

Supplementary Materials

Supplementary Materials and Methods

Animals and testis samples

Pig testis samples were acquired from Guanzhong black pigs aged 7, 30, 60, and 90 days old from the farm of Northwest A&F University (Yangling, China). All protocols and experiments were conducted in accordance with the Northwest A&F University Guide for the Care and Use of Laboratory Animals.

Generation of cell suspensions

Single-cell suspensions from seminiferous tubules were generated from the testes of Guanzhong black pigs using a two-step enzymatic digestion approach (Zhang et al., 2017). Briefly, testicular parenchyma was digested with 1 mg/mL Collagenase Type IV (Invitrogen) for 30 min at 34 °C and washed with Dulbecco's phosphate-buffered saline (DPBS; Invitrogen) to remove interstitial cells. The seminiferous tubules were then digested with 0.25% trypsin/EDTA (HyClone) for 5 min at 34 °C, followed by termination using 10% fetal bovine serum (FBS). Cell suspensions were filtered with 40 µm strainers, washed using DPBS with 0.5 mg/mL DNase I (Sigma) twice, then suspended in 1640/Dulbecco's Modified Eagles Medium (DMEM, HyClone) containing 5% FBS.

Hematoxylin and eosin staining

Testis tissues were fixed in Bouin's solution for 12 h before being transferred to 70% ethanol. The fixed tissues were embedded in paraffin and sectioned (5 µm thick). Coronal sections were de-paraffinized in xylene and re-hydrated in ethanol. The sections were then stained with hematoxylin for 5 min and incubated with 1% hydrochloric acid alcohol for 30 s. Slides were then rinsed in tap water for 10 min, followed by eosin staining for 2 min. Subsequently, sections were dehydrated and mounted with coverslips.

Single-cell testis transcriptomes using 10× Genomics Chromium platform

The cells were loaded into Chromium microfluidic chips using 10× Genomics 3' v3 chemistry to generate single-cell gel bead emulsions (GEMs) using the Chromium Controller according to the manufacturer's protocols. Libraries were sequenced on the Illumina NovaSeq6000 sequencing system by LC-Bio-Technology Co. Ltd., (Hangzhou, China). For gene expression data, FASTQ files were generated using CellRanger mkfastq. Alignment, filtering, and unique molecular identifier (UMI) counting were performed using CellRanger count. Sscrofa11.1 genome assembly/annotation was used as a reference. Further analysis used the UMI count tables.

Quality control, normalization, and batch-effect correction of scRNA-seq data

We filtered the data by retaining cells expressing >200 but <4 000 genes with <20% mitochondrial content and UMI counts between 2 000–100 000. In addition, we identified germ cells based on the well-established gene marker DDX4 and retained 9 307 germ cells for further analysis. We performed normalization by deconvolution with the scran package (Lun et al., 2016). We used the quickCluster function to pre-cluster cells following size factors. Expression values were scaled with the size factors, then log-transformed using the logNormCounts function. Unsupervised cell clustering and t-SNE analysis were performed on statistically significant principal

components. Batch effects were removed by the fast mutual nearest-neighbor correction (fastMNN) function in the batchelor R package (Haghverdi et al., 2018). Further downstream analysis was performed using the Scanpy (v1.7.1) tool, and SingleCellExperiment data were exported to anndata using the anndata2ri tool.

Cell cycle analysis

Cell cycle phase scores were calculated for each cell using the Scanpy function `sc.tl.score_gene_cell_cycle`. Genes associated with the S and G2M phases used in our datasets were from Tirosh et al. (2016).

Enrichment analysis

Metascape was used to perform enrichment analysis of marker genes. The pig genes were converted into orthologous human genes. Gene Ontology (GO) Biological Processes, Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways, Reactome Gene Sets, and CORUM were chosen as ontology sources. Based on membership similarities, the significantly enriched terms were grouped into clusters ($P < 0.01$, minimum counts > 3 , and enrichment factor > 1.5).

Iteration of human (GEO: GSE120508 and GSE134144) and pig datasets

Human testicular datasets from two independent studies using 10× Genomics sequencing (SRR9670686, SRR9670688, SRR9670690, SRR9670692, and SRR6860521) (Guo et al., 2018, 2020) were processed in a similar way to our porcine datasets. Raw human count matrices were filtered by retaining cells expressing > 200 genes with $< 20\%$ mitochondrial content. Human clusters were annotated using similar parameters.

Orthologous genes were obtained between humans and pigs using Ensembl. The one2one_ortholog-annotated orthologous genes in pigs were retained with the human reference. We identified 14 960 one-to-one orthologous genes for integration analysis. Pig-specific gene symbols were replaced by human orthologous gene symbols to combine the datasets.

Immunofluorescence assay

Sections were deparaffinized and rehydrated, with heat-mediated antigen retrieval then performed in 10 mM sodium citrate buffer. Sections were cooled for 30 min and blocked in 3% bovine serum albumin (BSA) for 2 h. The following primary antibodies were used: PCNA (1:200; Cell Signaling Technology, Catalog No. #2586), VASA (1:100; Abcam, Catalog No. ab13840), DBA (1:100; VECTOR), UCHL1 (1:200; Abcam, Catalog No. ab8189), CDH1 (1:200; Proteintech, Catalog No. 20874-1-AP), and SETDB1 (1:200; Cell Signaling Technology, Catalog No. #93212). Antigen reactions were conducted using the corresponding donkey anti-rabbit/mouse secondary antibodies (1:400; Alexa Fluor 488/594, Yeason) for 2 h at 4 °C in the dark. The sections were then incubated with 4',6-diamidino-2-phenylindole (DAPI) (Bioworld Technology) to facilitate nuclear visualization.

REFERENCES

Ernst C, Eling N, Martinez-Jimenez CP, Marioni JC, Odom DT. 2019. Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nature Communications*, **10**(1): 1251.

- Fayomi AP, Orwig KE. 2018. Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. *Stem Cell Research*, **29**: 207–214.
- França LR, Silva VA Jr, Chiarini-Garcia H, Garcia SK, Debeljuk L. 2000. Cell proliferation and hormonal changes during postnatal development of the testis in the pig. *Biology of Reproduction*, **63**(6): 1629–1636.
- Guo JT, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie XC, et al. 2018. The adult human testis transcriptional cell atlas. *Cell Research*, **28**(12): 1141–1157.
- Guo JT, Nie XC, Giebler M, Mlcochova H, Wang YQ, Grow EJ, et al. 2020. The dynamic transcriptional cell atlas of testis development during human puberty. *Cell Stem Cell*, **26**(2): 262–276.e4.
- Haghverdi L, Lun ATL, Morgan MD, Marioni JC. 2018. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature Biotechnology*, **36**(5): 421–427.
- Hirota T, Blakeley P, Sangrithi MN, Mahadevaiah SK, Encheva V, Snijders AP, et al. 2018. SETDB1 Links the Meiotic DNA Damage Response to Sex Chromosome Silencing in Mice. *Developmental Cell*, **47**(5): 645-659 e646.
- Law NC, Oatley JM. 2020. Developmental underpinnings of spermatogonial stem cell establishment. *Andrology*, **8**(4): 852–861.
- Liu Y, Zhang L, Li W, Huang Q, Yuan S, Li Y, et al. 2019. The sperm-associated antigen 6 interactome and its role in spermatogenesis. *Reproduction*, **158**(2): 181-197.
- Lun ATL, Bach K, Marioni JC. 2016. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biology*, **17**: 75.
- Meistrich ML, Mohapatra B, Shirley CR, Zhao M. 2003. Roles of transition nuclear proteins in spermiogenesis. *Chromosoma*, **111**(8): 483-488.
- Nestorowa S, Hamey FK, Sala BP, Diamanti E, Shepherd M, Laurenti E, et al. 2016. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood*, **128**(8): e20–e31.
- Park HJ, Lee WY, Park C, Hong K, Song H. 2019. CD14 is a unique membrane marker of porcine spermatogonial stem cells, regulating their differentiation. *Scientific Reports*, **9**(1): 9980.
- Perleberg C, Kind A, Schnieke A. 2018. Genetically engineered pigs as models for human disease. *Disease Models & Mechanisms*, **11**(1): dmm030783.
- Tan K, Song HW, Wilkinson MF. 2020. Single-cell RNAseq analysis of testicular germ and somatic cell development during the perinatal period. *Development*, **147**(3): dev183251.
- Tirosh I, Izar B, Prakadan SM, Wadsworth II MH, Treacy D, Trombetta JJ, et al. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell

RNA-seq. *Science*, **352**(6282): 189–196.

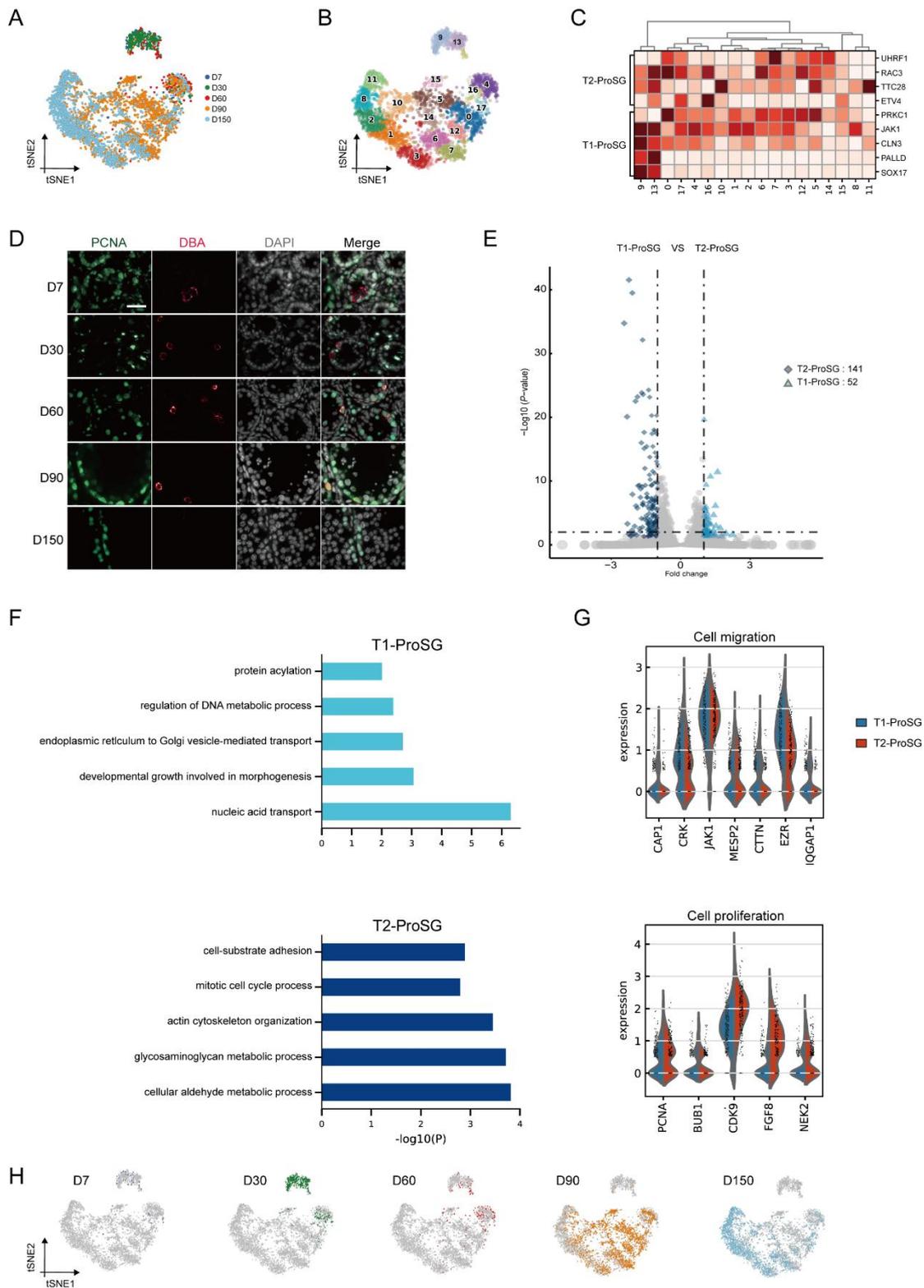
Turner JMA. 2007. Meiotic sex chromosome inactivation. *Development*, **134**(10): 1823–1831.

Zhang LK, Li FY, Lei PP, Guo M, Liu RF, Wang L, et al. 2021. Single-cell RNA-sequencing reveals the dynamic process and novel markers in porcine spermatogenesis. *Journal of Animal Science and Biotechnology*, **12**(1): 122.

Zhang PF, Chen XX, Zheng Y, Zhu JS, Qin YW, Lv YH, et al. 2017. Long-term propagation of porcine undifferentiated spermatogonia. *Stem Cells Development*, **26**(15): 1121–1131.

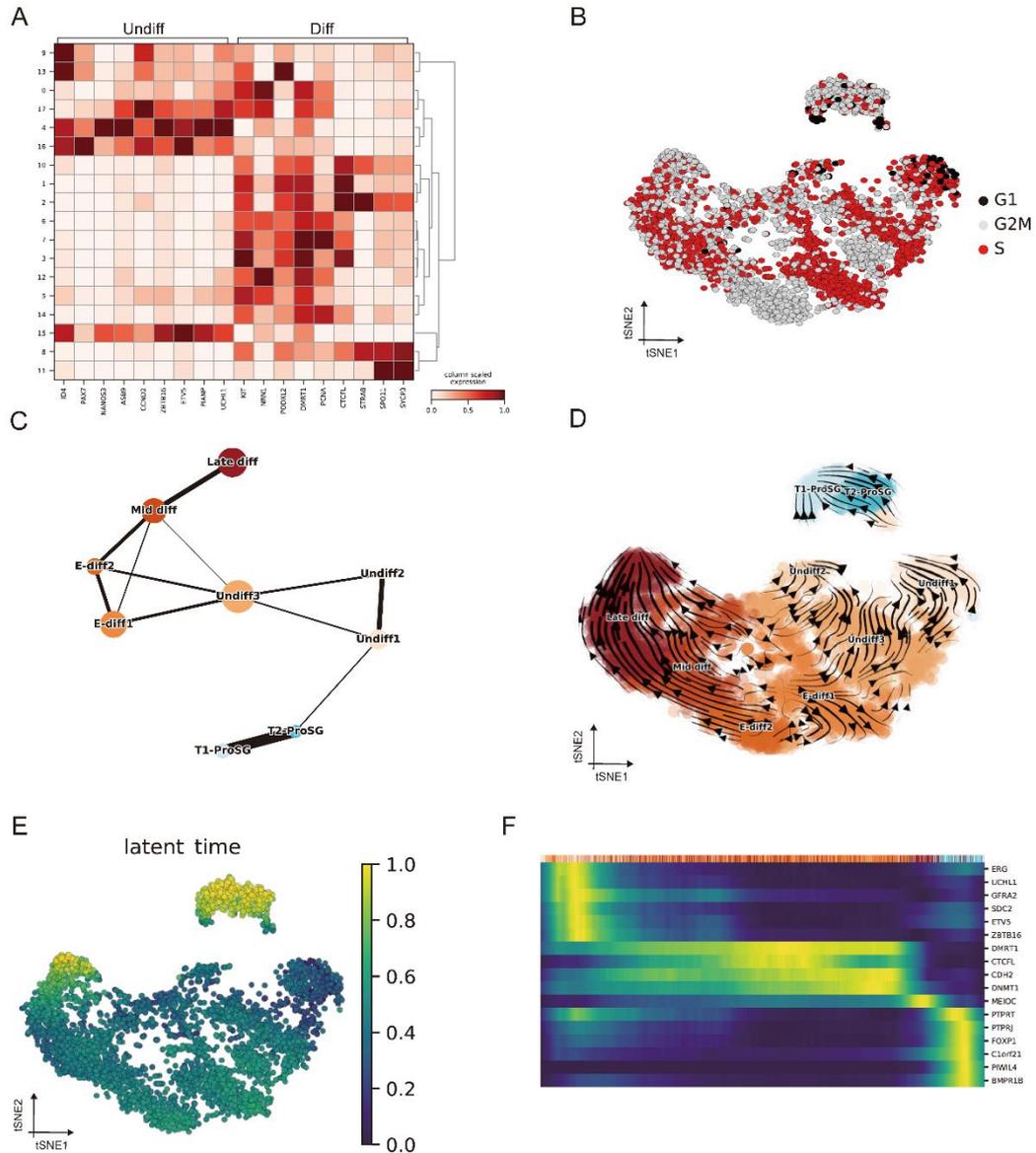
Zhang PF, Li FY, Zhang LK, Lei PP, Zheng Y, Zeng WX. 2020. Stage-specific embryonic antigen 4 is a membrane marker for enrichment of porcine spermatogonial stem cells. *Andrology*, **8**(6): 1923–1934.

coefficients for germ cell clusters with hierarchical clustering. H: Within each broad cluster, Pearson correlation coefficients for cells derived from each pig sample were plotted in heatmap format. I: Distribution of cell attributes from each cell type in porcine dataset. J: Pie chart showing distribution of germ cells (ProSG & spermatogonia, spermatocytes, and spermatids) from different pig samples. Cell types are colored according to legend.

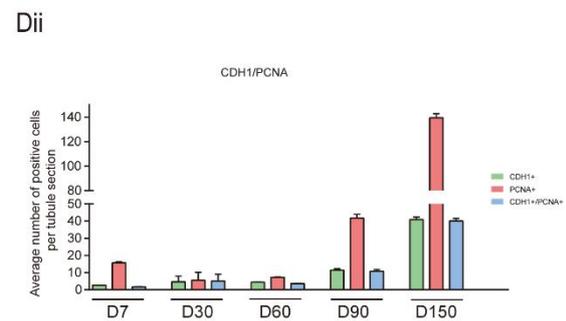
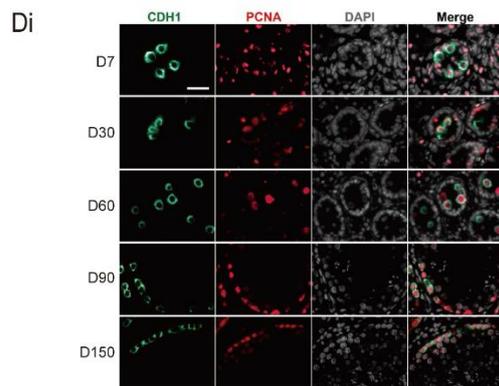
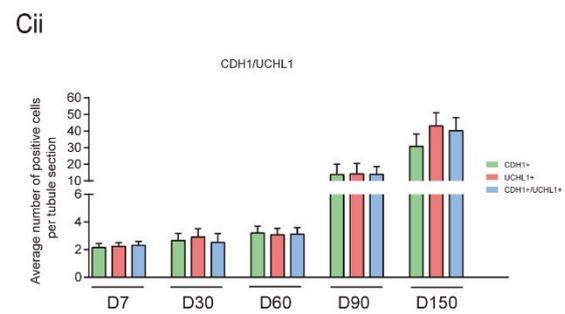
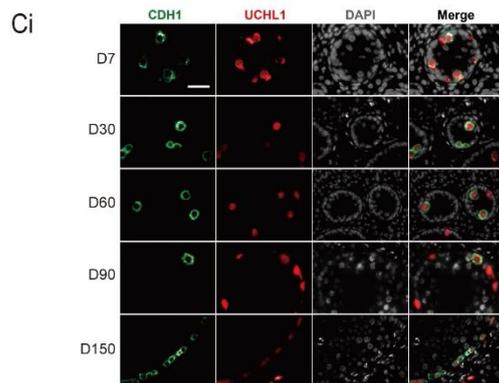
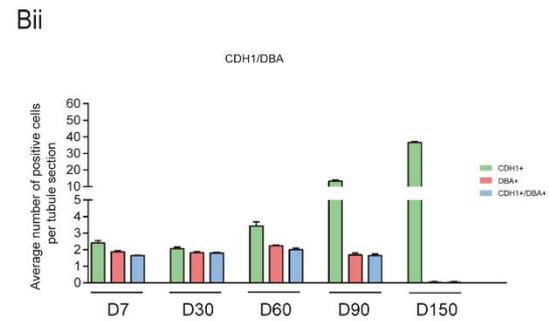
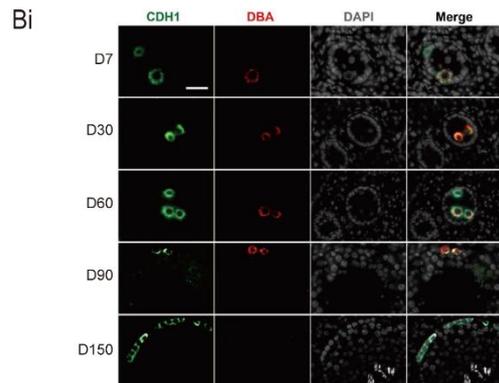
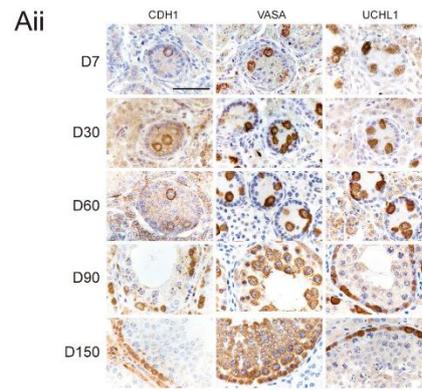
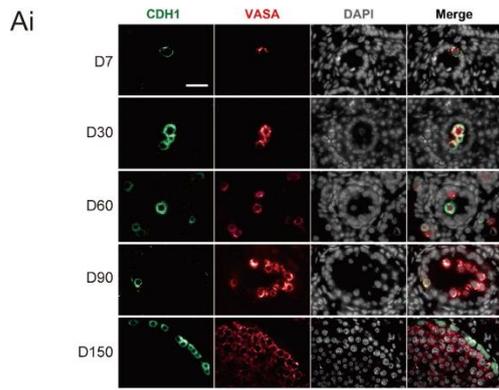


Supplementary Figure S2 Distinct phases of proliferation and migration in ProSG during porcine puberty. A: Focused analysis (t-SNE and clustering) of ProSG and spermatogonia showing developmental progression during puberty. Cells are colored based on ages. B: t-SNE plot showing unbiased Leiden clusters from ProSG and spermatogonia. C: Heatmap showing expression of T1-ProSG and

T2-ProSG marker genes in unbiased cell clusters. Differential gene expression levels are based on Z scores, as defined by color key. D: Immunolocalization co-staining of PCNA and DBA at different ages (7–150 days old), showing ProSG disappeared around D90. Scale bar=50 μ m. E: Volcano plot showing differentially expressed genes (DEGs) between T1-ProSG (n=52) and T2-ProSG (n=141). F: Enrichment terms and *P*-values for DEGs of T1-ProSG and T2-ProSG. G: Violin plot showing expression of genes associated with cell migration and proliferation in both T1-ProSG and T2-ProSG. H: Deconvolution of cell distribution plot according to ages/individuals.



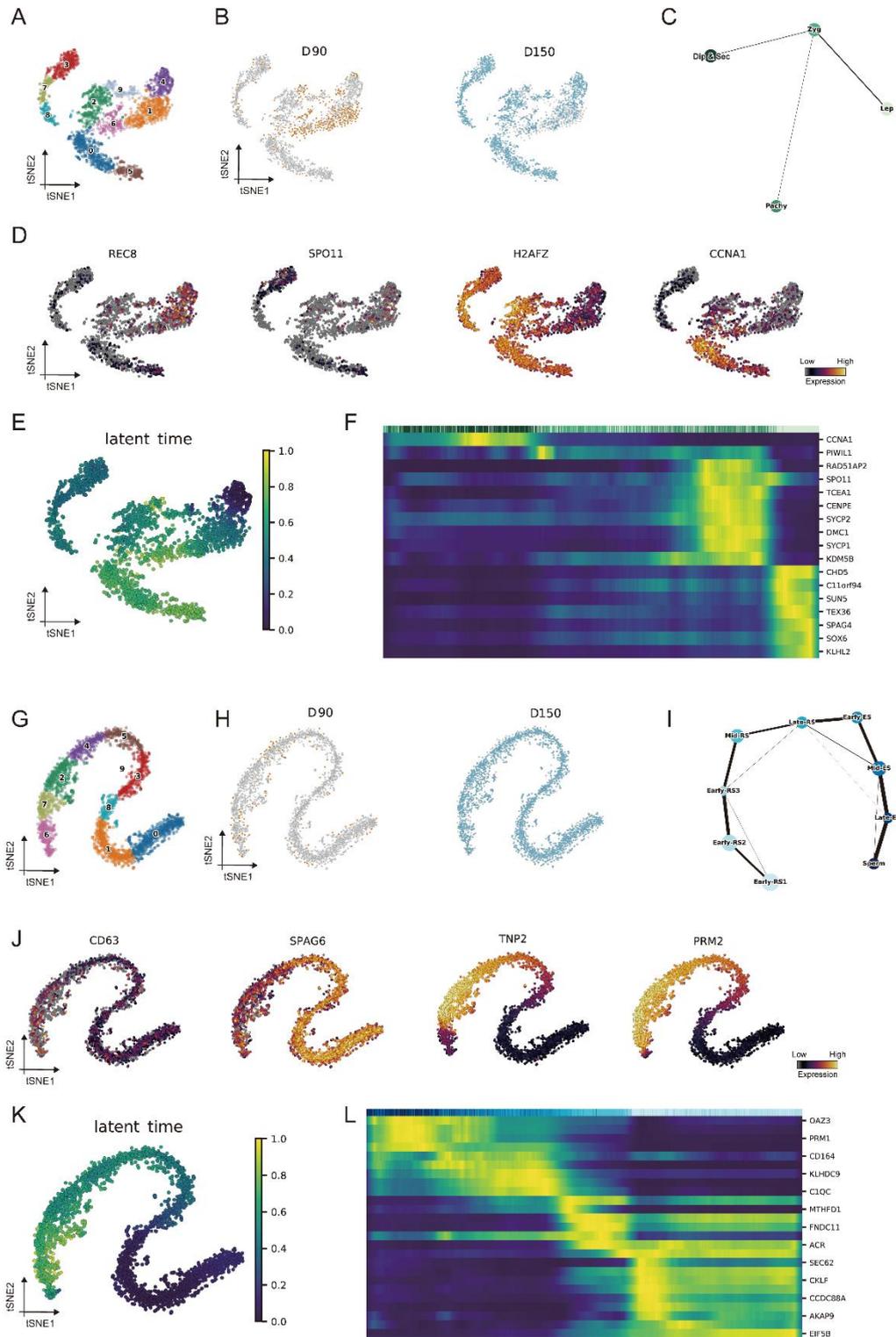
Supplementary Figure S3 Single-cell porcine spermatogonia trajectories reveal transitions with differentiation. A: Heatmap showing expression of undifferentiated and differentiated spermatogonia marker genes in Leiden cell clusters. B: Annotation of cell cycle score and cell cycle phase for ProSG and spermatogonia in t-SNE plots (G1-phase, black; S-phase, red; G2M, gray). C: PAGA graph showing ProSG and spermatogonia clusters. Size of clusters is proportional to cell number, and thickness of lines between clusters is proportional to connection strength. D: t-SNE plot showing *scVelo*-derived dynamics in ProSG and spermatogonia clusters. E: Single-cell transcriptomes of ProSG and spermatogonia were used for cell trajectories, ordered in latent time. F: Heatmap showing variable genes across latent time from identified cell types of ProSG and spermatogonia.



Supplementary Figure S4 Expression of CDH1 in porcine testes at different ages.

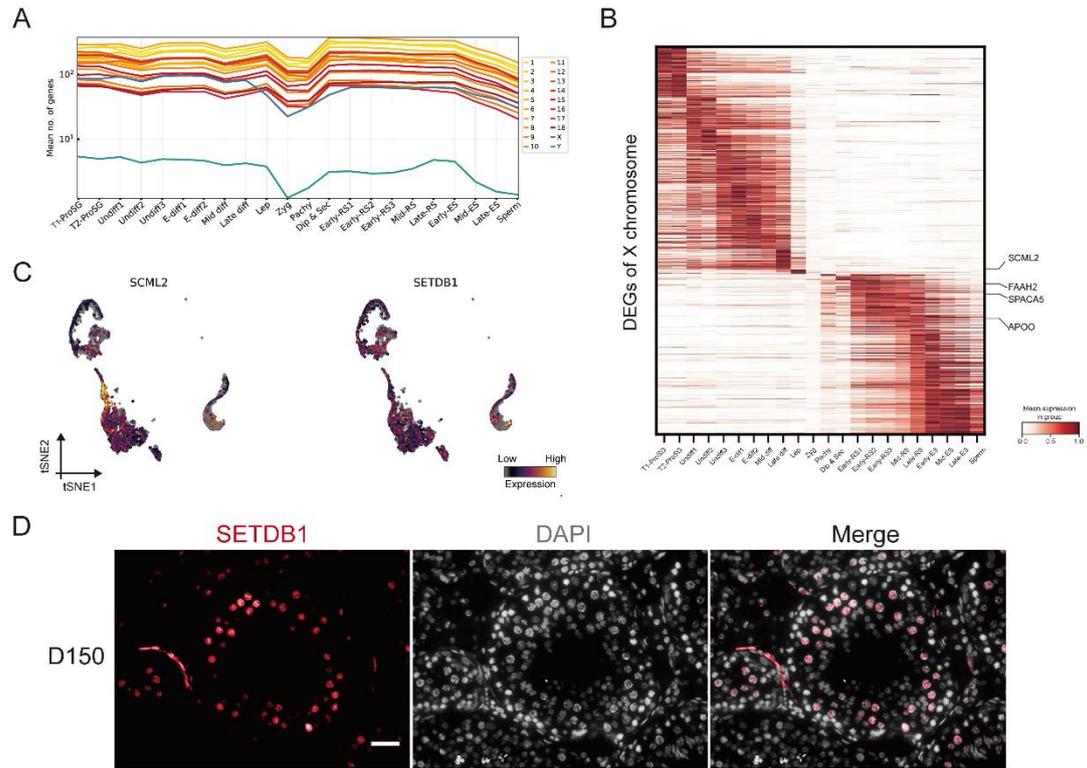
Ai: Co-immunofluorescence of CDH1 and VASA expression. Scale bars=50 μ m. **Aii:** Immunohistochemical staining of CDH1, VASA, and UCHL1 during porcine testicular development. Scale bars=50 μ m. **Bi:** Immunolocalization co-staining of CDH1 and DBA. Scale bar=50 μ m. **Bii:** Average number of CDH1⁺, DBA⁺, and

CDH1⁺/DBA⁺ germ cells per tubule in porcine testes at different ages. Data are mean±SD of three independent experiments. Ci: Immunolocalization co-staining of CDH1 and UCHL1. Scale bar=50 μm. Cii: Average number of CDH1⁺, UCHL1⁺, and CDH1⁺/UCHL1⁺ germ cells per tubule in porcine testes at different ages. Data are mean±SD of three independent experiments. Di: Co-immunofluorescence of CDH1 and PCNA expression. Scale bars=50 μm. Dii: Average number of CDH1⁺, PCNA⁺, and CDH1⁺/PCNA⁺ germ cells per tubule in porcine testes at different ages. Data are mean±SD of three independent experiments.

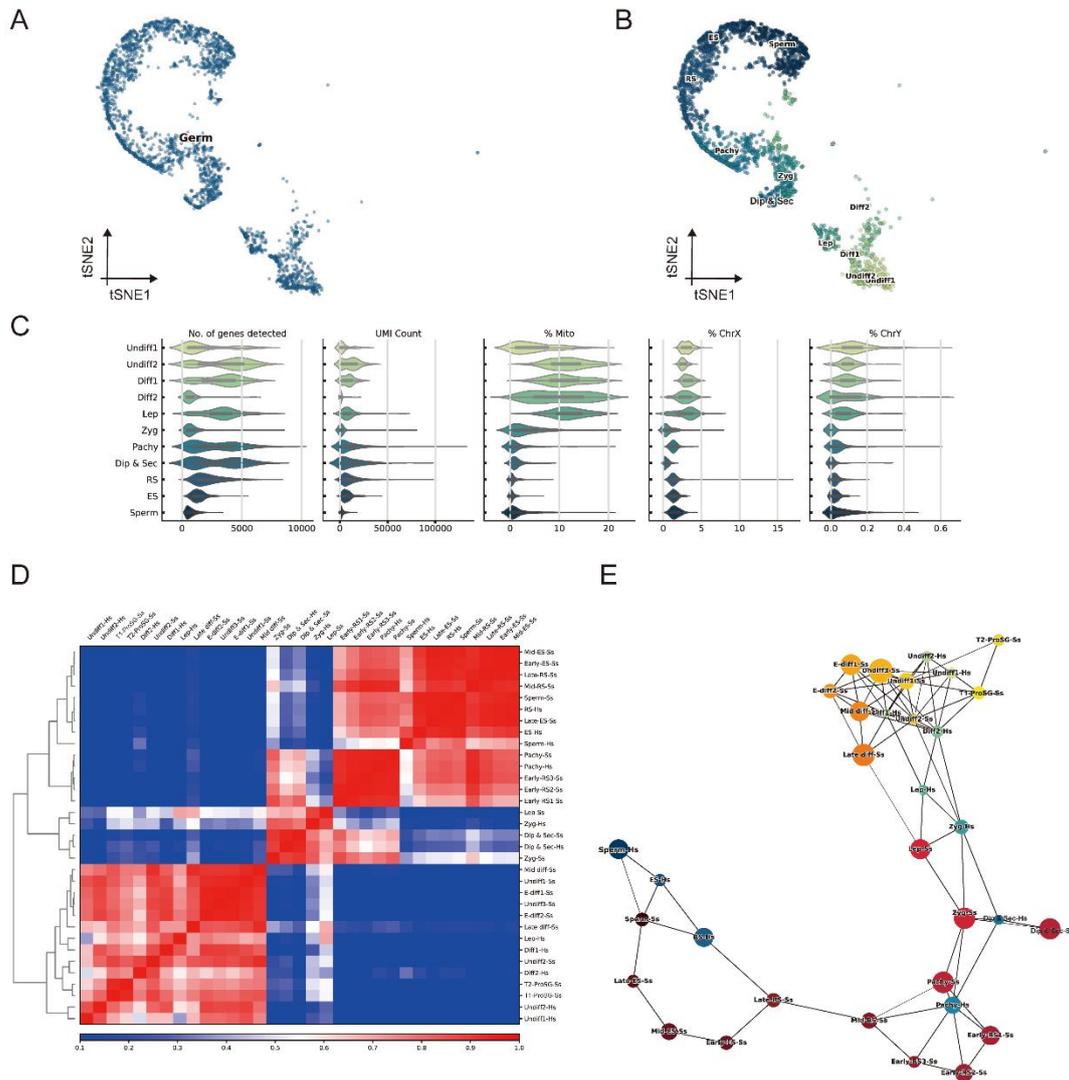


Supplementary Figure S5 Identification of spermatocyte and spermatid subtypes during puberty. A: t-SNE plot showing unbiased cell clusters from spermatocytes. B: t-SNE plot showing distribution of spermatocytes at D90 and D150. C: PAGA graph showing spermatocyte clusters. D: Visualization of selected spermatocyte marker gene expression levels across all single cells in t-SNE plot. E: Single-cell transcriptomes from spermatocytes were used for cell trajectories, ordered in latent

time. F: Heatmap showing variable genes across latent time from identified spermatocyte cell types. G: t-SNE plot showing unbiased cell clusters from spermatids. H: t-SNE plot showing distribution of spermatids at D90 and D150. I: PAGA graph showing spermatid clusters. J: Visualization of selected spermatid marker gene expression levels across all single cells in t-SNE plot. K: Single-cell transcriptomes from spermatids were used for cell trajectories, ordered in latent time. L: Heatmap showing variable genes across latent time from identified spermatid cell types.



Supplementary Figure S6 X chromosome dynamics during porcine spermatogenesis. A: Line plot showing mean number of genes expressed from every chromosome in each cell type during porcine spermatogenesis. (Chromosomes are colored according to legend) B: Heatmap showing all expressed genes on pig X chromosome during spermatogenesis. C: t-SNE plot showing gene expression of SCML2 and SETDB1. D: Immunofluorescence staining of SETDB1 in porcine testis, suggesting SETDB1 plays a critical role in porcine meiosis. Scale bars=50 μ m.



Supplementary Figure S7 Comparative analysis of germ cells in pigs and humans. A: t-SNE plot showing distribution of germ cells in human datasets. B: t-SNE plot showing identified germ cell types in humans (colored by 11 cell types, i.e., Undiff1, Undiff2, Diff1, Diff2, Lep, Zyg, Pachy, Dip & Sec, RS, ES, and Sperm). C: Distribution of cell attributes from each cell type in human datasets. D: Correlation analysis and hierarchical clustering of germ cell clusters across two species. E: PAGA graph of cluster relationships from integrated dataset.