


RESEARCH

Open Access



Single-cell RNA-sequencing reveals the dynamic process and novel markers in porcine spermatogenesis

Lingkai Zhang¹, Fuyuan Li¹, Peipei Lei¹, Ming Guo¹, Ruifang Liu¹, Ling Wang², Taiyong Yu¹, Yinghua Lv³, Tao Zhang², Wenxian Zeng^{1*} , Hongzhao Lu^{2*} and Yi Zheng^{1*}

Abstract

Background: Spermatogenesis is the process by which male gametes are formed from spermatogonial stem cells and it is essential for the reliable transmission of genetic information between generations. To date, the dynamic transcriptional changes of defined populations of male germ cells in pigs have not been reported.

Results: To characterize the atlas of porcine spermatogenesis, we profiled the transcriptomes of ~ 16,966 testicular cells from a 150-day-old pig testis through single-cell RNA-sequencing (scRNA-seq). The scRNA-seq analysis identified spermatogonia, spermatocytes, spermatids and three somatic cell types in porcine testes. The functional enrichment analysis demonstrated that these cell types played diverse roles in porcine spermatogenesis. The accuracy of the defined porcine germ cell types was further validated by comparing the data from scRNA-seq with those from bulk RNA-seq. Since we delineated four distinct spermatogonial subsets, we further identified CD99 and PODXL2 as novel cell surface markers for undifferentiated and differentiating spermatogonia, respectively.

Conclusions: The present study has for the first time analyzed the transcriptome of male germ cells and somatic cells in porcine testes through scRNA-seq. Four subsets of spermatogonia were identified and two novel cell surface markers were discovered, which would be helpful for studies on spermatogonial differentiation in pigs. The datasets offer valuable information on porcine spermatogenesis, and pave the way for identification of key molecular markers involved in development of male germ cells.

Keywords: Marker, Pig, scRNA-seq, Spermatogenesis

Background

A boar produces several billions of spermatozoa daily [1]. The highly efficient production of spermatozoa relies on spermatogenesis that comprises three stages, i.e., mitosis of spermatogonia, meiosis of

spermatocytes, and spermiogenesis (transformation of spherical spermatids) [2]. Mitotic divisions of spermatogonia include self-renewal of spermatogonial stem cells (SSCs) and cell differentiation that gives rise to differentiating type A spermatogonia, intermediate and B spermatogonia. Type B spermatogonia differentiate into spermatocytes that undergo two consecutive rounds of cell divisions and form round spermatids [3, 4]. Subsequently, round spermatids undergo morphological and biochemical changes to form the elongated spermatids that are released into the lumen of seminiferous tubules as spermatozoa [2, 5].

* Correspondence: zengwenxian2015@126.com; lhz780823@snut.edu.cn; yzheng@nwafu.edu.cn

¹Key Laboratory for Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling 712100, Shaanxi, China

²School of Biological Science and Engineering, Shaanxi University of Technology, Hanzhong 723001, Shaanxi, China

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

The complicated process of spermatogenesis requires timely coordinated gene expression [6]. Genome-wide microarray and RNA-sequencing (RNA-seq) techniques have been applied to analyze the enriched male germ cell populations or testis samples to obtain the transcriptomes [7–9]. As the process of spermatogenesis is coordinated radially within a cross section of seminiferous tubules and occurs asynchronously along the tubules, there are multiple germ cells at different stages of spermatogenesis, which makes it challenging to obtain stage-specific molecular resolution of germ cell differentiation. As a result, the molecular characterization of spermatogenic cells at a defined stage remains unclear.

A few approaches have been applied to separate the different subtypes of male germ cells. The STA-PUT technique has been used to isolate germ cells in mice, rats, pigs, and cattle, but it usually fails to obtain distinct cell types with high purity [10–14]. The magnetic-activated cell sorting (MACS) and fluorescence-activated cell sorting (FACS) depend on biochemical markers of cells and cell purity is largely affected by the specificity of the chosen antibodies. Moreover, MACS and FACS are unsuitable for separation of spermatocytes or spermatids due to the lack of suitable markers. Recently, the laser capture micro-dissection approach has been used to precisely capture subtypes of male germ cells from paraffin tissue sections, but it leads to RNase contamination after staining [15]. Fortunately, 10× Genomics single-cell RNA-sequencing (scRNA-seq) is an unbiased approach with high efficiency and sensitivity in cell capture, and it has been used to decipher the dynamic process of spermatogenesis in mouse, primate and sheep testis [16–20]. However, the gene expression profiling of spermatogenic cells at each sub-stage during spermatogenesis and their regulatory networks in pigs have not been uncovered.

Pigs (*Sus scrofa*) are a leading domestic species for meat production. Meanwhile, they are increasingly exploited as an animal model in physiological and pharmacological research [21, 22]. In this study, we profiled ~ 16,966 single-cell transcriptomes from the testis of a 150-day-old pig. This large number of cells allowed us to well identify the populations of germ cells and somatic cells. Using bulk RNA-seq, we provided further supportive evidence on the definition of germ cell clusters. We delineated four distinct spermatogonial subsets in porcine testes and identified novel cell surface markers for undifferentiated and differentiating spermatogonia. The datasheets from the present study could enrich our understanding of spermatogenesis and male reproduction in pigs.

Methods

Animals

All experiments utilizing animals were acquired from the farm, Northwest A&F University, Yangling, China.

Adult pig testicular tissue was castrated for following experiments. Procured tissue was transported to the laboratory on ice phosphate-buffered saline (PBS; Invitrogen, Carlsbad, CA, USA) with a double dose penicillin/streptomycin (Invitrogen). In all cases, a portion of the testicular tissue was fixed in 4% paraformaldehyde and Bouin's solution. The bulk of the tissue was used for cell isolation as noted below.

Generation of cell suspensions

Single-cell suspensions of seminiferous tubules were generated from a 150-day-old Guanzhong pig testis using a two-step enzymatic digestion approach [23]. Briefly, testicular parenchyma was digested with 1 mg/mL Collagenase Type IV (Invitrogen) for 30 min at 34 °C, washed with Dulbecco's phosphate-buffered saline (DPBS; Invitrogen) to remove interstitial cells, following, the seminiferous tubules were digested with 0.25% trypsin/EDTA (HyClone, Logan, UT, USA) for 5 min at 34 °C, and terminated with 10% FBS. Cell suspensions were obtained by filtering through 40 µm strainers and washed up by DPBS with 0.5 mg/mL DNase I (Sigma, San Francisco, CA, USA) for twice, then suspended in 1640/DMEM (Hyclone) containing 5% FBS.

Porcine germ cell enrichment by STA-PUT

Cells from adult pig seminiferous tubules were enriched for spermatogonia, spermatocyte or spermatid based on sedimentation velocity at unit gravity [24]. Three adult pigs were used to obtain three independent biological replicates for spermatogonia, spermatocytes and spermatids. Briefly, testis cells suspended in 50 mL buffer plus 0.5% bovine serum albumin (BSA) were loaded onto a 600 mL gradient of 2–4% BSA and allowed to sediment for 3 h at 4 °C. Approximately one hundred 6 mL fractions were collected in plastic tubes and analyzed for content of spermatogonia, spermatocyte or spermatid on the basis of morphology under contrast optics (which typically yields ≥ 80% purity). Fractions containing spermatogonia, spermatocyte or spermatid were pooled separately, concentrated (to ~ 2 × 10⁶ cells/mL) and stored in buffer containing FBS on ice until use.

Single-cell transcriptomes

Cell suspensions were loaded into Chromium microfluidic chips with 3' v3 chemistry and used to generate single-cell gel bead emulsions (GEMs) using the Chromium controller (10 × Genomics) per manufacturer recommendations. In all cases, suspensions containing ~ 20,000 cells were loaded on the instrument with the expectation of collecting up to 16,971 GEMs containing single cells. GEM-RT was performed in a T100 Thermal

cycler (Bio-Rad) and all subsequent steps to generate single-cell libraries were performed according to manufacturer recommendations. Libraries were sequenced at the Genome Sequencing Facility (GSF) by LC-Bio Technology Co., Ltd (HangZhou, China) on an Illumina NovaSeq 6000 sequencing system. Trimmed FASTQ files (26 bp cell barcode and UMI Read1, and 100 bp Read2), were generated using the CellRanger mkfastq command (a 10 × Genomics wrapper around BCL2Fastq). Primary data analysis (alignment, filtering, and UMI counting) to determine gene transcript counts per cell (producing a gene-barcode matrix) were performed using CellRanger count (10 × Genomics) and Sscrofa11.1 genome assembly/annotation references. All visualization results were completed by R.

Single-cell RNA-seq analysis

Raw count matrices (10 × Genomics) (Fluidigm C1) were imported to Seurat (v.3.2.0) [25], filtered (cells expressing >500 detected genes, genes expressed in >3 cells) and gene expression values were log normalized and scaled, then analyzed by principal component analysis (PCA). The first 15 principal components (PCs) were chosen to construct a K-nearest neighbors (KNN) graph and refine the edge weights between any two cells. Then we used “RunTSNE” function with the resolution parameter set as 0.7 to cluster cells, which identified 20 clusters and these clusters were renamed by accepted well known marker genes. The top 10 differentially-expressed genes (marker genes) of each cell cluster were determined by log fold change > 0.25 using default Wilcoxon rank-sum test [26]. The raw count matrices with cluster subsets (e.g., without testicular somatic cells) were imported to Monocle (v 2.10.0) [27] and used for additional combined cluster info. Differentially expressed genes or significantly variable genes among cells were identified and used for dynamic trajectory analysis which ordered cells in “pseudotime analysis”. With the gene count matrix as input, applied the function of “reduceDimension” and “orderCells” were carried out to generate the cell trajectory based on pseudotime. Pseudotime DEG analysis was performed using Monocle with a *P*-value cutoff of <0.01. All heatmaps were generated in pseudotime order, and line plots were plotted in pseudotime order with fitted curves in the method of “auto” by ggplot2 package [28]. For further analysis, we isolated spermatogonia, differentiating spermatogonia and leptotene/Zygotene spermatocyte cells and performed these steps again; we obtained 10 clusters and renamed by accepted well known marker genes. Seurat was then used to perform differential expression analysis of merged matrices based on the non-parametric Wilcoxon rank-sum test.

RNA-seq analysis

Total RNA was isolated and purified using Trizol reagent (Invitrogen) following the manufacturer’s procedure. The RNA quantities were measured using NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA, USA). The RNA integrity was assessed by Aligent 2100 with RIN number >7.0. RNA high throughput sequencing was performed by Cloud-Seq Biotech (Shanghai, China). Briefly, total RNA was used for removing rRNAs with NEBNext rRNA Depletion Kit (New England Biolabs) following the manufacturer’s instructions. RNA libraries were constructed by using NEBNext® Ultra™ II Directional RNA Library Prep Kit (New England Biolabs) according to the manufacturer’s instructions. Libraries were controlled for quality and quantified using the BioAnalyzer 2100 system (Agilent Technologies, Palo Alto, CA, USA). Library sequencing was performed on an illumine Hiseq instrument with 150 bp paired end reads. Paired-end reads were harvested from Illumina Hiseq-4000 sequencer, and were quality controlled by Q30. After 3’ adaptor-trimming and low qualities reads removing by cutadapt software (v1.9.3), the high qualities clean reads were aligned to the reference genome (Sscrofa11.1) with hisat2 software (v2.0.4). Then, guided by the Ensembl gtf gene annotation file, HTSeq software (v0.9.1) was used to get the raw count, and edgeR was used to perform normalization [29]. The differentially expressed mRNAs were identified by *P*-value and fold change. Enrichment analysis was performed based on the differentially expressed genes.

Functional enrichment analysis with Metascape

We performed enrichment analysis of marker genes of each cell type in Metascape (<http://metascape.org>). For the given gene list, we converted these genes to human homology and performed with Metascape database based on human. The following ontology sources were chosen: GO Biological Processes, KEGG Pathway, Reactome Gene Sets and CORUM. All human genes had been used as the enrichment background. Cutoffs for significantly enriched terms were *P*<0.01, minimum count of 3 and an enrichment factor >1.5. The terms were grouped into clusters based on their membership similarities.

Immunofluorescence assay

The isolated spermatogonia, pachytene spermatocytes and round spermatids were fixed with 4% paraformaldehyde for 25 min at 4 °C and washed with PBS for three times. Then, the cells were permeabilized for 10 min using 0.1% Triton-X 100 (Sigma-Aldrich) followed by washing with PBS for three time. The cells were further blocked with 10% donkey serum for 2 h at room

temperature, and incubated with primary antibodies, including UCHL1 (Abcam, Catalog No. ab8189), SYCP3 (Abcam, Catalog No. ab15093) and CD63 (Proteintech, Catalog No. 25682-1-AP) at a dilution with 1:400 overnight at 4 °C. On the next day, the cells were washed with PBS for four times and incubated with secondary antibody (Yeason) at a dilution with 1:400 for 1 h at room temperature. The nuclei were labelled with DAPI (Bioworld Technology, St. Louis Park, MN, USA). A fluorescence microscope (Tokyo) was used for fluorescence observation and photographing.

The testis tissue was fixed in Bouin's solution for 16 h, embedded in paraffin and sectioned. Sections were deparaffinized, rehydrated and heat-mediated antigen retrieval in 10 mmol/L sodium citrate buffer solution. After treatment with PBS for four times in 10 min, individual sections were incubated overnight at 4 °C. The following primary antibodies were used: CD99 (Proteintech, Catalog No.23079-1-AP), PODXL2 (R&D Systems, Catalog No.AF3534), KIT (Cell Signaling technology, Catalog No.3074S), UCHL1 (Abcam, Catalog No. ab8189), SYCP3 (Abcam, Catalog No. ab15093). Antigen reaction was conducted using with secondary antibody (Yeason) for 2 h at 4 °C in the dark. Then, the sections were incubated with DAPI (Bioworld Technology) to facilitate nuclear visualization. Images were obtained with Nikon Eclipse 80i fluorescence microscope camera (Tokyo).

Ethical statements

All experimental procedures involving animals were approved by the Northwest A&F University's Institutional Animal Care and Use Committee.

Data availability

The datasets generated in the current study are available in the NCBI GEO database under the accession number GSE174782.

Results

Cell partitioning through the analysis of porcine single cell transcriptomes

Previous research on gene expression during spermatogenesis has relied largely on RNA-seq analysis of bulk cell populations. However, the heterogeneity in each population remains elusive. To uncover the heterogeneity, we dissociated porcine testis tissue from a 150-day-old Guanzhong black pig into single cells and performed 10 × Genomics scRNA-seq (Fig. 1A). We isolated single cells using a standard two-step enzymatic digestion procedure and removed the erythrocytes. The viability of the cells was about 95%. The cells were then used to construct the single-cell library. From a total of ~ 16,966 cells, 16,791 (99.97%) passed the standard quality control

(QC) dataset filters and were retained for downstream analysis. The 842 Mb reads were obtained from the libraries, and the sequencing saturation rate was 30.40%. We obtained ~ 49.62 K reads/cell, and each cell had 19,690 unique molecular identifiers (UMIs) and 2759 genes detected (supplementary Table1). The cell-level quality metrics for the data were computed and shown in Fig. S1A and Fig. S1B.

To overcome the extensive technical noise in single genes of scRNA-seq data, we performed principal component analysis (PCA) on the matrix of scRNA-seq expression and selected the top 15 components for subsequent analysis. Cell partitioning via t-distributed stochastic neighbor embedding (t-SNE) analysis identified 20 clusters (Fig. 1B), and the number of the cells in each cluster was shown in supplementary Table1. As the markers of porcine male germ cells were not fully available, we drew on the markers in humans. The porcine protein coding genes were mapped to their human 1:1 orthologs. Based on these markers, 20 clusters were defined as 6 broad cell types, while other clusters were classified as unknown cell types (Fig. 1C). Specifically, germ cells and somatic cells could first be distinguished by the mutually exclusive expression of two genes, *DDX4* in germ cells and *VIM* in somatic cells as previously reported [19]. The germ cells with *DDX4* expression were further sub-classified as follows: spermatogonia (*UCHL1*, *ZBTB16*, *KIT*, *STRA8* and *DMRT1*), spermatocytes (*DMC1*, *RAD51AP2*, *HORAD1*, *SYCP1*, *PIWIL1* and *NME8*) and spermatids (*FHL5*, *TPPP2*, *PRMI* and *TNPI*). Somatic cells were identified: Sertoli cells (*GATA4*, *SOX9*, *AR* and *WT1*), Leydig cells (*CYP11A1*, *CYP17A1* and *STAR*), and myoid cells (*MYH11* and *ACTA2*) (Fig. 1E, Fig. S1D).

The marker genes were calculated by Seurat, which revealed the underlying features of gene expression in each cell type. The numbers of cells and genes in each cell type were shown in Fig. 1F. Dynamic changes in expression of the marker genes were presented Fig. S1C, in accordance with the previous reports in humans [17, 18]. To identify the functional categories associated with each cell type, we utilized the Metascape (<http://metascape.org>) for functional enrichment analysis, which provided further insights into the functional theme and regulatory network in porcine somatic cells and germ cells. For instance, genes in spermatogonia were enriched in cell cycle, DNA repair and transcriptional regulation by *TP53*, indicating the active proliferation of spermatogonia. The genes in spermatocytes were involved in meiotic cell cycle and regulation of chromosome organization. The genes in spermatids represented more downstream processes, such as spermatogenesis, fertilization and cilium organization (Fig. 1D). As to the testicular somatic cells, genes in myoid cells were

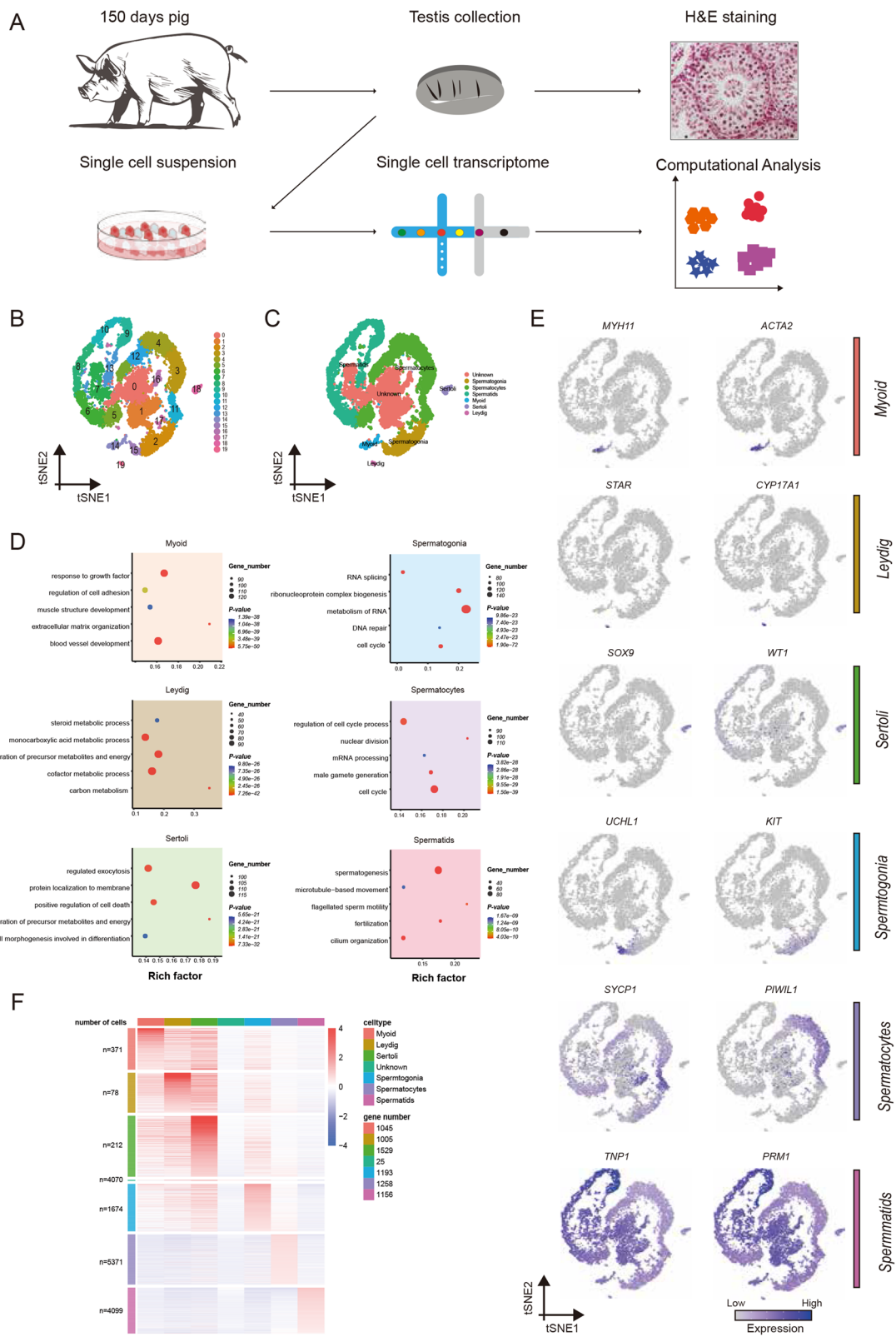


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 Single-cell transcriptome profiling and the cluster identification of porcine testicular cells. **(A)** Overview of the experimental design yielding droplet-based scRNA-Seq from porcine testis while using the contralateral testis for matched histology. **(B)** t-SNE plots show the 10 × Genomics profiling of unselected spermatogenic cells from porcine testis. Unbiased cell clusters are distinguished by color according to the key. **(C)** t-SNE plots show the identified cell types (cells are colored by the 6 broad cell types, and unknown cell type). **(D)** Enrichment terms and *P* values for the identified cell types (terms are colored by the 6 broad cell types). **(E)** Visualization of selected marker gene expression across all single cells in the t-SNE plot. For each cell cluster, 2 cell markers are shown in the Fig. **(F)** Scaled, normalized expression counts of the marker genes per cell type. Column and Row labels represent the cell type and the number of genes

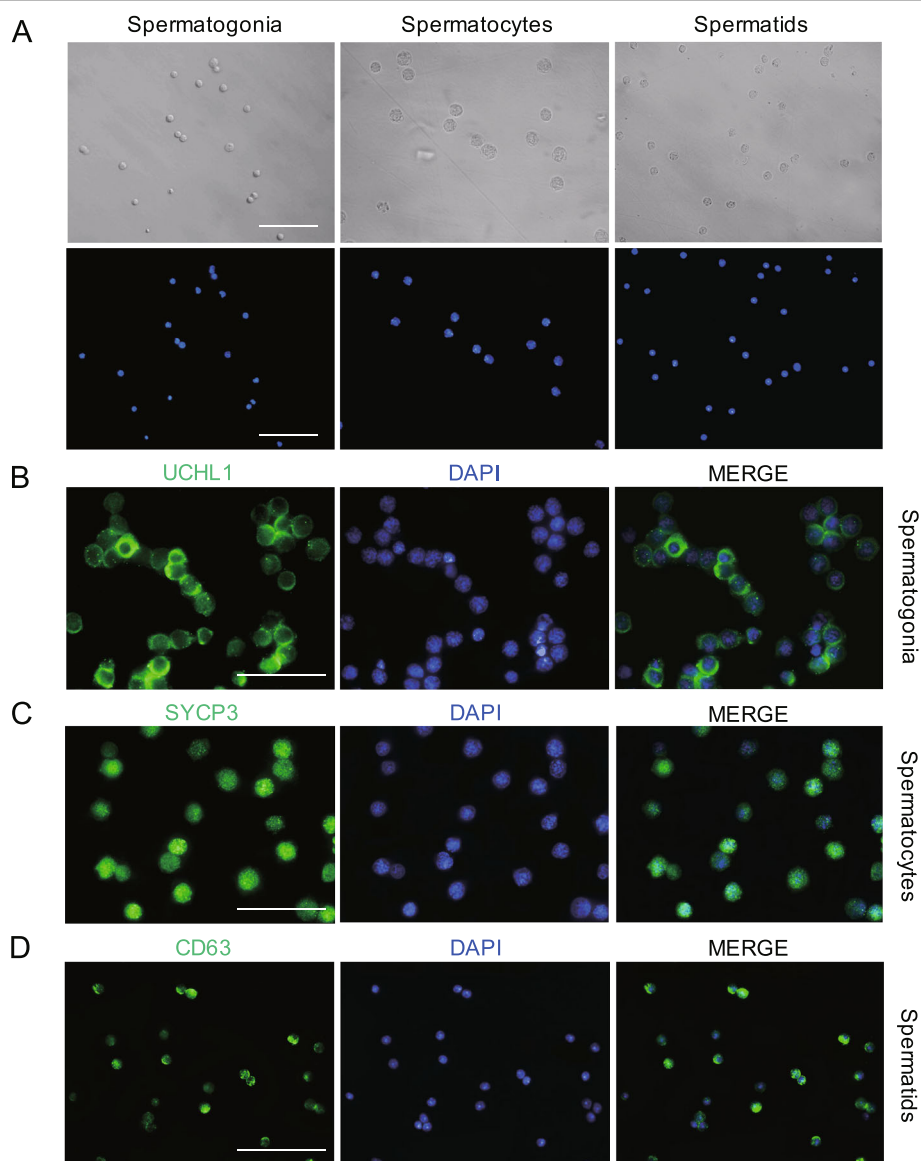


Fig. 2 Identification of porcine male germ cells. **(A)** The morphology of the freshly isolated porcine spermatogonia, spermatocytes and spermatids. Scale bar = 50 μm. **(B)** Immunocytochemistry staining for UCHL1 in spermatogonia. **(C)** Identification of SYCP3 the freshly isolated spermatocytes. **(D)** CD63 was identified to express in the spermatids. Scale bar = 50 μm

enriched with response to growth factor, extracellular matrix organization and regulation of cell adhesion. The genes in Leydig cells were involved in monocarboxylic acid metabolic process and steroid metabolic process, consistent with the main function of Leydig cells for testosterone synthesis. The genes in Sertoli cells were related to regulation of cell death and exocytosis. Thus, the enrichment analysis enables more knowledge about the potential function of these major cell types.

Isolation and assay of porcine spermatogonia, spermatocytes and spermatids

To link our computationally-defined clusters to morphologically-defined germ cell types, we collected porcine spermatogonia, spermatocytes and spermatids using STA-PUT, a velocity sedimentation approach. Different cell types can be distinguished based on the cell size and nuclear morphology. Spermatogonia are 10 μm in diameter and have ovoid nuclei, and spermatocytes are about 15 μm in diameter with patchy condensed chromatin and large nuclei, while spermatids are approximately 5 μm in diameter and have a round nucleus with a high nucleus-to-cytoplasm ratio and highly condensed DNA (Fig. 2A). The viability of the freshly isolated porcine spermatogonia, spermatocytes and spermatids were over 95%.

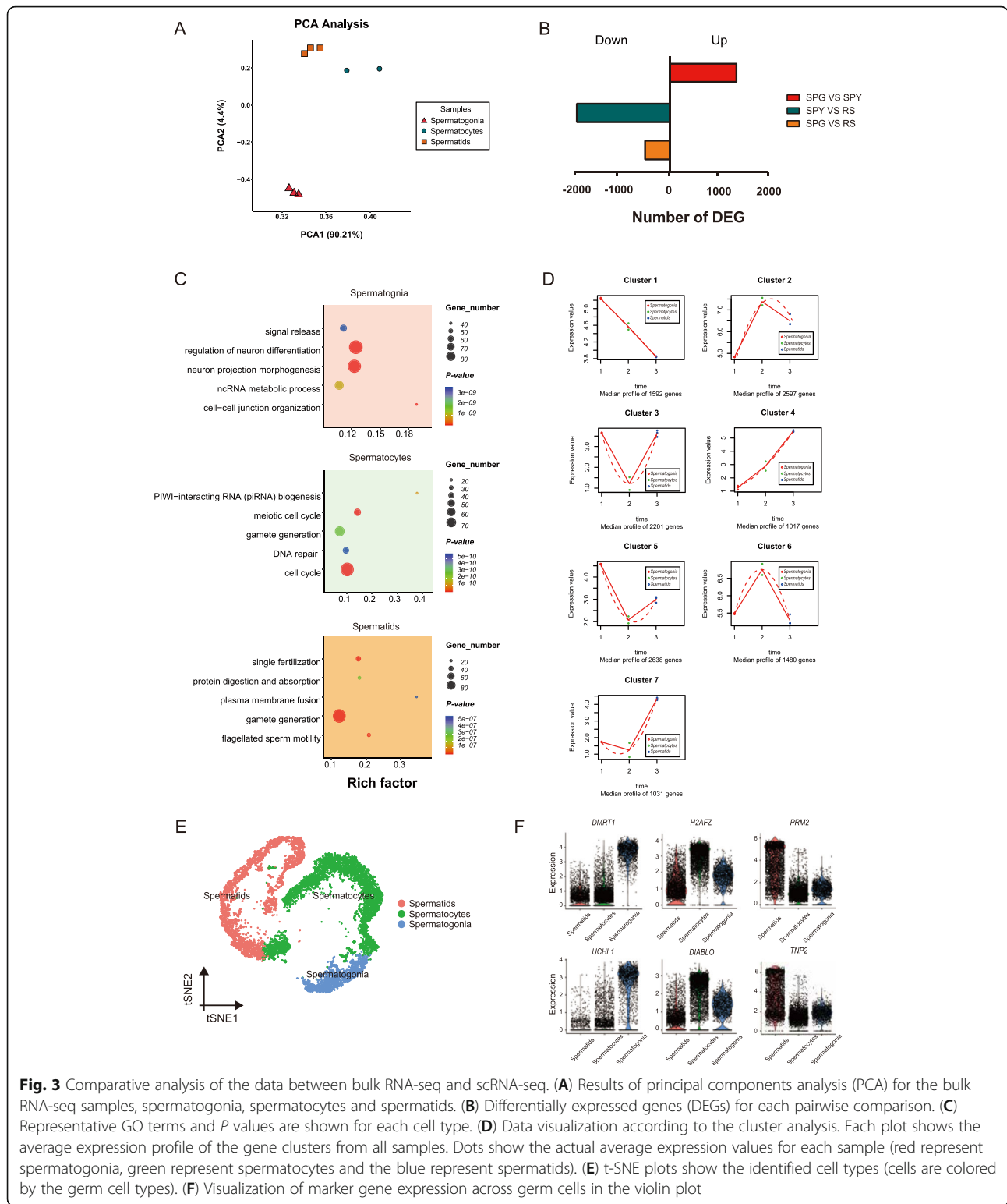
Subsequently, immunofluorescence staining was performed to detect the cell purity. UCHL1 was selected as the marker of spermatogonia (Fig. 2B). SYCP3, the axial/lateral elements of synaptonemal complex proteins, was selected as the marker of spermatocytes (Fig. 2C). Interestingly, we found that CD63, a cell surface protein, could be used as a new marker for porcine spermatids, as confirmed by immunofluorescence and immunohistochemistry analyses (Fig. S2A, Fig. S2B). Consequently, CD63 was chosen to identify the purity of the collected spermatids (Fig. 2D). Overall, the purity of isolated germ cells was more than 80%, suggesting that spermatogonia, spermatocytes and spermatids were successfully enriched.

Verification of porcine germ cell types by bulk RNA-seq

Upon successful enrichment of spermatogenic populations, we performed bulk RNA-seq to validate the scRNA-seq result, which could exhibit the similar gene expression and GO results between scRNA-seq and bulk RNA-seq in spermatogonia, spermatocytes and spermatids. For bulk RNA-seq analysis, we constructed transcriptomic libraries from 3 biological replicates per cell type. We calculated the Pearson correlation coefficient between the samples and the annotated cell types in scRNA-seq dataset, and one sample of spermatocytes was excluded due to the low repeatability (Fig. S3B).

The reproducibility for other samples was validated by PCA analysis (Fig. 3A). The total number of expressed genes exhibited dynamic changes during porcine spermatogenesis, with 12,974 expressed in spermatogonia, 11,456 in spermatocytes, and 13,220 in spermatids. The abundance of protein coding genes exhibited considerable variations in each cell type (Fig. S3A). Pairwise differential gene expression analysis between cell types generated lists of differentially expressed genes (DEGs; fold change > 2, FDR < 0.05). A total of 9919 DEGs were identified between spermatogonia and spermatocytes, of which 5626 were upregulated and 4293 were downregulated. From spermatocytes to spermatids, there were 2799 upregulated genes and 4742 downregulated genes. From spermatogonia to spermatids, 3680 were upregulated and 4126 were downregulated. Consistent with previous reports in mice [30], gene expression of the protein coding genes had global shutdown during meiosis, and the pairwise comparisons were more negative in spermatocytes, suggesting the lower transcriptional activity in spermatocytes (Fig. 3B).

To explore the possible functions of stage-specific genes, maSigPro was used to perform time series analysis [31]. The analysis revealed that 12,556 variable genes were clustered into seven groups (Fig. S3C). Among them, the trends of gene expression manifested that the genes in cluster 1 were specifically expressed in spermatogonia, while cluster 6 and cluster 7 matched spermatocytes and spermatids, respectively (Fig. 3D). In addition, there were 1592, 1480, and 1031 stage-specific genes in spermatogonia, spermatocytes, and spermatids, respectively. Subsequently, we performed functional enrichment analysis for the genes in these clusters. The upregulated genes in spermatogonia were enriched with ncRNA metabolic process. The 1480 upregulated genes in spermatocytes were involved in meiotic cell cycle, PIWI-interacting RNA (piRNA) biogenesis and gamete generation. Finally, Metascape analysis revealed that the upregulated genes in spermatids were enriched with single fertilization and flagellated sperm motility (Fig. 3C). Taken together, the functional enrichment results showed the similarity of the identified cell types by scRNA-seq to those by bulk RNA-seq. In addition, we clustered porcine germ cells in scRNA-seq (Fig. 3E), and matched the bulk and scRNA-seq data by using classification of the bulk samples based on stage-specific marker genes obtained from the defined cell types in scRNA-seq data (Fig. 3F). The Pearson correlation results further verified our correct classification (Fig. S3B). Thus, by comparing bulk RNA-seq data to those of scRNA-seq, we validated the accuracy of the defined porcine germ cell types.



Cellular heterogeneity during porcine spermatogonial differentiation

To capture the fate transition in porcine germ cells, we re-clustered the clusters 15, 2, 17 and 11 that

represented spermatogonia and early meiotic spermatocytes to analyze the subgroups of spermatogonial differentiation and meiotic entry. Cells were ordered via the pseudo-time method, and the analysis determined

(See figure on previous page.)

Fig. 4 Single-cell porcine spermatogonia trajectories reveal transitions with differentiation and initiation of meiosis. **(A)** Focused analysis (t-SNE and clustering) of the mitosis and initiation of meiotic germ cells reveals differentiation progression of spermatogonia. **(B)** t-SNE plot reveal the states in spermatogonia differentiation. **(C)** Dot plot depicts the expression of marker genes in each state. The dot size represents the percentage of cells expressed and the color intensity represents relative expression level. **(D)** The correlation across each state and illustrated by Pearson correlation analysis. **(E)** Expression patterns of new defined marker genes for undifferentiated and differentiating spermatogonia visualized in t-SNE plots. **(F)** Heatmap of marker genes from each state. Top, cell types; Left, number of marker genes; Right, representative GO terms and representative genes

trajectories of differentiation according to the marker gene expression patterns. The pseudo-time analysis provided an arrow vector which aligned with the developmental order of spermatogonia (Fig. S4A). The top 50 variable genes across pseudo-time were shown in Fig. S4B. Next, we performed re-clustering of these cell types, which revealed nine sub-clusters (Fig. 4A). Based on the known markers (Fig. S4C), we detected one cluster (Un-diff) that pertained to undifferentiated spermatogonia based on the expression of *UCHL1* and a small number of cells that expressed SSC markers (*ZBTB16*, *ETV5*, *ID4* and *FGFR3*) [32–34]. Moreover, based on expression of *CD81*, *PCNA*, *KIT*, *STRA8*, *TSPAN33* and *DMRT1*, we assigned five clusters to three stages of differentiating spermatogonia (Diff-1, Diff-2 and Diff-3) [18, 35–37]. The expression of *PCNA* and *DMRT1* suggests that the differentiating spermatogonia are highly proliferative, which mediates the transition of mitosis to meiosis [38, 39]. Besides, marker genes involved in the synaptonemal complex (*SYCP1*, *SYCP2* and *SYCP3*) and meiotic recombination (*DMC1*) identified three clusters as pre-leptotene spermatocytes (Pre-Leptotene), leptotene spermatocytes (Leptotene) and zygotene spermatocytes (Zygotene) [40]. Thus, we obtained cell-type identities for spermatogonial subpopulations and revealed the early stages of meiosis (Fig. 4B). The expression of marker genes in each cell types was shown in Fig. 4C. In addition, we analyzed the correlation of the identified cell types by Pearson analysis, and found low correlation between mitosis and meiosis (Fig. 4D).

Next, we selected marker genes in each cluster and the dynamic changes of gene expression were shown in Fig. 4F. To investigate the different functions of these clusters, functional enrichment analysis of marker genes was performed (Fig. 4F). The marker genes in Un-diff were significantly enriched with generation of precursor metabolites and energy and transcriptional regulation by *TP53*. Diff-1 genes-enriched terms included cell cycle, nucleoside monophosphate metabolic process and metabolism of RNA, consistent with the notion that differentiating spermatogonia are undergoing proliferative expansion [41]. Significantly enriched terms for Diff-2 genes included eukaryotic translation elongation and translation initiation complex formation. Remarkably, the marker genes in Diff-3 were involved in meiotic

synapsis and chromosome segregation, which suggested that the genes at this stage point to subsequent entry into meiosis. In addition, the marker genes in Pre-Leptotene, Leptotene and Zygotene were significantly enriched with meiotic process and gamete development. Together, these results revealed that many biological processes are active in porcine spermatogonial subsets.

To further explore cell-type specific genes for porcine undifferentiated and differentiating spermatogonia, we filtered a series of marker genes in each cluster. We analyzed the dynamic expression changes of these genes during spermatogonial differentiation (Fig. 4E). The analysis indicated that *SDC2*, *PTPRJ*, *PIANP*, *NDRG2*, *CD99* and *CCND2* were specifically expressed in Un-Diff, while *RARG*, *PODXL2*, *CTCFL* and *CDH2* were highly expressed in Diff-1 ~ 3. In addition, *NRN1* and *TCEA3* were highly expressed in both Un-Diff and Diff-1 ~ 3. Remarkably, these genes were specifically expressed in spermatogonia (Fig. S4D). Thus, our analysis confirmed many known genes and provided a list of key candidate genes with undefined functions in porcine spermatogonial development, thereby facilitating future studies in this respect.

Identification of cellular localization for CD99 and PODXL2 in porcine testes

To identify the novel markers from the candidate genes (Fig. 4E), we performed immunohistochemistry staining for CD99 and PODXL2. Immunohistochemical analysis revealed that CD99 and PODXL2 localized in spermatogonia adjacent to the basement membrane of seminiferous tubules (Fig. 5A). According to the single-cell RNA-seq data, CD99 and PODXL2 were expressed in different subpopulations of porcine spermatogonia. Of this, CD99 was primarily enriched in undifferentiated spermatogonia, while PODXL2 is preferentially expressed in differentiating spermatogonia. To validate this, we performed double immunofluorescence staining for CD99/*UCHL1* and PODXL2/*KIT* (Fig. 5B, C). The result showed that CD99 staining partly overlapped with that of *UCHL1*, an undifferentiated spermatogonial marker (96.03% ± 1.02%), indicating that CD99⁺ cells represent a subset of undifferentiated spermatogonia (Fig. 5D). The colocalization analysis for PODXL2 and *KIT* showed that

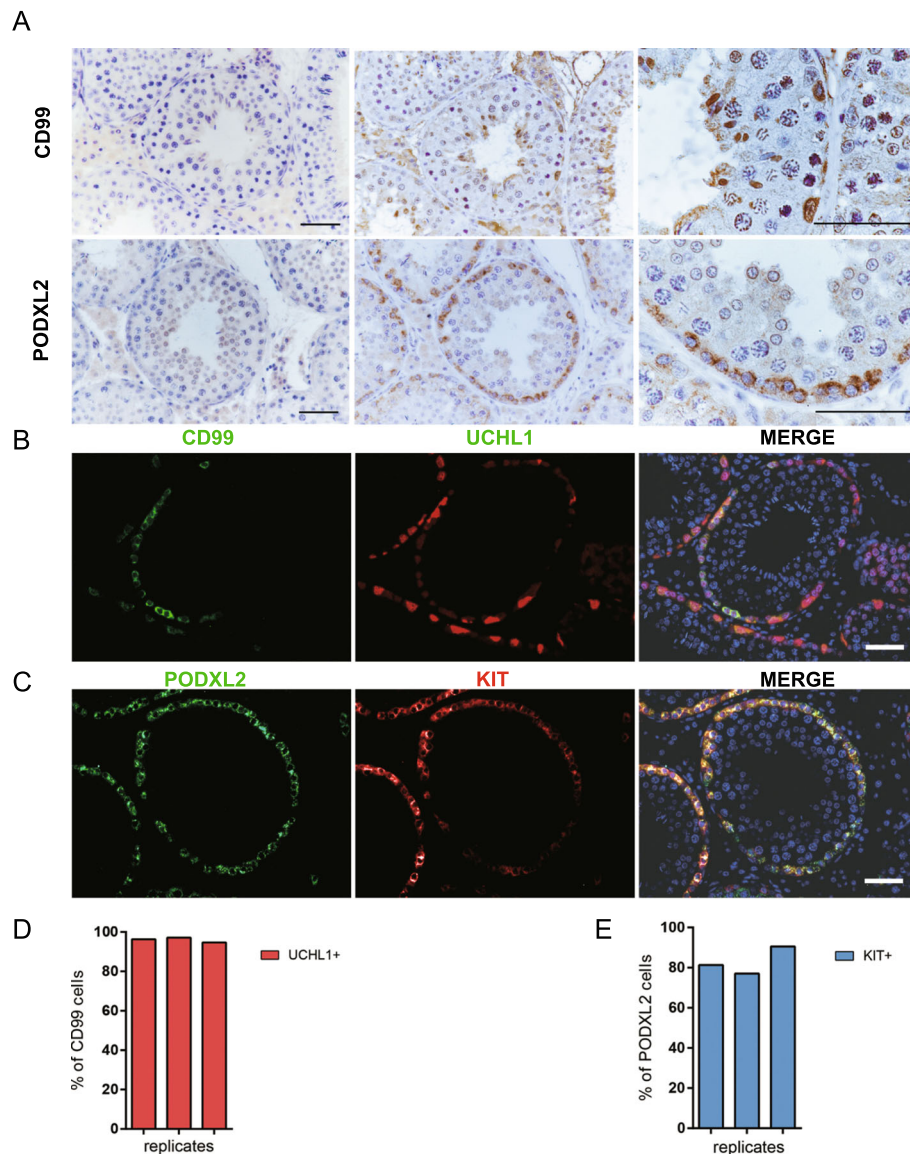


Fig. 5 Expression of CD99 and PODXL2 in porcine testis. **(A)** Immunohistochemical staining of CD99 and PODXL2 in porcine testis. Scale bars = 50 μ m. Coimmunofluorescence examination of the expression of **(B)** CD99 with UCHL1 and **(C)** PODXL2 with c-KIT in porcine testis. Scale bars = 50 μ m. Data are normalized by the number of double-positive cells for three independent experiments **(D)** CD99 with UCHL1 and **(E)** PODXL2 with c-KIT

the majority of PODXL2⁺ cells were also positive for KIT, a differentiating spermatogonial marker (82.94% \pm 5.66%), suggesting that PODXL2 marks the differentiating spermatogonial population in porcine testes (Fig. 5E). Taken together, our result revealed that CD99 and PODXL2 could serve as reliable markers for porcine spermatogonial subpopulations.

Discussion

Mammalian spermatogenesis is a complicated and dynamic cellular differentiation process. Gene expression profiles of male germ cells in mice and humans have

been reported previously [16, 18, 42]. However, the information in pigs remains largely unclear. In this study, we used scRNA-seq and performed a comprehensive analysis of gene expression during porcine spermatogenesis. Our analysis of \sim 16,966 individual testicular cells from the adult porcine testis afforded the maximum resolution of gene expression patterns associated with the full collection of cells within the spermatogenic cell lineage. We identified the subtypes of male germ cells and testicular somatic cells. Our single-cell data corroborate that spermatogenesis represents a continual process of ever-changing gene expression profiles.

Catalog of germ cells and somatic cells

Based on the germ cell marker *DDX4* and somatic cell marker *VIM* [19], the testicular cells could be first separated. Overall, we identified all major male germ cell types in porcine testes, including spermatogonia, spermatocytes and spermatids. The dynamic changes in expression of the marker genes were consistent with the previous reports in humans [17, 18]. In addition, we identified Sertoli cells, Leydig cells and myoid cells (Fig. 1E, Fig. S1D). The marker genes in myoid cells were enriched with response to growth factor, extracellular matrix organization and regulation of cell adhesion. The genes in Leydig cells were involved in synthesis for testosterone synthesis [43]. In Sertoli cells, the marker genes were involved in regulation of cell death and exocytosis.

In the present study, we identified spermatogonia, spermatocytes, spermatids, Sertoli cells, Leydig cells and myoid cells through scRNA-seq. Recently, Yang et al. reported characterization of sheep spermatogenesis via scRNA-seq [20]. However, spermatogonia cluster was not identified in their study. Here, by comparing scRNA-seq and bulk RNA-seq data, we further validated the accuracy of the defined porcine germ cell types.

Spermatogonial subsets

Our scRNA-seq analysis revealed the presence of 4 spermatogonial subsets: Un-diff, Diff-1, Diff-2, and Diff-3 SPG (Fig. 4B). Un-diff SPG matched undifferentiated spermatogonia based on expression of *UCHL1*, *ZBTB16*, *ETV5*, *ID4* and *FGFR3* [32–34]. In Diff-1 subset, compared to Un-diff, *UCHL1* expression decreased and expression of *KIT* (a marker for differentiating SPG) and *PCNA* (a marker for cell proliferation) increased. Diff-1 spermatogonia were enriched with genes involved in cell cycle, nucleoside monophosphate metabolic process and metabolism of RNA, which is consistent with the notion that differentiating spermatogonia are undergoing proliferative expansion [41]. Thus, Diff-1 subset may correspond to type A1-A4 spermatogonia [4, 44]. In Diff-2 subset, *KIT* expression remained at a high level, and *DMRT1* expression decreased slightly, while *STRA8* expression initiated (Fig. 4C, supplementary Table2). *DMRT1* promotes spermatogonial differentiation and controls the mitosis/meiosis switch [45, 46]. In mouse spermatogonia, *DMRT1* directly represses *STRA8* transcription and activates spermatogonial differentiation [47]. In Diff-3 subset, *DMRT1* expression decreased sharply, while *STRA8* expression further increased, and *SYCP2/3* essential for the assembly of the synaptonemal complex was upregulated, in agreement with that in humans [18]. As *DMRT1* and *STRA8* are reciprocally expressed in B-spermatogonia and preleptotene

spermatocytes [45], the Diff-2 subset is likely to represent intermediate and early type B spermatogonia and Diff-3 subset may represent type B spermatogonia at the late stage and ready for transition to preleptotene spermatocytes. Thus, this study has for the first time systematically investigated the spermatogonial subtypes in livestock. Our refined assignments of spermatogonial subtypes will facilitate future research on the mechanisms regulating the balance between SSC self-renewal and differentiation.

Spermatogonial markers

In the present study, the scRNA-seq analysis revealed twelve candidate spermatogonial markers. We explored the localization of CD99 and PODXL2 in porcine testes and found that they were localized in spermatogonia at the periphery of the seminiferous tubules (Fig. 5A). Double immunofluorescence staining for CD99/*UCHL1* revealed that 96.03 of CD99⁺SPG overlapped with *UCHL1*⁺ cells (Fig. 5D). CD99 is a cell surface glycoprotein that functions as a cell-cell adhesion factor [48]. Previous studies indicate that CD99 is a marker of stem cells in acute myeloid leukemia (AML) and myelodysplastic syndrome, and is also a marker for prospective separation of leukemic stem cells (LSCs) from functionally normal hematopoietic stem cells in AML [49]. Intriguingly, CD99 has been reported to be expressed in testes [50]. As CD99 colocalized with *UCHL1*, CD99⁺ cells could represent a subset of undifferentiated spermatogonia in pigs. It would be interesting to study the role of CD99 during spermatogonial proliferation and differentiation in future studies.

PODXL2 is a transmembrane type I sialomucin protein [51]. It promotes epithelial-to-mesenchymal (EMT) transition by disrupting cell-cell junctions [52, 53]. *PODXL2* depletion led to reduction of *OCT4* and *Nanog* expression and affected cell proliferation in BT474 cells [54]. Here, we demonstrated that the majority of *PODXL2*⁺ cells were also *KIT*⁺, suggesting that *PODXL2* marks the differentiating spermatogonial population in porcine testes. *KIT* is a well-established marker that can be used for isolation and enrichment of differentiating spermatogonia via fluorescence-activated cell sorting in mice and humans [55, 56]. However, in our preliminary study, several commercial c-kit antibodies could not be used for sorting porcine early germ cells. Thus, *PODXL2*, as a novel marker for differentiating spermatogonia, may be conducive to characterization of spermatogonial differentiation in pigs, and probably in other farm animals as well, which warrants future investigation. Future studies can also be carried out to explore the role of *PODXL2* in regulation of spermatogonial differentiation.

Conclusions

We identified four subsets of spermatogonia, spermatocytes, spermatids and three somatic cell types in porcine testes through scRNA-seq. Our scRNA-seq revealed twelve candidate spermatogonial markers, and two novel cell surface markers for porcine spermatogonia were validated, which would be conducive to characterization of spermatogonial differentiation in pigs. Collectively, our datasets offer valuable information on porcine spermatogenesis, and pave the way for identification of key molecular markers involved in development of male germ cells.

Abbreviations

AML: Acute myeloid leukemia; BSA: Bovine serum albumin; DEGs: Differentially expressed genes; FACS: Fluorescence-activated cell sorting; GEMs: Gel bead emulsions; KNN: K-nearest neighbors; Leptotene: Leptotene spermatocytes; LSCs: Leukemic stem cells; MACS: Magnetic-activated cell sorting; PCA: Principal component analysis; Pre-Leptotene: Pre-leptotene spermatocytes; QC: Quality control; RNA-seq: RNA-sequencing; scRNA-seq: Single-cell RNA sequence; SSCs: Spermatogonial stem cells; t-SNE: T-distributed stochastic neighbor embedding; Zygotene: Zygotene spermatocytes

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40104-021-00638-3>.

Additional file 1 supplementary Table1. List of data summary and the number of cells in each cluster.

Additional file 2 supplementary Table2. The expression of the genes involved in spermatogonia proliferation and differentiation in spermatogonia subsets

Additional file 3 Fig. S1 Data quality and cell type annotation by other genes (Relates to Fig. 1). (A) Distribution of cell attributes in the pig dataset. (B) The correlation between the number of genes and UMI counts. (C) Heatmap showed expression of 45 marker genes in 6 cell types. (D) Visualization of marker gene expression across unselected clusters in the violin plot. Unbiased cell clusters are distinguished by color according to the key.

Additional file 4 Fig. S2 Expression of CD63 in porcine testis. (A) Coimmunofluorescence examination of the expression of CD63 with PNA. Scale bars = 50 μ m. (B) Immunohistochemical staining of CD63 in porcine testis. Scale bars = 50 μ m.

Additional file 5 Fig. S3 (Relates to Fig. 3) (A) Density plot show the distribution of gene expression in spermatogonia, spermatocytes and spermatids. (B) The correlation between bulk and scRNA-seq data illustrated by Pearson correlation analysis. (C) Data visualization according to the cluster analysis. Each line shows a single gene expression from all samples.

Additional file 6 Fig. S4 (Relates to Fig. 4) (A) Single-cell transcriptomes from porcine undifferentiated spermatogonia, differentiating spermatogonia, Leptotene and Zygotene spermatocytes were used for cell trajectories ordered in pseudotime (right) and cells colored according to cell cluster (left). (B) Heatmaps show the top 50 variable genes across pseudotime from porcine undifferentiated spermatogonia, differentiating spermatogonia, Leptotene and Zygotene spermatocytes (scaled expression according to legend). (C) Visualization of marker gene expression across the differentiation states of spermatogonia in the violin plot. Unbiased cell clusters are distinguished by color according to the key. (D) Expression patterns of new defined marker genes for undifferentiated and differentiating spermatogonia visualized in t-SNE plots for all cell types.

Acknowledgements

We thank the farm, Northwest A&F University (Yangling, Shaanxi, China) for providing the porcine testis.

Authors' contributions

YZ, WXZ conceived the study. LKZ analyzed the data. FYL, PPL, RFL, MG, HZL and LW performed the experiments. TYY provided the resources. TZ and WXZ supported the funding. LKZ, WXZ and YZ wrote the manuscript. WXZ, HZL and YZ supervised the study and approved the final version. All authors read and approved the final manuscript.

Funding

This study was supported in part by the National Natural Science Foundation of China (Grant No. 31772605) to WXZ; Research Project of Shaanxi Science and Technology Department (2020NY-003) to TZ.

Availability of data and materials

All data supporting our findings are included in the manuscript.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Key Laboratory for Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling 712100, Shaanxi, China. ²School of Biological Science and Engineering, Shaanxi University of Technology, Hanzhong 723001, Shaanxi, China. ³College of Chemistry and Pharmacy, Northwest A&F University, Yangling 712100, Shaanxi, China.

Received: 14 May 2021 Accepted: 1 October 2021

Published online: 07 December 2021

References

- Okwun OE, Igboeli G, Ford JJ, Lunstra DD, Johnson L. Number and function of Sertoli cells, number and yield of spermatogonia, and daily sperm production in three breeds of boar. *J Reprod Fertil.* 1996;107(1):137–49. <https://doi.org/10.1530/jrf.0.1070137>.
- Oakberg EF. 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.* 1956;99(3):391–413. <https://doi.org/10.1002/aja.1000990303>.
- de Rooij DG, Russell LD. All you wanted to know about spermatogonia but were afraid to ask. *J Androl.* 2000;21(6):776–98.
- Fayomi AP, Orwig KE. Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. *Stem Cell Res.* 2018;29:207–14. <https://doi.org/10.1016/j.scr.2018.04.009>.
- de Kretser DM, Loveland KL, Meinhardt A, Simorangkir D, Wreford N. Spermatogenesis. *Hum Reprod.* 1998;13(Suppl 1):1–8. https://doi.org/10.1093/humrep/13.suppl_1.1.
- Lui WY, Cheng CY. Transcription regulation in spermatogenesis. *Adv Exp Med Biol.* 2008;636:115–32. https://doi.org/10.1007/978-0-387-09597-4_7.
- Ball RL, Fujiwara Y, Sun F, Hu J, Hibbs MA, Handel MA, et al. Regulatory complexity revealed by integrated cytological and RNA-seq analyses of meiotic substages in mouse spermatocytes. *BMC Genomics.* 2016;17(1):628. <https://doi.org/10.1186/s12864-016-2865-1>.
- Margolin G, Khil PP, Kim J, Bellani MA, Camerini-Otero RD. Integrated transcriptome analysis of mouse spermatogenesis. *BMC Genomics.* 2014; 15(1):39. <https://doi.org/10.1186/1471-2164-15-39>.
- Gong W, Pan L, Lin Q, Zhou Y, Xin C, Yu X, et al. Transcriptome profiling of the developing postnatal mouse testis using next-generation sequencing. *Sci China Life Sci.* 2013;56(1):1–12. <https://doi.org/10.1007/s11427-012-4411-y>.

10. Geisinger A, Rodriguez-Casuriaga R. Flow cytometry for gene expression studies in mammalian spermatogenesis. *Cytogenet Genome Res.* 2010; 128(1-3):46–56. <https://doi.org/10.1159/000291489>.
11. Liu Y, Niu M, Yao C, Hai Y, Yuan Q, Liu Y, et al. Fractionation of human spermatogenic cells using STA-PUT gravity sedimentation and their miRNA profiling. *Sci Rep.* 2015;5(1):8084. <https://doi.org/10.1038/srep08084>.
12. Chen X, Che D, Zhang P, Li X, Yuan Q, Liu T, et al. Profiling of miRNAs in porcine germ cells during spermatogenesis. *Reproduction.* 2017;154(6):789–98. <https://doi.org/10.1530/REP-17-0441>.
13. Shah MA, Xu C, Wu S, Zhao W, Luo H, Yi C, et al. Isolation and characterization of spermatogenic cells from cattle, yak and cattleyak. *Anim Reprod Sci.* 2018; 193:182–90. <https://doi.org/10.1016/j.anireprosci.2018.04.067>.
14. Da Ros M, Lehtiniemi T, Olotu O, Meikar O, Kotaja N. Enrichment of pachytene spermatocytes and spermatids from mouse testes using standard laboratory equipment. *J Vis Exp.* 2019;(151):e60271.
15. Mahalingam M. Laser capture microdissection: insights into methods and applications. *Methods Mol Biol.* 2018;1723:1–17. https://doi.org/10.1007/978-1-4939-7558-7_1.
16. Green CD, Ma Q, Manske GL, Shami AN, Zheng X, Marini S, et al. A comprehensive roadmap of murine spermatogenesis defined by single-cell RNA-seq. *Dev Cell.* 2018;46(5):651–67 e610. <https://doi.org/10.1016/j.devcel.2018.07.025>.
17. Hermann BP, Cheng K, Singh A, Roa-De La Cruz L, Mutoji KN, Chen IC, et al. The mammalian spermatogenesis single-cell transcriptome, from spermatogonial stem cells to spermatids. *Cell Rep.* 2018;25(6):1650–67 e1658. <https://doi.org/10.1016/j.celrep.2018.10.026>.
18. Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, et al. The adult human testis transcriptional cell atlas. *Cell Res.* 2018;28(12):1141–57. <https://doi.org/10.1038/s41422-018-0099-2>.
19. Lau X, Munusamy P, Ng MJ, Sangrithi M. Single-cell RNA sequencing of the cynomolgus macaque testis reveals conserved transcriptional profiles during mammalian spermatogenesis. *Dev Cell.* 2020;54(4):548–66 e547. <https://doi.org/10.1016/j.devcel.2020.07.018>.
20. Yang H, Ma J, Wan Z, Wang Q, Wang Z, Zhao J, et al. Characterization of sheep spermatogenesis through single-cell RNA sequencing. *FASEB J.* 2021; 35(2):e21187. <https://doi.org/10.1096/fj.202001035RRR>.
21. Sullivan TP, Eaglstein WH, Davis SC, Mertz P. The pig as a model for human wound healing. *Wound Repair Regen.* 2001;9(2):66–76. <https://doi.org/10.1046/j.1524-475x.2001.00066.x>.
22. Valent D, Arroyo L, Fábrega E, Font IFM, Rodríguez-Palmero M, Moreno-Munoz JA, et al. Effects of a high-fat-diet supplemented with probiotics and omega3-fatty acids on appetite regulatory neuropeptides and neurotransmitters in a pig model. *Benef Microbes.* 2020;11(4):347–59. <https://doi.org/10.3920/BM2019.0197>.
23. Zhang P, Chen X, Zheng Y, Zhu J, Qin Y, Lv Y, et al. Long-term propagation of porcine undifferentiated spermatogonia. *Stem Cells Dev.* 2017;26(15): 1121–31. <https://doi.org/10.1089/scd.2017.0018>.
24. Bryant JM, Meyer-Ficca ML, Dang VM, Berger SL, Meyer RG. Separation of spermatogenic cell types using STA-PUT velocity sedimentation. *J Vis Exp.* 2013;80:e50648. <https://doi.org/10.103791/50648>.
25. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive integration of single-cell data. *Cell.* 2019;177(7):1888–902 e1821. <https://doi.org/10.1016/j.cell.2019.05.031>.
26. Lin T, Chen T, Liu J, Tu XM. Extending the Mann-Whitney-Wilcoxon rank sum test to survey data for comparing mean ranks. *Stat Med.* 2021;40(7): 1705–17. <https://doi.org/10.1002/sim.8865>.
27. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods.* 2017; 14(10):979–82. <https://doi.org/10.1038/nmeth.4402>.
28. Ito K, Murphy D. Application of ggplot2 to pharmacometric graphics. *CPT Pharmacometrics Syst Pharmacol.* 2013;2(10):e79. <https://doi.org/10.1038/psp.2013.56>.
29. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
30. Sassone-Corsi P. Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science.* 2002;296(5576):2176–8. <https://doi.org/10.1126/science.1070963>.
31. Nueda MJ, Tarazona S, Conesa A. Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. *Bioinformatics.* 2014;30(18): 2598–602. <https://doi.org/10.1093/bioinformatics/btu333>.
32. Zhang P, Li F, Zhang L, Lei P, Zheng Y, Zeng W. Stage-specific embryonic antigen 4 is a membrane marker for enrichment of porcine spermatogonial stem cells. *Andrology.* 2020;8(6):1923–34. <https://doi.org/10.1111/andr.12870>.
33. Simon L, Ekman GC, Garcia T, Carnes K, Zhang Z, Murphy T, et al. ETV5 regulates sertoli cell chemokines involved in mouse stem/progenitor spermatogonia maintenance. *Stem Cells.* 2010;28(10):1882–92. <https://doi.org/10.1002/stem.508>.
34. von Kopylow K, Staeger H, Schulze W, Will H, Kirchhoff C. Fibroblast growth factor receptor 3 is highly expressed in rarely dividing human type a spermatogonia. *Histochem Cell Biol.* 2012;138(5):759–72. <https://doi.org/10.1007/s00418-012-0991-7>.
35. Tres LL, Kierszenbaum AL. The ADAM-integrin-tetraspanin complex in fetal and postnatal testicular cords. *Birth Defects Res C Embryo Today.* 2005;75(2): 130–41. <https://doi.org/10.1002/bdrc.20041>.
36. Liao J, Ng SH, Luk AC, Suen HC, Qian Y, Lee AWT, et al. Revealing cellular and molecular transitions in neonatal germ cell differentiation using single cell RNA sequencing. *Development.* 2019;146(6):e174953. <https://doi.org/10.1242/dev.174953>.
37. Guo J, Nie X, Giebler M, Mlcochova H, Wang Y, Grow EJ, et al. The dynamic transcriptional cell atlas of testis development during human puberty. *Cell Stem Cell.* 2020;26(2):262–76 e264. <https://doi.org/10.1016/j.stem.2019.12.005>.
38. Zhang T, Murphy MW, Gearhart MD, Bardwell VJ, Zarkower D. The mammalian doublesex homolog DMRT6 coordinates the transition between mitotic and meiotic developmental programs during spermatogenesis. *Development.* 2014;141(19):3662–71. <https://doi.org/10.1242/dev.113936>.
39. Lei N, Hornbaker KI, Rice DA, Karpova T, Agbor VA, Heckert LL. Sex-specific differences in mouse DMRT1 expression are both cell type- and stage-dependent during gonad development. *Biol Reprod.* 2007;77(3):466–75. <https://doi.org/10.1095/biolreprod.106.058784>.
40. Takemoto K, Tani N, Takada-Horisawa Y, Fujimura S, Tanno N, Yamane M, et al. Meiosis-specific C19orf57/4930432K21Rik/BRME1 modulates localization of RAD51 and DMC1 to DSBs in mouse meiotic recombination. *Cell Rep.* 2020;31(8):107686. <https://doi.org/10.1016/j.celrep.2020.107686>.
41. Suzuki H, Ahn HW, Chu T, Bowden W, Gassei K, Orwig K, et al. SOHLH1 and SOHLH2 coordinate spermatogonial differentiation. *Dev Biol.* 2012;361(2): 301–12. <https://doi.org/10.1016/j.ydbio.2011.10.027>.
42. Ernst C, Eling N, Martinez-Jimenez CP, Marioni JC, Odom DT. Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat Commun.* 2019;10(1):1251. <https://doi.org/10.1038/s41467-019-09182-1>.
43. Griswold SL, Behringer RR. Fetal Leydig cell origin and development. *Sex Dev.* 2009;3(1):1–15. <https://doi.org/10.1159/000200077>.
44. Kangawa A, Otake M, Enya S, Yoshida T, Kangawa Y, Shibata M. Spermatinipig. *Toxicol Pathol.* 2016;44(7):974–86. <https://doi.org/10.1177/0192623316654586>.
45. Don J, Nir U, Breitbart H. DMRT1 at the border between mitosis and meiosis. *Asian J Androl.* 2011;13(2):189–90. <https://doi.org/10.1038/aja.2010.168>.
46. Zhang T, Zarkower D. DMRT proteins and coordination of mammalian spermatogenesis. *Stem Cell Res.* 2017;24:195–202. <https://doi.org/10.1016/j.scr.2017.07.026>.
47. Matson CK, Murphy MW, Griswold MD, Yoshida S, Bardwell VJ, Zarkower D. The mammalian doublesex homolog DMRT1 is a transcriptional gatekeeper that controls the mitosis versus meiosis decision in male germ cells. *Dev Cell.* 2010;19(4):612–24. <https://doi.org/10.1016/j.devcel.2010.09.010>.
48. Kasinrerker W, Tokrasinwit N, Moonsom S, Stockinger H. CD99 monoclonal antibody induce homotypic adhesion of Jurkat cells through protein tyrosine kinase and protein kinase C-dependent pathway. *Immunol Lett.* 2000;71(1):33–41. [https://doi.org/10.1016/S0165-2478\(99\)00165-0](https://doi.org/10.1016/S0165-2478(99)00165-0).
49. Chung SS, Eng WS, Hu W, Khalaj M, Garrett-Bakelman FE, Tavakkoli M, et al. CD99 is a therapeutic target on disease stem cells in myeloid malignancies. *Sci Transl Med.* 2017;9(374):eaaj2025. <https://doi.org/10.1126/scitranslmed.aaj2025>.
50. Verajakorva E, Laato M, Pollanen P. CD 99 and CD 106 (VCAM-1) in human testis. *Asian J Androl.* 2002;4(4):243–8.
51. Doyonnas R, Kershaw DB, Duhme C, Merckens H, Chelliah S, Graf T, et al. Anuria, omphalocele, and perinatal lethality in mice lacking the CD34-related protein podocalyxin. *J Exp Med.* 2001;194(1):13–27. <https://doi.org/10.1084/jem.194.1.13>.

52. Perrais M, Pigny P, Buisine MP, Porchet N, Aubert JP, Van Seuningem-Lempire I. Aberrant expression of human mucin gene MUC5B in gastric carcinoma and cancer cells. Identification and regulation of a distal promoter. *J Biol Chem*. 2001;276(18):15386–96. <https://doi.org/10.1074/jbc.M010534200>.
53. Nicolini A, Ferrari P, Rossi G. Mucins and cytokeratins as serum tumor markers in breast cancer. *Adv Exp Med Biol*. 2015;867:197–225. https://doi.org/10.1007/978-94-017-7215-0_13.
54. Lin YY, Wang CY, Phan NN, Chiao CC, Li CY, Sun Z, et al. PODXL2 maintains cellular stemness and promotes breast cancer development through the Rac1/Akt pathway. *Int J Med Sci*. 2020;17(11):1639–51. <https://doi.org/10.7150/ijms.46125>.
55. Bai Y, Ye Z, Zeng F. Isolation of subtype spermatogonia in juvenile rats. *J Huazhong Univ Sci Technol Med Sci*. 2008;28(4):435–8. <https://doi.org/10.1007/s11596-008-0413-1>.
56. Guo J, Grow EJ, Yi C, Mlcochova H, Maher GJ, Lindskog C, et al. Chromatin and single-cell RNA-seq profiling reveal dynamic signaling and metabolic transitions during human spermatogonial stem cell development. *Cell Stem Cell*. 2017;21(4):533–46 e536. <https://doi.org/10.1016/j.stem.2017.09.003>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

