiotic initiation; thus, RA acts instructively on germ cells at both transitions. Third, the competencies of germ cells to undergo spermatogonial differentiation or meiotic initiation in response to RA were found to be distinct, periodic, and limited to particular seminiferous stages. Competencies for both transitions begin while RA levels are low, so that the germ cells respond as soon as RA levels rise. Together with other findings, our results demonstrate that periodic RA–STRA8 signaling intersects with periodic germ-cell competencies to regulate two distinct, cell-type-specific responses: spermatogonial differentiation and meiotic initiation. This simple mechanism, with one signal both starting and ending the amplifying divisions, contributes to the prodigious output of spermatozoa and to the elaborate organization of spermatogenesis.

spermatogenesis | Stra8 | mouse | retinoic acid | testis

The adult mammalian testis is among the body's most proliferative tissues, producing millions of highly specialized gametes, or spermatozoa, each day. Spermatogenesis (the program of sperm production) is carefully regulated, ensuring that spermatozoa are produced at a constant rate. We used the mouse as a model to understand how mammalian spermatogenesis is organized at the cellular and molecular level. We focused on two key transitions: spermatogonial differentiation, which occurs cyclically and begins a series of programmed mitotic divisions, and meiotic initiation, which ends these divisions and marks the beginning of the meiotic program (Fig. 1*A*).

Like other proliferative tissues (e.g., blood, intestine, and skin), the testis relies on a modest number of stem cells (1, 2). The undifferentiated spermatogonia (also known as the A_{single}/A_{paired}/$A_{aligned}$ spermatogonia), which encompass these stem cells, have a remarkable capacity for self-renewal and differentiation: They can reconstitute spermatogenesis upon transplantation to a germ-cell-depleted testis (3, 4). In vivo, undifferentiated spermatogonia ultimately give rise to a single cell type, spermatozoa, yet these undifferentiated spermatogonia express pluripotency-associated genes such as *Lin28a* (*Lin-28 homolog A*) (5) and *Pou5f*/*Oct4* (6, 7) and are the only postnatal mammalian cells from which functionally pluripotent cells have been derived in vitro without introduction of exogenous transcription factors or miRNAs (8). Undifferentiated spermatogonia periodically undergo spermatogonial differentiation (also known as the $A_{aligned}$-to-A_1 transition) to become differentiating spermatogonia (also known as A_1/A_2/A_3/A_4/intermediate/B spermatogonia). During spermatogonial differentiation, the spermatogonia down-regulate pluripotency-associated genes (5, 9), lose capacity for self-renewal (4), and accelerate their cell cycle (10) to begin a series of six transit-amplifying mitotic ivisions. At the conclusion of these mitotic divisions, germ cells become spermatocytes, and undergo meiotic initiation (Fig. 1*A*). This begins the meiotic program of DNA replication and reductive cell divisions, ensuring that spermatozoa contribute exactly one of each chromosome to the zygote.

Meiotic initiation is precisely coordinated with spermatogonial differentiation: The six mitotic divisions separating the two transitions occur over a span of exactly 8.6 d (11). Moreover, spermatogonial differentiation and meiotic initiation occur in close physical proximity. The testis comprises structures known as seminiferous tubules (Fig. S1*A*); while one generation of germ cells is initiating meiosis, a younger generation is simultaneously undergoing spermatogonial differentiation, within the same tubule

Significance

As male sex cells mature into sperm, two pivotal transitions are spermatogonial differentiation (exit from the stem cell pool) and meiotic initiation. These transitions occur in physical proximity, with 8.6-d periodicity. We report that the gene *Stra8*, essential for meiotic initiation, also promotes (but is not required for) spermatogonial differentiation. Moreover, injected RA induces both transitions to occur precociously. We conclude that a periodic RA signal, acting instructively through the common target *Stra8*, coordinates these transitions. This RA signal intersects with two distinct windows of sex-cell competency, which both begin while RA levels are low; sex cells respond quickly to rising RA. These mechanisms help account for the elaborate organization of sperm production, and its prodigious output.

Author contributions: T.E., K.A.R., E.L.A., D.G.d.R., and D.C.P. designed research; T.E., K.A.R., E.L.A., and A.E.B. performed research; T.E., K.A.R., E.L.A., and D.G.d.R. analyzed data; and T.E., K.A.R., and D.C.P. wrote the paper.

Reviewers included: W.W.W., Johns Hopkins Bloomberg School of Public Health.

The authors declare no conflict of interest.

Data deposition: The mRNA-sequencing dataset reported in this paper has been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE67169).

[1]T.E., K.A.R, and E.L.A. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: endo@wi.mit.edu or dcpage@wi.mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1505683112/-/DCSupplemental.

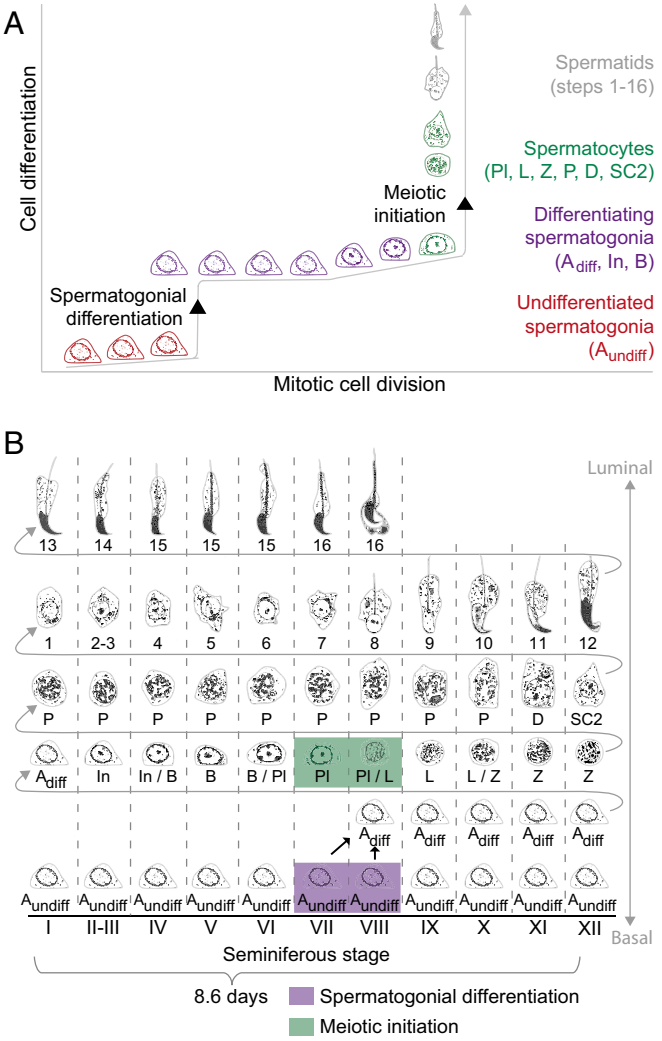


Fig. 1. Overview of spermatogenesis. (*A*) Diagram of mouse spermatogenesis. Germ cells develop from undifferentiated, mitotic spermatogonia (red, bottom left), to differentiating spermatogonia (purple, middle left), to spermatocytes (germ cells that undergo meiosis, green, middle right), to haploid spermatids (gray, top right). We highlight two cellular differentiation steps: spermatogonial differentiation and meiotic initiation. (*B*) Diagram of germ-cell associations (stages) in the mouse testis. Oakberg (33) identified 12 distinct cellular associations, called seminiferous stages I–XII. In the mouse, it takes 8.6 d for a section of the seminiferous tubule, and the germ cells contained within, to cycle through all 12 stages (41). Four turns of this seminiferous cycle are required for a germ cell to progress from undifferentiated spermatogonium to spermatozoon that is ready to be released into the tubule lumen. A_{undiff}, A_{diff}, In, and B: undifferentiated type A, differentiating type A, intermediate, and type B spermatogonia, respectively. D, L, P, Pl, SC2, and Z: diplotene, leptotene, pachytene, preleptotene, secondary spermatocytes, and zygotene, respectively. Steps 1–16: steps in spermatid differentiation. Purple: germ cells undergoing spermatogonial differentiation; green: meiotic initiation.

cross-section (Fig. 1*B* and Fig. S1*A*). Spermatogonial differentiation and meiotic initiation occur in close association not only in mice but also in other mammals, including rats (12, 13), hamsters, and rams (14). This precise coordination of different steps of spermatogenesis is called the "cycle of the seminiferous epithelium" (or "seminiferous cycle"); it has fascinated biologists for over a century (15). We sought to explain the cooccurrence of spermatogonial differentiation and meiotic initiation, to better understand the regulation of these two transitions and the overall organization of the testis.

Both of these transitions require RA, a derivative of vitamin A. In vitamin A-deficient (VAD) mice and rats, most germ cells arrest as undifferentiated spermatogonia (16, 17). In VAD rat testes, some germ cells instead arrest just before meiosis, as preleptotene spermatocytes (16, 18). When VAD animals are injected with RA or vitamin A, the arrested spermatogonia differentiate (17–19), and the arrested preleptotene spermatocytes initiate meiosis (18). During spermatogonial differentiation, RA is believed to act, at least in part, directly on germ cells: Spermatogonia express RA receptors (RARs) (20), and genetic ablation of RARs in germ cells modestly impairs spermatogonial differentiation (21).

To understand how RA might coregulate these two transitions, we needed to understand its target genes. During meiotic initiation, RA acts instructively, through the target gene *Stra8* (*Stimulated by retinoic acid gene 8*). RA induces *Stra8* expression in germ cells—not in somatic cells—in both males and females (22, 23). *Stra8* is required for meiotic initiation in both sexes: *Stra8*-deficient germ cells in postnatal males and fetal females arrest just before meiosis, without entering meiotic prophase (24, 25). In contrast, no specific RA target genes have been implicated in spermatogonial differentiation: RA could either instruct the germ cells or simply be permissive for this transition. We considered *Stra8* as a candidate regulator of spermatogonial differentiation: STRA8 protein is expressed in spermatogonia as well as in preleptotene spermatocytes in vivo (26, 27), and in vitro studies suggest that RA can act directly on early spermatogonia to increase expression of *Stra8* (28). However, the functional role, if any, of *Stra8* in spermatogonia was not previously known. Using two complementary perturbations of RA–STRA8 signaling—genetic disruption of *Stra8* function and chemical manipulation of RA levels—we demonstrated that RA acts instructively, and at least in part through STRA8, at spermatogonial differentiation as well as at meiotic initiation. The shared RA–STRA8 signal helps to coordinate these two transitions in time and space.

Results

Massive Accumulations of Type A Spermatogonia in Testes of Aged *Stra8*-Deficient Males. As we previously reported, *Stra8*-deficient testes lacked meiotic and postmeiotic cells (24, 25); thus, at 8 wk of age, *Stra8*-deficient testes were much smaller than wild-type testes (25). However, we observed that, after 6 mo, some *Stra8*-deficient testes were grossly enlarged (>400 mg) (44%; 11 of 25 mice) compared with wild-type testes (91 ± 7 mg; average of testes from three mice) (Fig. 2*A*). Both small and large aged *Stra8*-deficient testes (88%; 22 of 25 mice) contained accumulated cells that resembled spermatogonia and expressed the germ-cell marker DAZL (Fig. 2*B*); these accumulations were absent in aged wild-type and heterozygous mice (0%, 0 of 10 mice). In wild-type testes, spermatogonia were confined to the basal lamina of seminiferous tubules, but even in small *Stra8*-deficient testes (<50 mg) occasional tubules were filled with presumptive spermatogonia, which sometimes spilled into the testicular interstitium. Large *Stra8*-deficient testes were composed almost entirely of presumptive spermatogonia, with few remnants of tubule structure. Spermatogonial morphology was very similar between small and large *Stra8*-deficient testes. We used mRNA sequencing (mRNA-Seq) to confirm that spermatogonia had accumulated in *Stra8*-deficient testes; known spermatogonial marker genes (29) were up-regulated in *Stra8*-deficient testes vs. wild-type testes (Fig. S2*A*).

We classified more precisely the spermatogonia in these massive accumulations. Based on nuclear morphology, spermatogonia can be classified as type A, intermediate, or type B (Fig. 1 *A* and *B*) (11). Type A includes the undifferentiated and early differentiating spermatogonia, whereas intermediate and type B encompass the later differentiating spermatogonia. The

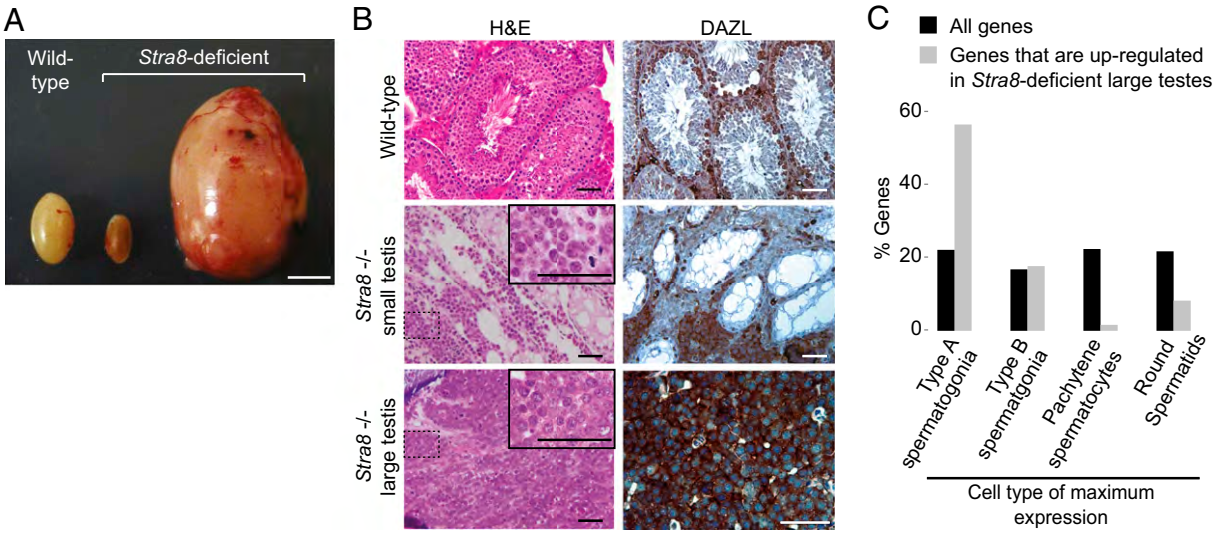


Fig. 2. Aged *Stra8*-deficient testes accumulate type A spermatogonia. (*A*) Wild-type (left) and *Stra8*-deficient small (center) and large (right) testes from 1-y-old mice. (Scale bar, 5 mm.) (*B*) Testis sections from 1-y-old mice: wild-type (*Top*), *Stra8*-deficient small testis (*Middle*), and *Stra8*-deficient grossly enlarged testis (*Bottom*). (*Left*) H&E staining (*Right*) DAZL immunostaining. Insets enlarge the boxed regions. (Scale bars, 50 μm.) (*C*) Percentage of genes whose highest expression is found in type A spermatogonia, type B spermatogonia, pachytene spermatocytes, or round spermatids. Black bars (control), all analyzable genes (17,345 genes). Gray bars, 100 genes most significantly up-regulated in *Stra8*-deficient large testes relative to wild-type testes (Table S1).

Stra8-deficient spermatogonia had type A morphology (Fig. 2*B*). To confirm this, we used mRNA-Seq to identify the 100 genes most significantly up-regulated in *Stra8*-deficient vs. wild-type testes (Table S1). We analyzed their expression among different cell types in wild-type testes, using a published microarray dataset (30); 59% of these 100 genes were most highly expressed in type A spermatogonia, vs. 22.7% of a control gene set (all 17,345 genes on the microarray) (Fig. 2*C*) ($P < 0.001$, Fisher's exact test). A genome-wide clustering analysis of these data and other publically available datasets confirmed that the expression patterns of *Stra8*-deficient testes were overall quite similar to those of type A spermatogonia (Fig. S2*B* and *SI Results and Discussion*). We conclude that type A spermatogonia accumulate in *Stra8*-deficient testes. This suggests that STRA8 has a functional role in type A spermatogonia, distinct from its role in meiotic initiation.

Early Postnatal *Stra8*-Deficient Testes Contain Accumulations of Undifferentiated Spermatogonia and Are Depleted for Differentiating Spermatogonia. We considered our findings in light of published observations that (*i*) a subset of type A spermatogonia undergo spermatogonial differentiation, (*ii*) RA is required for spermatogonial differentiation (17, 18), and (*iii*) RA can act directly on spermatogonia to induce *Stra8* expression (28). We postulated that in the unperturbed wild-type testis, RA induction of *Stra8* promotes spermatogonial differentiation, whereas in the absence of *Stra8*, impaired spermatogonial differentiation leads to accumulation of undifferentiated spermatogonia, accounting for the massive accumulations of type A spermatogonia observed in aged *Stra8*-deficient males.

This hypothesis predicts that even in very young males *Stra8*-deficient testes should contain more undifferentiated spermatogonia than wild-type testes. To test this, we counted undifferentiated spermatogonia in testes from 10-d-old (p10) animals (Fig. 3 *A* and *B* and Fig. S3 *A* and *B*), using the markers LIN28A and PLZF (promyelocytic leukemia zinc finger, a.k.a. ZBTB16) (9, 31). As predicted, LIN28- and PLZF-positive spermatogonia were enriched in *Stra8*-deficient testes ($P < 10^{-15}$ for LIN28A, $P < 10^{-4}$ for PLZF, one-tailed Kolmogorov–Smirnov test) (Fig. 3*B* and Fig. S3*B*). *Stra8*-deficient testes had 9.4 ± 2.3 LIN28A-positive spermatogonia per tubule cross-section, vs. 4.4 ± 0.8 in wild-type testes ($P = 0.026$, one-tailed Welch's *t* test). As the testis matured, the number of undifferentiated spermatogonia per tubule cross-section declined in both wild-type and *Stra8*-deficient testes, but undifferentiated spermatogonia remained significantly enriched in *Stra8*-deficient testes at p30 (Fig. 3*C* and Fig. S3*C*). Indeed, some testis tubules in p30 *Stra8*-deficient mice contained large clusters of LIN28A-positive and PLZF-positive type A spermatogonia (Fig. 3*D* and Fig. S3*E*). Spermatogonia in these clusters were densely packed in multiple layers, whereas in wild-type testes type A spermatogonia were widely spaced in a single layer (Fig. S3*D*). Thus, we conclude that undifferentiated spermatogonia progressively accumulate in *Stra8*-deficient animals. mRNA-Seq and immunohistochemical data from testes of aged *Stra8*-deficient mice were consistent with such an accumulation (Fig. S2 *B–E* and *SI Results and Discussion*).

If this progressive accumulation were due to a defect in spermatogonial differentiation, the proportion of differentiating spermatogonia should be smaller in *Stra8*-deficient testes than in wild-type testes. We thus counted type B (differentiating) spermatogonia in *Stra8*-deficient and wild-type testes at p30 (Fig. 3*C* and Fig. S3*F*). Indeed, compared with wild-type testes, *Stra8*-deficient testes contained significantly fewer type B spermatogonia per tubule cross-section. As predicted, the ratio of differentiating-to-undifferentiated (LIN28-positive) spermatogonia was decreased to 2.8 in *Stra8*-deficient testes, vs. 7.9 in wild-type testes ($P = 0.05$ by Mann–Whitney *U* test). We conclude that STRA8 promotes (but is not strictly required for) spermatogonial differentiation.

STRA8 Expression Begins Shortly Before Spermatogonial Differentiation. Previous reports showed that STRA8 is expressed in spermatogonia and spermatocytes but did not distinguish between different subtypes of spermatogonia (26, 27). We tested our model's prediction that STRA8 must be expressed before or during spermatogonial differentiation, immunostaining intact testis tubules for STRA8 and for PLZF (Fig. 4*A* and Fig. S4*A*). A subset of PLZF-positive (undifferentiated) spermatogonia expressed STRA8. We next immunostained for GFRα1 (GDNF family receptor alpha 1), a marker of early undifferentiated spermatogonia (Fig. S4*A*) (32). GFRα1 did not overlap with STRA8. We thus hypothesized that STRA8 expression begins immediately before spermatogonial differentiation.

To confirm this, we immunostained testis sections for STRA8 and then classified tubules by stage of the cycle of the seminif-

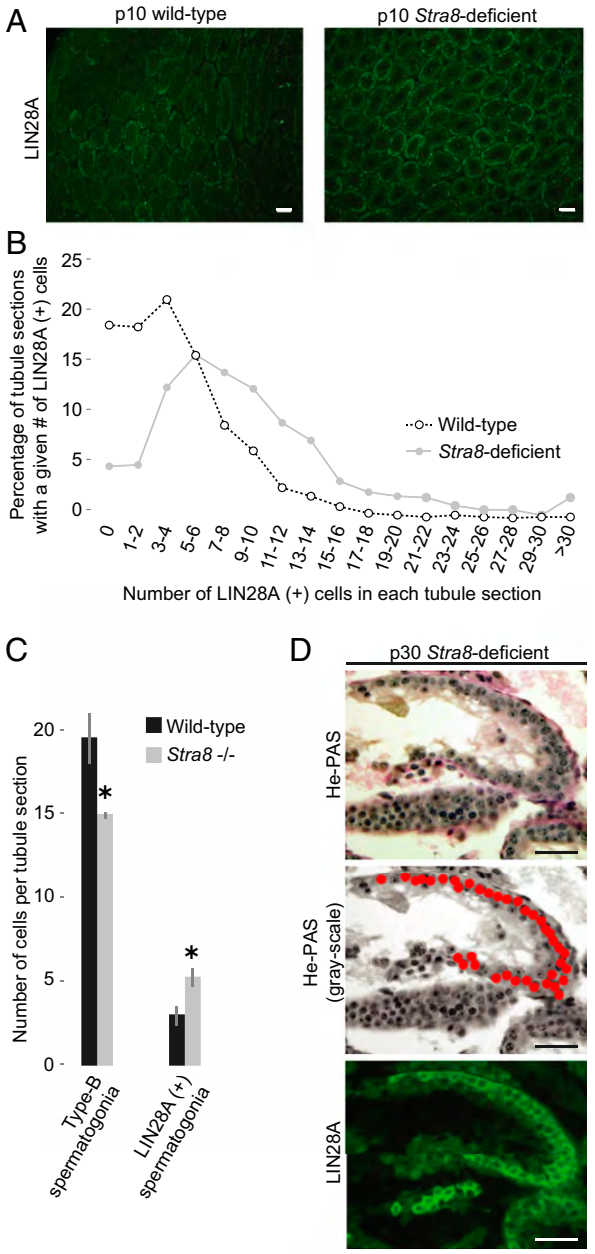


Fig. 3. Testes of young *Stra8*-deficient males are enriched for LIN28A-positive spermatogonia. (*A*) Immunostaining for LIN28A on postnatal day 10 (p10) testis sections: wild-type (*Left*) and *Stra8*-deficient (*Right*). (Scale bars, 50 μm.) (*B*) Histogram of number of LIN28A-positive cells per tubule cross-section, in p10 wild-type and *Stra8*-deficient testes. For each genotype, testes from three mice were counted and averaged. (*C*) Number of type B and LIN28A-positive spermatogonia per tubule cross-section, in p30 wild-type and *Stra8*-deficient testes. Only tubules containing ≥1 type B spermatogonia were counted. Error bars, mean ± SD $*P < 0.01$ (one-tailed Welch's *t* test). (*D*) Clusters of LIN28A-positive undifferentiated spermatogonia in p30 *Stra8*-deficient testes. (*Top*) Hematoxylin and periodic acid-Schiff (He-PAS) staining. (*Middle*) Grayscale version of top panel, with red dots indicating type A spermatogonia. (*Bottom*) LIN28A immunostaining on adjacent section. (Scale bars, 50 μm.)

erous epithelium (hereafter referred to as "seminiferous stage") (Fig. 1*B* and Fig. S1*B*). In the testis, particular germ-cell types are always found in the same tubule cross-section; Oakberg (33) identified 12 such stereotypical associations of germ cells, called stages I–XII. Any given section of a seminiferous tubule cycles through all 12 stages in order, with a full seminiferous cycle encompassing these stages taking 8.6 d. Spermatogonial differentiation and meiotic initiation occur together in stages VII/VIII. Consistent with our hypothesis, STRA8 expression was rare in stages II–VI (before spermatogonial differentiation) then increased rapidly in stages VII–VIII (during spermatogonial differentiation) and remained high thereafter, in stages IX–I (Fig. S1*C*). [Similar increases in STRA8 expression in stage VII or VIII have been previously reported (34, 35)]. We further find that STRA8 expression overlapped with PLZF in stages VII–VIII and was limited to PLZF-low and -negative spermatogonia thereafter, in stages IX–X (Fig. 4 *A* and *B* and Fig. S4*A*). We thus conclude that STRA8 expression begins in late undifferentiated spermatogonia and persists in differentiating spermatogonia.

Intriguingly, in stages VII–VIII, STRA8 expression increased in preleptotene spermatocytes (premeiotic cells) as well as in undifferentiated spermatogonia (Fig. S1 *D* and *E*) (27, 35). Because STRA8 promotes spermatogonial differentiation and is required for meiotic initiation, precisely timed increases in STRA8 expression might coordinate both transitions, ensuring their cooccurrence in stages VII–VIII.

STRA8 expression is induced by RA (22, 23, 26, 28). We hypothesized that RA, acting through *Stra8*, coordinates spermatogonial differentiation with meiotic initiation. To test this, we perturbed RA signaling in the testes of wild-type mice, by injecting either RA or WIN18,446, an inhibitor of RA synthesis (36, 37). We first predicted that injected RA would induce ectopic STRA8 expression both in undifferentiated spermatogonia and in premeiotic cells (differentiating spermatogonia/preleptotene spermatocytes). Furthermore, injected RA should induce both precocious spermatogonial differentiation and precocious meiotic initiation. In contrast, WIN18,446 should inhibit STRA8 expression, spermatogonial differentiation, and meiotic initiation. We proceeded to test these predictions.

Injected RA Induces Precocious STRA8 Expression in Both Undifferentiated Spermatogonia and Premeiotic Germ Cells. We first verified that injected RA induced precocious STRA8 protein expression in the spermatogonial population, which encompasses both undifferentiated spermatogonia and premeiotic cells. In the unperturbed wild-type testis, spermatogonia began to express STRA8 in stages VII–VIII (during spermatogonial differentiation/meiotic initiation) (Fig. 4*C*). At 1 d after RA injection, STRA8 expression was strongly induced in spermatogonia in stages II–VI (Fig. 4*C*), as previously reported (34). In contrast, when we treated mice for 2 d with the RA synthesis inhibitor WIN18,446, spermatogonial STRA8 expression was almost completely eliminated in all stages (Fig. 4 *C* and *D*).

We next showed that injected RA induced precocious STRA8 expression specifically in undifferentiated spermatogonia, by staining testis sections and intact testis tubules for STRA8 and PLZF (Fig. 4 *A*, *B*, and *E* and Fig. S4 *A* and *B*). Indeed, after RA injection, STRA8 expression was strongly induced in a subset of undifferentiated spermatogonia. STRA8 was not induced in any undifferentiated spermatogonia in stages IX–X, but only in late undifferentiated spermatogonia in stages II–VI. We conclude that a very specific subset of undifferentiated spermatogonia is competent to express STRA8.

Finally, we demonstrated that injected RA could induce precocious STRA8 expression in premeiotic cells, which we identified by nuclear morphology and by absence of PLZF expression (Fig. S4*C*). In the unperturbed testis, STRA8 was expressed in preleptotene spermatocytes in stages VII–VIII (during meiotic initiation) but was otherwise absent in premeiotic cells (Fig. 4*F* and Fig. S4*D*). After RA injection, STRA8 was strongly induced in all premeiotic cells, including intermediate and type B spermatogonia and preleptotene spermatocytes, in stages II–VI. We conclude that both undifferentiated spermatogonia and premeiotic cells precociously express STRA8 when exposed to RA in vivo.

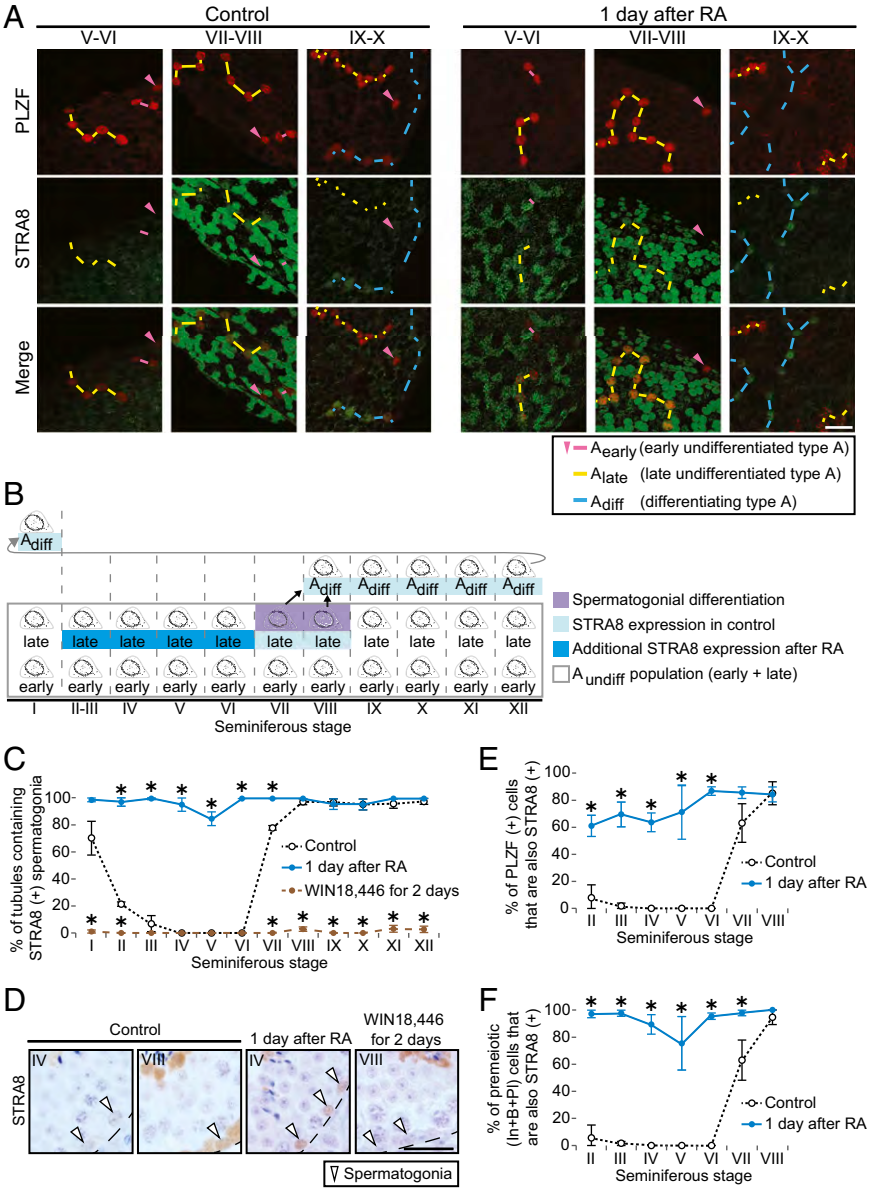


Fig. 4. STRA8 protein is normally present in late undifferentiated spermatogonia and can be precociously induced by injected RA. (*A*) Whole-mount immunostaining of intact wild-type testis tubules for PLZF (red) and STRA8 (green), in controls (*Left*) and 1 d after RA injection (*Right*). Arrowheads: isolated (single) spermatogonia. Dashed lines: putative interconnected chains of spermatogonia. Magenta labels: early undifferentiated type A spermatogonia (A_{early}). Yellow labels: late undifferentiated type A spermatogonia (A_{late}). Blue labels: differentiating type A spermatogonia (A_{diff}). (Scale bar, 30 μm.) (*B*) Diagram of STRA8 expression in type A spermatogonia, in controls (light blue) and 1 d after RA injection (light blue + dark blue). Diagram is based on observations in *A* and *C*. A_{undiff}, early, late, and A_{diff}: undifferentiated type A, early undifferentiated type A, late undifferentiated type A, and differentiating type A spermatogonia. (*C*) Percentage of testis tubule cross-sections containing STRA8-positive spermatogonia, in controls, 1 d after a single RA injection, and after 2 d of WIN18,446 treatment. Control data are duplicated from Fig. S1*C*. Error bars, mean ± SD *$P < 0.01$ (one-tailed Welch's *t* test). (*D*) Immunostaining for STRA8 on testis cross-sections in stages IV and VIII. Dashed lines: basal laminae. Arrowheads: spermatogonia. (Scale bar, 30 μm.) (*E* and *F*) Percentage of PLZF-positive cells (*E*) or PLZF-negative premeiotic cells (*F*) that are also positive for STRA8 in testis cross-sections, in controls or 1 d after RA injection. Premeiotic germ cells (ln+B+Pl): intermediate and type B spermatogonia and preleptotene spermatocytes. Error bars, mean ± SD *$P < 0.01$ (one-tailed Welch's *t* test).

Injected RA Induces Precocious Spermatogonial Differentiation. Because injected RA induced precocious STRA8 expression in stages II–VI, and STRA8 promotes spermatogonial differentiation, we hypothesized that RA would also induce spermatogonial differentiation in these stages. In the unperturbed testis, as a consequence of their differentiation in stages VII–VIII, spermatogonia express KIT protooncogene and enter mitotic S phase. They eventually develop into type B spermatogonia, then become preleptotene spermatocytes, and then initiate meiosis to become leptotene spermatocytes. We predicted that RA injection would cause undifferentiated spermatogonia to precociously begin this developmental progression.

We first confirmed that RA injection induced precocious KIT expression in spermatogonia. In control testis sections, KIT expression was absent in type A spermatogonia in stages II–VI and present in stages VII–VIII (Fig. S4 *E* and *G*) (7). As predicted, at 1 d after RA injection, KIT was strongly induced in stages II–VI. We next tested for precocious entry into S phase, using PLZF to identify undifferentiated and newly differentiating spermatogonia, and BrdU incorporation to assay for S phase (Fig. S4 *F* and *H*). Indeed, at 1 d after RA injection, many PLZF-positive spermatogonia in stages II–VIII incorporated BrdU, whereas in control testes BrdU incorporation did not begin until stage VIII (10, 38).

If injected RA had induced precocious spermatogonial differentiation, the spermatogonia should develop into type B spermatogonia, preleptotene spermatocytes, and leptotene spermatocytes after 7, 8.6, and 10.6 d, respectively (Figs. 1*B* and 5*A*) (11). Thus, we should see transient increases in these cell types. As predicted, at 7 d after RA injection, type B spermatogonia were present in an increased fraction of testis tubules, in a much broader range of stages (XII–VI) than in control testes (IV–VI) (Fig. 5 *B* and *C* and Fig. S5*A*). Preleptotene spermatocytes were similarly increased at 8.6 d after RA injection (in stages II–VIII, vs. VI–VIII in control testes) (Fig. 5 *D* and *E* and Fig. S5*B*); throughout these stages, most of the premeiotic cells in the tubule cross-sections were preleptotene spermatocytes (Fig. S5*D*). Finally, at 10.6 d after RA injection, leptotene spermatocytes were present in an increased fraction of tubules, throughout stages VI–X (vs. VIII–X in control testes) (Fig. 5 *F* and *G* and Fig. S5*C*). We confirmed our identification of leptotene spermatocytes throughout this broad range of stages by immunostaining for meiotic markers: γH2AX (phosphorylated H2A histone family member X, a marker of DNA double strand breaks)

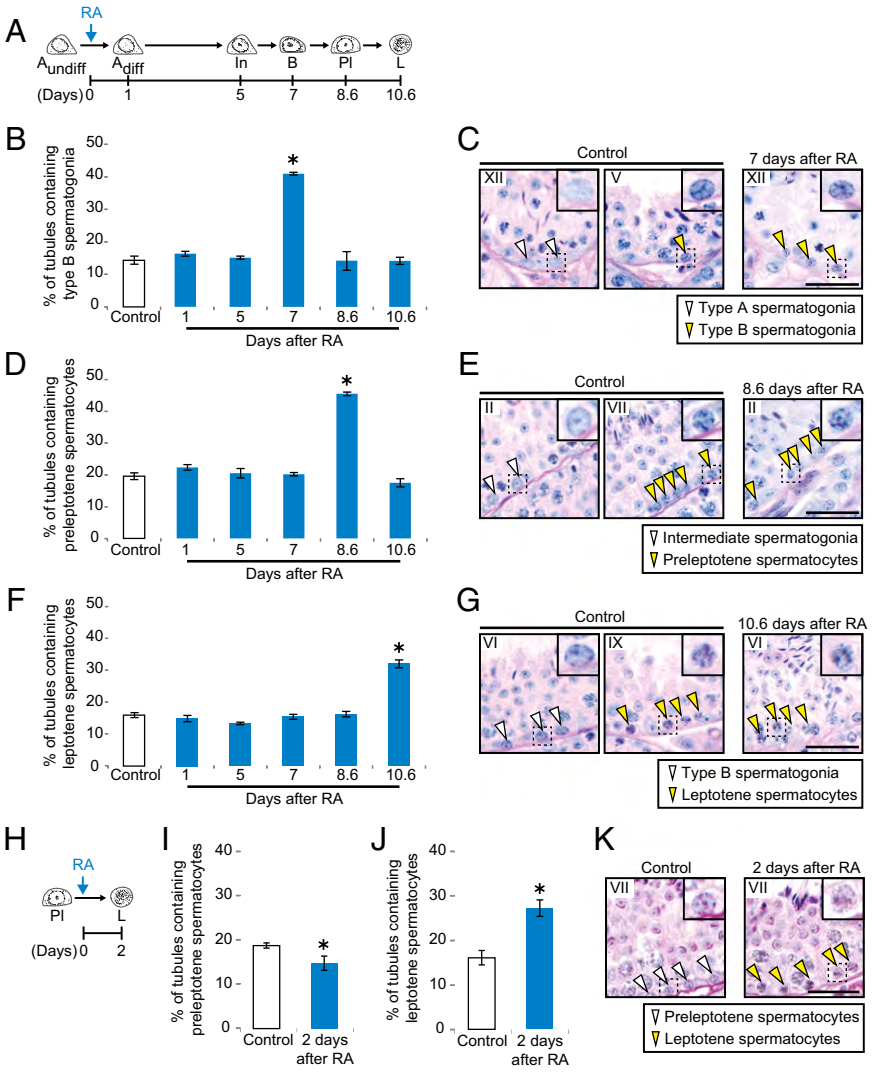


Fig. 5. Injected RA induces precocious spermatogonial differentiation and meiotic initiation. (*A*) Diagram of predicted germ-cell development after RA-induced spermatogonial differentiation. (*B*, *D*, and *F*) Percentage of tubules containing type B spermatogonia (*B*), preleptotene spermatocytes (*D*), or leptotene spermatocytes (*F*), in control or RA-injected testis cross-sections. Error bars, mean ± SD **P* < 0.01 compared with control (Dunnett's test). (*C*, *E*, and *G*) Control and RA-injected testis cross-sections, stained with hematoxylin and periodic acid-Schiff (He-PAS). Roman numerals indicate stages. Insets enlarge the boxed regions. Arrowheads in *C*: type A (white) and type B (yellow) spermatogonia. Arrowheads in *E*: intermediate spermatogonia (white) and preleptotene spermatocytes (yellow). Arrowheads in *G*: type B spermatogonia (white) and leptotene spermatocytes (yellow). (Scale bars, 30 μm.) (*H*) Diagram of predicted germ-cell development after RA-induced meiotic initiation. (*I* and *J*) Percentage of tubules containing preleptotene (*I*) or leptotene (*J*) spermatocytes, in control or RA-injected testis cross-sections. Error bars, mean ± SD **P* < 0.01 compared with control (Dunnett's test). (*K*) Control and RA-injected testis cross-sections, stained with He-PAS. Roman numerals indicate stages. Insets enlarge boxed regions. Arrowheads: preleptotene (white) and leptotene (yellow) spermatocytes. (Scale bars, 30 μm.)

(39) and SYCP3 (synaptonemal complex protein 3) (40). Indeed, at 10.6 d after RA injection, leptotene spermatocytes in stages VI–X were γH2AX- and SYCP3-positive (Fig. 6*A*). The stages at which type B spermatogonia, preleptotene spermatocytes, and leptotene spermatocytes appeared after RA injection were completely consistent with spermatogonial differentiation having occurred throughout stages II–VIII (Fig. S5 *E*–*G* and Table S2).

We conclude that injected RA induced precocious spermatogonial differentiation. The precociously differentiated spermatogonia then progressed into meiotic prophase, ahead of schedule. Spermatogonial differentiation was limited to stages II–VIII, whereas undifferentiated spermatogonia in stages IX–I were seemingly unaffected by RA.

Injected RA Induces Precocious Meiotic Initiation. Because injected RA induces precocious STRA8 expression in both premeiotic cells and undifferentiated spermatogonia, and STRA8 is required for meiotic initiation, we hypothesized that injected RA would also induce precocious meiotic initiation. In the unperturbed testis, germ cells initiate meiosis in late stage VII and stage VIII and then develop into leptotene spermatocytes 2 d later. Thus, we expect a transient increase in leptotene spermatocytes at 2 d after RA injection. Indeed, leptotene spermatocytes were present in an increased fraction of testis tubules, in a broader range of stages (VII–X) than in control testes (VIII–X). The percentage of tubules containing preleptotene spermatocytes was correspondingly decreased (Fig. 5 *H*–*K* and Fig. S5*H*). The precocious leptotene cells had normal meiotic γH2AX and SYCP3 expression patterns (Fig. 6*B*). To confirm that precocious meiotic initiation was a specific effect of RA–STRA8 signaling, we used WIN18,446 to chemically block RA synthesis and inhibit STRA8 expression (Fig. S5*I*). As expected, WIN18,446 prevented meiotic initiation in preleptotene spermatocytes (Fig. S5 *J* and *K*). We also confirmed that the precocious leptotene spermatocytes could progress normally through meiosis (Fig. S6 and *SI Results and Discussion*).

Our results show that premeiotic cells initiated meiosis precociously in response to injected RA and then progressed normally through meiosis, ahead of their usual schedule. However, precocious meiotic initiation occurred in fewer tubules than precocious spermatogonial differentiation (Fig. 5 *B* and *J*), strongly suggesting that the window of competence for meiotic initiation was narrower than that for spermatogonial differentiation. Based on the stages in which leptotene spermatocytes, zygotene spermatocytes, and meiotically dividing cells appeared after RA injection, we calculate that precocious meiotic initiation occurred in stage VI, and perhaps also in stages IV–V. This contrasts with precocious spermatogonial differentiation, which occurred throughout stages II–VI. Moreover, only premeiotic cells, not undifferentiated spermatogonia, were able to initiate

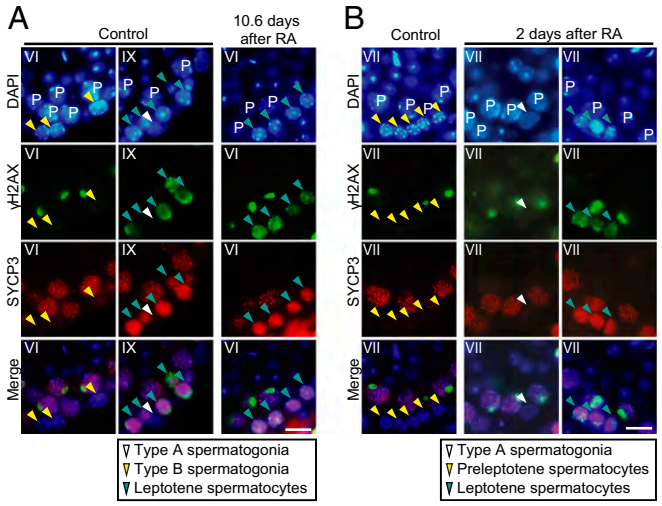


Fig. 6. Injected RA induces precocious expression of meiotic markers. (*A* and *B*) Immunostaining for γH2AX (green) and SYCP3 (red), with DAPI counterstain (blue), on control and RA-injected testis cross-sections. Roman numerals indicate stages. Arrowheads in *A*: type A spermatogonia (white), type B spermatogonia (yellow), and leptotene spermatocytes (green). Arrowheads in *B*: type A spermatogonia (white), preleptotene spermatocytes (yellow), and leptotene spermatocytes (green). P: pachytene spermatocytes. (Scale bars, 10 μm.)

meiosis directly in response to RA, as judged by the absence of γH2AX and SYCP3 signals in type A spermatogonia, 2 d after RA injection (Fig. 6*B*). Thus, the competencies of germ cells to interpret the RA–STRA8 signal are distinct between undifferentiated spermatogonia and premeiotic cells.

Competencies for Spermatogonial Differentiation and Meiotic Initiation Are Limited to Distinct Subsets of Germ Cells. We set out to verify these distinct competencies using more stringent criteria. In the unperturbed testis, the seminiferous cycle lasts 8.6 d (i.e., in a given tubule section, spermatogonial differentiation and meiotic initiation occur once every 8.6 d) (Fig. 1*B*). We administered successive RA injections, once per 8.6-d cycle (Fig. 7*A*). We then predicted when different germ-cell types should appear, based on our findings that competence for spermatogonial differentiation was limited to stages II–VIII, that competence for meiotic initiation was limited to a subset of these stages, and that germ cells developed normally after precocious spermatogonial differentiation/meiotic initiation.

Competence for meiotic initiation. We first predicted that premeiotic cells in stages VI–VIII (and possibly also in stages IV/V), having initiated meiosis, would develop into step 7–8 spermatids after two 8.6-d intervals of RA injection (2 × 8.6 d) (Fig. 1*B*). Indeed, an increased percentage of tubules contained step 7–8 spermatids; step 6 spermatids were correspondingly depleted (Fig. S7*A*). Step 2–5 spermatids were virtually unchanged, demonstrating that meiotic initiation occurred specifically in stages VI–VIII, not in stages IV–V.

Competence for spermatogonial differentiation. We next predicted that spermatogonia in stages II–VIII, having differentiated, would develop into step 7–8 spermatids after 3 × 8.6 d of RA injection, then develop into spermatozoa after 4 × 8.6 d, and finally be released into the tubule lumen (Fig. 1*B*). Indeed, we saw increases in step 7–8 spermatids and spermatozoa after 3 × 8.6 d and 4 × 8.6 d, respectively (Fig. 7 *B* and *C* and Fig. S7*A*). Successive RA injections were able to repeatedly induce spermatogonial differentiation; at 4 × 8.6 d, spermatozoa combined with younger RA-induced generations of germ cells to produce an excess of stage VII/VIII germ-cell associations, with a corresponding depletion of stages II–VI (Fig. 7*B*). Finally, after 4 × 8.6 plus 2 d of RA injection (36.4 d total), an increased percentage of tubules had released their spermatozoa (Fig. S7 *B–D*). All these results are entirely consistent with competence for spermatogonial differentiation being limited to stages II–VIII.

Competence for neither. Finally, because germ cells in stages IX–I are competent for neither spermatogonial differentiation nor meiotic initiation, they should be unaffected by successive RA

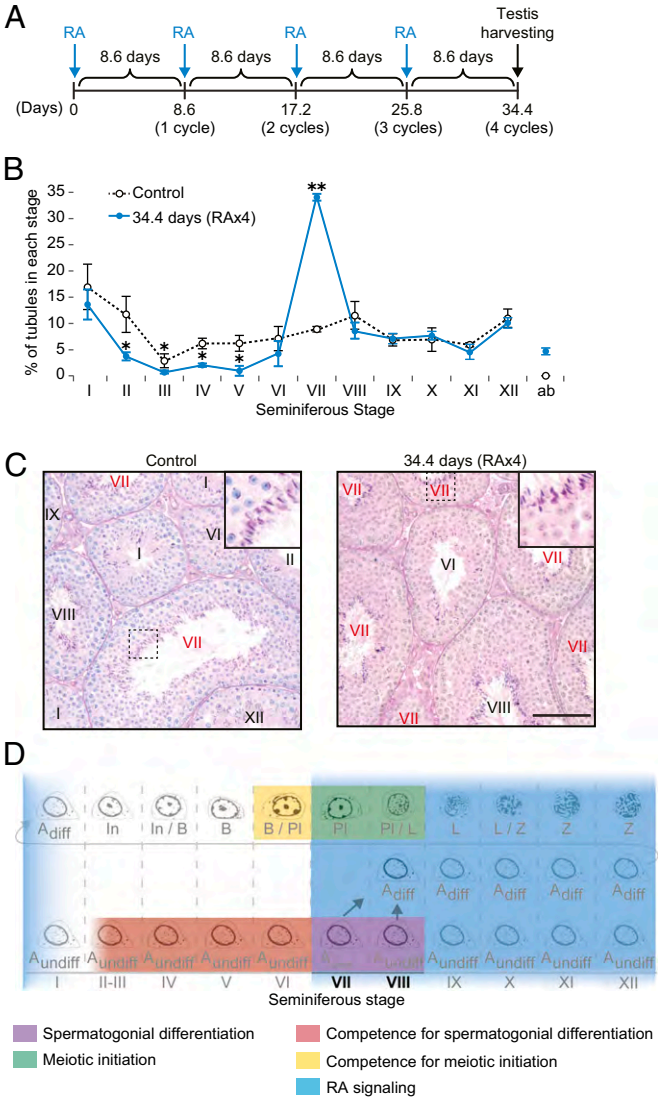


Fig. 7. Periodic RA–STRA8 signaling intersects with periodic germ-cell competencies to regulate spermatogenesis. (*A*) Diagram of four successive RA injections. Testes were harvested 34.4 d after the first RA injection. (*B*) Percentage of tubules in each stage of seminiferous cycle, in controls and after four RA injections. Abnormal germ-cell associations indicated by "ab"; all other germ-cell associations are fully normal. Error bars, mean ± SD $*P < 0.05$, $**P < 0.01$ (two-tailed Welch's *t* test). (*C*) Testis cross-sections, stained with He-PAS, in controls (*Left*) or after four RA injections (*Right*). Cross-sections with the highest frequency of stage VII tubules (red) were selected. Stage VII tubules contain spermatozoa, clustered around the tubule lumen. Roman numerals indicate stages. Insets enlarge boxed regions. (Scale bar, 100 μm.) (*D*) Model: Periodic RA–STRA8 signaling and periodic germ-cell competencies regulate spermatogonial differentiation and meiotic initiation. Red: competence for spermatogonial differentiation. Yellow: competence for meiotic initiation. Blue: stages in which we infer RA signaling is active. Competence for spermatogonial differentiation intersects with RA signaling to induce spermatogonial differentiation (purple). Competence for meiotic initiation intersects with RA signaling to induce meiotic initiation (green).

injections. Indeed, after 4 × 8.6 d of RA injections, the frequency of stage IX–I tubules was the same as in controls (Fig. 7*B*).

We found that, when germ cells are provided with RA, competence to undergo spermatogonial differentiation is strictly limited to stages II–VIII, whereas competence to undergo meiotic initiation is strictly limited to stages VI–VIII. The accuracy of our predictions, over long time scales, demonstrated that these windows of competence are precise. Furthermore, germ cells were able to develop at their normal pace after precocious spermatogonial differentiation/meiotic initiation. This development occurred even when germ cells were outside of their usual cell associations. Injected RA is thus able to accelerate spermatogenesis. Finally, we note that successive RA injections, combined with intrinsic germ-cell competencies, repeatedly induced spermatogonial differentiation; four successive injections were thus able to reestablish normal stage VII/VIII germ-cell associations (Fig. 7 *B* and *C*). We conclude that spermatogonial differentiation and meiotic initiation are regulated by a shared RA–STRA8 signal intersecting with two distinct germ-cell competencies (Fig. 7*D*).

Discussion

RA–STRA8 Signaling Coordinates Spermatogonial Differentiation and Meiotic Initiation. Spermatogenesis in rodents is elaborately organized, with multiple generations of germ cells developing in stereotypical cell associations. This organization was first reported in 1888 (15) and by the 1950s had been comprehensively described (13, 33, 41). To understand spermatogenesis, we must systematically perturb its organization. Here, we used two complementary perturbations, genetic ablation of *Stra8* function and chemical manipulation of RA levels, to probe the coordination of two key transitions: spermatogonial differentiation and meiotic initiation.

We report that RA–STRA8 signaling plays an instructive role in both spermatogonial differentiation and meiotic initiation, inducing these transitions to occur together. Specifically, we provide the first functional evidence to our knowledge that STRA8, an RA target gene, promotes spermatogonial differentiation (as well as being required for meiotic initiation) (24, 25). In the absence of *Stra8*, spermatogonial differentiation was impaired: Undifferentiated spermatogonia began to accumulate as early as p10, ultimately giving rise to massive accumulations of type A spermatogonia in aged testes. These findings show that RA acts instructively at spermatogonial differentiation, by altering gene expression in spermatogonia. Genetic ablation of *Stra8* did not completely block spermatogonial differentiation, indicating that RA must act through additional targets at this transition. Additional targets could be activated either directly by RARs in spermatogonia, or indirectly, by the action of RA on the supporting somatic (Sertoli) cells of testis. Indeed, indirect RA signaling, via RARα in Sertoli cells, is critical for the first round of spermatogonial differentiation (42). We also report that, in wild-type mice, RA injection induced precocious spermatogonial differentiation and meiotic initiation. We infer that, in the unperturbed wild-type testis, a single pulse of RA signaling drives STRA8 expression in both undifferentiated spermatogonia and premeiotic spermatocytes and induces two distinct, cell-type-specific responses. This shared RA–STRA8 signal helps to ensure that spermatogonial differentiation and meiotic initiation occur at the same time and place (Fig. 7*D*).

Evidence of Elevated RA Concentration in Stages VII–XII/I. In any given tubule cross-section, spermatogonial differentiation and meiotic initiation occur periodically, once every 8.6 d. Sugimoto et al. (43) and Hogarth et al. (44) have hypothesized that RA concentration also varies periodically over the course of this 8.6-d cycle. This hypothesis is supported by expression data, functional studies, and direct measurements of RA levels (34, 43, 45, 46). However, the pattern of RA periodicity was previously unclear. Hogarth et al. (34, 44) suggested a sharp RA peak in stages VIII–IX. In contrast, based on expression patterns of RA-responsive genes and the functional consequences of inhibiting RA signaling, Hasegawa and Saga (45) suggested that RA levels rise in stage VII and remain high through stage XII.

Our data support the latter model, of a prolonged elevation of RA levels (Fig. 7*D*). We and others (35) have demonstrated that, in the unperturbed testis, STRA8 is periodically expressed and is present for the majority of the seminiferous cycle. Specifically, we show that STRA8 protein is present in spermatogonia in stages VII–XII/I and absent in II–VI. Furthermore, we show that, at the level of the tubule cross-section, spermatogonial STRA8 expression marks the presence of RA: When we increased RA levels by injecting RA or decreased them by injecting WIN18,446, STRA8 expression was immediately induced or repressed in all seminiferous stages (Fig. 4*C*). We thus agree with and extend the model of Hasegawa and Saga (45): In the unperturbed testis, RA levels rise in stage VII, rapidly inducing STRA8 and then inducing spermatogonial differentiation and meiotic initiation. RA levels remain high until stages XII/I. This model of a long RA–STRA8 pulse is consistent with additional published data. First, the enzyme *Aldh1a2,* which increases RA levels, is strongly expressed in stages VII–XII, whereas the enzymes *Lrat* and *Adfp*, which reduce RA levels, are expressed in stages I–VI/VII (43, 46). Second, although measured RA levels seem to peak in stages VIII–IX, they remain elevated for an extended period (2–4 d in pubertal animals, and through stage XII in adults) (34). Despite these persistently elevated RA levels, neither spermatogonial differentiation nor meiotic initiation recurs in stages IX–I. As we will now discuss, germ cells at these later stages lack competence for these transitions.

Undifferentiated Spermatogonia and Premeiotic Cells Have Different Competencies to Respond to RA–STRA8 Signaling. By examining responses to exogenous RA, we provide functional evidence that germ cells have periodic, stage-limited competencies to undergo spermatogonial differentiation and meiotic initiation. These competencies intersect with instructive, periodic RA–STRA8 signaling. Specifically, undifferentiated spermatogonia are competent for spermatogonial differentiation in stages II–VIII, and premeiotic cells are competent for meiotic initiation in stages VI–VIII (Fig. 7*D*). Competencies for both transitions begin while RA levels are low, so that the germ cells respond as soon as RA levels rise. Competencies for both transitions end simultaneously, while RA levels are still high. Thus, germ-cell competencies and high RA levels intersect briefly, causing spermatogonial differentiation and meiotic initiation to occur at the same time and place (in stages VII–VIII).

We also conclude that undifferentiated spermatogonia and premeiotic cells enact different molecular and cellular programs in response to RA–STRA8 signaling. In response to injected RA, only preleptotene spermatocytes (and possibly late type B spermatogonia, the immediate precursors of preleptotene spermatocytes) began to express meiotic markers such as SYCP3 and γH2AX. Undifferentiated spermatogonia were not competent to initiate meiosis directly. Instead, in response to injected RA, most late undifferentiated spermatogonia began a program of spermatogonial differentiation, followed by six mitotic cell divisions. The early undifferentiated spermatogonia and a fraction of the late undifferentiated spermatogonia were seemingly unaffected by RA; they did not express STRA8 and did not differentiate. We believe that these undifferentiated spermatogonia are able to self-renew and proliferate even in the presence of RA, preventing the pool of undifferentiated spermatogonia from becoming depleted. Indeed, a normal complement of germ cells remained after repeated RA injections, indicating that injected RA did not eliminate the pool of undifferentiated spermatogonia (which includes the spermatogonial stem cells). Thus, distinct germ-cell competencies enable a single RA signal to induce both

spermatogonial differentiation and meiotic initiation and ensure that a subset of spermatogonia are able to self-renew and proliferate despite exposure to the RA signal.

We do not yet know the molecular mechanism behind the stage- and cell-type-specific competencies to differentiate in response to RA. These competencies cannot simply be explained by RAR expression, because the RARs do not have precise stage-specific expression patterns. For instance, RARγ expression can be observed in all stages of the seminiferous cycle (21, 46). The competencies must therefore result from other aspects of germ-cell state. We note that competence for spermatogonial differentiation is closely correlated with proliferative activity. Specifically, undifferentiated spermatogonia in stages II–VIII, which are competent for differentiation, are arrested in the G0/G1 phase of the cell cycle, whereas undifferentiated spermatogonia in stages IX–I are actively proliferating (10, 38). Further studies are needed to identify the mechanisms by which competencies to undergo spermatogonial differentiation, and then meiotic initiation, are achieved.

The critical role for intrinsic germ-cell competence during spermatogenesis is in some respects analogous to its role of competence during oogenesis. In adult ovaries, immature oocytes are arrested at the diplotene stage of meiotic prophase. Some of these arrested oocytes grow and acquire intrinsic competence to resume meiosis and then acquire competence to mature (i.e., to progress to metaphase II arrest) (47, 48). These serially acquired competencies intersect with extrinsic, hormonal signals. We suggest that, in both oogenesis and spermatogenesis, properly timed differentiation depends on the intersection of extrinsic chemical cues and intrinsic competence.

RA–STRA8 Signaling Can Both Perturb and Reestablish the Complex Organization of the Testis. We find that injected RA can induce spermatogonial differentiation and meiotic initiation to occur precociously and ectopically, outside of their normal context. In the unperturbed testis, these two transitions occur together at stages VII/VIII, but, following a single RA injection, they occurred in different stages, with spermatogonial differentiation as early as stage II and meiotic initiation as early as stage VI. Then, when provided with RA at 8.6-d intervals, these precociously advancing germ cells were able to complete meiosis and develop into spermatozoa, ahead of schedule. This developmental flexibility is surprising, given the seemingly rigid organization of spermatogenesis (11). In the unperturbed testis, multiple generations of germ cells occur together in stereotypical associations; these associations are conserved across mammals and, before this study, had proven difficult to chemically disrupt (43, 49, 50). Nevertheless, when provided with RA, germ cells proceeded through spermatogenesis, outside of their usual environs, with no apparent guidance from the neighboring germ cells.

Why is spermatogenesis so precisely organized, if the stereotypical associations are not required for germ-cell development? We posit that this precise organization is in part a by-product of RA–STRA8 signaling (and germ-cell competencies): Cooccurrence of spermatogonial differentiation and meiotic initiation nucleates the stereotypical germ-cell associations. In support of this idea, when we administered successive RA injections at 8.6-d intervals, to repeatedly drive precocious spermatogonial differentiation and meiotic initiation, we were able to perturb and reestablish the characteristic germ-cell associations in vivo. The stereotypical associations, established by RA–STRA8 signaling, may ensure the efficiency of spermatogenesis.

We conclude that a simple regulatory mechanism helps to explain the testis's extraordinary capacity for proliferation and differentiation. Periodic RA signaling repeatedly induces spermatogonial differentiation and meiotic initiation, driving germ cells toward becoming highly specialized haploid spermatozoa. Meanwhile, distinct germ-cell competencies enforce that every spermatogonium undergoes programmed amplifying divisions before initiating meiosis, guaranteeing a prodigious output of spermatozoa. Moreover, a fraction of spermatogonia undergo neither spermatogonial differentiation nor meiotic initiation in response to RA, ensuring that a reservoir of undifferentiated spermatogonia is maintained throughout the animal's reproductive lifetime. This basic understanding of the organization of spermatogenesis, derived from genetic and chemical perturbations, will facilitate future studies of germ-cell development, RA-driven differentiation, and cell competence, both in vivo and in vitro.

Materials and Methods

Mice. Three types of mice were used: wild-type (C57BL/6NtacfBR), *Stra8*-deficient (extensively back-crossed to C57BL/6) (26, 27), and *Dmc1*-deficient (B6.Cg-Dmc1^{tm1Jcs}/JcsJ) (51). See *SI Materials and Methods* for strain and genotyping details. Unless otherwise noted, experiments were performed on 6- to 8-wk-old male mice, fed a regular (vitamin A-sufficient) diet. All experiments involving mice were approved by the Committee on Animal Care at the Massachusetts Institute of Technology.

Statistics. Data are represented as mean ± SD of three biological replicates. To compare two groups, Welch's *t* test (one- or two-tailed as indicated) or the Mann–Whitney *U* test were used. To compare three or more groups, one-way ANOVA with the Tukey–Kramer post hoc test was used. To compare multiple experimental groups with a control group, one-way ANOVA with Dunnett's post hoc test was used. To compare distributions, the Kolmogorov–Smirnov test was used. When performing genome-wide analysis of mRNA-Seq data, the Benjamini–Hochberg procedure was used to control the false discovery rate.

mRNA-Seq Sample Preparation. Testes were stripped of the tunica albuginea, placed in TRIzol (Invitrogen), homogenized, and stored at −20 °C. Total RNAs were prepared according to the manufacturer's protocol. Total RNAs were then DNase-treated using DNA Free Turbo (Ambion). Libraries were prepared using the Illumina mRNA-Seq Sample Preparation Kit according to the manufacturer's protocol. Libraries were validated with an Agilent Bioanalyzer. Libraries were diluted to 10 pM and applied to an Illumina flow cell using the Illumina Cluster Station. The Illumina Genome Analyzer II platform was used to sequence 36-mers (single end) from the mRNA-Seq libraries.

mRNA-Seq and Microarray Data Analysis. For mRNA-Seq data, reads were aligned to the mouse genome using TopHat (52). Analysis was performed using edgeR (53), Cufflinks (54), and custom R scripts. Microarray data were normalized with the GCRMA package from Bioconductor, and replicates were averaged using limma (55). Comparison mRNA-Seq and microarray datasets were downloaded from National Center for Biotechnology Information GEO and Sequence Read Archive (SRA). See *SI Materials and Methods* for details on mRNA-Seq and microarray data processing and comparison.

Histology. Testes were fixed overnight in Bouin's solution, embedded in paraffin, sectioned, and stained with hematoxylin and eosin, or with hematoxylin and periodic acid-Schiff (PAS). All sections were examined using a light microscope. Germ-cell types were identified by their location, nuclear size, and chromatin pattern (11). See *SI Materials and Methods* for details on identification of the stages of the seminiferous cycle.

Chemical Treatments. For RA injection experiments, mice received i.p. injections of 100 μL of 7.5 mg/mL all-*trans* RA (Sigma-Aldrich) in 16% (vol/vol) DMSO–H_2O. For BrdU incorporation experiments, mice received i.p. injections of 10 μL/g body weight of 10 mg/mL BrdU (Sigma-Aldrich) in PBS, 4 h before they were killed. For WIN18,446 injection experiments, mice received i.p. injections of 100 μL of 20 mg/mL WIN18,446 (sc-295819A; Santa Cruz Biotechnology) in 16% DMSO–H_2O; mice were dosed at intervals of 12 h for a total of 2 or 4 d.

Immunostaining on Testis Sections. Testes were fixed overnight in Bouin's solution or 4% (wt/vol) paraformaldehyde, embedded in paraffin, and sectioned at 5-μm thickness. Slides were dewaxed, rehydrated, and heated in 10 mM sodium citrate buffer (pH 6.0). Sections were then blocked, incubated with the primary antibody, washed with PBS, incubated with the secondary antibody, and washed with PBS. Detection was fluorescent or

colorimetric. Antibodies and incubation conditions are provided in *SI Materials and Methods* and in Table S3.

Immunostaining on Intact Testis Tubules. Testes were stripped of the tunica albuginea, dispersed in PBS, fixed overnight in 4% paraformaldehyde at 4 °C, and washed with PBS. Testes were blocked with 2.5% (vol/vol) donkey serum, incubated with the primary antibody, washed with PBS, incubated with the secondary antibody, and washed with PBS. Finally, seminiferous tubules were dissected from testes and mounted with SlowFade Gold antifade reagent with DAPI (S36939; Life Technologies). Antibodies and incubation conditions are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank H. Skaletsky for statistical advice; M. E. Gill for helpful discussions; and D. W. Bellott, T. Bhattacharyya, M. Carmell, J. Hughes, M. Kojima, B. Lesch, S. Naqvi, P. Nicholls, T. Shibue, S. Soh, and L. Teitz for critical reading of the manuscript. This work was supported by the Howard Hughes Medical Institute and by NIH Pre-Doctoral Training Grant T32GM007287.

1. Nakagawa T, Nabeshima Y, Yoshida S (2007) Functional identification of the actual and potential stem cell compartments in mouse spermatogenesis. *Dev Cell* 12(2): 195–206.
2. Tegelenbosch RA, de Rooij DG (1993) A quantitative study of spermatogonial multiplication and stem cell renewal in the C3H/101 F1 hybrid mouse. *Mutat Res* 290(2): 193–200.
3. Brinster RL, Zimmermann JW (1994) Spermatogenesis following male germ-cell transplantation. *Proc Natl Acad Sci USA* 91(24):11298–11302.
4. Shinohara T, Orwig KE, Avarbock MR, Brinster RL (2000) Spermatogonial stem cell enrichment by multiparameter selection of mouse testis cells. *Proc Natl Acad Sci USA* 97(15):8346–8351.
5. Zheng K, Wu X, Kaestner KH, Wang PJ (2009) The pluripotency factor LIN28 marks undifferentiated spermatogonia in mouse. *BMC Dev Biol* 9:38.
6. Pesce M, Wang X, Wolgemuth DJ, Schöler H (1998) Differential expression of the Oct-4 transcription factor during mouse germ cell differentiation. *Mech Dev* 71(1-2):89–98.
7. Tokuda M, Kadokawa Y, Kurahashi H, Marunouchi T (2007) CDH1 is a specific marker for undifferentiated spermatogonia in mouse testes. *Biol Reprod* 76(1):130–141.
8. Kanatsu-Shinohara M, et al. (2004) Generation of pluripotent stem cells from neonatal mouse testis. *Cell* 119(7):1001–1012.
9. Buaas FW, et al. (2004) Plzf is required in adult male germ cells for stem cell self-renewal. *Nat Genet* 36(6):647–652.
10. Lok D, de Rooij DG (1983) Spermatogonial multiplication in the Chinese hamster. III. Labelling indices of undifferentiated spermatogonia throughout the cycle of the seminiferous epithelium. *Cell Tissue Kinet* 16(1):31–40.
11. Russell LD, Ettlin RA, Sinha Hikim AP, Clegg ED (1990) *Histological and Histopathological Evaluation of the Testis* (Cache River, Clearwater, FL).
12. Huckins C (1971) The spermatogonial stem cell population in adult rats. I. Their morphology, proliferation and maturation. *Anat Rec* 169(3):533–557.
13. Leblond CP, Clermont Y (1952) Spermiogenesis of rat, mouse, hamster and guinea pig as revealed by the periodic acid-fuchsin sulfurous acid technique. *Am J Anat* 90(2): 167–215.
14. Lok D, Weenk D, De Rooij DG (1982) Morphology, proliferation, and differentiation of undifferentiated spermatogonia in the Chinese hamster and the ram. *Anat Rec* 203(1):83–99.
15. von Ebner V (1888) Zur spermatogenese bei den säugethieren. *Arch f Mikr Anat* 31(1): 236–292. German.
16. Van Pelt AMM, De Rooij DG (1990) The origin of the synchronization of the seminiferous epithelium in vitamin A-deficient rats after vitamin A replacement. *Biol Reprod* 42(4):677–682.
17. van Pelt AMM, de Rooij DG (1990) Synchronization of the seminiferous epithelium after vitamin A replacement in vitamin A-deficient mice. *Biol Reprod* 43(3):363–367.
18. van Pelt AMM, de Rooij DG (1991) Retinoic acid is able to reinitiate spermatogenesis in vitamin A-deficient rats and high replicate doses support the full development of spermatogenic cells. *Endocrinology* 128(2):697–704.
19. Morales C, Griswold MD (1987) Retinol-induced stage synchronization in seminiferous tubules of the rat. *Endocrinology* 121(1):432–434.
20. Baleato RM, Aitken RJ, Roman SD (2005) Vitamin A regulation of BMP4 expression in the male germ line. *Dev Biol* 286(1):78–90.
21. Gely-Pernot A, et al. (2012) Spermatogonia differentiation requires retinoic acid receptor γ. *Endocrinology* 153(1):438–449.
22. Bowles J, et al. (2006) Retinoid signaling determines germ cell fate in mice. *Science* 312(5773):596–600.
23. Koubova J, et al. (2006) Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proc Natl Acad Sci USA* 103(8):2474–2479.
24. Anderson EL, et al. (2008) Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. *Proc Natl Acad Sci USA* 105(39): 14976–14980.
25. Baltus AE, et al. (2006) In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. *Nat Genet* 38(12):1430–1434.
26. Oulad-Abdelghani M, et al. (1996) Characterization of a premeiotic germ cell-specific cytoplasmic protein encoded by Stra8, a novel retinoic acid-responsive gene. *J Cell Biol* 135(2):469–477.
27. Zhou Q, et al. (2008) Expression of stimulated by retinoic acid gene 8 (Stra8) in spermatogenic cells induced by retinoic acid: An in vivo study in vitamin A-sufficient postnatal murine testes. *Biol Reprod* 79(1):35–42.
28. Zhou Q, et al. (2008) Expression of stimulated by retinoic acid gene 8 (Stra8) and maturation of murine gonocytes and spermatogonia induced by retinoic acid in vitro. *Biol Reprod* 78(3):537–545.
29. Wang PJ, McCarrey JR, Yang F, Page DC (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* 27(4):422–426.
30. Namekawa SH, et al. (2006) Postmeiotic sex chromatin in the male germline of mice. *Curr Biol* 16(7):660–667.
31. Costoya JA, et al. (2004) Essential role of Plzf in maintenance of spermatogonial stem cells. *Nat Genet* 36(6):653–659.
32. Nakagawa T, Sharma M, Nabeshima Y, Braun RE, Yoshida S (2010) Functional hierarchy and reversibility within the murine spermatogenic stem cell compartment. *Science* 328(5974):62–67.
33. Oakberg EF (1956) A description of spermiogenesis in the mouse and its use in analysis of the cycle of the seminiferous epithelium and germ cell renewal. *Am J Anat* 99(3): 391–413.
34. Hogarth CA, et al. (2015) Processive pulses of retinoic acid propel asynchronous and continuous murine sperm production. *Biol Reprod* 92(2):37.
35. Mark M, Teletin M, Vernet N, Ghyselinck NB (2015) Role of retinoic acid receptor (RAR) signaling in post-natal male germ cell differentiation. *Biochim Biophys Acta* 1849(2):84–93.
36. Amory JK, et al. (2011) Suppression of spermatogenesis by bisdichloroacetyldiamines is mediated by inhibition of testicular retinoic acid biosynthesis. *J Androl* 32(1): 111–119.
37. Hogarth CA, et al. (2011) Suppression of Stra8 expression in the mouse gonad by WIN 18,446. *Biol Reprod* 84(5):957–965.
38. Oakberg EF (1971) Spermatogonial stem-cell renewal in the mouse. *Anat Rec* 169(3): 515–531.
39. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* 273(10): 5858–5868.
40. Yuan L, et al. (2000) The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. *Mol Cell* 5(1):73–83.
41. Oakberg EF (1956) Duration of spermatogenesis in the mouse and timing of stages of the cycle of the seminiferous epithelium. *Am J Anat* 99(3):507–516.
42. Raverdeau M, et al. (2012) Retinoic acid induces Sertoli cell paracrine signals for spermatogonia differentiation but cell autonomously drives spermatocyte meiosis. *Proc Natl Acad Sci USA* 109(41):16582–16587.
43. Sugimoto R, Nabeshima Y, Yoshida S (2012) Retinoic acid metabolism links the periodical differentiation of germ cells with the cycle of Sertoli cells in mouse seminiferous epithelium. *Mech Dev* 128(11-12):610–624.
44. Hogarth CA, et al. (2013) Turning a spermatogenic wave into a tsunami: Synchronizing murine spermatogenesis using WIN 18,446. *Biol Reprod* 88(2):40.
45. Hasegawa K, Saga Y (2012) Retinoic acid signaling in Sertoli cells regulates organization of the blood-testis barrier through cyclical changes in gene expression. *Development* 139(23):4347–4355.
46. Vernet N, et al. (2006) Retinoic acid metabolism and signaling pathways in the adult and developing mouse testis. *Endocrinology* 147(1):96–110.
47. Mehlmann LM (2005) Stops and starts in mammalian oocytes: Recent advances in understanding the regulation of meiotic arrest and oocyte maturation. *Reproduction* 130(6):791–799.
48. Sorensen RA, Wassarman PM (1976) Relationship between growth and meiotic maturation of the mouse oocyte. *Dev Biol* 50(2):531–536.
49. Meistrich ML (1986) Critical components of testicular function and sensitivity to disruption. *Biol Reprod* 34(1):17–28.
50. Snyder EM, Small C, Griswold MD (2010) Retinoic acid availability drives the asynchronous initiation of spermatogonial differentiation in the mouse. *Biol Reprod* 83(5):783–790.
51. Pittman DL, et al. (1998) Meiotic prophase arrest with failure of chromosome synapsis in mice deficient for Dmc1, a germline-specific RecA homolog. *Mol Cell* 1(5):697–705.
52. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.
53. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1): 139–140.
54. Trapnell C, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28(5):511–515.
55. Smyth GK (2005) limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, eds Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S (Springer, New York), pp 397–420.

Cell Research

www.nature.com/cr
www.cell-research.com

ARTICLE OPEN

The adult human testis transcriptional cell atlas

Jingtao Guo[1,2], Edward J. Grow[1], Hana Mlcochova[3], Geoffrey J. Maher[3], Cecilia Lindskog[4], Xichen Nie[1], Yixuan Guo[1], Yodai Takei[5], Jina Yun[5], Long Cai[5], Robin Kim[6], Douglas T. Carrell[2], Anne Goriely[3], James M. Hotaling[2] and Bradley R. Cairns[1]

Human adult spermatogenesis balances spermatogonial stem cell (SSC) self-renewal and differentiation, alongside complex germ cell-niche interactions, to ensure long-term fertility and faithful genome propagation. Here, we performed single-cell RNA sequencing of ~6500 testicular cells from young adults. We found five niche/somatic cell types (Leydig, myoid, Sertoli, endothelial, macrophage), and observed germline-niche interactions and key human-mouse differences. Spermatogenesis, including meiosis, was reconstructed computationally, revealing sequential coding, non-coding, and repeat-element transcriptional signatures. Interestingly, we identified five discrete transcriptional/developmental spermatogonial states, including a novel early SSC state, termed State 0. Epigenetic features and nascent transcription analyses suggested developmental plasticity within spermatogonial States. To understand the origin of State 0, we profiled testicular cells from infants, and identified distinct similarities between adult State 0 and infant SSCs. Overall, our datasets describe key transcriptional and epigenetic signatures of the normal adult human testis, and provide new insights into germ cell developmental transitions and plasticity.

Cell Research (2018) 28:1141–1157; https://doi.org/10.1038/s41422-018-0099-2

INTRODUCTION

Human spermatogenesis involves the differentiation of adult spermatogonial stem cells (SSCs) into mature sperm through a complex developmental process, regulated by the testis niche. Human SSCs must carefully balance their self-renewal and differentiation, and then undergo niche-guided transitions between multiple cell states and cellular processes—including a commitment to mitosis, meiosis, and the subsequent stages of sperm maturation, which are accompanied by chromatin repackaging and major morphological changes.[1,2] Through a wide range of approaches, considerable progress in understanding gametogenesis and germline-niche communication has been achieved in mice.[3,4] In contrast, in humans, although adult testis physiology is well described,[5–7] much less is known about SSCs and their regulation. Ultimately, a full understanding will require the integration of molecular, genomic, proteomic and physiological approaches.

Toward this goal, single cell RNA-seq (scRNA-seq) approaches can effectively delineate cell types, uncover heterogeneity, and infer developmental trajectories.[8] These approaches have recently been applied to human fetal germ cells, providing important new biological insights.[9] Single-cell approaches are well suited for addressing fundamental questions about SSCs, differentiating spermatogonia and gametogenesis. For example, what are the main molecular features that enable SSCs to serve as the long-term adult germline stem cells? How do SSCs transition from their initial, most naïve and quiescent states to spermatogonia that will eventually commit to meiosis? Are these transitions irreversible, or do spermatogonia possess bidirectional plasticity that helps ensure a lifelong pool of SSCs? Beyond spermatogonia, what are the subsequent sequential transcription and signaling programs that accompany gametogenesis? How are these processes influenced by communication with niche cells—what are the specific signaling and transcription pathways that regulate self-renewal, proliferation rates, metabolism, and transitions between differentiation states? Importantly, these questions overlap conceptually with other stem cell systems. Here, we aim to utilize single-cell transcriptome analysis from the full repertoire of germline and niche cells to address these questions.

Prior scRNA-seq efforts characterizing spermatogonia enriched via cell surface markers have provided initial insights into human spermatogenesis.[10] However, thanks to new technological advances, it is now possible to use unbiased approaches to assess germline and niche cell transcriptional profiles. Here, we performed extensive scRNA-seq characterization of unselected human testicular cells of young adults using the 10× Genomics Chromium platform—yielding a transcriptional cell atlas of all cell types in the testis, including germline and niche cells. We delineate five distinct spermatogonial states in adults, including a novel early SSC state, termed State 0, which displays high similarity to infant SSCs. We further describe the genic and non-coding RNA expression programs that accompany spermatogenesis. Intriguingly, combining RNA 'velocity' analyses[11] with chromatin mapping and DNA methylation (DNAme), we provide computational and molecular evidence that human spermatogonia possess considerable transcriptional/state

[1] Department of Oncological Sciences and Huntsman Cancer Institute, Howard Hughes Medical Institute, University of Utah School of Medicine, Salt Lake City, UT 84112, USA; [2]Department of Surgery (Andrology/Urology), Center for Reconstructive Urology and Men's Health, University of Utah Health Sciences Center, Salt Lake City, UT 84122, USA; [3]Radcliffe Department of Medicine, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX39DS, UK; [4]Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, SE-751 85 Uppsala, Sweden; [5]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA and [6]Section of Transplantation, Department of Surgery, University of Utah School of Medicine, Salt Lake City, UT 84132, USA
Correspondence: Bradley R. Cairns (brad.cairns@hci.utah.edu)
Senior authors: Anne Goriely, James M. Hotaling, Bradley R. Cairns

Received: 30 August 2018 Revised: 7 September 2018 Accepted: 19 September 2018
Published online: 12 October 2018

plasticity, suggesting a conceptual framework for human spermatogonial homeostasis, similar to that described in other stem cell systems.

RESULTS

Cell partitioning through the analysis of single cell transcriptomes

We isolated single cells from whole-testis of 3 individuals using a standard two-step procedure of enzymatic digestion and physical filtering.[7,10] For each donor, two separate technical replicates were performed (Fig. 1a), resulting in six datasets. From a total of ~7000 cells, 6,490 passed standard quality control (QC) dataset filters and were retained for downstream analysis. We obtained ~250 K reads/cell which enabled the analysis of ~2500 genes/cell. The sequencing saturation rate was >83%, and technical replicates were highly similar ($r > 0.96$; Supplementary information, Fig. S1a).

Cell partitioning via t-distributed stochastic neighbor embedding (tSNE) analyses[12] identified 13 clusters (Fig. 1b; Supplementary information, Table S1), with only minor variation based on batch/experiment or donor origin (Fig. 1c; Supplementary information, Fig. S1b and c). Cluster identity was assigned based on known cell-type marker expression (Fig. 1d; Supplementary information, Fig. S2). Clusters 9–13 correspond to macrophage, endothelial, myoid, Sertoli and Leydig cells, respectively (Fig. 1d; Supplementary information, Fig. S2a). Germline-specific markers were expressed solely in Clusters 1–8 (e.g., *DAZL* and *MAGEA4*; Fig. 1d; Supplementary information, Fig. S2b). Moreover, known SSC markers (e.g., *UTF1*, *ID4* and *FGFR3*), differentiating markers (e.g., *KIT* and *DMRT1*), meiosis markers (e.g., *SYCP3*, *SPO11*, and *MLH3*), spermatid structure proteins (e.g., *SPAG6*, *ZPBP*, *CAMK4* and *CREM*) and nuclear condensation/protamine repackaging factors (e.g., *TNP1* and *PRM2*) showed sequential expression peaks in Clusters 1 to 8, respectively—mirroring the temporal order of gametogenesis.

Human-mouse comparisons in intra-niche and niche-germline interaction

We began by describing the niche cell datasets (Fig. 2). Testicular macrophages (Cluster 9) promote spermatogonia maintenance,[13,14] and were identified by multiple specific markers (i.e., *CD14*, *CD163*, *C1QA*; Fig. 1d and 2a). Previous work showed that mouse Sertoli cells help maintain CXCR4$^+$ spermatogonia population by secreting CXCL12, the ligand for CXCR4.[15] Interestingly, in humans, RNA encoding CXCL12 was observed in Leydig cells, while the *CXCR4* receptor was expressed in both macrophages and spermatogonia (Fig. 2a), suggesting that CXCL12-CXCR4 promotes co-localization of macrophages and spermatogonia in humans. Furthermore, *CSF1R*, the receptor for *CSF1*, was specifically expressed in macrophages, whereas in mice its expression is confined to spermatogonia.[16]

In endothelial cells (Cluster 10, marked by *VWF* and *PECAM1*), the receptor (*NOTCH4*) and the downstream signaling factors (*JAG1*, *HES1* and *MAML1*) for NOTCH signaling were specifically upregulated (Fig. 2b). Hedgehog signaling is important for mouse fetal myoid and Leydig cell development;[9,17] but the receptors (*PTCH1*, *PTCH2*), and downstream signaling components (*GLI* and *IGFBP6*) of the Hedgehog pathway were highly expressed in human adult myoid (Cluster 11, marked by *MYH11* and *ACTA2*) and Leydig cells (Fig. 2c), indicating that both Hedgehog and NOTCH signaling activity persists through adulthood in human testes. Sertoli cells (Cluster 12, marked by *SOX9* and *AMH*) express *ITGA6*, an integrin found in the basal membrane of seminiferous tubules in humans (Fig. 2d). Notably, Sertoli cells express *WFDC2*, which is known to encode an epididymis protein that may promote sperm maturation,[18] and *PRND*, which encodes a glycosylphosphatidylinositol-anchored glycoprotein with a putative interaction role with receptors from germ cells.[19]

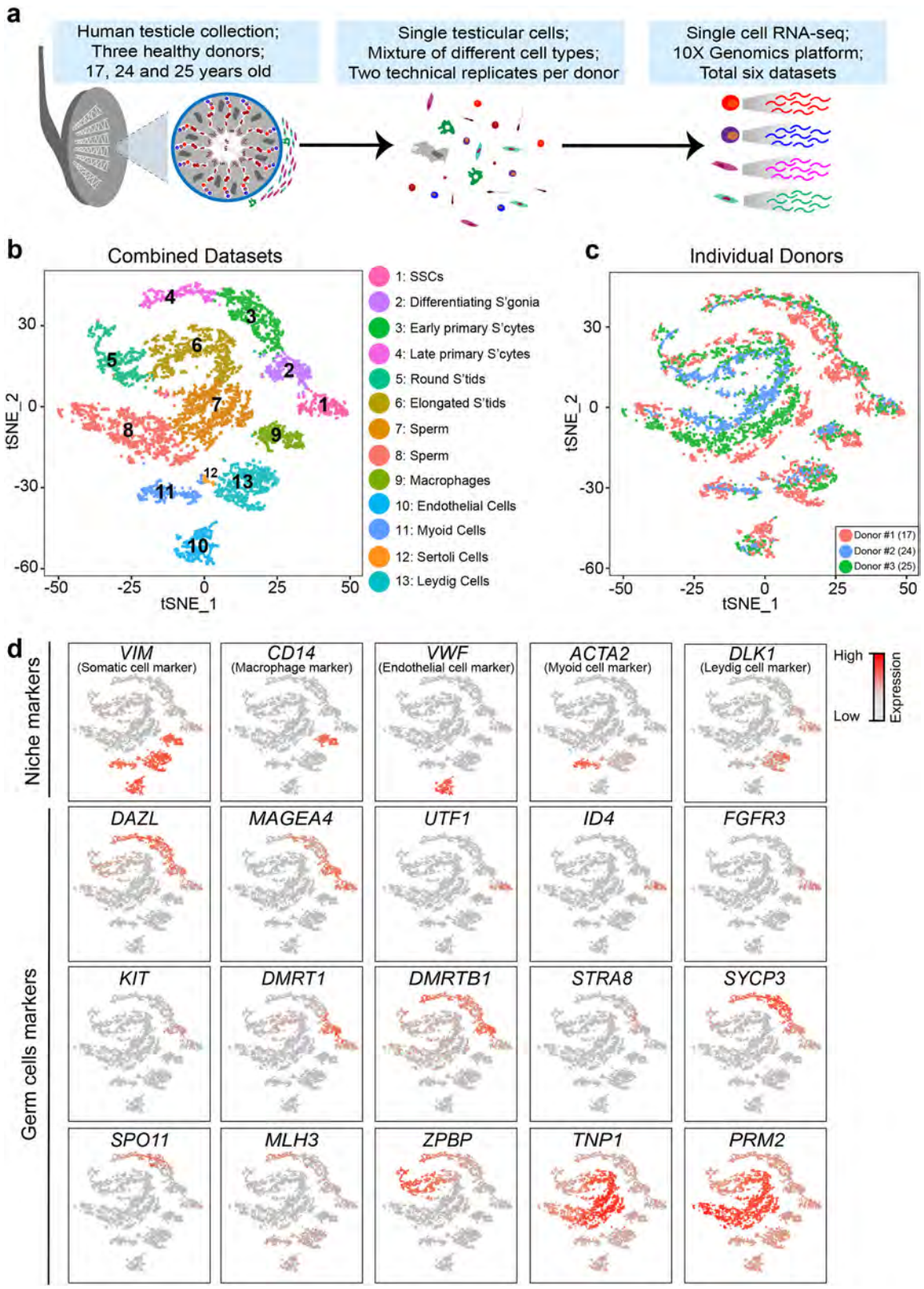
Leydig cells (Cluster 13, marked by *DLK1* and *IGF1*) also expressed specific genes (Fig. 2e), including those encoding IGF binding proteins (*IGFBP5* and *IGFBP3*) and INHBA, a subunit of both inhibin and activin, and the extracellular matrix protein VIT. Interestingly, key genes for testosterone biosynthesis, *STAR* and *HSD17B3*, were expressed in both Sertoli and Leydig cells, while the expression of the responsive genes, *SHBG* and *SRD5A2*, were observed in maturing sperm (Fig. 2f). Retinoic acid (RA) induces germ cell differentiation, and enzymes for RA synthesis, *ALDH1A1* and *ALDH1A3*, were specifically expressed in Leydig and myoid cells; while *STRA8*, an RA target gene, was only observed during the transition of spermatogonia to spermatocytes (Fig. 2f). Interestingly, we found the WNT ligand, WNT2B, was expressed primarily in myoid cells, while WNT2B receptors were confined to primary spermatocytes, suggesting a role for *WNT2B* in human meiosis (Fig. 2f). *PDGFB* was expressed in endothelial cells, and its receptors *PDGFRA* and *PDGFRB* were found in Leydig and myoid cells, indicating that endothelial cells may indirectly affect germ cell development, via cross-talk mechanisms with other niche cells. Taken together, our data highlight both similarities and notable differences in germline-niche interactions in humans and mice that warrant further detailed functional investigations.

Pseudotime and clustering analyses reveal dynamic gene expression patterns during spermatogenesis

Noticeably, the germ cell clusters (1–8) formed a wave-like continuum, sometimes separated by distinct bottlenecks, that recapitulated the temporal order of spermatogenesis. Pseudotime analysis[20] provided an arrow vector which aligned with the developmental order of gametogenesis (Fig. 3a). Clustering analysis of genes (rows) while fixing the order of cells (columns) along pseudotime, revealed 12 distinct gene cohorts. Gene ontology (GO) analysis of these clusters yielded a dynamic developmental, cellular and metabolic sequence of events, consistent with well-organized germline development (Fig. 3b; Supplementary information, Table S2). Next, differential analysis (bimodal test; adjusted p-value < 0.01; $|\log FC| > 0.25$) identified differentially-expressed genes (Fig. 3c; Supplementary information, Table S3). As expected, 'cell cycle', 'meiosis' and 'spermatogenesis' were significantly-enriched GO terms during spermatogenic progression. Interestingly, >4600 genes were differentially expressed (2525 up and 2101 down) during the transition from spermatocytes to round spermatids, displaying the most dramatic transcriptomic change.

Germline expression dynamics of transposable elements, lncRNAs and *XIST*

Single-cell datasets allow a refined examination of transposable elements (TE) and long non-coding RNAs (lncRNAs) (Supplementary information, Figs. S3a–c and S4a–b). Pseudotime and clustering analyses identified dynamic TE and lncRNA programs during spermatogenesis. Notably, LTR12C/D/E and the active TEs SVA_D and AluYa5 showed high expression during early spermatogenic stages. By contrast, LTR10A and LTR40c expression peaks during late spermatogonial or post-meiotic stages; and satellite and multiple MLT-family TEs at spermatid and sperm stages (Supplementary information, Fig. S3d). Moreover, we explored X chromosome inactivation and meiotic sex chromosome inactivation (MSCI) during spermatogenesis.[21,22] Notably, we observed *XIST* expression during spermatogonial stages (Supplementary information, Fig. S4c and d), and unexpectedly selective attenuation of genes that are near the X-inactivation center during spermatogonial stages (Supplementary information, Fig. S4e and f), suggesting a role for *XIST*-mediated silencing in this process. Overall, our datasets and analyses provide a comprehensive resource to study TEs and lncRNA expression dynamics during male germline development.



Analysis of meiotic cells reveals dynamic transcriptional programs and key factors during meiotic transition

Next, we singled out Clusters 3–4 (Fig. 1b) and performed re-clustering of the meiotic cells, which revealed five sub-clusters (Fig. 4a). Using known markers[23,24] (Fig. 4b), we assigned pre-leptotene, leptotene, zygotene/early pachytene, late pachytene, and diplotene cell types, consistent with the pseudotime developmental order. Secondary spermatocytes were under-represented consistent with their rapid progression into round spermatids.[25] Gene clustering analysis (Fig. 4c) identified five

Fig. 1 Single cell transcriptome profiling from healthy adult whole testes. **a** Schematic illustration of the experimental workflow. **b** tSNE and clustering analysis of combined single-cell transcriptome data from human testes ($n = 6490$). Each dot represents a single cell and is colored according to its cluster identity as indicated on the figure key. The 13 cluster identities were assigned based on marker gene expression shown in Fig. 1d and Supplementary information, Fig. S2. tSNE: t-distributed stochastic neighbor embedding. Note: the 40 µm filtering step likely limits capture of the large Sertoli cells. **c** tSNE plot of single cell transcriptome data with cells colored based on their donors of origin, as indicated on the figure key. **d** Expression patterns of selected markers projected on the tSNE plot. Red indicates high expression and gray indicates low or no expression, as shown on the figure key. Top row represents somatic/niche cell markers; bottom three rows are representative germ cell markers. For each cell type, we show one marker in the main figures and a gallery in supplementary information, Fig. S2

distinct molecular signatures (4594 genes) suggesting striking transcriptional changes during meiotic entry and exit, but only gradual changes during meiosis. Notably, several RNA binding proteins were upregulated at zygotene/early pachytene, and some HOX genes (e.g., *HOXB4* and *HOXC6*) showed late pachytene-specific expression (Fig. 4c, d).

We observed a dynamic expression pattern of DMRT and SOX family members (Fig. 4e): Consistent with their role in meiotic entry inhibition in mice,[26,27] *DMRT1* and *SOX4* were only expressed in pre-leptotene cells. Although the function of *DMRTC2* and *DMRT3* is largely unknown, *Sox30* knockout causes murine germ cell development arrest at round spermatid stage, and reduces expression of *Sox5*.[28,29] Overall, our data are consistent with findings in mice, but also provide evidence for candidate genes with human-specific functions during meiosis.

Identification of known and novel spermatogonial stem cell states

To further characterize the spermatogonial 'States', we re-clustered the early germ cells Clusters 1–2 (Fig. 5a). This analysis yielded five distinct clusters: while four showed high similarity to the clusters/states previously described (States 1–4 in ref.[10]) an additional state, hereafter termed 'State 0' was identified (Supplementary information, Table S4). None of the clusters/states (including State 0) consisted of cells derived from a particular donor or within a specific cell cycle phase (Supplementary information, Fig. S5a–d). Pseudotime analysis revealed a wave-like progression from State 0 to State 4, and clustering analyses defined gene expression signatures associated with each State (Fig. 5b; Supplementary information, Table S5). Notably, we observed a striking shift in transcriptional programs between State 1 and State 2, dominated by expression of cell cycle/proliferation genes (e.g., *MKI67*), suggesting that this transition represents a critical developmental node (see Discussion).

While both State 0 and State 1 cells co-express many key stem cell signaling factors and TFs (Fig. 5c, d), our analyses identify 490 genes that are either most highly expressed, or specifically expressed, in State 0 (e.g., *PIWIL4*, *EGR4, TSPAN33, PHGDH*, *PPP1R36*, *ICA1L* (Fig. 5e, f)). Known early SSC markers (*ID4*, *FGFR3, TCF3* and *UTF1*) were expressed in States 0 and 1, whereas known markers of differentiation (*KIT)* or proliferation (*MKI67)* were specifically expressed during or after State 2, suggesting State 0 and State 1 may represent two distinct quiescent SSC states. Interestingly, most State 0 cells displayed low expression of *ST3GAL2*, the enzyme that catalyzes the formation of SSEA4 (Supplementary information, Fig. S5g), suggesting that State 0 cells do not express this spermatogonial cell surface marker.

RNA velocity analysis and chromatin profiling suggest SSC plasticity

Next, we applied RNA 'velocity' analysis, a computational approach that utilizes nascent transcription in scRNA-seq datasets to infer developmental trajectories.[11] Here, the ratio of unspliced to spliced reads for each transcript is used as a proxy measurement of new transcription. By comparison with steady state (spliced) transcripts in the other cells, a velocity vector representing the future transcriptional state of each individual cell can be defined. Within each cell of the tSNE plot (Fig. 6a), the amplitude and direction of the vector reflects a transcriptional trajectory. This analysis revealed two unexpected features. First, within the State 0 cluster, we observed two sub-populations: one proximal to State 1, bearing long vectors, indicating an apparent progression towards State 1—and a second sub-population lacking long vectors. This pattern suggests that the former cell sub-population is actively progressing/committing towards State 1, in response to specific developmental cues/signals, while the latter is not. Second, we also observed a sub-group of State 2 cells displaying long velocity vectors pointing back toward State 1. This forward and backward movement between States 0–1–2 raises the possibility that human spermatogonia display dynamic plasticity and metastable/uncommitted behaviors.

We then explored whether methylation or chromatin status may provide further evidence of plasticity. First, although > 8,000 genes show differential expression during spermatogonial development and spermiogenesis (Fig. 3b), we observed almost no differences in the DNAme profiles of SSEA4-enriched human SSCs and that of mature human sperm—which parallels a similar finding in mice.[10,30] Thus, no DNAme barrier exists that might deter spermatogonial de-differentiation. Next, we profiled open chromatin from differentiating/committed spermatogonia (enriched using c-KIT), and compared it to profiles of self-renewing SSCs (enriched with SSEA4) (Fig. 6b). Notably, the open chromatin maps of c-KIT or SSEA4-enriched spermatogonia were highly similar ($r > 0.83$), and their nearest peak summits were typically overlapping (within a distance of ~120 bp), indicating that very few changes in the open chromatin landscape occurred during the commitment of undifferentiated SSEA4$^+$ SSCs into differentiating c-KIT$^+$ spermatogonia, in spite of the activation and repression of hundreds of genes. Taken together, the scRNA-seq analysis, the RNA velocity trends and evidence derived from the analysis of open chromatin landscape and DNAme are consistent with the proposal that SSCs follow a developmental progression involving 5 sequential transcriptional states characterized by a 'flat' chromatin/DNAme landscape, strongly suggestive of dynamic behaviors of spermatogonial cells (Fig. 6c).

The adult State 0 is most similar to infant germ cells

To assess whether State 0 may represent the earliest/naïve SSC in adults, we profiled testicular cells from infants (12–13 months old). After QC filtering, we obtained ~1300 single cells and assigned cell identities based on known markers. This analysis identified four somatic cell types, i.e., Sertoli, Leydig, endothelial and macrophage (Fig. 7a, b; Supplementary information, Fig. S7), and one tight cluster of 37 germ cells. Comparison to adult SSC States via tSNE and pseudotime analysis, positioned the infant germ cells adjacent to adult State 0, at the 'beginning' of the developmental trajectory (Fig. 7c). Moreover, most State 0 markers were highly expressed in infant germ cells (Fig. 7d). A small set of factors (e.g., *TBX3*, *HOXA3*) showed specific expression in infant spermatogonia,[31] which may specify their germline identity. In addition, transcriptomic data of infant somatic cells should provide a useful resource for future analysis.

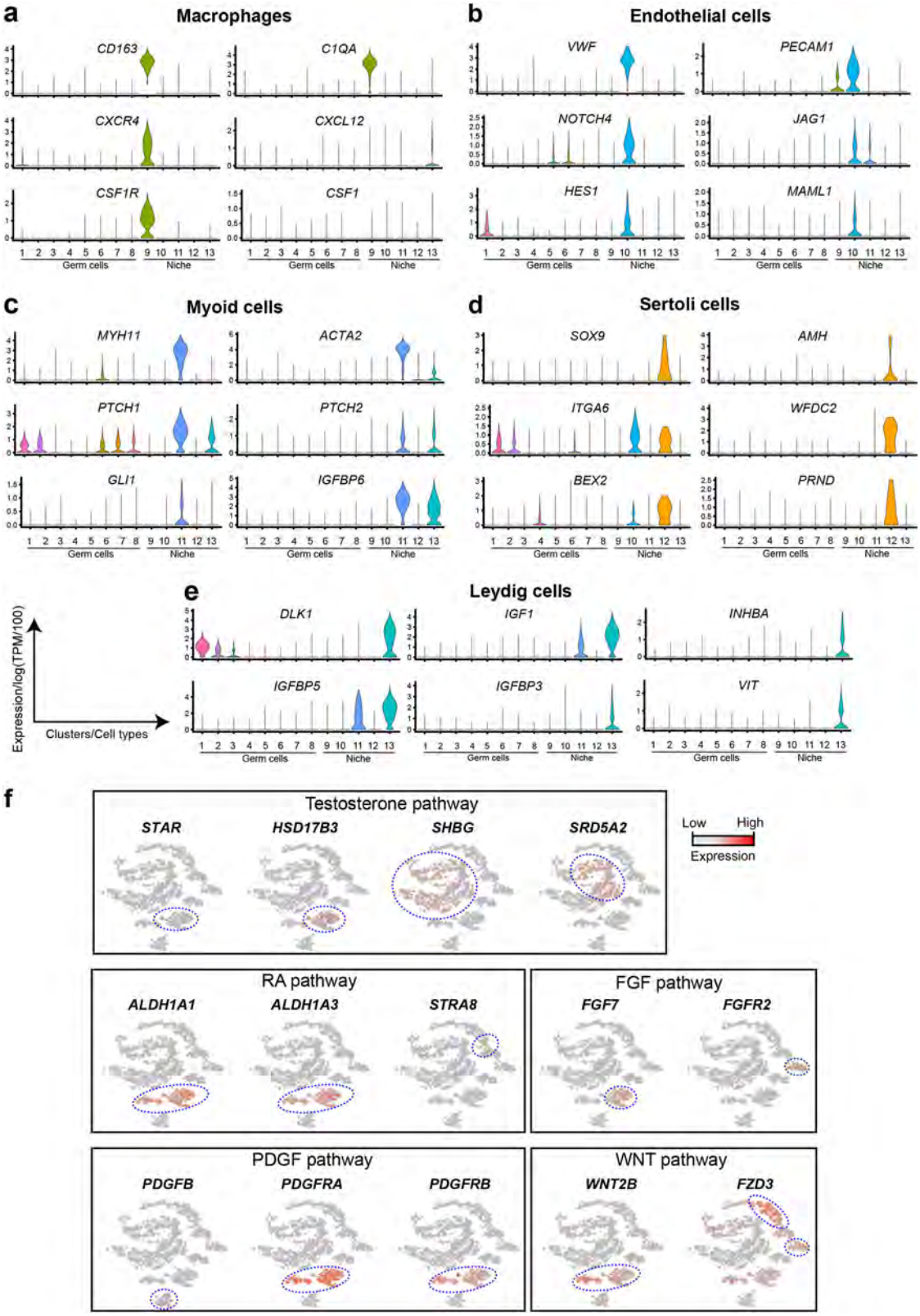


Fig. 2 Expression patterns of representative genes marking niche cells, and Niche-Germline interactions. **a** Expression patterns (violin plot) of macrophage-specific genes across the 13 different Clusters (Clusters 1–8 = germ cells; Clusters 9–13 = Niche/somatic cells). **b** Expression patterns (violin plot) of endothelial cell-specific genes across the different clusters. **c** Expression patterns (violin plot) of myoid cell-specific genes across the different clusters. **d** Expression patterns (violin plot) of Sertoli cell specific genes across different clusters. **e** Expression patterns (violin plot) of Leydig cell specific genes across different clusters. **f** Relative expression levels of representative genes from different key signaling pathways projected onto the tSNE plot from Fig. 1b. Stage-specific expression is highlighted by blue dotted circles

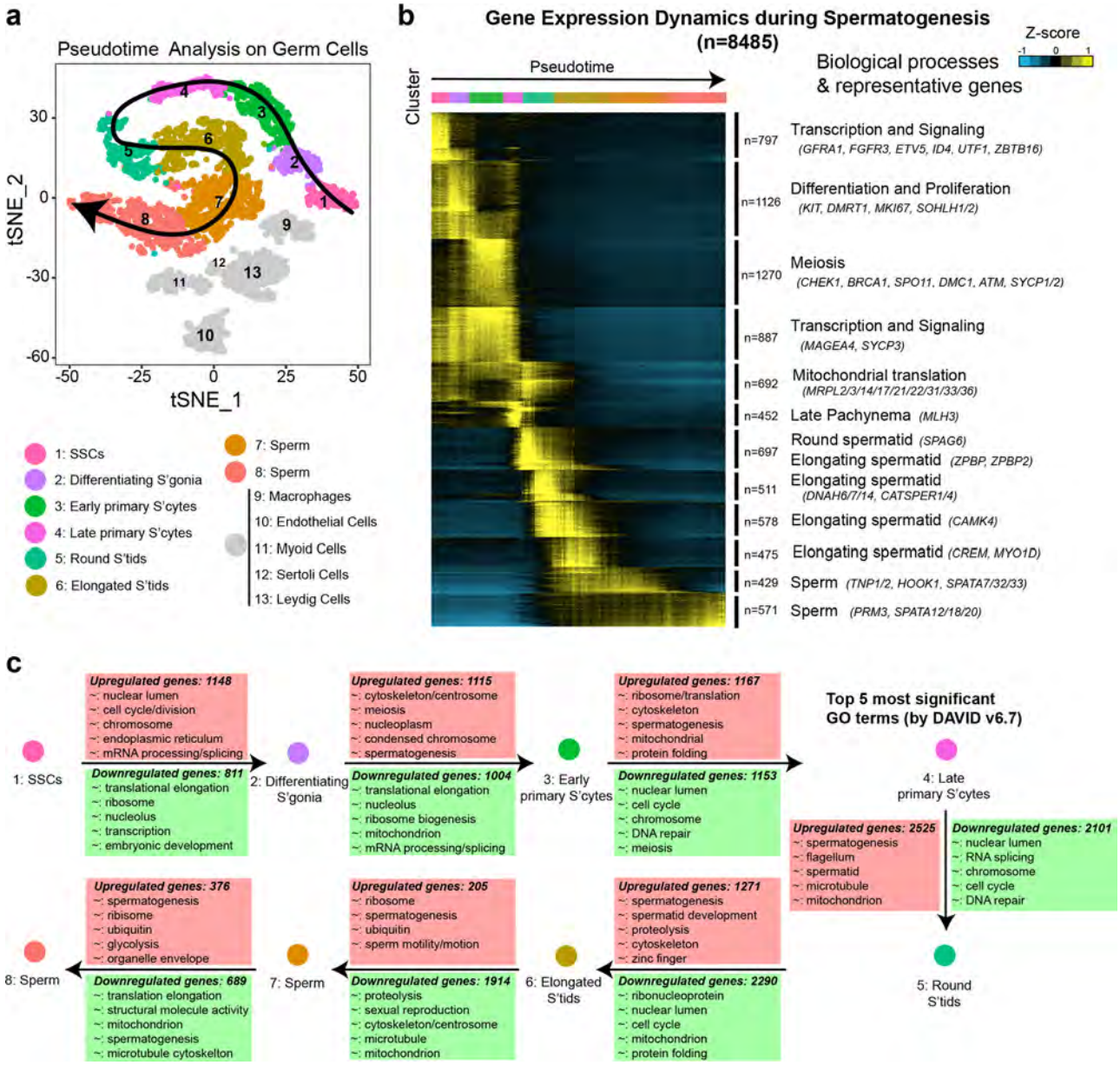


Fig. 3 Gene expression dynamics during spermatogenesis. **a** Pseudotime analysis on germ cells (Clusters 1–8). Cluster 1 represents the start of pseudotime, with Cluster 8 at the end. **b** K-means clustering of genes exhibiting differential expression ($n = 8485$) across germ cell populations. Note: each row represents a gene, and each column represents a single cell, with columns/cells placed in pseudotime order as defined in Fig. 3a and depicted by a thick colored line (top, color code as in Fig. 3a). Differential gene expression levels utilize a *Z* score, which represents the variance from the mean, as defined on the color key in the right top corner. **c** Differentially-expressed genes and associated GO terms (using DAVID v6.7) characteristic of germ cell developmental transitions, based on the 8 germ cell Clusters represented in Fig. 2a. The 5 most significant up-regulated GO terms are annotated in pink boxes, and down-regulated GO terms in green boxes. GO: gene ontology

Validation of SSC States via sequential mRNA fluorescence in situ hybridization

For validation, we performed sequential single molecule RNA fluorescence in situ hybridization (seqFISH)[32] on 5 key genes (Fig. 8a; Supplementary information, Table S6). In cells expressing moderate/high *TCF3* (State 0 and State 1), only 27% (9/33) displayed high expression of *PIWIL4* (State 0 marker) and *ETV5*/*L1TD1* (State 1), while the remainder expressed either *PIWIL4* (27%) or *ETV5*/*L1TD1* (39%). This yielded significant non-overlap between State 0 and State 1 markers (hypergeometric test; $p = 0.03$), consistent with scRNA-seq data.

The State 0 marker TSPAN33 shows co-localization with high FGFR3

We next sought to enrich State 0 cells using cell surface markers, and assess their SSEA4 status. Here, we used TSPAN-family receptors, which show a strong enrichment in State 0 (e.g., TSPAN33; Fig. 5f). Flow cytometry analysis of cells expressing high levels of the SSC marker FGFR3 (peaks in State 0, present in State 1; Fig. 5d), show that FGFR3^{high} cells have high TSPAN33 but low SSEA4 (Fig. 8b)—characterizing State 0 as FGFR3^{high} TSPAN33^{high} SSEA4^{low}.

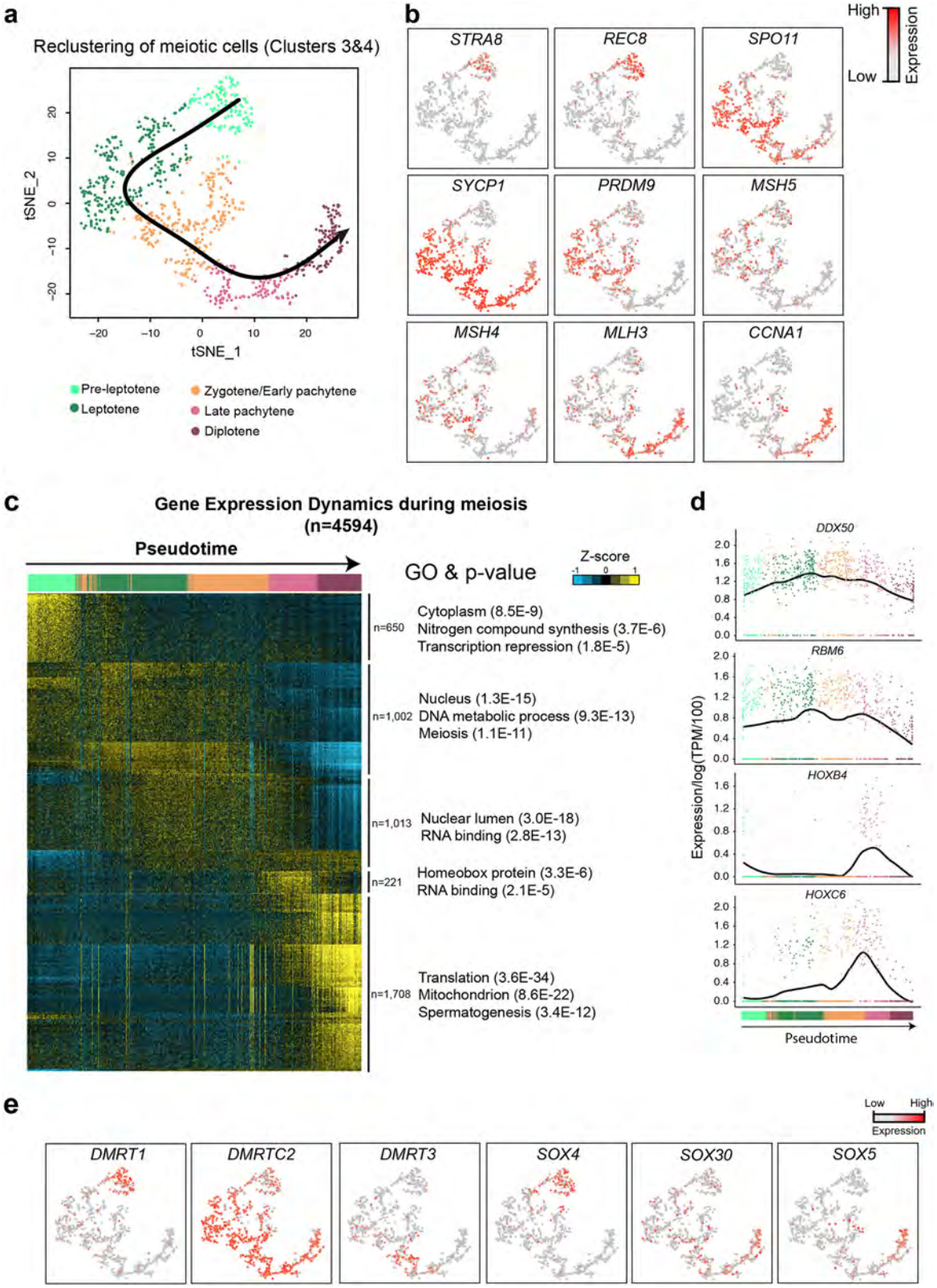
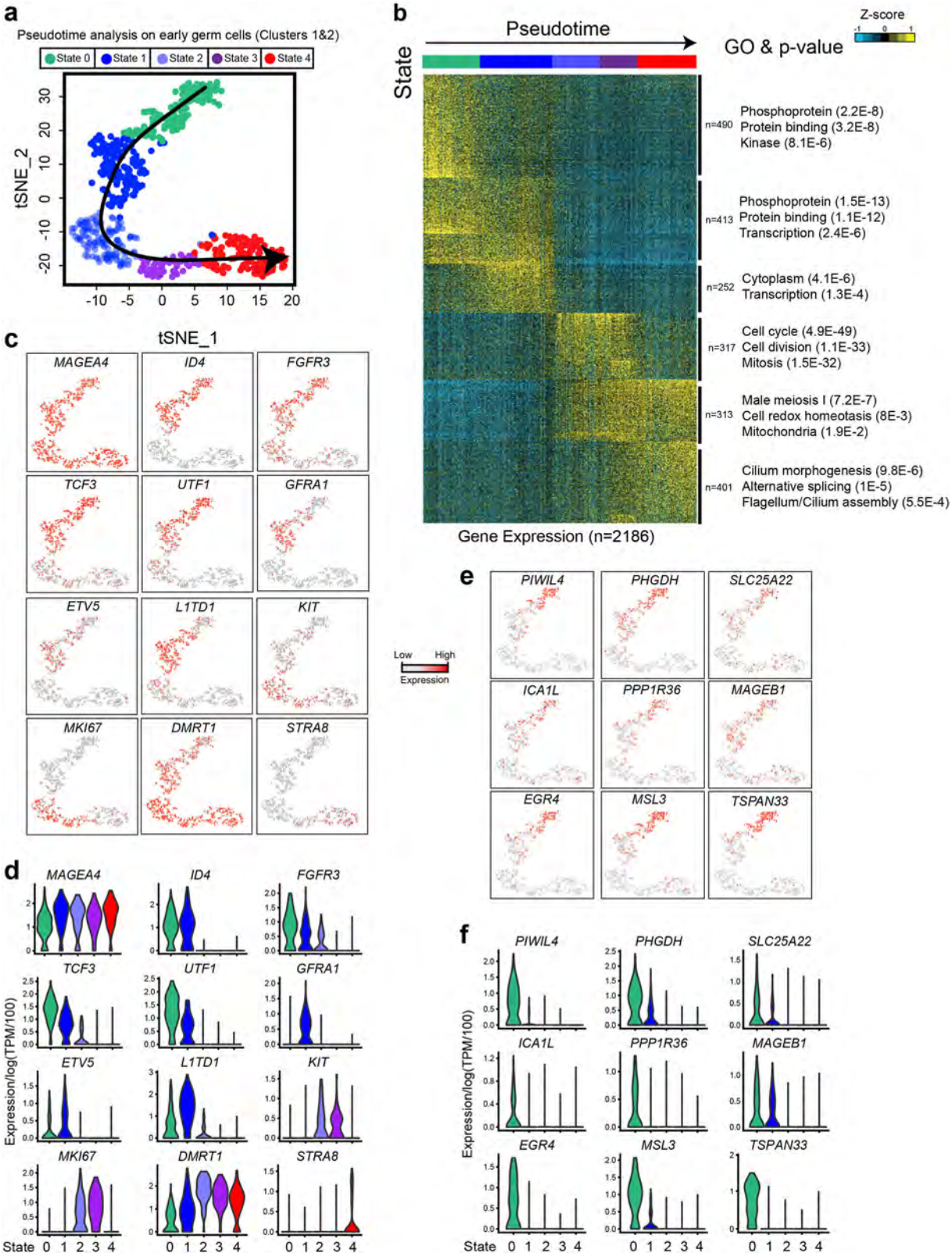


Fig. 4 Gene Expression Dynamics during Meiosis. **a** Focused analysis (tSNE, clustering and pseudotime ordering) of the cells from Clusters 3 and 4 (from Figs. 1b and 3a) reveals developmental progression during meiosis I. **b** Expression patterns of known meiotic markers projected onto the tSNE plot. Red indicates high expression and gray indicates low or no expression, key on figure. **c** K-means clustering of genes exhibiting differential expression ($n = 4594$) during meiosis I. Note: each row represents a gene, and each column represents a single cell, with columns/cells placed in pseudotime order as defined in **a** and depicted by a thick colored line (top, color code as in **a**). Gene expression levels utilize a Z score, which depicts variance from the mean, as defined on the color key in the right top corner. **d** Expression levels of representative genes during meiosis progression. x-axis represents pseudotime (as defined on **a**), and *y*-axis represents gene expression levels. **e** Expression patterns of key transcription factors during meiosis, with their expression projected onto the tSNE plot (**a**)



In situ observation of early SSC States via protein immunofluorescence

We characterized the protein expression of early spermatogonial markers (UTF1, GFRA1, FGFR3, and TCF3)[33] that exhibit differential expression across early States (0–1–2) by performing triple immunofluorescence (IF) staining. *UTF1* expression peaked at State 0, and only partially overlaped with *GFRA1*, which peaks in State 1 (Fig. 5d). We observed that ~66% of cells located at the periphery of the seminiferous tubules expressed either UTF1 and/or GFRA1, as previously reported.[34] Here, we distinguished two

Fig. 5 Identification of five discrete transcriptional states for SSCs. **a** Focused analysis (tSNE, clustering and pseudotime ordering) of Clusters 1 and 2 (from Fig. 1b and 3a) reveals five discrete cellular states (States 0 to 4) during SSC development. **b** K-means clustering ($k = 6$) of genes exhibiting differential gene expression in States 0–4. Six gene clusters (termed S1-S6) were identified. Gene ontology associated with each gene block is shown on the right. Note: each row represents a gene, and each column represents a single cell, with columns/cells placed in pseudotime order (depicted by different colors on the top of the figure) as defined in **a**. Gene expression levels utilize a *Z* score, which depicts variance from the mean, as defined on the color key in the right top corner. **c** Relative expression levels of selected SSC markers projected on the tSNE plot represented in **a**. **d** Violin plots representing the expression levels of the selected markers shown in **c** in States 0–4 (*x*-axis). **e** Relative expression levels of selective State 0-specific markers projected on the tSNE plot represented in **a**. **f** Violin plots representing the expression levels of selective State 0-specific markers shown in **e** in States 0–4 (*x*-axis)

broad spermatogonial phenotypes, characterized by either UTF1^{high}/GFRA1^{low} or GFRA1^{high}/UTF1^{low}, that recapitulate the temporal progression from State 0 to State 1. Furthermore, the spermatogonial markers FGFR3 and TCF3 were expressed in a subset of cells expressing either UTF1 and/or GFRA1, but were never observed on their own (Fig. 8c). One simple explanation is that RNAs strongly expressed in State 0 (e.g., *UTF1*) produce a protein with a longer half-life than its RNA, and therefore the protein persists into the State 1 *GFRA1*-expressing cells—providing heterogeneity. As expected, the proliferating marker MKI67 showed almost no overlap with GFRA1 or UTF1 (Fig. 8c and data not shown), suggesting that transition to the proliferative State 2 occurs with the loss of SSC markers, which is consistent with our computational analyses (Figs. 5 and 6c).

Finally, we further characterized the expression of genes specific to State 0 or State 1. Using the Human Protein Atlas resource (http://www.proteinatlas.org/) to review the pattern of 490 State 0 genes,[35] 16 candidate early SSC markers which displayed specific expression in cells located along the periphery of the seminiferous tubule were chosen. Triple IF staining with UTF1 and GFRA1 showed that each of these 16 markers is expressed in GFRA1-positive and/or UTF1-positive cells. Antibodies to PHGDH and PPP1R36 display strong staining in UTF1^{high} (State 0) cells, as predicted by the scRNA-seq analysis (Fig. 8d). Other State 0 markers showed expression in early spermatogonia (Fig. 8d; Supplementary information, Fig. S8b), an expected result for instances where the protein half-life is greater than that of the transcripts. In several instances, including PIWIL4 and MAGEB1, we observed differences in staining intensity and/or sub-nuclear localization in cells co-expressing either UTF1^{high} or GFRA1^{high} marker (Fig. 8d; Supplementary information, Fig. S8b). These observations are consistent with the proposal that although State 0 and State 1 define transcriptionally discrete states, they likely represent metastable/heterogeneous cellular phenotypes that afford SSCs the ability to adapt to a dynamic niche environment and ensure homeostatic regulation within the testis.

DISCUSSION

Human adult spermatogenesis is a complex process, and a full understanding will involve the integration of multiple data types—including those from rodents, where genetic tools and SSC culturing systems are already available—to determine both shared and unique mechanism in mice and men. Here, we aimed to provide foundational scRNA-seq data of all cells contained within the normal human young adult testis, complemented by computational analysis and validation studies—to offer new insights into the regulation of male gametogenesis in humans.

Signaling features in the human testis niche

A major area of current interest involves communication between the niche and germline, and how these interactions mediate the changes observed during puberty and ageing. Here, we report new data that reveal potential differences between mice and men, and changes during development, for future functional investigation. First, for CXCR4-CXCL12 signaling (for SSC homing to the niche), *CXCL12* was primarily expressed at the RNA level by Leydig cells in humans rather than by Sertoli cells in mice. Furthermore, mouse SSCs express *Csf1r* for Csf1 response, while human *CSF1R* expression appears to be specific to macrophages. Moreover, we identified novel markers for niche cells, like *WFDC2* and *PRND* for Sertoli cells, and *VIT* for Leydig cells. Beyond these examples, our datasets provide a resource for additional analyses of niche cells and niche-germline communication. Furthermore, our transcriptome data on infant testes concur that major differences exist between infant somatic cells and their counterparts in the adult (Supplementary information, Fig. S7), and exploration of the mechanisms mediating these changes prior to and during puberty will be an important future research focus.

Computational analysis of spermatogonial development and spermatogenesis

Here, we demonstrate that the dynamic and temporal trajectory of germline development could be reconstructed using scRNA-seq data and dedicated computational/analytical tools. The distribution of cell types for all three donors largely overlap during SSC stages and the early and mid stages of gametogenesis, establishing overall consistency between donors. However, modest differences were observed in late stages. Here, we note that the RNA content of maturing/mature sperm is both low and of poor integrity, a known and conserved property in mammals, conferring low number of genes/UMIs (unique molecular identifiers) in the maturing sperm datasets (Supplementary information, Fig. S1d). Therefore, although two distinct clusters of sperm were computationally obtained (Clusters 7 and 8), an alternative interpretation is that these sperm populations are similar, and simply contain different levels of RNA degradation/removal during this unique, transcriptionally inactive stage.

We also provide several new insights into spermatogonial development, most importantly the identification of a novel and early quiescent state, termed State 0. This interpretation is based on the properties/categories of the 490 differentially-regulated genes and the high similarities to the transcriptomic profile of infant germ cells. Moreover, State 0 cells (in contrast to State 1 cells) show low expression of the cell surface marker SSEA4, likely explaining why they escaped characterization in our prior study, which relied on SSEA4-mediated enrichment of testicular cells. Thus, we suggest that State 0 cells represent the undifferentiated and quiescent 'reserve' stem cell pool in the adult germline that is largely maintained from infants to adults.

Our data also call for a re-examination of prior work on human spermatogonial development. GFRA1 marks early undifferentiated spermatogonia,[36] but exhibits a heterogeneous expression in human[34] and mouse[37] SSCs. Recently, using immunostaining, DiPersio et al.,[34] proposed that 'early' (slow proliferating) GFRA1^{high} UTF1^{-} SSCs progress to GFRA1^{low} UTF1^{+}, before committing to differentiation (c-KIT^{+}). However, our scRNA-seq data and pseudotime analysis suggest an alternative sequence of events—where *UTF1* precedes the expression of *GFRA1*—which, interestingly aligns with recent work in the mouse.[37] Beyond these two markers, we additionally provided data on hundreds of candidate spermatogonial markers (and validated 16 new protein markers) that define State 0 and State 1 SSCs in humans.

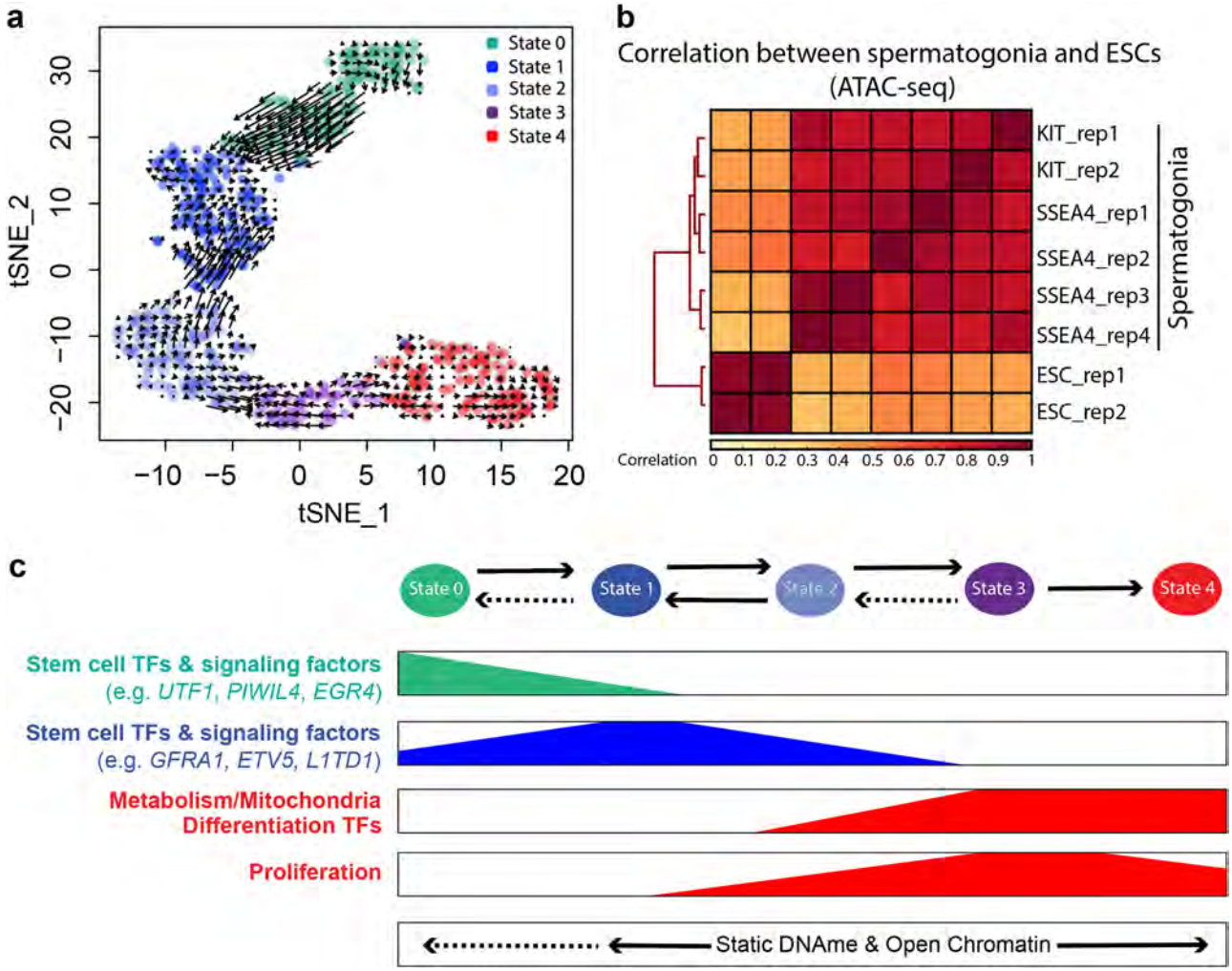


Fig. 6 Computational and molecular examination of spermatogonial plasticity. **a** Visualization of the RNA velocity analysis results on the tSNE plot of SSCs (see main text for details on vectors). **b** Heatmap and hierarchical clustering of ATAC-seq data from KIT$^+$ spermatogonia (two replicates), SSEA4$^+$ SSCs (four replicates) and ESCs (two replicates). Note: SSEA4$^+$ SSC and ESC data are from ref.[10] **c** Schematic summarizing the combinatorial gene expression programs and cellular events promoting five distinct SSC states (States 0–4) and depiction of the proposed spermatogonial dynamics/kinetics and behavioral plasticity of States with main cellular events and molecular pathways. Dotted arrows are speculative

Computational and molecular evidence for plasticity within early spermatogonia

Plasticity and stochastic behaviors within spermatogonial stem cell populations have been reported in several species including *Drosophila* and mice, in which differentiating spermatogonia can dedifferentiate and regain their early self-renewing properties.[38–41] Our work provides two lines of evidence consistent with developmental plasticity in early human spermatogonia. First, RNA velocity analysis singled out a population of State 2 spermatogonia tending to 'dedifferentiate' into State 1-like cells. Notably, the transition from State 1 to State 2 is marked by an upregulation of proliferative markers—and thus might be a critical node for homeostasis by which developmental plasticity allows to balance early versus differentiating SSC populations.[42] Second, we observe very limited changes in open chromatin and DNAme along the developmental trajectory of spermatogonia, which may enable transcriptional plasticity to take place by lowering epigenetic barriers to transcriptional changes, and dedifferentiation. Overall, we propose a spermatogonial developmental trajectory that involves 5 sequential transcriptional States, which generate moderately heterogeneous (metastable) proteomes, to enable State transitions and maintain a constant SSC pool; properties which might be essential to maintain life-long fertility, and critical for the germline replenishment in case of damage.

A resource for future investigation of spermatogenesis

Recently, scRNA-seq has emerged as a highly useful approach for the study of human and mouse spermatogenesis.[29,43,44] Here, our data reveals > 8000 genes that undergo significant differential regulation during male gametogenesis, and our results align well with a very recent scRNA-seq study of gametogenesis.[44] Beyond coding genes, our work has uniquely explored transposable elements (TE) and long non-coding RNAs—which are shown to display remarkable stage-specific expression. Of particular interest are the expression of active LTR12C/D/E, SVA_D and AluYa5 elements during early spermatogenesis, LTR10A and LTR40c elements during spermatogonial stages. We also described *XIST* expression during spermatogonial stages—which we show coincides with the selective attenuation of genes near the XIC (Supplementary information, Fig. S4), suggesting an unexpected role for *XIST* spreading and silencing in this process. Taken together, our datasets and analyses provide a comprehensive resource for the study of both niche cells and germline cells—including coding genes, TEs and lncRNA expression dynamics—

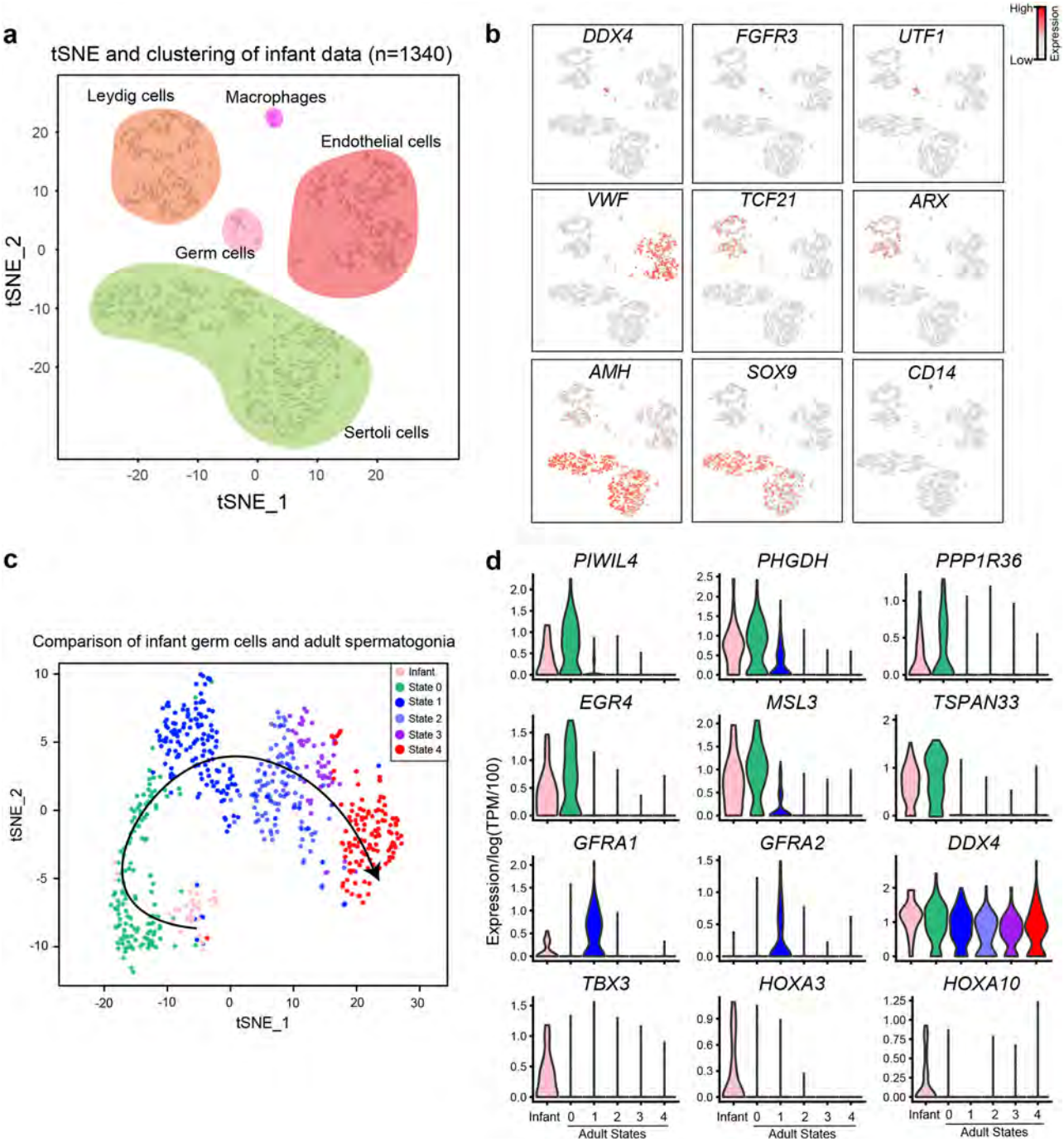


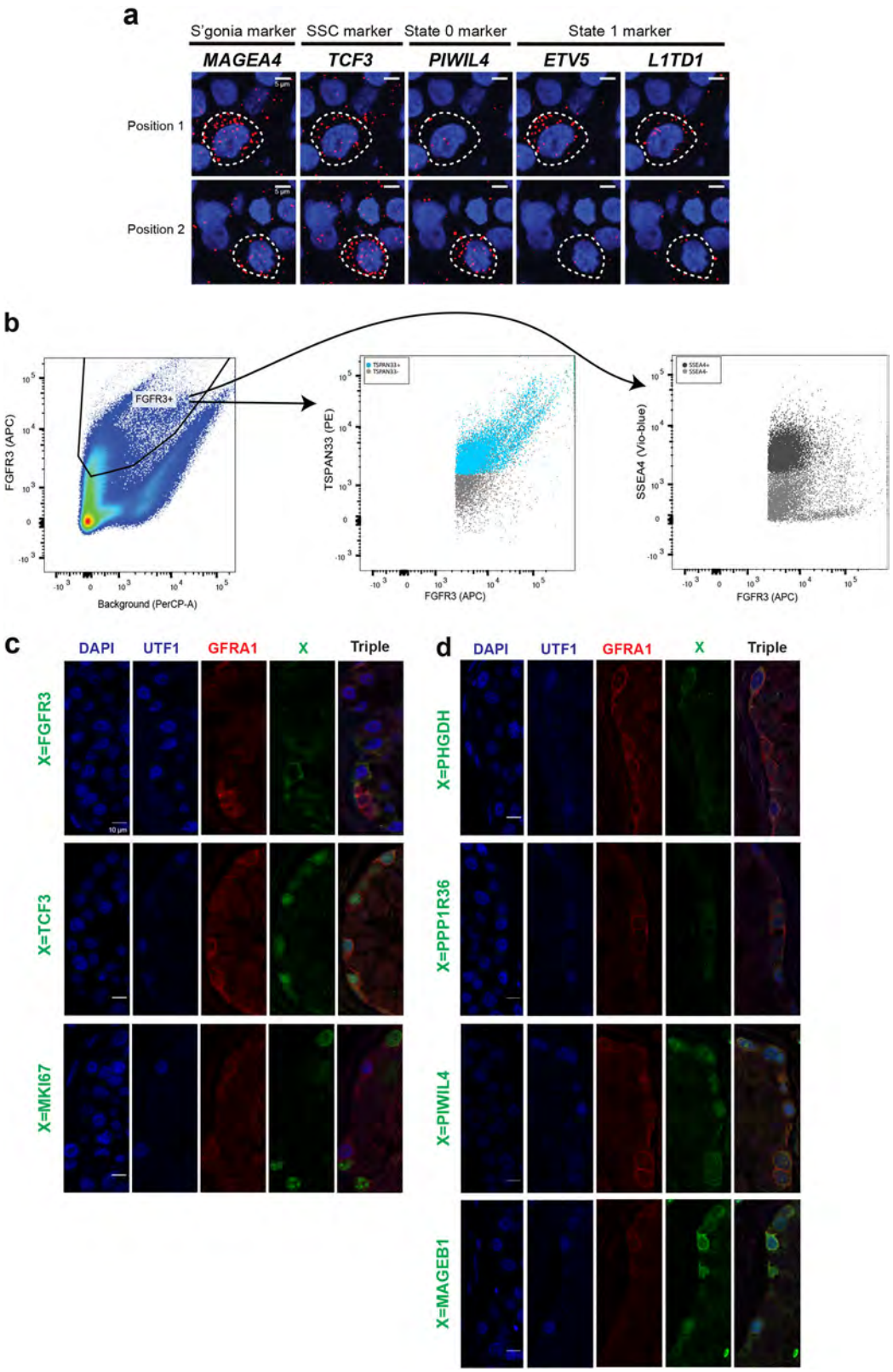
Fig. 7 Single cell RNA profiling from infant testis and comparison to adult scRNA-seq data. **a** tSNE and clustering analysis of single-cell transcriptome from infant testis ($n = 1340$). **b** Expression patterns of representative markers to help assign cell identities. **c** tSNE and pseudotime analysis of infant germ cells and adult spermatogonia. **d** Expression patterns (violin plot) of representative genes in infant germ cells and adult spermatogonia

which will also serve as a useful reference dataset for comparisons to younger and older men, infertile men, and testicular cancer.

MATERIALS AND METHODS

Experimental model and subject details

Adult human testicular samples for scRNA-seq and immunostaining were from three healthy men (donor #1: 17 years old; donor #2: 24 years old; donor #3: 25 years old); sample for mRNA seqFISH was from a healthy man (donor #4: 23 years old). Infant testicular samples for scRNA-seq were from two infant donors (13 months old). All six samples were obtained through the University of Utah Andrology laboratory and Intermountain Donor Service. Those samples were removed from deceased individuals who consented to organ donation for transplantation and research. Sample used for ATAC-seq was obtained through the University of Utah Andrology laboratory consented for research (IRB approved protocol #00075836: understanding the transcriptional and epigenetic dynamics in human spermatogonial stem cell self-renewal, proliferation and differentiation).



Sample storage by cryopreservation

Once collected, the pair of whole testis samples was transported to the research laboratory on ice in saline or Hank's Balanced Salt Solution (HBSS; GIBCO cat # 24020117) and processed within 1 h of removal by surgery. Around 90% of each testis was divided into smaller portions (~0.5–1 g for each) using scissors and directly transferred into cryovials (Corning cat # 403659) in DMEM medium (Life Technologies cat # 11995073) containing 10% DMSO (Sigma-Aldrich cat # D8779), 15% fetal bovine serum/FBS (Gibco cat # 10082147) and cryopreserved in a Mr. Frosty Freezing container

Fig. 8 RNA and protein staining to validate state 0. **a** Sequential RNA FISH of SSC markers on tubular sections. Two SSCs are highlighted as representative examples. Blue is the DAPI signal; red detects RNA FISH signal for gene as indicated on the figure. White dashed line circles represent the cell boundaries. Scale bar: 5 μm. **b** Expression patterns of FGFR3 (marks State 0 and 1), TSPAN33 (marks State 0) and SSEA4 (marks State 1) in testicular cells via flow cytometry. Human single testicular cells were used for staining with the markers indicated; non-stained cells were used as control for gating purposes. Left: $FGFR3^+$ cells were identified and used for analysis in the middle and right panels. Middle: co-staining pattern between FGFR3 (*x*-axis) and TSPAN33 (*y*-axis), with blue as $TSPAN33^+$ and red and gray as $TSPAN33^-$. Right: co-staining pattern between FGFR3 (*x*-axis) and SSEA4 (*y*-axis), with black as $SSEA4^+$ and gray as $SSEA4^-$. **c** Immunolocalization of UTF1 (State 0 marker, in blue), GFRA1 (State 1 marker, in red) and FGFR3 or TCF3 or MKI67 (in green). Each combination of single or triple antigen (named in green on the left side) is represented by 5 panels. Scale bar: 10 μm. **d** Immunolocalization of UTF1 (State 0 marker, in blue), GFRA1 (State 1 marker, in red) and 4 new candidate markers (in green). Each antigen (named in green on the left side) is represented by 5 panels. Scale bar: 10 μm

(Thermo Fisher Scientific cat # 5100–0001) ensuring a slow controlled freezing rate at −80 °C for overnight. Cryovials were then transferred to liquid nitrogen for long-term storage.

Sample fixation for immunostainings

Around 10% of the remaining testis tissues were incubated in 40 mL of 1× PBS containing 4% paraformaldehyde/PFA (Thermo Fisher Scientific cat # 28908) overnight at 4 °C with agitation on a rotor (60 rpm). Fixed samples were then washed three times in cold PBS and stored in PBS at 4 °C until processing for immunostaining.

Human adult testis sample preparation for single cell RNA sequencing

For each single cell sequencing experiment (technical replicate for one donor), ~5 cryovials were thawed in ~3 min. Tissues were washed twice in 1× PBS, and subjected to a standard two-step digestion procedure, as described previously.[10] Briefly, testicular tissues were digested with collagenase type IV (Sigma Aldrich cat # C5138–500MG) for 5 min at 37 °C with gentle agitation (250 rpm), then shaken vigorously and incubated for another 3 min. The tubules were sedimented by centrifugation at 200× *g* for 5 min and washed with HBSS before digestion with 4.5 mL 0.25% trypsin/ethylenediaminetetraacetic acid (EDTA; Invitrogen cat # 25300054) and 4 kU DNase I (Sigma Aldrich cat # D4527–500ku). The suspension was triturated vigorously three to five times and incubated at 37 °C for 5 min. The process was repeated in 5 min increments for up to 15 min total. The digestion was stopped by adding 10% FBS (Gibco cat # 10082147). Single testicular cells were obtained by filtering through strainers with mesh size 70 μm (Fisher Scientific cat # 08–771–2) and 40 μm (Fisher Scientific cat # 08–771–1). The cells were pelleted by centrifugation at 600× *g* for 15 min, and washed twice with 1× PBS. Cell number was measured using hemocytometer, and cells were then re-suspended in 1× PBS + 0.4% BSA (Thermo Fisher Scientific cat # AM2616) at the concentration of ~1000 cells/uL, ready for single cell sequencing.

Human infant testis sample preparation for single cell RNA sequencing

We performed two technical replicates for the infant donor. A quarter of the testis was thawed in ~5 min. Tissues were washed twice in 1× PBS, and minced into small pieces for better digestion outcome. Tissues were then treated with trypsin EDTA for ~25 min at 37 °C. The digestion was then stopped by adding 10% FBS (Gibco cat # 10082147). Single testicular cells were obtained by filtering through strainers with mesh size 70 μm (Fisher Scientific cat # 08–771–2) and 40 μm (Fisher Scientific cat # 08–771–1). The cells were pelleted by centrifugation at 600× *g* for 15 min, and washed twice with 1× PBS. Cell number was measured using hemocytometer, and cells were then re-suspended in 1× PBS + 0.4% BSA (Thermo Fisher Scientific cat # AM2616) at the concentration of ~1000 cells/uL, ready for single cell sequencing.

Single cell RNA-seq performance, library preparation and sequencing

scRNA-Seq was performed using the 10× Genomics system. Briefly, each experiment captured ~1500 single cells, in order to obtain ~0.8% multiplex rate. Cells were diluted following manufacturer recommendations, and mixed with 33.8 μL of total mixed buffer before being loaded into 10× Chromium Controller using Chromium Single Cell 3′ v2 reagents. Each sequencing library was prepared following the manufacturer's instructions, with 13 cycles used for cDNA amplification. Then ~100 ng of cDNA were used for library amplification by 12 cycles. The resulting libraries were then sequenced on a 26 × 100 cycle paired-end run on an Illumina HiSeq 2500 instrument.

Process of single cell RNA-seq data

Raw sequencing data were demultiplexed using the mkfastq application (Cell Ranger v1.2.1). Three types of fastq files were generated: I1 contains 8 bp sample index; R1 contains 26 bp (10 bp cell-BC + 16 bp UMI) index and R2 contains 100 bp cDNA sequence. Fastq files were then run with the cellranger count application (Cell Ranger v1.2.1) using default settings, to perform alignment (using STAR v2.5.4a), filtering and cellular barcode and UMI counting. The UMI count tables of each cellular barcode were used for further analysis.

Sequential RNA florescence in situ hybridization

Non-barcoded seqFISH (sequential FISH) probes were designed by targeting consensus of all constitutive exons (Supplementary information, Table S6) present in the masked hg38 human genome with 35-nucleotide (nt). All probes were blasted against the human transcriptome, and expected copy numbers of off-target probe hits were calculated using predicted RNA counts from RNA-seq dataset.[10] Probes were then attached with one of the DNA hybridization chain reaction (HCR) initiator sequences (B1, B2, B3, B4 or B5) at 5′ end with 4-nt 'ATAT' space in between. Initiator sequences were specific to genes in each round of hybridizations.

Non-barcoded seqFISH was performed by following the previous protocol.[32] Human testis tissues were first perfused with RNase-free PBS, and embedded in 30% RNase free sucrose (VWR cat # 97061–430). After the tissues sank, they were frozen using a dry ice/isopropanol bath in OCT media and stored at −80 °C. 15-μm sections were cut using cryotome and immediately placed on an aminosilane modified coverslip. The generated human testis sections mounted to coverslips (Thermo Scientific cat # 152450) were permeabilized at 4 °C in 70% EtOH for 12–18 h. Tissue sections were further permeabilized by adding RNase-free 8% SDS (10% SDS Ambion cat # AM9822) for 20 min. Samples were rinsed with 70% EtOH to remove SDS, and air-dried. The hybridization chambers (Grace Bio-Labs cat # 621505) were adhered around the tissue sections. Then samples were washed once with 2× SSC (Invitrogen cat # 15557–036) diluted in Ultrapure water (Invitrogen cat # 10977–015), and hybridized with 2.5 nM probes per incubation for overnight at 37 °C in Hybridization Buffer (50%

HB: 2× SSC, 50% Formamide (v/v) (Ambion cat # AM9344), 10% Dextran Sulfate (w/v) (Sigma cat # D8906), in Ultrapure water). Samples were washed in 50% Wash Buffer (50% WB: 2× SSC, 50% Formamide (v/v), 0.1% Triton-X 100 (Sigma X-100) in Ultrapure water) for 30 min at room temperature. While washing, aliquoted HCR hairpins (Molecular Instruments Inc.) were heated to 95 °C for 1.5 min and allowed to cool to room temperature for 30 min in the dark. The HCR hairpins were diluted to a concentration of 120 nM per hairpin in amplification buffer (2× SSC, 10% Dextran Sulfate (w/v)), and incubated with the samples for 45 min at room temperature in the dark. Following amplification, samples were washed in the 30% Wash Buffer (30% WB: 2× SSC, 30% Formamide (v/v), 0.1% Triton-X 100 in Ultrapure water) for 30 min to remove non-specifically bound hairpins. Samples were then stained with 5 μg/mL DAPI (Sigma cat # D8417) in 2× SSC and imaged as described below. After imaging, samples were digested with DNase I (10 units of DNase I, 1× buffer (Roche cat # 04716728001) in Ultrapure water) for 2 h at 37 °C on the microscope using custom made heat pad. Following DNase I treatment, the samples were washed with 30% WBT at 37 °C for 30 min, and hybridized with the following round of probe set for overnight with 2.5 nM probes per each in 50% HB at 37 °C on the microscope. Samples were then washed, amplified with HCR hairpins and imaged as before. The above steps were iterated at each hybridization round.

Following the last non-barcoded seqFISH, immunofluorescence was performed. The samples were washed with 1× PBS (Ambion cat # AM9624) for a few times, blocked with 5% BSA blocking solution (5% BSA (Gemini cat # 700–106 P), 1× PBS, and 0.3% Triton-X 100 in Ultrapure water), and then incubated at room temperature for 1 h. The primary antibody, Anti-beta Catenin (Abcam cat # ab6301), was 100-fold diluted in 1% BSA solution (1% BSA, 1× PBS, 0.3% Triton-X), incubated with the samples at room temperature for 3 h. The samples were then washed with 1× PBS for three times for 15 min each. The secondary antibody (anti-mouse Alexa Fluor 647 Invitrogen cat # A31571) was 500-fold diluted in 1% BSA buffer, and incubated with the samples at room temperature for 1 h. Samples were washed with 1× PBS for three times, stained with DAPI and imaged as described below.

Samples were imaged in an anti-bleaching buffer (14 mM Tris-HCl, pH = 8.0, 35 mM NaCl, 0.8% D-Glucose (Sigma cat # G7528), 100-fold diluted Catalase (Sigma cat # C3155), Pyranose oxidase with OD_{405} of 0.05 (Sigma cat # P4234), and saturated amount of Trolox (Sigma cat # 238813)) with the microscope (Leica, DMi8) equipped with a confocal scanner unit (Yokogawa CSU-W1), a sCMOS camera (Andor Zyla 4.2 Plus), 63× oil objective lens (Leica 1.40 NA), and a motorized stage (ASI MS2000). Lasers from CNI and filter sets from Semrock were used. Snapshots were acquired with 0.5 μm z steps across 15 μm with more than 10 positions per sample.

seqFISH signals were visualized using ImageJ software. Firstly, all images were aligned manually in xy and z by using DAPI channel signals. Each channel of HCR signals was background subtracted using ImageJ's subtract background function with rolling ball radius of 3 pixels. Images were applied by ImageJ's mean integral image filter with block radius of 3 pixels, and then contrasted. Each image was visualized with an overlay of DAPI signals of the first round of hybridization.

Immunostainings of testis tissues

The triple immunofluorescence stainings were performed on 5 μm formalin-fixed paraffin-embedded (FFPE) sections from portions of the testis from Donor 2 and 3 (24 and 25 years old respectively) following deparaffinisation, rehydratation and heat-mediated antigen retrieval in 10 mM sodium citrate buffer solution (pH 6). After treatment with Superblock (PBS) Blocking Buffer (Thermo Fisher Scientific, cat # 37515) for 30 min, individual sections were incubated overnight at 4 °C with a mix of three diluted antibodies (UTF1 (mouse monoclonal), GFRA1 (goat polyclonal) and a third rabbit polyclonal antibody (for antibodies details and dilutions, see the Table below). Antigen detection was conducted using the appropriate combination of Alexa Fluor 488, 555 and 647 secondary antibodies (all 1:500; Thermo Fisher Scientific, cat # A21202, cat # A21432, cat # A31573 respectively) for 2 h at room temperature in the dark. All primary/secondary antibodies were diluted in SignalBoost™ Immunoreaction Enhancer Kit (Calbiochem, cat # 407207–1KIT). After three washes in PBS, sections were incubated with DAPI (4′,6-Diamidine-2-phenylindole dihydrochloride) (Roche, cat # 10 236 276 001) to facilitate nuclear visualization (dilution: 1 μg/mL). Specificity of the antibody staining was confirmed using the same protocol but with omission of primary antibodies. Following multiple washes in PBS, slides were mounted using Vectashield mounting medium for fluorescence (Vector Laboratories, Inc., Burlingame, CA, cat # H-1000). Images were obtained under 25× objective (LD LCI PA 25×/0.8 DIC WD = 0.57 mm Imm Corr (UV)VIS-IR (Oil-Immersion) with a Zeiss LSM 780 Upright Multi-Photon Confocal Microscope and analyzed using Image J software (NIH, Bethesda, MD, USA).

Ab name	Antibody host	Ab ID	Dilution	Company
UTF1	Mouse monoclonal	MAB4337 (5G10.2)	(1:1000)	Millipore
GFRα1	Goat polyclonal	AF560	(1:25)	R&D systems
FGFR3	Rabbit mAb	C51F2 (#4574)	(1:50)	Cell signaling technology
Ki67	Rabbit polyclonal	ab16667	(1:200)	Abcam
MAGEB1	Rabbit polyclonal	HPA001193	(1:300)	Human protein atlas
PPP1R36	Rabbit polyclonal	HA077492	(1:2000)	Human protein atlas
CAMK2B	Rabbit polyclonal	HPA051783	(1:275)	Human protein atlas
PHGDH	Rabbit polyclonal	HPA24031	(1:500)	Human protein atlas
ERICH5	Rabbit polyclonal	HPA025070	(1:500)	Human protein atlas
PIWIL4	Rabbit polyclonal	HPA036588	(1:100)	Human protein atlas
TCF3	Rabbit polyclonal	HPA062476	(1:150)	Human protein atlas
APBB1	Rabbit polyclonal	HPA038521	(1:300)	Human protein atlas
C19orf81	Rabbit polyclonal	HPA060238	(1:100)	Human protein atlas
GPRC5C	Rabbit polyclonal	HPA029776	(1:135)	Human protein atlas
ICA1L	Rabbit polyclonal	HPA042507	(1:100)	Human protein atlas
LMNTD2	Rabbit polyclonal	HPA058474	(1:300)	Human protein atlas
MAGEC1	Rabbit polyclonal	HPA004622	(1:500)	Human protein atlas
TUBA1A	Rabbit polyclonal	HPA043684	(1:100)	Human protein atlas
MAP2K5	Rabbit polyclonal	HPA027347	(1:400)	Human protein atlas
HLA-DPA1	Rabbit polyclonal	HPA017967	(1:35)	Human protein atlas
SLC25A22	Rabbit polyclonal	HPA014662	(1:300)	Human protein atlas

Human c-KIT$^+$ spermatogonia isolation using MACS
c-KIT$^+$ cells were enriched using magnetic activated cell sorting (MACS) protocols (Miltenyi Biotec, Inc.). Single testicular cell suspensions were incubated with anti-c-KIT microbeads (Miltenyi Biotec cat # 130–098–571) at 4 °C. Following microbead binding, cells were re-suspended in autoMACS running buffer (Miltenyi Biotec cat # 130–091–221) and ran through LS columns (Miltenyi Biotec cat # 130–042–401) placed in a magnetic field. Columns were rinsed three times with buffer in autoMACS running buffer (Miltenyi Biotec cat # 130–091–221) before being removed from the magnetic field. MACS running/separation buffer (Miltenyi Biotec cat # 130–091–221) was then applied to the column before magnetically-labeled cells were flushed out by firmly pushing the plunger into the column. Cells were then centrifuged and re-suspended to a desired concentration.

ATAC-seq library preparation and sequencing
The ATAC-seq libraries were prepared as previously described[45] on ~30k sorted KIT$^+$ spermatogonia, SSEA4$^+$ SSCs or cultured ESCs.[10] Briefly, collected cells were lysed in cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ and 0.1% IGEPAL CA-630) and the nuclei were pelleted and resuspended in Transposase buffer. The Tn5 enzyme was made in-house and the transposition reaction was carried out for 30 min at 37 °C. Following purification, the Nextera libraries were amplified for 12 cycles using the NEBnext PCR master mix (NEB cat # M0541L) and purified using the Agencourt AMPure XP–PCR Purication (Beckman Coulter cat # A63881). All libraries were sequenced on a 125-cycle paired-end run on an Illumina HiSeq 2500 instrument.

Flow cytometry analysis
Cells were analyzed by flow cytometry using Aria Analyzer. For FGFR3 staining, cells were firstly incubated with anti-FGFR3 antibody (mouse monoclonal; Santa Cruz cat # sc-13121), washed and then incubated with Alexa Fluor-647 (anti-mouse; Thermo Fisher cat # Z25008); for TSPAN33 staining, cells were incubated with TSPAN33 PE-conjugated antibody (R&D Systems cat # FAB8405P-015); for SSEA4 staining, cells were incubated with SSEA4 VioBlue-conjugated antibody (Miltenyi Biotec cat # 130–098–366). Gating was based on unstained and single stained samples. FACS data were analyzed using FlowJo software (Ashland).

Quantifications and statistical analysis
Cell type identification and clustering analysis using Seurat program. The Seurat program (http://satijalab.org/seurat/, R package, v.2.0.0) was firstly applied for analysis of RNA-Sequencing data. To start with, UMI count tables from each replicates and donors were loaded into R using Read10X function, and Seurat objects were built from each experiment. Each experiment was filtered and normalized with default settings. Specifically, cells were retained only when they had greater than 500 genes expressed, and less than 20% reads mapped to mitochondrial genome. We first ran t-SNE and the clustering analysis for each replicate, which resulted in similar t-SNE map. Next, to minimize variation between technical replicates, we normalized and combined technical replicates from the same donor using the 10× Genomics built-in application from Cell Ranger "cellrange aggr". Data matrices from different donors were then loaded into R using Seurat. Next, cells were normalized to the total UMI read count as well as mitochondrial read percentage, as instructed in the manufacturer's manual (http://satijalab.org/seurat/). Seurat objects (matrices from different donors) were then combined using RunCCA function. t-SNE and clustering analyses were then performed on the combined dataset using the top 5000 highly variable genes and PCs 1–15, which showed most significant *p*-values. Given the low number of Sertoli cells (underrepresented due to size filtering), the initial clustering analysis did not identify them as a separate cluster. We performed deeper clustering of somatic cells, identified the Sertoli cell cluster, and projected it back to the overall clusters, which resulted in 13 discrete cell clusters. Correlation of different replicates was calculated based on average expression (normalized UMIs by Seurat) in each experiment.

Pseudotime and clustering analysis. Germ cells (Clusters 1–8) from t-SNE plot were used for pseudotime analysis by slingshot (https://github.com/kstreet13/slingshot, R package, v0.1.2–3). Cluster 1 (SSCs) was used as start, cluster 4 (secondary spermatocytes) as middle, and cluster 8 (sperm) as end of pseudotime. After pseudotime time was determined, gene clustering analysis was performed to determine the fidelity of pseudotime. Here, cells (in columns) were ordered by their pseudotime, and genes (in rows) were clustered by k-means clustering using Cluster 3.0. Different k-mean numbers were used to reach the optimal clustering number. Genes within each gene clusters were then used to perform Gene Ontology analysis by David (v6.7).

Transposable element and lncRNA analysis. First, gtf files for TEs and lncRNAs were downloaded from UCSC and lncipedia, respectively. TE gtf was treated and filtered using the same approach as described.[46] These gtf files were then used to replace the default gtf files (for genes) in Cell Ranger, and UMI count tables were generated using the same approach as described above. For downstream analysis, TE and lncRNA expression patterns were cast to the gene expression based clustering and pseudotime.

Reclustering of spermatogonia/SSCs (Cluster 1 and 2). We parsed out cells in Cluster 1 and Cluster 2 (total number $n = 614$), and loaded their gene expression matrices into R through Seurat. Clustering and t-SNE analyses were performed, and a small cluster (containing 24 cells) was identified as outlier and excluded from further analysis. The remaining cells ($n = 590$) were re-clustered and analyzed using t-SNE using the top 5000 highly variable genes and PCs 1–5, which showed the most significant *p*-values. Pseudotime was performed as mentioned above using slingshot (v0.1.2–3).

Cell cycle analysis. Cell cycle analysis was performed using scran program (https://bioconductor.org/packages/3.7/bioc/vignettes/scran/inst/doc/scran.html, R Package; v1.6.5). Briefly, cell cycle genes were obtained from scran program, and their expression in States 0–4 were loaded into scran. Cell cycle phases (G1, S, and G2/M) were then assigned to each single cell.

Regulon analysis. Regulon analysis was performed using SCENIC program (https://github.com/aertslab/SCENIC, R Package; v0.1.7). State 0–4 cells were used to generate the regulon activity score of transcription factors as instructed by their manual. The regulon activity scores were then projected onto the t-SNE plot based on gene expression levels.

RNA velocity analysis. Cell velocity analysis was performed using Velocyto.R program (http://velocyto.org, v0.5), as instructed.[11] Briefly, at first, Velocyto used raw data to count the spliced (mRNA) and unspliced (Intron) reads for each gene, and generated a.loom file. Those.loom files were then loaded intro R (v3.4) using read.loom.matrices function to generate count tables for splicing and unsplicing reads. To generate RNA velocity map for spermatogonia, splicing and unspliced reads from States 0–4 were further used, and coordinates of the cells in the t-SNE plot were also provided. Lastly, the RNA velocity map was projected onto the t-SNE plot.

ATAC-seq comparison analysis. SAM alignments were generated from the Illumina Fastq files aligned to human hg19 genome using Novocraft's novoalign aligner (http://www.novocraft.com) with the following parameters: –o SMA –r ALL 50. Peak calling was performed using macs2 (https://github.com/taoliu/MACS, v2.1.2.20160309) using the following settings: –g 2.7e9–call-summit –f BAMPE –nomodel –B –SPMR –extsize 200. Generated bedgraph file was then transformed to bw format using UCSC bedGraphToBigWig application (v4). Correlation was generated using deepTools (v3) by firstly using multiBigwigSummary bins application (with default settings) and then plotCorrelation application (with the following parameters: --skipZeros --removeOutliers). Distance of peak summit was calculated using bedtool (http://bedtools.readthedocs.io/en/latest/content/tools/makewindows.html, v2.25.0) closestBed application.

Data and software availability

The accession number for all sequencing data reported in this paper is GEO: GSE120508. Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Bradley R. Cairns (brad.cairns@hci.utah.edu).

ACKNOWLEDGEMENTS

We thank Brian Dalley and Opal Allen for sequencing expertise, Chris Conley and Tim Parnell for bioinformatics assistance, James Marvin for flow cytometry assistance, and Intermountain Donor Service staff for sample handling. Financial support was from Howard Hughes Medical Institute to B.R.C.; P30CA042014 to Huntsman Cancer Institute core facilities from the National Cancer Institute. Flow cytometry core was supported by the National Center for Research Resources of the National Institutes of Health under Award Number 1S10RR026802–01. The Wellcome, UK (Senior Investigator Award 102731 to Prof. Andrew Wilkie) supports A.G., G.J.M., and H.M.; the Wolfson Imaging Center (Oxford) is supported by the WIMM Strategic Alliance (G0902418 and MC_UU_12025). C.L. is supported by the Knut and Alice Wallenberg Foundation. L.C. is supported by NIH U01EB021240.

AUTHOR CONTRIBUTIONS

B.R.C. and J.G. conceived the project, designed the genomics and analysis experiments, and advised on other sections. J.G. conducted genomics experiments and computational analyses with help from E.J.G. Sample acquisition was led by J.M. H. and R.K. with input from D.T.C., B.R.C., and J.G. Samples were transferred and processed by J.G., Y.G., X.N. and E.J.G. Y.T., J.Y., and L.C. led the design and execution mRNA seqFISH. C.L. provided HPA antibodies. A.G. contributed to project design and led protein validation efforts. H.M. and G.J.M. performed protein immunostaining. J. G., A.G., and B.R.C. wrote the manuscript, with input from all authors.

ADDITIONAL INFORMATION

Supplementary information accompanies this paper at https://doi.org/10.1038/s41422-018-0099-2.

Competing interests: The authors declare no competing interests.

REFERENCES

1. Clermont, Y. Spermatogenesis in man. *Fertil. Steril.* **17**, 705–721 (1966).
2. Griswold, M. D. Spermatogenesis: the commitment to meiosis. *Physiol. Rev.* **96**, 1–17 (2015).
3. Brinster, R. L. & Zimmermann, J. W. Spermatogenesis following male germ-cell transplantation. *Proc. Natl Acad. Sci.* **91**, 11298–11302 (1994).
4. Kanatsu-Shinohara, M. & Shinohara, T. Spermatogonial stem cell self-renewal and development. *Annu. Rev. Cell. Dev. Biol.* **29**, 163–187 (2013).
5. Clermont, Y. Kinetics of spermatogenesis in mammals: seminiferous epithelium cycle and spermatogonial renewal. *Physiol. Rev.* **52**, 198–236 (1972).
6. Izadyar, F. et al. Identification and characterization of repopulating spermatogonial stem cells from the adult human testis. *Hum. Reprod.* **26**, 1296–1306 (2011).
7. Valli, H. et al. Fluorescence- and magnetic-activated cell sorting strategies to isolate and enrich human spermatogonial stem cells. *Fertil. Steril.* **102**, 566–580.e7 (2014).
8. Wu, A. R., Wang, J., Streets, A. M. & Huang, Y. Single-cell transcriptional analysis. *Annu. Rev. Anal. Chem.* **10**, 439–462 (2017).
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 (2017).
10. Guo, J. et al. Chromatin and single-cell RNA-seq profiling reveal dynamic signaling and metabolic transitions during human spermatogonial stem cell development. *Cell Stem Cell* **21**, 533–546 (2017).
11. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
12. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
13. Bhushan, S. & Meinhardt, A. The macrophages in testis function. *J. Reprod. Immunol.* **119**, 107–112 (2017).
14. DeFalco, T. et al. Macrophages contribute to the spermatogonial niche in the adult testis. *Cell Rep.* **12**, 1107–1119 (2015).
15. Yang, Q.-E., Kim, D., Kaucher, A., Oatley, M. J. & Oatley, J. M. CXCL12–CXCR4 signaling is required for the maintenance of mouse spermatogonial stem cells. *J. Cell. Sci.* **126**, 1009–1020 (2013).
16. Oatley, J. M., Oatley, M. J., Avarbock, M. R., Tobias, J. W. & Brinster, R. L. Colony stimulating factor 1 is an extrinsic stimulator of mouse spermatogonial stem cell self-renewal. *Development* **136**, 1191–1199 (2009).
17. Yao, H. H.-C., Whoriskey, W. & Capel, B. Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Genes Dev.* **16**, 1433–1440 (2002).
18. Chhikara, N. et al. Human Epididymis Protein-4 (HE-4): a novel cross-class protease inhibitor. *PLoS ONE* **7**, e47672 (2012).
19. Mastrangelo, P. & Westaway, D. The prion gene complex encoding PrPC and Doppel: insights from mutational analysis. *Gene* **275**, 1–18 (2001).
20. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
21. Turner, J. M. A. Meiotic sex chromosome inactivation. *Development* **134**, 1823–1831 (2007).
22. McCarrey, J. R. & Dilworth, D. D. Expression of Xist in mouse germ cells correlates with X–chromosome inactivation. *Nat. Genet.* **2**, 200–203 (1992).
23. Jan, S. Z. et al. Molecular control of rodent spermatogenesis. *Biochim. Biophys. Acta* **1822**, 1838–1850 (2012).
24. Handel, M. A. & Schimenti, J. C. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nat. Rev. Genet.* **11**, 124–136 (2010).
25. Mays-Hoopes, L. L., Bolen, J., Riggs, A. D. & Singer-Sam, J. Preparation of spermatogonia, spermatocytes, and round spermatids for analysis of gene expression using fluorescence-activated cell sorting. *Biol. Reprod.* **53**, 1003–1011 (1995).
26. Zhang, T. & Zarkower, D. DMRT proteins and coordination of mammalian spermatogenesis. *Stem Cell Res.* **24**, 195–202 (2017).
27. Zhao, L. et al. SOX4 regulates gonad morphogenesis and promotes male germ cell differentiation in mice. *Dev. Biol.* **423**, 46–56 (2017).
28. Feng, C.-W. A. et al. SOX30 is required for male fertility in mice. *Sci. Rep.* **7**, 17619 (2017).
29. Chen, Y. et al. Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. *Cell Res.* **28**, 879–896 (2018).
30. Hammoud, S. S. et al. Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. *Cell Stem Cell* **15**, 239–253 (2014).
31. Han, J. et al. Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature* **463**, 1096–1100 (2010).
32. Shah, S., Lubeck, E., Zhou, W. & Cai, L. In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* **92**, 342–357 (2016).
33. von Kopylow, K. & Spiess, A.-N. Human spermatogonial markers. *Stem Cell Res.* **25**, 300–309 (2017).
34. Persio, S. D. et al. Spermatogonial kinetics in humans. *Development* **144**, 3430–3439 (2017).
35. Lindskog, C. The Human Protein Atlas–an important resource for basic and clinical research. *Expert. Rev. Proteom.* **13**, 627–629 (2016).
36. Meng, X. et al. Regulation of cell fate decision of undifferentiated spermatogonia by GDNF. *Science* **287**, 1489–1493 (2000).
37. Garbuzov, A. et al. Purification of GFRα1+and GFRα1–spermatogonial stem cells reveals a niche-dependent mechanism for fate determination. *Stem Cell Rep.* **10**, 553–567 (2018).
38. Brawley, C. & Matunis, E. Regeneration of male germline stem cells by spermatogonial dedifferentiation in vivo. *Science* **304**, 1331–1334 (2004).
39. Sheng, X. R., Brawley, C. M. & Matunis, E. L. Dedifferentiating spermatogonia outcompete somatic stem cells for niche occupancy in the drosophila testis. *Cell Stem Cell.* **5**, 191–203 (2009).
40. Nakagawa, T., Sharma, M., Nabeshima, Y., Braun, R. E. & Yoshida, S. Functional hierarchy and reversibility within the murine spermatogenic stem cell compartment. *Science* **328**, 62–67 (2010).

41. Hara, K. et al. Mouse spermatogenic stem cells continually interconvert between equipotent singly isolated and syncytial states. *Cell Stem Cell* **14**, 658–672 (2014).
42. Feng, L. & Chen, X. Epigenetic regulation of germ cells—remember or forget? *Curr. Opin. Genet. Dev.* **31**, 20–27 (2015).
43. Wang, M. et al. Single-cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. *Cell Stem Cell* **23**, 1–16 (2018).
44. Green, C. D. et al. A comprehensive roadmap of murine spermatogenesis defined by single-cell RNA-seq. *Dev. Cell.* **46**, 651–667.e10 (2018).
45. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
46. Jin, Y., Tam, O. H., Paniagua, E. & Hammell, M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics* **31**, 3593–3599 (2015).