Lin *et al. BMC Genomics* (2025) 26:459
https://doi.org/10.1186/s12864-025-11636-4

BMC Genomics

RESEARCH Open Access

Isoform-resolution single-cell RNA sequencing reveals the transcriptional panorama of adult Baoshan pig testis cells

Wan Lin[1†], Xia Zhang[2†], Zhipeng Liu[1†], Hailong Huo[3†], Yongcheng Chang[1], Jiading Zhao[4], Shaorong Gong[4], Guiying Zhao[1*] and Jinlong Huo[1*]

Abstract

Background As the primary organ of the male reproductive system, the testis facilitates spermatogenesis and androgen secretion. Due to the complexity of spermatogenesis, elucidating cellular heterogeneity and gene expression dynamics within the porcine testis is critical for advancing reproductive biology. Nevertheless, the cellular composition and regulatory mechanisms of porcine testes remain insufficiently characterized. In this study, we applied integrated long-read (Nanopore) and short-read (Illumina) scRNA-seq to Baoshan pig testes, establishing a comprehensive transcriptional profile to delineate cellular heterogeneity and molecular regulation.

Results Through systematic analysis of testicular architecture and the temporal progression of spermatogenesis, we characterized 11,520 single cells and 23,402 genes, delineating germ cell developmental stages: proliferative-phase spermatogonia (SPG), early-stage spermatocytes (Early SPC) and late-stage spermatocytes (Late SPC) during meiosis, and spermiogenic-phase round spermatids (RS) followed by elongating/elongated spermatids (ES), culminating in mature spermatozoa (Sperm). We further identified nine distinct testicular cell types, with germ cells spanning all developmental stages and somatic components comprising Sertoli cells, macrophages, and peritubular myoid cells as microenvironmental constituents, revealing the cellular heterogeneity of testicular tissue and dynamic characteristics of spermatogenesis. We obtained the dynamic expression changes of 16 vital marker genes during spermatogenesis and performed immunofluorescence validation on 7 marker genes. Gene ontology analysis revealed that germ cells at various stages were involved in specific biological processes, while cell communication networks highlighted eight pivotal signaling pathways, including MIF, NRG, WNT, VEGF, BMP, CCL, PARs, and ENHO pathways. Long-read sequencing further captured the full integrity and diversity of RNA transcripts, identifying 60% of the novel annotated isoforms and revealing that FSM isoforms exhibited longer transcript lengths, longer coding sequences,

[†]Wan Lin, Xia Zhang, Zhipeng Liu and Hailong Huo contributed equally to this work.

*Correspondence:
Guiying Zhao
zhaoguiying2006@163.com
Jinlong Huo
jinlonghuo973@163.com

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

longer open reading frames, and a great number of exons, suggesting the complexity of isoforms within the testicular microenvironment.

Conclusions Our results provide insight into the cellular heterogeneity, intercellular communication, and gene expression/transcript diversity in porcine testes, and offer a valuable resource for understanding the molecular mechanisms of porcine spermatogenesis.

Keywords Single-cell RNA sequencing, Porcine testis, Long-read sequencing, Spermatogenesis, Cellular heterogeneity

Background

The Baoshan pig, indigenous to Baoshan City in Yunnan Province, China, exhibits remarkable adaptability, crude feed tolerance, efficient utilization of green forage, and superior feed conversion efficiency. Its meat is characterized by tender texture, rich umami flavor, and elevated concentrations of amino acids and unsaturated fatty acids, rendering it an exceptional raw material for cured meats and ham production, with nutritional profiles significantly surpassing those of exotic breeds [1]. However, the breed's suboptimal reproductive efficiency, as demonstrated by lower live birth rates relative to commercial counterparts, constrains its large-scale implementation in modern intensive production systems [2]. Consequently, elucidating the regulatory mechanisms of genes governing spermatogenesis, sperm maturation, and fertilization in Baoshan pigs holds critical scientific implications for both genetic conservation and optimized exploitation of this germplasm resource. The sexual maturity period of Baoshan pigs is 6 months. To ensure stable reproductive functionality, fully sexually mature individuals at 18 months of age were selected for this investigation.

The porcine testis, central to the boar reproductive system, is responsible for spermatozoa production and androgen secretion [3]. Male fertility depends on spermatogenesis [4], a complex process encompassing spermatogonia mitosis, meiosis to generate spermatids, and spermiogenesis to produce mature spermatozoa [5]. During spermatogenesis, spermatogonia undergo mitosis to produce type B spermatogonia [6], which differentiate into preleptotene spermatocytes and enter meiosis [7]. Meiosis involves homologous recombination and reductive division, resulting in haploid round spermatids [8]. In the final phase, spermatogenesis, termed spermiogenesis, round spermatids undergo significant morphological and biochemical transformations to become spermatozoa, which are then released into the seminiferous tubule lumen [9]. A thorough understanding of these cellular dynamics and the diversity of individual cells is crucial for comprehending spermatogenesis in the porcine testis.

Traditional bulk RNA sequencing (RNA-seq) depends on total RNA extracted from numerous cells, obscuring cellular-scale nuances [10]. In contrast, single-cell RNA sequencing (scRNA-seq) enables an unbiased, comprehensive assessment of individual cell states and types [11, 12]. scRNA-seq is uniquely valuable for investigating spermatogenesis in the testis by providing a detailed insight into gene expression in individual cells, vital for comprehending the molecular mechanisms underlying germ cell differentiation and maturation, providing higher-resolution data essential for identifying stage-specific markers and regulatory pathways during the complex spermatogenic process [13–15]. Moreover, scRNA-seq can detect rare and transient cell states, such as progenitor or transitional populations, which may be overlooked in bulk RNA-seq [16, 17]. Long-read scRNA-seq takes advantage of its read length to more accurately capture the splicing diversity and sequence heterogeneity of RNA transcripts, enabling the precise quantification of RNA transcript isoforms, reducing the biases associated with read length limitation, improving the accuracy of cell mutation genotyping [18–20]. In addition, long-read scRNA-seq provides a clearer molecular portrait of cells and facilitates the discovery of novel genetic variations and regulatory mechanisms crucial for understanding the intricate differentiation process. Consequently, we utilize long-read scRNA-seq to study gene expression profiles and isoform diversity in spermatogenic cells at various stages of porcine spermatogenesis.

In this study, we developed a refined pipeline based on 10 × Genomics protocols to conduct single-cell RNA sequencing (scRNA-seq) on testicular tissue from Baoshan pigs. The transcriptomes of approximately 11,520 single cells were analyzed using both Illumina short-read data and Oxford Nanopore long-read data, yielding data for 23,402 genes. We identified nine different cell populations within the testis and characterized gene expression profiles and isoforms across various stages of porcine spermatogenesis. Cell communication networks highlighted eight pivotal signaling pathways. The findings from this study contribute to a deeper understanding of spermatogenesis and male reproductive biology in pigs.

Results

Testicular histomorphology of the Baoshan pigs

To evaluate the morphological characteristics and spermatogenic activity of testicular tissues of 18-month-old Baoshan pigs, hematoxylin-eosin (H&E) stained sections were analyzed (Fig. 1). Histological observation showed that the seminiferous tubules were structurally intact and the spermatogenic cell population was arranged in an orderly manner according to the developmental stage. Cell types were identified according to the classical morphological criteria of the mammalian testes [21]: spermatogonia are densely distributed adjacent to the basement membrane, exhibiting small, rounded nuclei with deeply stained nuclear; Spermatocytes occupy a luminal position relative to spermatogonia, characterized by large spherical nuclei containing coarse chromatin bundles; Spermatids localize near the tubular lumen, featuring spherical nuclei or elongated nuclei with densely packed chromatin; Spermatozoa are localized in the tubular lumen, displaying a streamlined nucleus devoid of cytoplasmic residues, with retained acrosome and flagellum; Sertoli cells are large columnar cells extending from the basement membrane to the luminal surface of seminiferous tubules. These cells are characterized by triangular or irregular nuclei with pale chromatin and prominent nucleoli; Leydig cells cluster in interstitial regions, exhibiting polygonal morphology, eosinophilic cytoplasm, and centrally positioned nuclei; Peritubular myoid cells positioned adjacent to the seminiferous tubule basement membrane, exhibiting an elongated fusiform morphology with nuclei aligned parallel to the tubular axis. These results indicated that 18-month-old Baoshan pigs have reached adult maturation.

The landscape of short-read single-cell transcriptome in Baoshan pig testicular tissues

We employed a 10 × Genomics workflow to analyze the cellular heterogeneity of Baoshan pig testicular tissues (Fig. 2A). Testis samples were processed to generate single-cell suspensions, ensuring precise extraction by removing red blood cells. The resulting live cell concentration was 878 cells/μl, with viability of 98.0%, an average cell diameter of 12.9 μm, and a clustering rate of 4.5%. We excluded cells with fewer than 200 genes, more than 8,000 unique molecular identifiers (UMIs), or mitochondrial gene expression exceeding 5% of total reads (Fig. 2B). These thresholds were validated by the observed correlations in the dataset: the strong positive correlation between nCount_RNA and nFeature_RNA ($r = 0.89$; Fig. 2C) confirmed that high-quality cells exhibited proportional increases in sequencing depth and gene detection, and the weak negative correlation between nCount_RNA and percent.mt (r = -0.16; Fig. 2D) indicated high-quality cellular integrity. Following stringent quality control, we retained expression data for 11,520 cells, identifying 23,402 genes. To mitigate technical noise in the single-cell RNA sequencing data, we performed principal component analysis (PCA) to distill biologically relevant variation. The ElbowPlot (Fig. 2E) confirmed that major biological variability was captured within the first 15–20 PCs, whereas higher components predominantly reflected technical noise. The "elbow point" at PC15 (marked by a distinct inflection in the curve) was selected for downstream dimensionality reduction. This approach optimized the retention of biologically meaningful variation, ensuring subsequent analyses focused on authentic cellular heterogeneity. Subsequently, we systematically analyzed the gene expression distribution associated with the first 15 principal components, identifying candidate

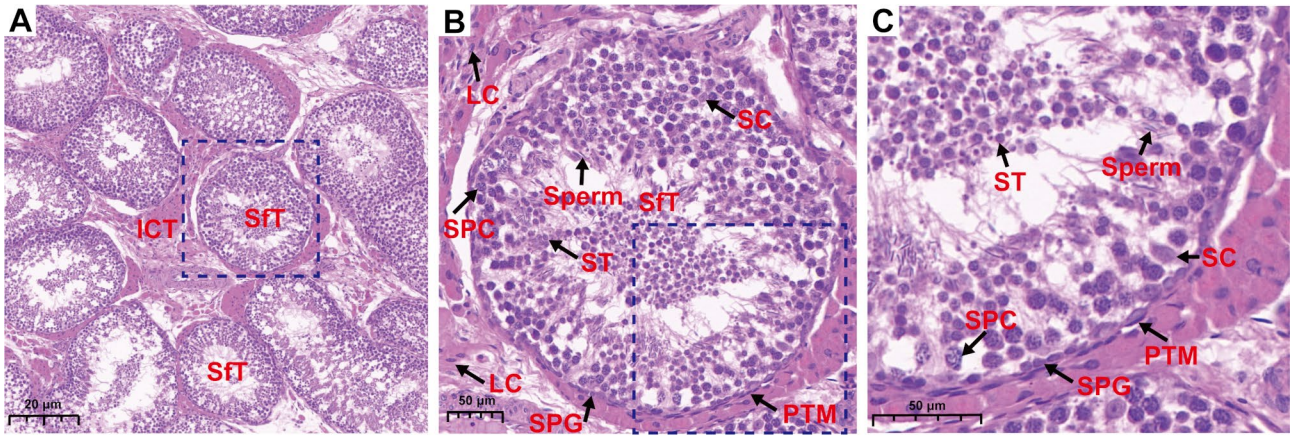

Fig. 1 H&E staining results of testicular tissue from Baoshan pigs. (**A**) Microscopic view of testicular tissue of Baoshan pigs at 20 × magnifications. (**B**) Microscopic view of testicular tissue of Baoshan pigs at 63 × magnifications. (**C**) Microscopic views of testicular tissue of Baoshan pigs at 120 × magnifications. SfT, seminiferous tubules; ICT, interstitial tissue of the testis; LC, Leydig cell; SC, Sertoli cell; SPG, spermatogonia; SPC, spermatocytes; ST, spermatids; Sperm, spermatozoa; PTM, peritubular myoid cells

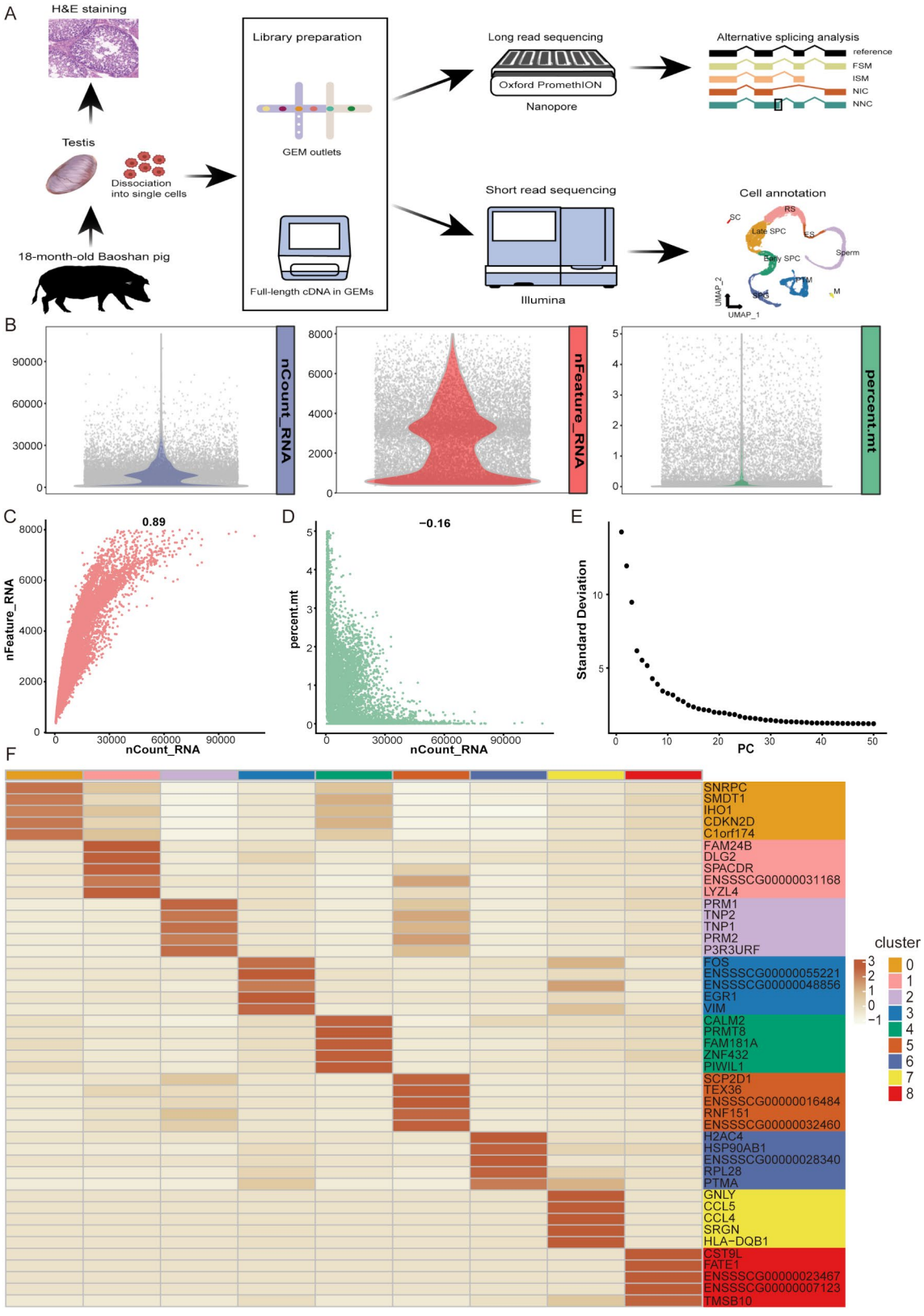

Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 scRNA-seq profiles in Baoshan pig testes. (**A**) The workflow for scRNA-seq analysis. (**B**) The violin plots illustrate the distribution of single-cell RNA-seq quality control metrics. Correlation between the number of genes and single-molecule labels. (**C**) Correlation between nFeature_RNA and nCount_RNA. (**D**) Correlation between percent_mt and nCount_RNA. (**E**) The ElbowPlot determines the optimal number of principal components (PCs) to retain after principal component analysis (PCA) in dimensionality reduction workflows. (**F**) Heatmap displaying expression levels of the top five genes in each cluster, with columns representing clusters and rows representing genes. Color intensity indicates gene expression levels, with darker shades reflecting higher expression

genes that may underlie the biological (Fig. S1). We applied Uniform Manifold Approximation and Projection (UMAP) for cell clustering, identifying nine distinct cell clusters (Fig. 3A). The heatmap summarizes the top five differentially expressed genes (DEGs) per cluster identified through FindAllMarkers function analysis (Fig. 2F), highlighting cellular heterogeneity of testicular tissue. Complete differentially expressed genes per cluster are documented in Supplementary Table S1.

Identification of cell types in Baoshan pig testicular tissues

To delineate the cellular trajectory in Baoshan pig testicular tissue, we implemented a single-cell transcriptomic analytical pipeline using Seurat (v4.3.0) for cell type identification and annotation. Due to the incomplete characterization of specific markers for porcine male germ cells, we annotated the nine identified clusters (Fig. 3A) by mapping porcine protein-coding genes to their human homologs and integrating insights from relevant mammalian testicular single-cell literature [22–26], as well as marker genes from the Protein Atlas database (https://www.proteinatlas.org), successfully identifying nine key cell types (Fig. 3B), the distribution and number of cells within each type were illustrated in Fig. 3C and Fig. S2A, including Sertoli cells (SC) (*ALDH1A1* [27], *FATE1* [28]), macrophages (M) (*CD74* [29], *MAFB* [30]), peritubular myoid cells (PTM) (*MYH11* [31], *ACTA2* [32]), spermatogonia (*UCHL1* [33], *DMRT1* [34]), early-stage primary spermatocytes (*PIWIL1* [35], *DMC1* [36]), late-stage spermatocyte (*CCNB1* [37], *YBX2* [38]), round spermatids (*ACRV1* [39], *LYZL4* [40]), elongating/elongated spermatids (*SPATA21* [41], *GARIN4* [42]) and spermatozoa (*TPPP2* [43], *PRM2* [44]) (Fig. 3D). In addition, we obtained the dynamic expression changes of 9 vital marker genes during spermatogenesis (Fig. 3E). To validate the spatial expression patterns of key germ cell markers in Baoshan pig testes, we performed immunofluorescence (IF) staining analysis. As illustrated in Fig. 3F, distinct spatial expression patterns were observed: *UCHL1* localized predominantly to SPG within the basal compartment of seminiferous tubules; *DMC1* marked Early SPC; *YBX2* marked Late SPC; *ACRV1* and *LYZL4* were robustly expressed in RS, with minor signals detected in ES; *TPPP2* and *PRM2* localized to Sperm within the seminiferous tubules lumen. We also assessed the number of UMIs, gene counts, and the percentage of mitochondrial genes across these nine cell types (Fig. S2B). Using the monocle2 package to examine the developmental trajectory of the germ cells, we identified six distinct stages in the Baoshan pig testis: Spermatogonia (SPG), early-stage primary Spermatocytes (Early SPC), late-stage Spermatocytes (Late SPC), Round Spermatids (RS), Elongating/elongated Spermatids (ES), culminating in spermatozoa (Sperm) (Fig. 3G).

Gene ontology analysis of germ cells in Baoshan pig testicular tissues

To characterize gene expression profiles of germ cells in Baoshan pig testis, we employed the Seurat package to analyze differentially expressed genes (DEGs) across six different germ cell types. The identified DEGs were as follows 5,797 in SPG, 4,297 in early SPC, 3,608 in late SPC, 3,084 in RS, 3,784 in ES, and 5,332 in sperm. We then utilized clusterProfiler for Gene ontology analysis to elucidate the biological processes associated with these germ cells, complete statistical findings are documented in Supplementary Table S2. Our findings indicated that SPG genes were primarily enriched in processes such as cell cycle, chromosome organization, and ribonucleoprotein complex biogenesis, highlighting the active proliferation of spermatogonial cells (Fig. 4A). In contrast, in early SPC and late SPC genes were involved in meiotic cell cycle and male gametes production (Fig. 4B and C). Additionally, genes in RS and ES were associated with downstream processes such as spermatogenesis, male gamete generation, and cilium assembly (Fig. 4D and E). Finally, genes in sperm were predominantly linked to acrosome assembly and flagellated sperm motility (Fig. 4F), indicating the completion of the sperm development trajectory.

Cell communication in the testis of Baoshan pigs

To explore cell communication among different cell types in the testicular tissue of adult Baoshan pigs, we employed the CellChat package to examine signaling dynamics across nine cell types. Our analysis revealed significant differences in both the quantity and strength of interactions (Fig. 5A). We identified eight key signaling pathways: macrophage migration inhibitory factor (MIF), neuregulin (NRG), wingless/integrated (WNT), vascular endothelial growth factor (VEGF), bone morphogenetic protein (BMP), C-C motif chemokine ligand (CCL), protease-activated receptors (PARs), and energy homeostasis associated (ENHO) (Fig. 5B). Notably, top colored bar plot displays Macrophages (M) and Round Spermatids

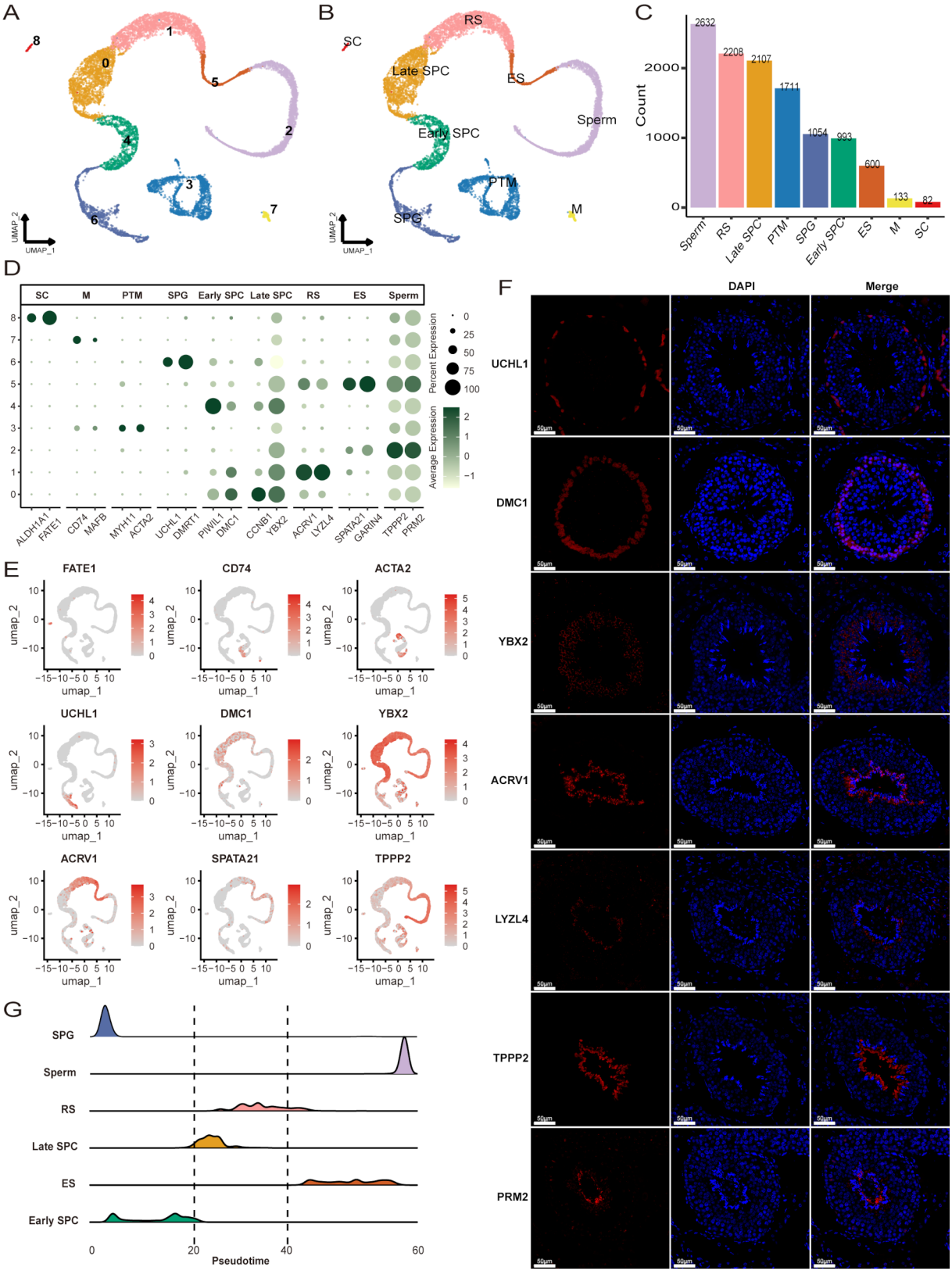

Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Cell type analysis of Baoshan pig testes. (**A**) UMAP visualization reveals nine distinct clusters identified through Seurat clustering. (**B**) UMAP plot illustrating the annotation of nine distinct cell types. (**C**) Histogram representing the cell counts for each of the nine cell types. (**D**) Dot plot displaying the expression levels of 18 marker genes across the nine cell types. (**E**) UMAP plot visualizing the expression of 9 critical marker genes. (**F**) Immunofluorescence (IF) staining analysis of marker gene expression and localization in testicular germ cells of Baoshan pigs. Scale bar = 50 μm. (**G**) Ridge plot from pseudotime analysis depicting the six essential stages of germ cell development

(RS) emerged as the primary signal receiver across these pathways, SPG was identified as the central hub for cellular communication, functioning as both a signal sender and receiver (Fig. 5B). Intriguingly, all pathways, except for PARs, involved two or more signal sender. Moreover, apart from the PARs and ENHO pathways, which featured multiple signal receivers, the remaining pathways each had a single primary receiver (Fig. 5B). In particular, the active receiver cells were M for MIF, RS for NRG, SPG for WNT, PTM for VEGF, SPG for BMP, and early SPC for CCL (Fig. 5B), suggesting that interactions among testicular cells in Baoshan pigs predominantly follow a many-to-one targeting pattern. Furthermore, robust cell communication was evident among the six reproductive cell types, particularly strong among SPG, Early SPC, Late SPC, and RS (Fig. 5C).

Characteristics of isoforms identified by long-read single-cell data

To delve into transcript diversity across different cell types in the testicular tissues of adult Baoshan pigs, we employed Nanopore single-cell sequencing, resulting in the identification of 19,620 non-redundant isoforms and 5,787 genes. We then categorized these isoforms into five types based on the Ensembl 111 *Sus scrofa* reference annotation. Among these, 21.98% matched known transcripts full splice match (FSM, isoforms exactly matching reference splice junctions), 15.94% were categorized as incomplete splice match (ISM, partial splice junction overlap), 13.98% were identified as a novel in the catalog (NIC, new isoforms of known genes), 47.01% were classified as a novel not in catalog (NNC, new isoforms from unannotated loci), and 1.10% fell into other categories, including Fusion and Genic (Fig. 6A), indicating the significant discovery of novel isoforms using Nanopore's third-generation single-cell sequencing technology. Supplementary Table S3 provides comprehensive annotations of isoforms identified in long-read sequencing data following stringent quality filtering. The identified isoforms exhibited a length distribution ranging from 500 to 2,500 bp, with a peak density around 1,000 bp (Fig. 6B). Structural analysis revealed over 7,000 transcripts with at least one new splice site (Fig. 6C). Notably, FSM isoforms exhibited longer transcript lengths (Fig. 6D), longer coding sequences (Fig. 6E), longer open reading frame (Fig. 6F), and a great number of exons (Fig. 6G) compared to the other three categories. Additionally, we analyzed the overlap rates of the 5' CAGE cap (Fig. 6H) and 3' polyA tail (Fig. 6I). With reference annotation ISM displayed lower overlap with Ensembl transcript at the 5' end compared to FSM, NIC, and NNC, potentially due to the presence of reverse transcription or mRNA degradation products.

Unique isoform profiles across cell classes in the testis of Baoshan pigs

To identify isoforms expression across various cell types and highlight the advantages of long-read sequencing in detecting low-expression isoforms, we employed the same methods used in short-read data for dimensionality reduction and cell clustering on long-read data, successfully identifying seven distinct cell types: SPG, Early SPC, Late SPC, RS, ES, Sperm, and PTM (Fig. 7A), resembling the cell annotation results from short-read data. Although transcript counts varied among different cell types (Fig. 7B), the overall distribution of isoform categories remained consistent, with novel transcripts accounting for approximately 60% of the total (Fig. 7C). Notably, the novel identified isoforms (NNC and NIC) exhibited lower expression levels (Fig. 7D), explaining their absence in previous studies. Therefore, single-cell long-read sequencing significantly increased the number of detected isoforms. Furthermore, approximately 50% of genes in each germ cell type expressed only one isoform, with similar distributions across different cell categories (Fig. 7E). Over 60% of expressed isoforms across all cell types were novel, and the proportions of different isoform categories varied due to distinct expression patterns (Fig. 7F). Intriguingly, most isoforms were detected in at least three cell types, with approximately 27% of isoforms uniquely expressed in a single cell type, including 1067 in Sperm, 737 in Late SPC, 607 in SPG, 400 in RS, 364 in Early SPC, and 7 in ES (Fig. 7F), complete statistical findings are documented in Supplementary Table S4.

Discussion

Spermatogenesis in mammals is a complex process involving multiple cell types and intricate molecular regulatory pathways [6, 25]. While significant advances have been made in understanding spermatogenesis in mice and humans, knowledge of this process in pigs remains limited [13, 15, 16, 45]. In this study, we employed integrative Illumina and Oxford Nanopore sequencing to perform single-cell RNA sequencing on Baoshan pig testicular tissues. We analyzed the approximately 11,520 cells, identifying distinct cell populations, and examined

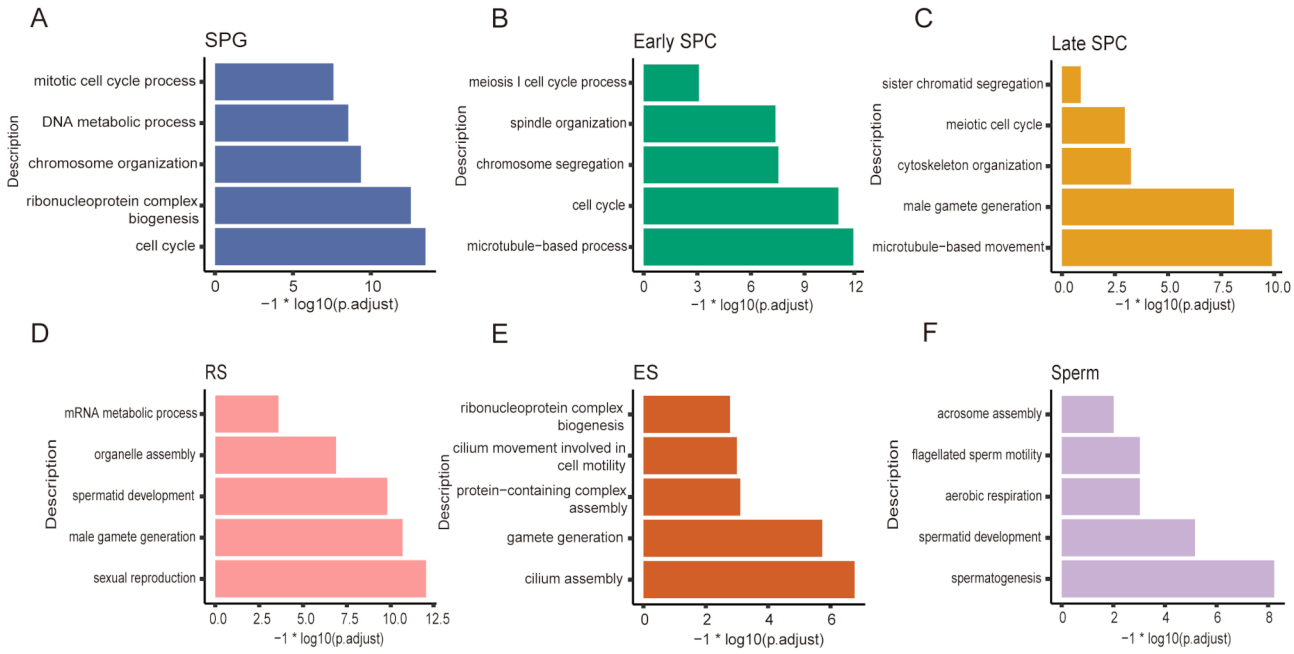

Fig. 4 Gene ontology (GO) enrichment analysis of testicular germ cells in Baoshan pigs. (**A-F**) display the GO enrichment terms and corrected *P*-values for SPG, Early SPC, Late SPC, RS, ES, and Sperm cells, respectively

gene expression profiles and isoform diversity in spermatogonial cells across various stages of spermatogenesis. These findings provided new insights into the molecular mechanisms underlying porcine spermatogenesis and male reproduction.

Although we leveraged single-cell RNA sequencing to comprehensively characterize the cellular landscapes of the Baoshan pig testis, identifying nine cell types critical for spermatogenesis, including germ cells (spermatogonia to spermatozoa) and somatic cells (Sertoli cells, macrophages, peritubular myoid cells), our analysis failed to distinctly identify Leydig cell clusters, and the number of Sertoli cells is the minority. In the testes of Guanzhong Black pigs [46], somatic cell types exhibit dynamic changes during development. Sertoli cells predominate before 90 days postnatal, followed by a gradual expansion of peritubular myoid cells surrounding the seminiferous tubules. By postnatal day 90 (post-puberty), myoid cells become the dominant somatic population. Additionally, Leydig cells display a similar pattern: their numbers progressively decline with age, consistent with observations in cynomolgus monkeys [47]. Guo et al. [48] and Wang et al. [49] demonstrated that peritubular myoid cells and Leydig cells in mammalian testes may originate from common progenitor cells, which could lead to misclassification of scarce Leydig cells into the peritubular myoid cell cluster during computational clustering. This limitation likely explains the absence of Leydig cells and the low abundance of Sertoli cells in our single-cell atlas of 18-month-old Baoshan pig testes. Future studies should incorporate multiple developmental time points to resolve transient or rare cell states and refine cell type annotations.

Cell communication in the mammalian testis is a complex and tightly regulated process involving diverse signaling pathways and cell types [22, 50]. Macrophage migration inhibitory factor (MIF), a multifunctional cytokine originally identified as a T cell-derived regulator of macrophage mobility, mediates immunomodulation by suppressing macrophage migration [51]. Emerging evidence highlights its critical role in the testicular microenvironment, particularly in governing spermatogonial migratory activity [52]. Specifically, MIF levels exhibit a correlation with the dynamics of spermatogonial migration [52], a finding corroborated by our study (Fig. 5B), which observes SPG as the primary source of MIF signaling and macrophages as its specific cellular targets. Bone morphogenetic proteins (BMPs) are pivotal regulators in male reproductive biology [53]. Specifically, BMP2 enhances juvenile spermatogonial proliferation, while BMP4 modulates spermatogonial differentiation [54]. Our data suggest that BMP4-driven differentiation may be mediated through negative feedback loops involving Early SPC, Late SPC, and ES (Fig. 5B). Moreover, peptide growth factors NRG1 (neuregulin 1) and KITL (KIT ligand) have been demonstrated as essential mediators for the survival of differentiating spermatogonia on laminin substrates in no somatic cells systems [55]. Mechanistically, NRG1 and KITL act downstream of retinoic acid signaling in germ cells to facilitate syncytial growth

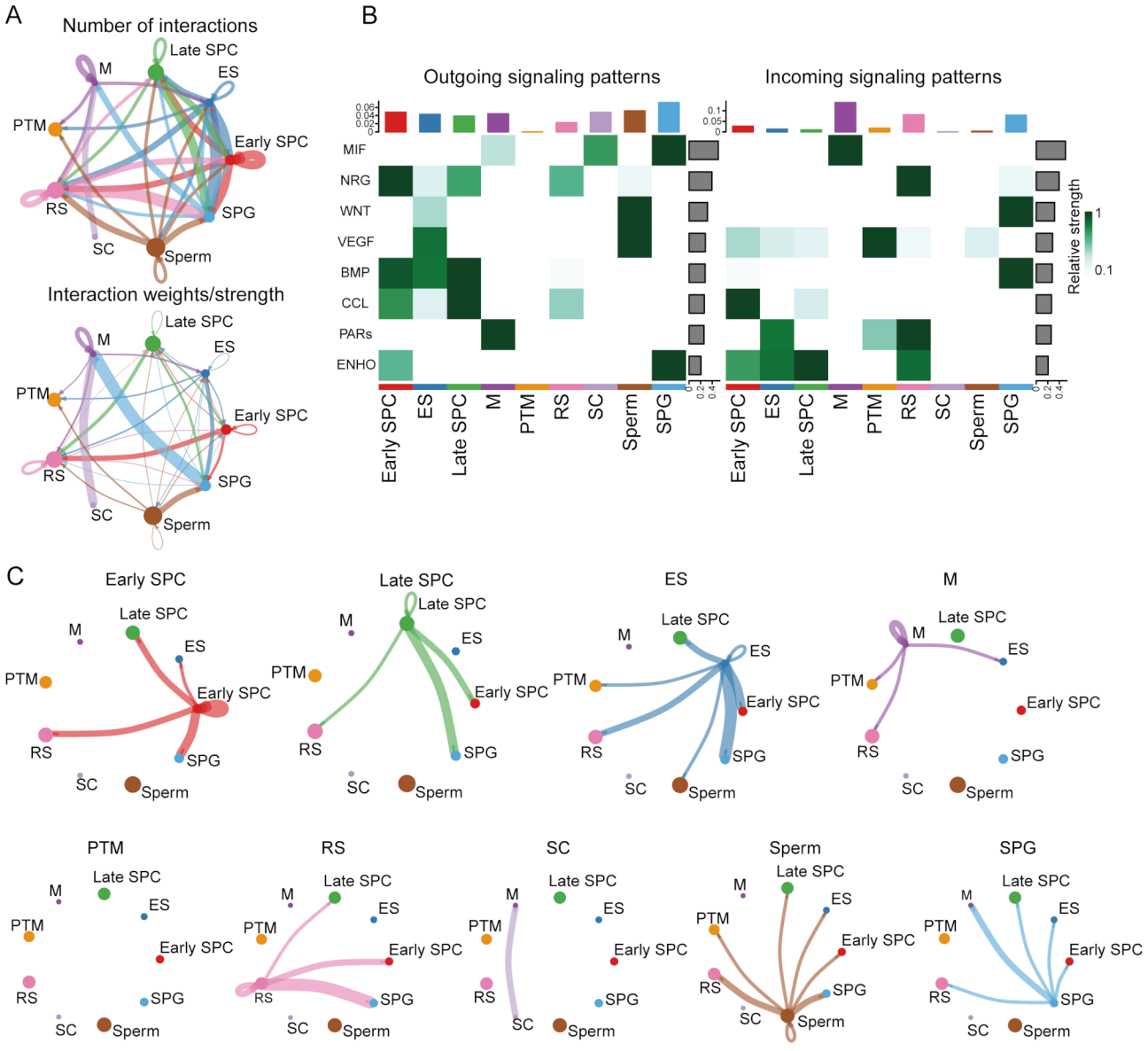

Fig. 5 Analysis of cell subpopulation interactions in the testis of Baoshan pigs. (**A**) The network diagram illustrates the quantity and strength of interactions across the nine cell types. (**B**) The heatmap depicts overall signal transduction, with cell types on the horizontal axis and pathways on the vertical axis. Outgoing signaling patterns represent signal-sending cell populations, Incoming signaling patterns indicate signal-receiving cell populations. The left panel shows the strength of signal emission for each pathway across different cell types, while the right panel indicates the strength of signal reception for each pathway. The top-colored bar plot displays the aggregated signaling strength of each cell group by summing the contributions of all pathways shown in the heatmap. This helps identify the most active cell groups in the signaling network. The right grey bar plot shows the total pathway-specific signaling strength across all cell groups, highlighting the most influential pathways. (**C**) The network diagram displays the number of interactions between cell types, where the size of the outer colored circles represents the number of cells, and the more significant the circle represents the greater the number of cells. Cells with outgoing arrows represent ligands, while those with incoming arrows represent receptors. The thickness of the lines connecting ligand-receptor pairs reflects more interactions

of differentiating spermatids independent of somatic support [56]. This pathway may underlie the biological functions of RS observed in our study. Furthermore, emerging evidence highlights the critical role of WNT signaling in governing spermatogonial fate determination [57]. Postnatal maintenance of mouse spermatogonia is tightly regulated by WNT pathway activity [58], which orchestrates spermatogonia fate by balancing self-renewal and differentiation. These findings revealed that intercellular communication in Baoshan pig testes is a finely orchestrated process involving multiple signaling pathways that regulate spermatogenesis, providing crucial insights into the molecular mechanisms of spermatogenesis, and predicting the activity of signaling pathways.

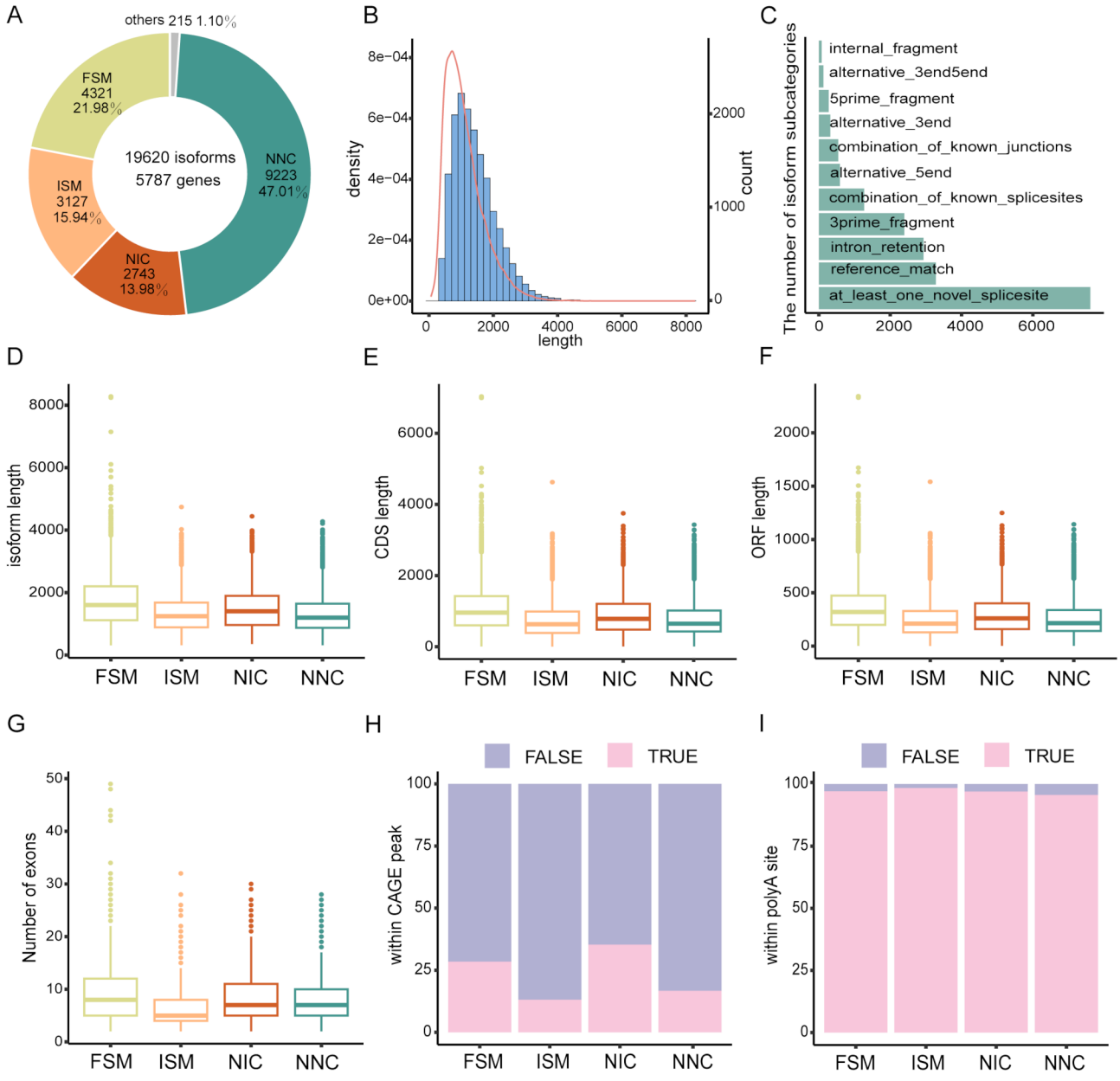

Fig. 6 Characterization of testicular tissue isoforms in Baoshan pigs. (**A**) Pie chart illustrating the percentage of identified isoform types. FSM, Full Splice Match; ISM, Incomplete Splicing Match; NIC, novel in Catalog; NNC, Novel Not in Catalog. (**B**) Histogram displaying the length distributions and density profiles of all non-redundant isoforms. (**C**) Horizontal histogram representing the structure types of detected isoforms. Characterization of identified isoforms, including isoform length (**D**), coding sequence length (**E**), open reading frame length (**F**), and exon count (**G**). Overlap of 5′-terminal CAGE caps (**H**) and 3′-terminal polyA tails (**I**) with reference annotations

Contemporary single-cell RNA sequencing (scRNA-seq) predominantly employs short-read platforms, wherein cDNA libraries generated from captured single cells undergo fragmentation, amplification, and library preparation prior to sequencing [59]. This approach enables robust genome-wide transcript quantification but fails to resolve transcript isoform diversity, structural variations, and tissue/cell subtype-specific RNA processing events [60]. These limitations arise because conventional short-read scRNA-seq methods are inherently restricted to sequencing 150–250 bp fragments from transcript termini (3’ or 5’ ends), which inadequately span full-length mRNA molecules [20]. In contrast, long-read scRNA-seq technologies overcome this constraint by capturing intact RNA molecules, thereby enabling precise characterization of splice variants, isoform-level quantification, and improved genotyping accuracy for somatic mutations [61]. In the context of

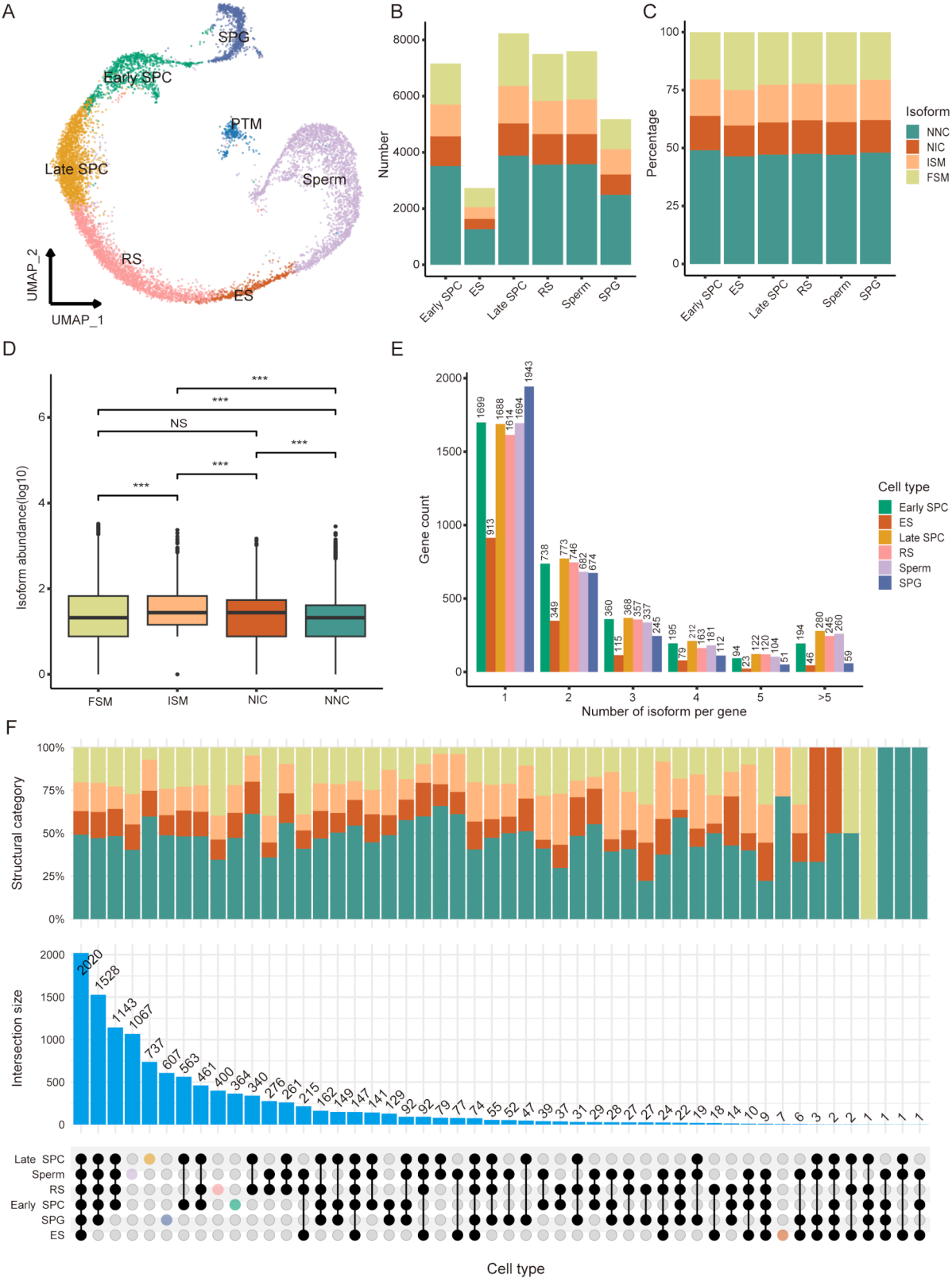

Fig. 7 (See legend on next page.)

(See figure on previous page.)

Fig. 7 Unique isoform profiles in testicular cell types of Baoshan pigs. (**A**) UMAP plot depicting the annotation of seven cell types. (**B**) Bar charts display the number of different isoform types in each cell category. (**C**) Bar charts representing the percentage of different isoform types in each cell category. (**D**) Box plots illustrating isoform abundance across different categories, with adjusted *p*-values indicating inter-group differences. (**E**) Bar charts displaying the number of isoforms per gene. (**F**) UpSet plot showing isoform overlap between cell types, with the number and percentage of shared isoforms highlighted in top bar chats

our study, short-read sequencing demonstrated superior performance in profiling repetitive and low-complexity genomic regions, providing consistent sequencing depth and base-level accuracy critical for reliable gene expression quantification, while long-read sequencing revealed previously undetected isoform diversity and structural heterogeneity [62].

Our findings demonstrate that long-read sequencing not only recapitulates the germ cell clustering patterns observed in short-read data (Fig. 7A) but also uncovers substantial previously undetected isoform diversity. The predominance of novel isoforms (NIC and NNC, accounting for ~60% of total isoforms; Fig. 7C) highlights the limitations of conventional short-read approaches in resolving full-length transcript structures, consistent with previous studies emphasizing isoform complexity during spermatogenesis [63]. Notably, the lower expression levels of novel isoforms (Fig. 7D) suggest their potential role in fine-tuning cellular functions rather than serving as bulk transcriptional outputs. For instance, sperm-specific isoforms (1,067 unique to sperm; Fig. 7F) may include truncated variants of chromatin packaging genes (e.g., PRM2 isoforms lacking N-terminal domains) that facilitate histone-to-protamine transition [64]. Conversely, the conserved expression of single isoforms in ~50% of genes across cell types (Fig. 7E) implies critical constraints on essential spermatogenic regulators, where isoform switching could disrupt vital interactions. Cell type-specific isoform distributions further reflect functional specialization, with late spermatocytes (Late SPC) undergoing meiotic recombination exhibiting 737 unique isoforms (Fig. 7F), potentially enriched for the meiotic cell cycle. Interestingly, our comparative analysis revealed the absence of macrophages and Sertoli cells in long-read single-cell annotations (Fig. 7A) compared to short-read data (Fig. 3B), likely attributable to the lower sensitivity of long-read sequencing and the stringent quality control implemented to prioritize high-quality germ cell transcriptomes, potentially excluding rare low-abundance somatic populations. Therefore, to obtain comprehensive and accurate information, a combined analysis of short-read and long-read single-cell sequencing technologies should be considered, highlighting the complementary strengths of these platforms in integrative transcriptomic profiling.

Conclusions

This study employed both long-read and short-read scRNA-seq technologies to comprehensively analyze the Baoshan pig testis, revealing cellular heterogeneity, complex intercellular communication, and diverse gene expression and isoforms. We identified six stages of spermatogenesis, nine cell types, 9 vital marker genes, and eight pivotal signaling pathways. The application of long-read sequencing identified 60% of the novel annotated isoforms, revealing that FSM isoforms exhibited longer transcript lengths, longer coding sequences, longer open reading frames, and a great number of exons, highlighting the complexity and diversity of gene expression during spermatogenesis, which will provide valuable insights into the molecular mechanisms underlying porcine spermatogenesis and serve as significant resources for future studies.

Materials and methods

Animals

Eighteen-month-old male Baoshan pigs were obtained from the Baoshan Pig Research Institute (Yunnan, China). Testicular samples were obtained by veterinary castration surgery, with local infiltration anesthesia using lidocaine injected at multiple points around the scrotum. After removing the surrounding fatty and fascial tissues, a small parenchymal tissue sample was excised from the middle region of the testis. A portion of the testicular tissue was fixed in 4% paraformaldehyde, while the remaining tissues were preserved in a 4 °C tissue preservation solution (Miltenyi Biotec, Cologne, NW, Germany) for cell separation.

Histological analysis

Porcine testicular tissues underwent sequential dehydration in a graded ethanol series, cleared in xylene, and embedded in paraffin wax. Tissue blocks were sectioned at 4–6 μm thickness using a Leica RM2016 rotary microtome (Leica, Shanghai, China), followed by hematoxylin and eosin (H&E) staining (Servicebio, Wuhan, China) to visualize tissue architecture and cellular organization. The histological morphology of Baoshan pig testes was caught using a Pannoramic MIDI FL whole-slide scanner (3DHISTECH, Budapest, Hungary) for high-resolution imaging and analysis.

Immunofluorescence of testicular tissues

Tissue sections were dewaxed in xylene, dehydrated through a graded ethanol series, and rehydrated in distilled water. Antigen retrieval was performed using preheated citrate-based buffer (pH 6.0) at 98 °C, followed by natural cooling and subsequent washing with phosphate-buffered saline (PBS, pH 7.4). After air-drying, tissue boundaries were demarcated using a hydrophobic barrier pen. Sections were blocked with 10% donkey serum for 30 min at room temperature. Primary antibodies (Supplementary Table S5) were applied and incubated overnight at 4 °C. Following three 5-minute PBS washes, sections were incubated in darkness for 50 min with CY3-conjugated goat anti-rabbit IgG secondary antibody (1:300 dilution, RGAR001; Servicebio, Wuhan, China). Nuclei were counterstained with DAPI for 10 min, and fluorescent images were acquired using a Nikon ECLIPSE C1 fluorescence microscope equipped with appropriate filter sets.

Single-cell suspension Preparation

Single-cell suspensions of seminiferous tubules were generated using the Miltenyi Tumor Dissociation Kit (Miltenyi Biotec, Bergisch Gladbach, Germany). Testicular tissues were washed twice with Dulbecco's Modified Eagle Medium (DMEM; Gibco, Carlsbad, USA) to remove residual preservative, minced into 2–3 mm^3 fragments, and enzymatically digested according to the manufacturer's protocol. The resulting cell suspension was filtered through a 40 μm nylon cell strainer and centrifuged at 300 × g for 5 min at 4 °C. Pelleted cells were resuspended in an ice-cold complete medium (DMEM supplemented with 10% FBS and 1% penicillin-streptomycin). Erythrocyte contamination was addressed by incubating the suspension with Red Blood Cell Lysis Buffer (Miltenyi Biotec) for 10 min at room temperature, with volume adjusted empirically based on visible hemolysis. The single-cell suspension was mixed with trypan blue and analyzed using the Countess II FL automated cell counter (Thermo Fisher Scientific, Waltham, MA, USA). Aggregate percentage was automatically calculated by setting a cell size range (4–30 μm) and shape factor threshold (0.7–0.9), maintained below 5% (4.5% in this study) to ensure single-cell resolution, while cell viability (98.0%) was assessed via trypan blue exclusion. Subsequent removal of residual dead cells was performed using the dead Cell Removal Kit (Miltenyi Biotec), which magnetically labels and eliminates membrane-compromised cells.

Single-cell transcriptome library construction

Single-cell transcriptome libraries were constructed using the Chromium Single Cell 3' Reagent Kit v3.1 (10X Genomics, Pleasanton, USA) following the manufacturer's protocol. Cell suspensions were assessed for viability (> 90%) and concentration (800–1,200 cells/μL) via 0.4% (w/v) Trypan Blue exclusion on a Countess II Automated Cell Counter (Thermo Fisher Scientific). Single-cell suspensions, master mix, and barcoded gel beads were co-encapsulated into Gel Bead-in-Emulsions (GEMs) using a Chromium Controller (10X Genomics). Post-encapsulation, GEMs were lysed to release RNA, followed by the removal of residual enzymes and primers via Dynabeads MyOne SILANE magnetic beads (Thermo Fisher Scientific) with a 5-minute incubation at room temperature. Amplified cDNA libraries were quantified using an Agilent High Sensitivity DNA Kit on a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, USA).

Illumina single-cell sequencing

For short-read sequencing data, splice sequences and adapters were removed using CellRanger (v7.1.0). The trimmed FASTQ files were then aligned to the pig reference genome (Sscrofa 11.1, Ensembl release 112) to assess the alignment accuracy, match quality, and gene expression levels. Low-quality data, unaligned reads, multi-mapped reads, and PCR duplicates were filtered out. Gene transcript counts were subsequently calculated using unique molecular identifiers (UMIs) and cell barcodes, enabling detailed analysis of gene expression and functional insights.

Nanopore single-cell sequencing

For Oxford Nanopore single-cell sequencing, quality-controlled cDNA amplicons underwent end-repair and 3' adenylation using NEBNext Ultra II reagents (New England Biolabs, Ipswich, USA). Adapter ligation was performed with the SQK-LSK114 Ligation Sequencing Kit (v1.0.3; Oxford Nanopore Technologies, Oxford, UK), followed by magnetic bead-based purification (AMPure XP; Beckman Coulter, Brea, USA) to remove unbound adapters. Final libraries were quantified via Qubit 4.0 Fluorometer (Thermo Fisher Scientific, Waltham, USA) and loaded onto a PromethION R10.4.1 flow cell (Oxford Nanopore Technologies) for real-time single-molecule sequencing.

Illumina short-read data preprocessing

Raw Illumina short-read sequencing data were processed using CellRanger (v7.1.0; 10X Genomics, USA) to remove adapter sequences and trim low-quality bases. Cleaned reads were aligned to the *Sus scrofa* reference genome (Sscrofa11.1, Ensembl release 112) using the STAR aligner (v2.7.10a) with default parameters. Post-alignment, low-quality reads, unmapped reads, multi-mapped reads, and PCR duplicates were filtered out. Gene expression matrices were generated by counting unique molecular identifiers (UMIs) and cell barcodes with

CellRanger's count function, incorporating Ensembl-annotated transcripts for quantification.

Nanopore long-read data preprocessing

Raw signal data (FAST5 files) from Oxford Nanopore sequencing were base-called using Guppy (v6.0.7) in high-accuracy mode. Cellular barcodes were demultiplexed by identifying pre-defined flanking sequences. To improve barcode assignment accuracy, we dynamically adjusted the Levenshtein edit distance threshold (range: 1–3) based on cross-correlation analysis between Illumina short-read-derived cellular barcodes and their corresponding gene expression profiles. Finally, demultiplexed barcodes and unique molecular identifiers (UMIs) were appended to the header of each processed read in the final FASTQ files, ensuring unambiguous assignment of long-read transcripts to individual cells while preserving original sequence integrity.

Cell clustering and annotation

To better understand the diversity, function, and organization of testicular cells in Baoshan pigs, we performed cell clustering and annotation analyses. The initial count matrix was imported into Seurat (v4.3.0) [65], where we applied filtering criteria: mitochondrial gene content below 5%, a minimum of 200 detectable genes per cell, and each gene expression in at least three cells. Doublets/multiplets were identified and removed using scDblFinder (v1.16.0) [66] to ensure data accuracy. We normalized the remaining gene expression values with the NormalizeData function, identifying highly heterogeneous genes to reduce expression differences. Principal component analysis (PCA) was performed using RunPCA, followed by dimensionality reduction with RunUMAP. Cell proximity relationships were calculated using FindNeighbors, and marker genes for each cluster were identified with FindAllMarkers. To further investigate gene expression dynamics and cell developmental trajectories, Monocle (v2.26.0) was used for trajectory analysis and cell sorting [67]. Finally, differentially expressed genes were subjected to novel isoforms using clusterProfiler (v4.6.0) to reveal the biological functions of the testicular cell population [68].

Cell-cell communication inference

To investigate the global communication among porcine testicular cell types, we employed CellChat (v1.6.1) to analyze cell-to-cell communication based on homologous genes shared between pigs and humans, revealing continuity in the developmental trajectory of testicular cells [69]. Additionally, we applied social network analysis and pattern recognition to characterize the communication networks among these cells, ascertaining key input and output signals to elucidate how cellular coordination derives functional outcomes.

Detection and quantification of isoforms

To accurately quantify isoforms in testicular cells and obtain high-quality barcodes, we used BLZAE (v2.0.7) to identify adaptor and barcode locations, thus obtaining superior barcodes [70]. To ensure accurate barcode and UMI assignment, short-read-derived cell annotations were cross-referenced to validate and trim barcode-UMI sequences in long-read data. Validated reads were aligned to the *Sus scrofa* reference genome (Sscrofa11.1) using minimap2 (v2.24). The sequences mapped to the genome were then polished and grouped into consensus transcript assemblies. Finally, we removed the redundant isoforms and filtered out degraded 5' isoforms using cDNA-Cupcake, ensuring excellent full-length isoforms.

Isoform classification and filtering

To ensure transcript accuracy and completeness, we performed classification and quality control using SQANTI3(v5.2.0), comparing identified transcripts with the reference genome [71]. Isoform categorization was extracted using SQANTI3 and different cell types were visualized using ComplexUpset (v1.3.3). We classified FSM and ISM as known isoforms, while NIC and NNC as novel transcripts of known genes. Fusion isoforms, gene isoforms, intergenic isoforms, and antisense isoforms were categorized as transcripts of novel isoforms. To ensure transcript integrity, isoforms containing non-canonical splice sites, potential nonsense-mediated decay (NMD) targets, or 3' termini deviating > 100 bp from annotated polyadenylation sites were excluded. This stringent curation pipeline yielded a high-confidence isoform catalog for downstream functional analyses.

Abbreviations

BMP	Bone Morphogenetic Protein
CCL	C-C Motif Chemokine Ligand
DEGs	Differentially Expressed Genes
Early SPC	Early-stage Primary Spermatocytes
ENHO	Energy Homeostasis Associated
ES	Elongating/elongated Spermatids
FSM	Full Splice Match
GO	Gene Ontology
GEMs	Gel Bead-In-EMlusion
H&E	Hematoxylin and Eosin
ICT	Interstitial tissue of the testis
ISM	Incomplete Splice Match
Late SPC	Late-stage Spermatocytes
LC	Leydig Cell
M	Macrophages
MIF	Macrophage Migration Inhibitory Factor
NIC	Novel In Catalog
NNC	Novel Not in Catalog
NRG	Neuregulin
PARs	Protease-Activated Receptors
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PTM	Peritubular Myoid Cell

QC	Quality Control
RS	Round Spermatids
SC	Sertoli Cell
scRNA-seq	Single-cell RNA sequencing
SPC	Spermatocyte
Sperm	Spermatozoa
SfT	Seminiferous Tubules
SPG	Spermatogonia
ST	Spermatid
tSNE	T-distributed Stochastic Neighbor Embedding
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
VEGF	Vascular Endothelial Growth Factor
WNT	Wingless/Integrated

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11636-4.

Supplementary Figure 1: Principal component analysis. Gene expression distribution analysis of the top 15 principal components

Supplementary Figure 2: Proportional distribution and quality control metrics analysis of nine cell types (**A**) The proportional distribution of cell types. (**B**) The number of UMIs, genes, and the percentage of mitochondrial genes in nine cell types

Supplementary Table S1: A comprehensive list of differentially expressed genes (DEGs) across all nine cell clusters

Supplementary Table S2: Full Gene Ontology (GO) enrichment results for spermatogonia (SPG), early/late spermatocytes (SPC), round/elongated spermatids (RS/ES), and spermatozoa

Supplementary Table S3: This table provides detailed annotations for each isoform

Supplementary Table S4: This table summarizes the expression abundance and types of isoforms across six germ cell types

Supplementary Table S5: Information on primary antibodies in immunofluorescence

Author contributions

H.J.L. and Z.G.Y. designed and supervised the experiments; L.W., Z.X., L.Z.P., H.H.L. and C.Y.C. analyzed the Baoshan pig testicular scRNA-seq data; Z.J.D. and G.S.R. prepared the Baoshan pig testicular samples; L.W., Z.X., L.Z.P., and H.H.L. wrote the manuscript. All authors read and approved the final version of the manuscript.

Funding

This work was supported by funds from Basic Research Project of Shanxi Province, China (no. 202403021222371 to X.Z.), Basic Research Key Project of Yunnan Province, China (no. 202501AS070041 to J.L.H.), the Agricultural Joint Project of the Yunnan Province Science and Technology Department (no. 202401BD070001-004 to J.L.H.), Science and Technology Major Project of the Yunnan Province Science and Technology Department (no. 202302AE090016 to G.Y.Z.), the National Natural Science Foundation of China (no. 32060733 to J.L.H.), Key Research and Development Project of Yunnan Province (no. 2018BB003 to G.Y.Z.), Science, Education and Culture Project of Yunnan Provincial Department of Finance (no. [2022]281-A3032022211 to G.Y.Z.), Yunnan Provincial academician and expert workstation (no.202405AF140009).

Data availability

All single-cell RNA sequencing data from this project were deposited at NCBI under BioProjectID PRJNA1172176.

Declarations

Ethics declarations

The animal study was reviewed and approved by the Life Science Ethics Committee of Yunnan Agricultural University (approval number YNAU202300805). All experimental protocols were conducted in accordance with the approved guidelines.

Competing interests

The authors declare no competing interests.

Author details

[1]College of Animal Science and Technology, Yunnan Agricultural University, Kunming 650201, Yunnan, China
[2]Department of Biological and Food Engineering, Lyuliang University, Lvliang 033001, Shanxi, China
[3]Yunnan Open University, Kunming 650500, Yunnan, China
[4]Baoshan Pig Research Institute, Baoshan 678200, Yunnan, China

Received: 27 October 2024 / Accepted: 24 April 2025
Published online: 08 May 2025

References

1. Wu X, Zhao J, Gong S, Zhang H, Tan L, Fu G, Yang Z, Zhu Y, Zhao G. Performance comparative research on Baoshan pigs with Duroc×Baoshan pigs, Berkshire×Baoshan pigs. J Yunnan Agricultural Univ (Natural Scsience). 2018;33(2):249–254. (in Chinese).
2. Li Z, Fu G, Xsin J, Gong S, Zhao J, Weng Y, Qiu B, Wu X, Zhao G. Correlation of polymorphism of PRLR, NOCA1 and HTRA3 gens with litter size of Baoshan pigs. Anim Husb &Veterinary Med. 2020;52(4):19–23. (in Chinese).
3. Raeside JI, Christie HL, Renaud RL, Sinclair PA. The Boar testis: the most versatile steroid producing organ known. Soc Reprod Fertil Suppl. 2006;62:85–97.
4. Wang P, Liu Z, Zhang X, Huo H, Wang L, Dai H, Yang F, Zhao G, Huo J. Integrated analysis of LncRNA, MiRNA and mRNA expression profiles reveals regulatory pathways associated with pig testis function. Genomics. 2024;116(2):110819.
5. de Kretser DM, Loveland KL, Meinhardt A, Simorangkir D, Wreford N. Spermatogenesis. Hum Reprod. 1998;13(suppl1):1–8.
6. Fayomi AP, Orwig KE. Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. Stem Cell Res. 2018;29:207–214.
7. de Rooij DG. Proliferation and differentiation of spermatogonial stem cells. Reproduction. 2001;121(3):347–354.
8. Handel MA, Schimenti JC. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. Nat Rev Genet. 2010;11(2):124–136.
9. O'Donnell L, Nicholls PK, O'Bryan MK, McLachlan RI, Stanton PG. Spermiation: the process of sperm release. Spermatogenesis. 2011;1(1):14–35.
10. Wang Y, Navin NE. Advances and applications of single-cell sequencing technologies. Mol Cell. 2015;58(4):598–609.
11. Kulkarni A, Anderson AG, Merullo DP, Konopka G. Beyond bulk: a review of single cell transcriptomics methodologies and applications. Curr Opin Biotechnol. 2019;58:129–136.
12. Zheng Z, Chen E, Lu W, Mouradian G, Hodges M, Liang M, Liu P, Lu Y. Single-Cell transcriptomic analysis. Compr Physiol. 2020;10(2):767–783.
13. Wang M, Liu X, Chang G, Chen Y, An G, Yan L, Gao S, Xu Y, Cui Y, Dong J, et al. Single-Cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. Cell Stem Cell. 2018;23(4):599–614.
14. Dong F, Ping P, Ma Y, Chen XF. Application of single-cell RNA sequencing on human testicular samples: a comprehensive review. Int J Biol Sci. 2023;19(7):2167–2197.
15. Kavarthapu R, Anbazhagan R, Pal S, Dufau ML. Single-Cell transcriptomic profiling of the mouse testicular germ cells reveals important role of phosphorylated GRTH/DDX25 in round spermatid differentiation and acrosome biogenesis during spermiogenesis. Int J Mol Sci 2023, 24(4).
16. Di Persio S, Tekath T, Siebert-Kuss LM, Cremers JF, Wistuba J, Li X, Meyer Zu Hörste G, Drexler HCA, Wyrwoll MJ, Tüttelmann F, et al. Single-cell RNA-seq unravels alterations of the human spermatogonial stem cell compartment in patients with impaired spermatogenesis. Cell Rep Med. 2021;2(9):100395.
17. Ulrich ND, Shen YC, Ma Q, Yang K, Hannum DF, Jones A, Machlin J, Randolph JF Jr., Smith YR, Schon SB, et al. Cellular heterogeneity of human fallopian

tubes in normal and hydrosalpinx disease States identified using scRNA-seq. Dev Cell. 2022;57(7):914–929.
18. Sun YH, Wang A, Song C, Shankar G, Srivastava RK, Au KF, Li XZ. Single-molecule long-read sequencing reveals a conserved intact long RNA profile in sperm. Nat Commun. 2021;12(1):1361.
19. Tian L, Jabbari JS, Thijssen R, Gouil Q, Amarasinghe SL, Voogd O, Kariyawasam H, Du MRM, Schuster J, Wang C, et al. Comprehensive characterization of single-cell full-length isoforms in human and mouse with long-read sequencing. Genome Biol. 2021;22(1):310.
20. Hazzard B, Sá JM, Ellis AC, Pascini TV, Amin S, Wellems TE, Serre D. Long read single cell RNA sequencing reveals the isoform diversity of plasmodium Vivax transcripts. PLoS Negl Trop Dis. 2022;16(12):e0010991.
21. Hess RA, De Franca LR. Spermatogenesis and cycle of the seminiferous epithelium. Mol Mech Spermatogenesis 2009:636:1–15.
22. Huang L, Zhang J, Zhang P, Huang X, Yang W, Liu R, Sun Q, Lu Y, Zhang M, Fu Q. Single-cell RNA sequencing uncovers dynamic roadmap and cell-cell communication during Buffalo spermatogenesis. Iscience 2023, 26(1):105733.
23. Zhang L, Li F, Lei P, Guo M, Liu R, Wang L, Yu T, Lv Y, Zhang T, Zeng W, et al. Single-cell RNA-sequencing reveals the dynamic process and novel markers in Porcine spermatogenesis. J Anim Sci Biotechnol. 2021;12(1):122.
24. Yang H, Ma J, Wan Z, Wang Q, Wang Z, Zhao J, Wang F, Zhang Y. Characterization of sheep spermatogenesis through single-cell RNA sequencing. FASEB J. 2021;35(2):e21187.
25. Hermann BP, Cheng K, Singh A, Roa-De La Cruz L, Mutoji KN, Chen I-C, Gildersleeve H, Lehle JD, Mayo M, Westernströer B. The mammalian spermatogenesis single-cell transcriptome, from spermatogonial stem cells to spermatids. Cell Rep. 2018;25(6):1650–1667.
26. Wang M, Liu X, Chang G, Chen Y, An G, Yan L, Gao S, Xu Y, Cui Y, Dong J. Single-cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. Cell Stem Cell. 2018;23(4):599–614.
27. Topping T, Griswold MD. Global deletion of ALDH1A1 and ALDH1A2 genes does not affect viability but blocks spermatogenesis. Front Endocrinol. 2022;13:871225.
28. Lecluze E, Rolland AD, Filis P, Evrard B, Leverrier-Penna S, Maamar MB, Coiffec I, Lavoué V, Fowler PA, Mazaud-Guittot S, et al. Dynamics of the transcriptional landscape during human fetal testis and ovary development. Hum Reprod. 2020;35(5):1099–1119.
29. Xia P, Ouyang S, Shen R, Guo Z, Zhang G, Liu X, Yang X, Xie K, Wang D. Macrophage-Related testicular inflammation in individuals with idiopathic Non-Obstructive azoospermia: A Single-Cell analysis. Int J Mol Sci. 2023;24(10):8819.
30. Cuevas VD, Anta L, Samaniego R, Orta-Zavalza E, de la Vladimir J, Baujat G, Domínguez-Soto Á, Sánchez-Mateos P, Escribese MM, Castrillo A, et al. MAFB determines human macrophage Anti-Inflammatory polarization: relevance for the pathogenic mechanisms operating in multicentric carpotarsal osteolysis. J Immunol. 2017;198(5):2070–2081.
31. Wen Q, Wang Y, Tang J, Cheng CY, Liu YX. Sertoli cell Wt1 regulates peritubular myoid cell and fetal Leydig cell differentiation during fetal testis development. PLoS ONE. 2016;11(12):e0167920.
32. Van Saen D, Vloeberghs V, Gies I, De Schepper J, Tournaye H, Goossens E. Characterization of the stem cell niche components within the seminiferous tubules in testicular biopsies of Klinefelter patients. Fertil Steril. 2020;113(6):1183–1195.
33. Zhang LK, Ma HD, Guo M, Wang L, Zheng Y, Wu XD, Li TJ, Lu HZ, Zeng WX, Zhang T. Dynamic transcriptional atlas of male germ cells during Porcine puberty. Zool Res. 2022;43(4):600–603.
34. 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):612–624.
35. Heyn H, Ferreira HJ, Bassas L, Bonache S, Sayols S, Sandoval J, Esteller M, Larriba S. Epigenetic disruption of the PIWI pathway in human spermatogenic disorders. PLoS ONE. 2012;7(10):e47892.
36. Cao D, Shi F, Guo C, Liu Y, Lin Z, Zhang J, Li RHW, Yao Y, Liu K, Ng EHY, et al. A pathogenic DMC1 frameshift mutation causes nonobstructive azoospermia but not primary ovarian insufficiency in humans. Mol Hum Reprod. 2021;27(9):gaab058.
37. Zhang Y, Wang Y, Liu Q, Wang H, Wang Q, Shao C. The role of Z chromosome localization gene psmd9 in spermatogenesis of Cynoglossus semilaevis. Int J Mol Sci. 2024;25(12):6372.
38. Snyder E, Soundararajan R, Sharma M, Dearth A, Smith B, Braun RE. Compound heterozygosity for Y box proteins causes sterility due to loss of translational repression. PLoS Genet. 2015;11(12):e1005690.
39. Huleihel M, Jorban A, Lunenfeld E, Estrogen Specifically decreased the expression of round spermatid, but not other markers of spermatogenesis under in vitro maturation conditions; possible regulation by sertoli and leydig cells. Fertil Steril. 2024;122(4):e312.
40. La Y, Li Z, Ma X, Bao P, Chu M, Guo X, Liang C, Yan P. Age-dependent changes in the expression and localization of LYZL4, LYZL6 and PCNA during testicular development in the Ashidan Yak. Animal Biotechnol. 2024;35(1):2344213.
41. Keyhan S, Burke E, Schrott R, Huang Z, Grenier C, Price T, Raburn D, Corcoran DL, Soubry A, Hoyo C,et al. Male obesity impacts DNA methylation reprogramming in sperm. Clin Epigenetics. 2021;13:1–14.
42. Wang H, Iida-Norita R, Mashiko D, Pham AH, Miyata H, Ikawa M. Golgi associated RAB2 interactor protein family contributes to murine male fertility to various extents by assuring correct morphogenesis of sperm heads. PLoS Genet. 2024;20(6):e1011337.
43. Zhu F, Yan P, Zhang J, Cui Y, Zheng M, Cheng Y, Guo Y, Yang X, Guo X, Zhu H. Deficiency of TPPP2, a factor linked to oligoasthenozoospermia, causes subfertility in male mice. J Cell Mol Med. 2019;23(4):2583–2594.
44. Amjad S, Mushtaq S, Rehman R, Zahid N, Munir A, Siddiqui PQR. Spermatozoa retrieval in azoospermia and expression profile of JMJD1A, TNP2, and PRM2 in a subset of the Karachi population. Andrology. 2021;9(6):1934–1942.
45. Wang HQ, Wu XL, Zhang J, Wang ST, Sang YJ, Li K, Yang CF, Sun F, Li CJ. Meiotic transcriptional reprogramming mediated by cell-cell communications in humans and mice revealed by scATAC-seq and scRNA-seq. Zool Res. 2024;45(3):601–616.
46. Zhang L, Guo M, Liu Z, Liu R, Zheng Y, Yu T, Lv Y, Lu H, Zeng W, Zhang T, et al. Single-cell RNA-seq analysis of testicular somatic cell development in pigs. J Genet Genomics. 2022;49(11):1016–1028.
47. 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):548–566.
48. Guo J, Nie X, Giebler M, Mlcochova H, Wang Y, Grow EJ, Kim R, Tharmalingam M, Matilionyte G, Lindskog C. The dynamic transcriptional cell atlas of testis development during human puberty. Cell Stem Cell. 2020;26(2):262–276.
49. Wang X, Pei J, Xiong L, Guo S, Cao M, Kang Y, Ding Z, La Y, Liang C, Yan P, et al. Single-cell RNA sequencing reveals atlas of Yak testis cells. Int J Mol Sci. 2023;24(9):7982.
50. Cao C, Ma Q, Mo S, Shu G, Liu Q, Ye J, Gui Y. Single-cell RNA sequencing defines the regulation of spermatogenesis by Sertoli-cell androgen signaling. Front Cell Dev Biology. 2021;9:763267.
51. Mitchell RA, Liao H, Chesney J, Fingerle-Rowson G, Baugh J, David J, Bucala R. Macrophage migration inhibitory factor (MIF) sustains macrophage proinflammatory function by inhibiting p53: regulatory role in the innate immune response.Proc Natl Acad Sci USA. 2002, 99(1):345–350.
52. Huleihel M, Abofoul-Azab M, Abarbanel Y, Einav I, Levitas E, Lunenfeld E. Production of macrophage inhibitory factor (MIF) by primary Sertoli cells; its possible involvement in migration of spermatogonial cells. J Cell Physiol. 2017;232(10):2869–2877.
53. Itman C, Loveland KL. SMAD expression in the testis: an insight into BMP regulation of spermatogenesis. Dev Dyn. 2008;237(1):97–111.
54. Pellegrini M, Grimaldi P, Rossi P, Geremia R, Dolci S. Developmental expression of BMP4/ALK3/SMAD5 signaling pathway in the mouse testis: a potential role of BMP4 in spermatogonia differentiation. J Cell Sci. 2003;116(Pt 16):3363–3372.
55. Chapman KM, Medrano GA, Chaudhary J, Hamra FK. NRG1 and KITL signal downstream of retinoic acid in the germline to support Soma-Free syncytial growth of differentiating spermatogonia. Cell Death Discov. 2015;1:15018.
56. Zhang J, Eto K, Honmyou A, Nakao K, Kiyonari H, Abé S. Neuregulins are essential for spermatogonial proliferation and meiotic initiation in neonatal mouse testis. Development. 2011;138(15):3159–3168.
57. Chang YF, Lee-Chang JS, Harris KY, Sinha-Hikim AP, Rao MK. Role of β-catenin in post-meiotic male germ cell differentiation. PLoS ONE. 2011;6(11):e28039.
58. Chassot AA, Le Rolle M, Jourden M, Taketo MM, Ghyselinck NB, Chaboissier MC. Constitutive WNT/CTNNB1 activation triggers spermatogonial stem cell proliferation and germ cell depletion. Dev Biol. 2017;426(1):17–27.
59. Deng E, Shen Q, Zhang J, Fang Y, Chang L, Luo G, Fan X. Systematic evaluation of single-cell RNA-seq analyses performance based on long-read sequencing platforms. J Adv Res. 2024; S2090-1232(24)00210-8.
60. Byrne A, Beaudin AE, Olsen HE, Jain M, Cole C, Palmer T, DuBois RM, Forsberg EC, Akeson M, Vollmers C. Nanopore long-read RNAseq reveals widespread

transcriptional variation among the surface receptors of individual B cells. Nat Commun. 2017;8:16027.
61. Volden R, Palmer T, Byrne A, Cole C, Schmitz RJ, Green RE, Vollmers C. Improving nanopore read accuracy with the R2C2 method enables the sequencing of highly multiplexed full-length single-cell cDNA. Proc Natl Acad Sci U S A. 2018;115(39):9726–9731.
62. Boldogkői Z, Moldován N, Balázs Z, Snyder M, Tombácz D. Long-Read Sequencing - A powerful tool in viral transcriptome research. Trends Microbiol. 2019;27(7):578–592.
63. Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, Barthès P, Kokkinaki M, Nef S, Gnirke A, et al. Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. Cell Rep. 2013;3(6):2179–2190.
64. Steger K, Balhorn R. Sperm nuclear protamines: A checkpoint to control sperm chromatin quality. Anat Histol Embryol. 2018;47(4):273–279.
65. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–3587.
66. Germain P, Lun A, Garcia Meixide C, Macnair W, Robinson M. Doublet identification in single-cell sequencing data using scDblFinder [version 2; peer review: 2 approved]. *F1000Research* 2022, 10(979).
67. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods. 2017;14(10):979–982.
68. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov (Camb). 2021;2(3):100141.
69. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using cellchat. Nat Commun. 2021;12(1):1088.
70. You Y, Prawer YDJ, De Paoli-Iseppi R, Hunt CPJ, Parish CL, Shim H, Clark MB. Identification of cell barcodes from long-read single-cell RNA-seq with BLAZE. Genome Biol. 2023;24(1):66.
71. Pardo-Palacios FJ, Arzalluz-Luque A, Kondratova L, Salguero P, Mestre-Tomás J, Amorín R, Estevan-Morió E, Liu T, Nanni A, McIntyre L et al. SQANTI3: curation of long-read transcriptomes for accurate identification of known and novel isoforms. Nat Methods 2024; 21(5):793-797.

Publisher's note