Biology Direct

RESEARCH

Open Access

Single-cell RNA sequencing reveals the critical role of alternative splicing in cattle testicular spermatagonia

Xiuge Wang[1,2,3†], Chunhong Yang[1,2†], Xiaochao Wei[1,2†], Yaran Zhang[1,2], Yao Xiao[1,2], Jinpeng Wang[1,2], Qiang Jiang[1,2], Zhihua Ju[1,2], Yaping Gao[1,2], Yanqin Li[1,2], Yundong Gao[1,2] and Jinming Huang[1,2,3*]

Abstract

Spermatogonial stem cells (SSCs) form haploid gametes through the precisely regulated process of spermatogenesis. Within the testis, SSCs undergo self-renewal through mitosis, differentiation, and then enter meiosis to generate mature spermatids. This study utilized single-cell RNA sequencing on 26,888 testicular cells obtained from five Holstein bull testes, revealing the presence of five distinct germ cell types and eight somatic cell types in cattle testes. Gene expression profiling and enrichment analysis were utilized to uncover the varied functional roles of different cell types involved in cattle spermatogenesis. Additionally, unique gene markers specific to each testicular cell type were identified. Moreover, differentially expressed genes in spermatogonia exhibited notable enrichment in GO terms and KEGG pathway linked to alternative splicing. Notably, our study has shown that the activity of the YY1 regulation displays distinct expression patterns in spermatogonia, specifically targeting spliceosome proteins including RBM39, HNRNPA2B1, HNRNPH3, CPSF1, PCBP1, SRRM1, and SRRM2, which play essential roles in mRNA splicing. These results emphasize the importance of mRNA processing in spermatogonia within cattle testes, providing a basis for further investigation into their involvement in spermatogonial development.

Keywords Alternative splicing, Cattle, Single-cell RNA sequencing, Spermatogonia

Introduction

Bulls are undeniably crucial in genetic enhancement within the realm of cattle breeding, as their fertility and reproductive capabilities significantly influence the profitability of the cattle industry [2, 47]. The testis of cattle, a pivotal reproductive organ, plays essential roles in the process of sperm production. Spermatogenesis, a complex process, requires coordination among different types of testicular cells, which can be classified into three primary stages: proliferation, meiosis, and differentiation [43]. Certain spermatogonia stem cells (SSCs) undergo mitosis to maintain the population in the proliferative phase, while others progress to the meiotic phase, where spermatocytes undergo two rounds of cell division [23].

†Xiuge Wang, Chunhong Yang and Xiaochao Wei contributed equally to this work.

*Correspondence:
Jinming Huang
huangjinm@sina.com
[1]Key Laboratory of Livestock and Poultry Multi-omics of MARA, Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan 250100, P. R. China
[2]Technical Innovation Center of Dairy Cattle Breeding Industry of Shandong Province, Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan 250100, P. R. China
[3]Shandong Key Laboratory of Animal Disease Control and Breeding, Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan 250100, P. R. China

In the process of differentiation, the spermatids produced during meiosis undergo a complex series of changes to mature into functional sperm cells, involving the formation of the acrosome, development of the flagellum, and elimination of excess cytoplasm [10, 53]. However, our comprehension of the cellular composition and underlying biology of the bovine testis remains limited.

While traditional bulk analysis methods have provided valuable insights, they often mask the inherent cellular heterogeneity present in tissues such as the testis, thus limiting the ability to capture the nuanced dynamics of individual cell types. In contrast, the emergence of 10× Genomics single-cell RNA sequencing (scRNA-seq) technology in recent years has proven to be a robust tool for revealing the diversity of cell populations and elucidating the biological functions within complex tissues [4, 20, 22]. Through the application of scRNA-seq, researchers have effectively elucidated the cellular composition and heterogeneity, as well as unraveled the dynamic process of spermatogenesis in testes of human, mouse, macaque, porcine, and sheep [15, 30, 31, 54, 55]. Furthermore, the identification of conserved and divergent molecular markers and cell states of mammalian spermatogenesis has been achieved through scRNA-seq analysis of testes from human, macaque, and mouse [44]. Despite the recent advancements in scRNA-seq research spanning various tissues and organisms, there is a noticeable dearth of comprehensive single-cell studies focusing on cattle testes.

Alternative splicing (AS) is known to have significant implications in tissue development and the pathogenesis of human diseases. It has been reported that approximately one-third of genetic disorders are linked to abnormal splicing events [33]. Recent research has highlighted the intricate nature of AS in testicular and brain tissues, with a particularly high level of AS regulation observed in testicular development, crucial for the intricate development of this organ [26, 34]. The regulation of pre-mRNA selective splicing plays a critical role in the development of testicular germ cells [40]. Specifically, the targeted deletion of the splicing factor Bud31 in reproductive cells results in aberrant self-renewal and differentiation of mouse spermatogonial stem cells, improper initiation of meiosis, and ultimately, male infertility [36]. UHRF1, SRSF1, and SRSF10 are involved in maintaining the homeostasis and functionality of spermatogonial cells through the regulation of mRNA AS [28, 56]. Additionally, many RNA splicing proteins exhibit higher expression in type A spermatogonia and preleptotene spermatocytes [14, 42]. Researchers have also utilized scRNA-seq technology to analyze testicular tissue, revealing significant enrichment of Gene Ontology terms related to mRNA splicing and processing in spermatogonial cells [16].

In our research, we characterized 13 distinct populations of germ and somatic cell types within bull testes through the analysis of approximately 27,000 single-cell RNA sequencing data sets obtained from the testes of five adult bulls. By delineating the unique characteristics and gene expression patterns of each cell type, we elucidated crucial regulatory elements and signaling pathways, providing insight into the intricate cellular processes underlying cattle testicular function. Additionally, we observed a notable enrichment of genes involved in alternative splicing within spermatogonia, highlighting their significance in this biological process. Furthermore, our research has shown that the YY1 regulon displays distinct expression profiles in spermatogonia, particularly targeting spliceosome proteins essential for mRNA splicing. This investigation not only advances our knowledge of cattle testicular biology but also has implications for the development of targeted interventions to improve fertility in livestock.

Materials and methods

Animal samples

Five adult testes from Chinese Holstein cows were sourced from the bull station located in Jinan City, Shandong Province, China. The cattle testicular tissues were promptly removed and transported to the laboratory on ice in Hanks Balanced Salt Solution (HBSS) within a two-hour timeframe. Each testis was immediately sectioned into small portions and preserved by freezing in liquid nitrogen. A portion of the testicular tissue was fixed in Bouin's solution and 4% paraformaldehyde for future analysis, while the remaining tissues were utilized for cell isolation.

Cell dissociation and single-cell RNA sequencing

Single-cell suspensions of testicular tissue were generated using a two-step enzymatic digestion approach. Briefly, after washing the testicular tissue with calcium-free and magnesium-free 1X phosphate-buffered saline (PBS) containing penicillin/streptomycin, the tissues were incubated in a dissociation solution comprising 1.5 mg/mL collagenase I (Sigma, C5138-100MG) and 0.01 mg/mL DNase I (Sangon Biotech, B300065-0001) in DMEM/F12 medium (Gibco, A4192001) for 10 min at 25 °C with gentle agitation to dissociate tubules. Subsequently, the cell suspensions were centrifuged at 500 g/min for 5 min, and the supernatant was treated with 0.25% trypsin-EDTA (Gibco, 25300054) for 5 min at 37 °C in a water bath with intermittent shaking. The reaction was terminated by adding DMEM/F12 supplemented with 10% FBS. Single testicular cells were obtained by filtration through a 40 μm strainer, followed by centrifugation at 500 g/min for 10 min and washing with 1X PBS. The resulting cells were resuspended in 1X PBS containing 0.04% cattle

serum albumin (BSA) for further processing. Cell viability, confirmed by trypan blue staining, was required to be above 80%, and the cell count was determined using a hemocytometer.

Single-cell suspensions were loaded into each channel of the 10x Chromium platform to capture over 5000 single cells, adhering to the manufacturer's guidelines. The Chromium Single Cell 3'Reagent Kits (V3) (10 × Genomics, PN-1000092) were utilized for single cell barcoding, cDNA amplification, and library construction, following the standard protocol. Subsequently, libraries underwent sequencing on the Illumina HiSeq2500 instrument through paired-end sequencing runs.

Processing of single-cell RNA sequencing data

Chromium single cell data was processed by the set of pipelines in Cell Ranger (v7.0) (10 × Genomics). Raw base call files from Illumina sequencers were demultiplexed using cellranger mkfastq to generate FASTQ files. These FASTQ files underwent alignment to the cow UMD3.1 genome, filtering, barcode counting, and UMI counting using cellranger count. The output files from Cell Ranger were then imported into Seurat (v4.0) for dimensional reduction, clustering, and scRNA-seq data analysis. A total of 26,889 single cells, derived from five bull testes, passed the quality control threshold. This was determined by ensuring that all genes were expressed in more than three cells, the number of genes detected per cell ranged between 200 and 4000, and the percentage of mitochondrial gene expression was less than 5%. For data pre-processing, Seurat was additionally employed to normalize expression values across different samples. Canonical correlation analysis (CCA) was performed to identify the top 1000 highly variable genes and reduce the dimensionality of the data. Subsequently, clustering and UMAP analysis were conducted on the combined dataset, employing the first 1–12 canonical components (CCs) and utilizing a weighted Shared Nearest Neighbor (SNN) graph-based clustering method. Marker genes for each cluster were identified using the Wilcoxon rank-sum test, with default parameters applied via the FindAllMarkers function in Seurat. This selection process targeted marker genes expressed in more than 20% of the cell cluster, with a log fold change greater than 0.25.

Pseudotime analysis was conducted for various cell types using the Monocle package (v2.10.1). Trajectory reconstruction was performed using the dpFeature method, utilizing the top 1000 genes showing significant differential expression. By applying reduce dimension to the gene expression matrix, cells were ordered to generate the dynamic cell trajectory.

Enrichment analysis of differentially expressed genes

Each cell type is compared against other type cells to identify the differentially expressed genes (DEGs). The criteria set for identifying genes with expression level differences were an average log2 fold-change ≥ 1 and an adjusted *P*-value ≤ 0.01. GO-term and KEGG enrichment analysis of the DEGs in each cell type was conducted using the online tool David. Up-regulated DEGs in each cell type were uploaded as a gene list, with Bos taurus selected as the background for analysis. GO terms and KEGG pathway meeting the thresholds of gene count ≥ 3 and *P*-value ≤ 0.01 were considered significant.

Antibodies

The primary antibodies which were used in this study for immunofluorescence were as listed below: anti-UCHL1 (ABclonal, A2131); anti-WDHD1 (ABclonal, A15396); anti-ACRV1 (ABclonal, A8098); anti-WT1 (ABclonal, A16298); anti-INHA (ABclonal, A1734); anti-MYH11 (ABclonal, A10827); anti-ITGA7 (ABclonal, A14246); anti-rabbit Alexa Fluor 488 (Thermo Fisher Scientific, A32790); antirabbit Alexa Fluor 555 (Thermo Fisher Scientific, A32794).

Analysis of transcription factors in germ cells

SCENIC [1] was used to conduct the regulatory network and regulon activity analysis by combing human and mouse transciption factor (TF) datasets. Co-expression modules between transcription factors and candidate target genes were built using GRNBoost. TF motif enrichment analysis and direct TF regulons were derived based on motif enrichment around TSS using the RcisTarget al.gorithm. AUCell was employed to measure and score the activity of each regulon in single cells.

Chromatin immunoprecipitation-sequencing (ChIP-Seq)

Testicular tissue samples were collected from adult mice, and YY1-bound chromatin fragments were enriched using ChIP-Seq methodology. In brief, testicular tissue cells were initially cross-linked to DNA and proteins, followed by cell lysis. Subsequently, DNA was fragmented via sonication, then subjected to immunoprecipitation using a specific YY1 antibody (Santa Cruz Biotechnology, sc-7341) to target sequences. Finally, DNA was purified, underwent second-generation library construction and sequencing, and was then analyzed to generate high-resolution maps of YY1 binding sites across the bovine genome.

Results

Cell clustering analysis of single-cell RNA sequencing

Testicular tissues from five Holstein bulls, aged 2, 4, 4, 5, and 5 years (designated as A1, D1, E1, F1, and G1, respectively), were utilized for scRNA-seq. The testis cells were

dissociated using a two-step mechanical and enzymatic digestion method, resulting in a viability rate exceeding 80% for each sample. Subsequently, the isolated single cells were used to generate a single-cell library through the 10× Genomics platform (Fig. 1A). A total of 40,975 cells were captured, with 26,888 cells meeting standard quality control (QC) criteria for further analysis. The libraries produced 3117 Mb reads, with an average sequencing saturation rate of 56.94%. Each cell demonstrated a mean of 75.49 K reads and identified around 1612 genes (Table S1). Furthermore, we analyzed the histological composition and cell types of the seminiferous tubules in cattle testis through hematoxylin-eosin (H&E) staining (Fig. 1B).

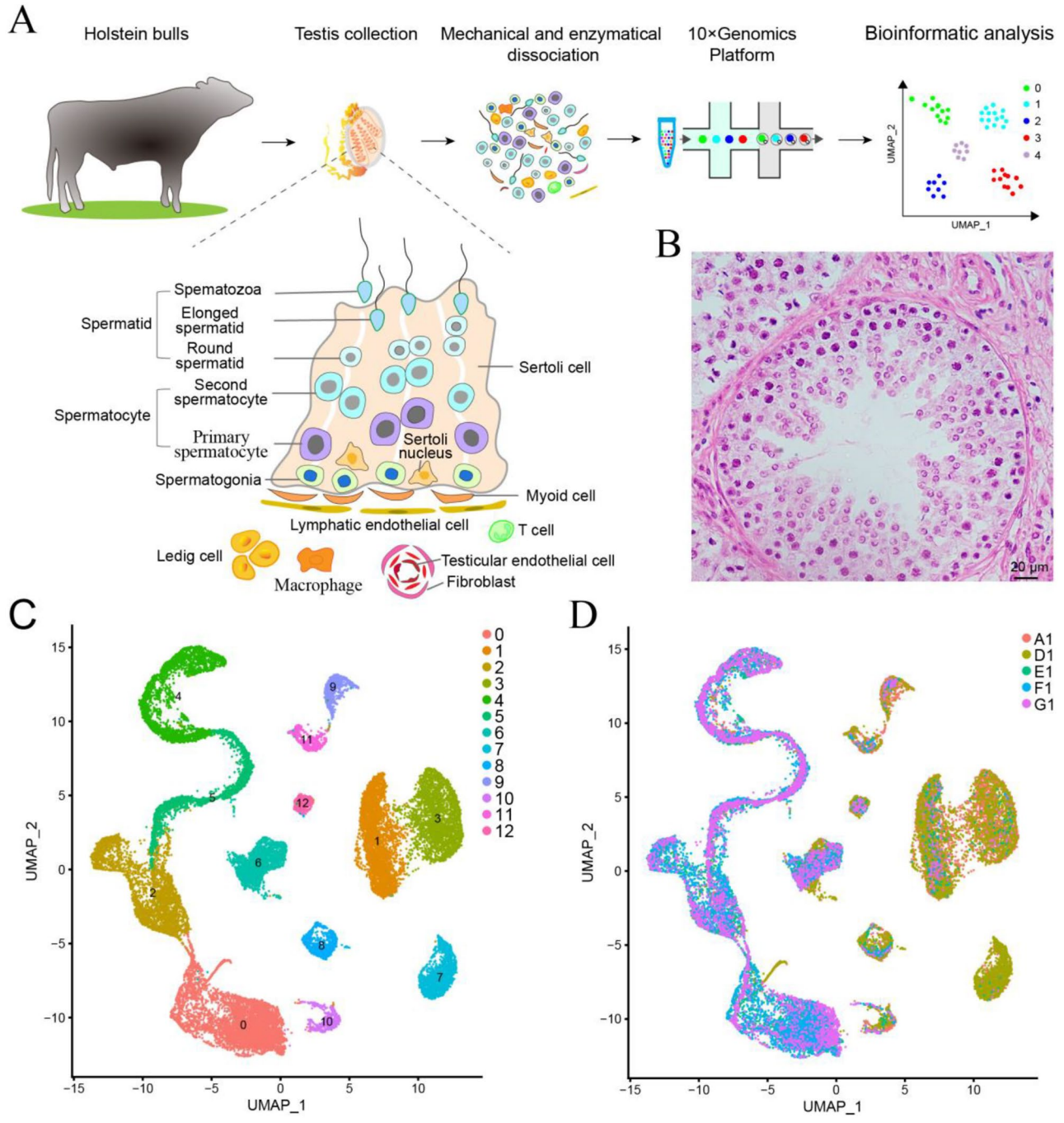

Fig. 1 Cell clustering analysis of scRNA-seq from cattle testes. (**A**) Experimental workflow of sample collection and analysis. (**B**) HE staining of cattle testis. (**C**) Cell clustering analysis of combined single-cell transcriptome data from cattle testes. Each dot represents a single cell and is colored according to cell clusters. The 13 cell clusters were visualized in UMAP. (**D**) The cells colored in each cluster indicating the origin of testis tissues

The vst method for dimensionality reduction in principal component analysis (PCA) was utilized to identify the most highly variable genes. Subsequently, the top principal components and a graph-based method were utilized to implement the uniform manifold approximation and projection (UMAP) algorithm, leading to the delineation of 13 distinct cell clusters visualized in a two-dimensional space through UMAP (Fig. 1C). Importantly, all identified clusters were consistently present in each testis sample (Fig. 1D).

Identification of germ and somatic cell types in cattle testis

Utilizing unsupervised clustering, a total of 13 discrete cell clusters were identified within the dataset, consisting of 5 germ cells and 8 somatic cells in cattle testes (Fig. 2A). Distinction between germ cells and somatic cells was accurately achieved based on *DDX4* and *VIM* markers (Figure S1), thereby validating the accuracy of our clustering outcomes [24, 41]. Established marker genes for specific cell types were employed for the identification and examination of cell clusters in cattle testes. Germ cell types could be further categorized as spermatogonia (*DAZL*, *DMRT1*, *UCHL1*, *SYCP3*, *STRA8*), spermatocytes (*GKAP1*, *TBPL1*, *UBE2C*, *EFHD1*), round spermatids (*TMEM190*, *ACRV1*), elongating spermatids (*OTUB2*, *TNP1*, *PRM2*, *KIF5C*), and sperm (*CA2*, *TSSK6*). Somatic cell types encompassed Sertoli (*FATE1*, *CITED1*), myoid (*MYH11*), lymphatic endothelial (*MMRN1*), testicular endothelial (*PRSS23*, *CD34*, *VWF*, *CD34*, *TIE1*), Leydig (*IGF1*, *DCN*, *SFRP2*), T cell (*CD52*, *CD69*), macrophage (*C1QA*, *C1QB*, *C1QC*, *CD14*, *CD163*, *CD74*, *CSF1R*, *LYZ*), and fibroblast (*CXCL12*) (Fig. 2B; Figure S2). The major cell types identified in this study were in accordance with those observed in the testes of other mammalian species [44], as anticipated and supported by subsequent analysis. To confirm the clustering outcomes, an expression analysis of cluster-specific marker genes was performed using Loupe software, demonstrating elevated expression levels in specific cell types (Fig. 2C; Figure S3).

Moreover, a statistical analysis was conducted on gene numbers, total transcripts, and cell numbers for each cell type in cattle testis (Fig. 2D). The findings revealed that germ cells, with the exception of elongating spermatids, displayed a lower number of genes and UMI transcripts sequenced per cell in comparison to somatic cells. Furthermore, it was observed that spermatogonia, macrophages, and myoid cells exhibited the lowest cell counts relative to other cell types. Moreover, somatic cells in cattle testis displayed an average of 2 K detected genes, a significantly higher number compared to other species [47], indicating unique gene expression profiles and robust cellular conditions.

Gene enriched analysis of testicular cell types in single-cell sequencing

Gene expression patterns dictate cellular functions; therefore, we conducted a comparative analysis of gene expression among cell groups. Through this analysis, differentially expressed genes (DEGs) were identified between various testicular cell types in the dataset (Fig. 3A) (Table S2). The findings suggest the presence of potentially novel cell-type-specific marker genes that may serve as distinguishing factors for different cell types in cattle testes. For instance, we identified *TKTL1*, *FMR1NB*, *ESX1*, *WDHD1*, and *RBBP8* as marker genes for spermatogonia, *GSN*, *APOD*, *OGN*, *COL1A2*, *APOE*, *COL1A1*, *GSTM4*, *LAMA1*, and *C4A* for Leydig cells, and *DEFB119*, *INHA*, *OXT*, *AQP8*, *WFDC15B*, and *AARD* for Sertoli cells, and *BMX* for fibroblast. Notably, *TKTL1* was also identified as a marker gene for spermatogonia in cattle testes, consistent with previous studies [49]. In addition, we have selected the antibodies of these newly - identified DEGs, and further verified their expression locations in different cell types within the testis through immunofluorescence (Figure S4) Thus, our analysis confirmed known marker genes in mammalian testicular cells and revealed candidate genes with potential functions in cattle spermatogonial development.

To explore the functional enrichment of individual testicular cell types within the single-cell sequencing dataset, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis on the upregulated Differentially Expressed Genes (DEGs) using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) [8]. Our examination revealed unique functional enrichment profiles across various germ and somatic cell types and emphasized the top five upregulated GO terms and KEGG pathways for differentially expressed genes associated with each cell type (Fig. 3B) (Table S3). For example, germ cells displayed significant enrichment of genes associated with spermatogenesis, particularly those involved in flagellated sperm motility and sperm flagellum. In contrast, genes highly enriched in spermatogonia primarily functioned in histone binding and RNA binding. Sertoli cells exhibited enrichment of genes related to mitochondrial structure and function, whereas Leydig cells showed enrichment of genes associated with extracellular matrix organization and integrin binding. Moreover, the enrichment of ribosome-related genes was observed in testicular endothelial cells, lymphatic endothelial cells, and T cells. Subsequent analysis of gene expression patterns in fibroblasts and myoids revealed a distinct cluster of genes associated with protein binding and actin binding. Additionally, a minority population of macrophages exhibited high enrichment for genes linked to immune response. The enrichment of these GO terms across different

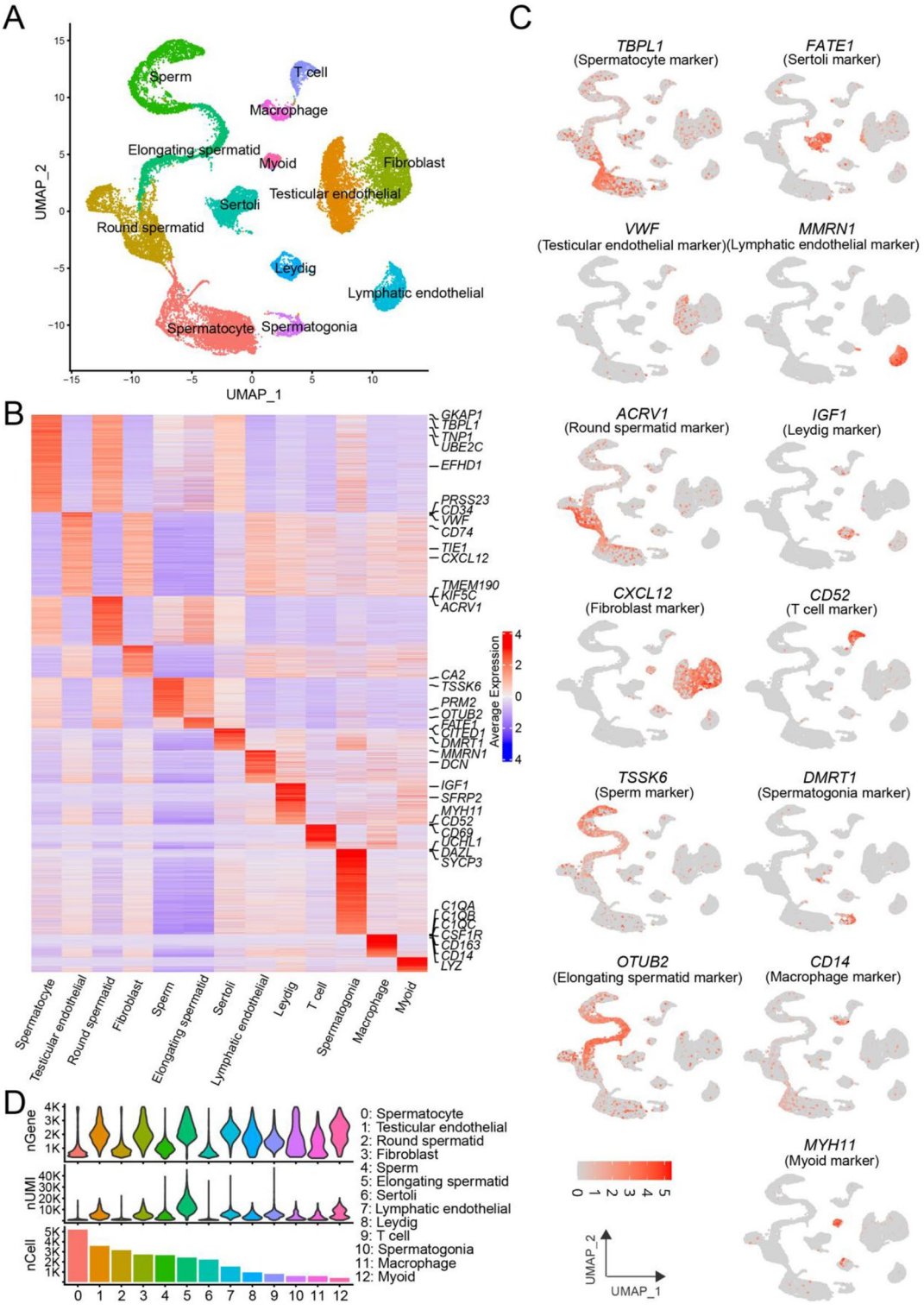

Fig. 2 Cell types and specific marker genes identified in scRNA-seq datasets. (**A**) Major testis cell types corresponding to cell clusters in cattle testes in UMAP. (**B**) Heatmap showing of the relative expression of representative cell-specific marker genes in the 13 major cell types for single-cell RNA-seq datasets. Color bar indicates the relative expression level for selected genes. (**C**) Expression patterns of selected markers used to assign cell types projected on the UMAP plots. One cell marker for each cell type is shown in the Figure Red (or gray) indicates a high (or low) expression level as shown on the color key. (**D**) Distribution profiles of per-cell attributes compared across the 13 cell types. nGene, the number of detected genes per cell; nUMI, the number of Unique Molecular Identifiers (UMI) per cell; nCell, the number of cells within each cell type

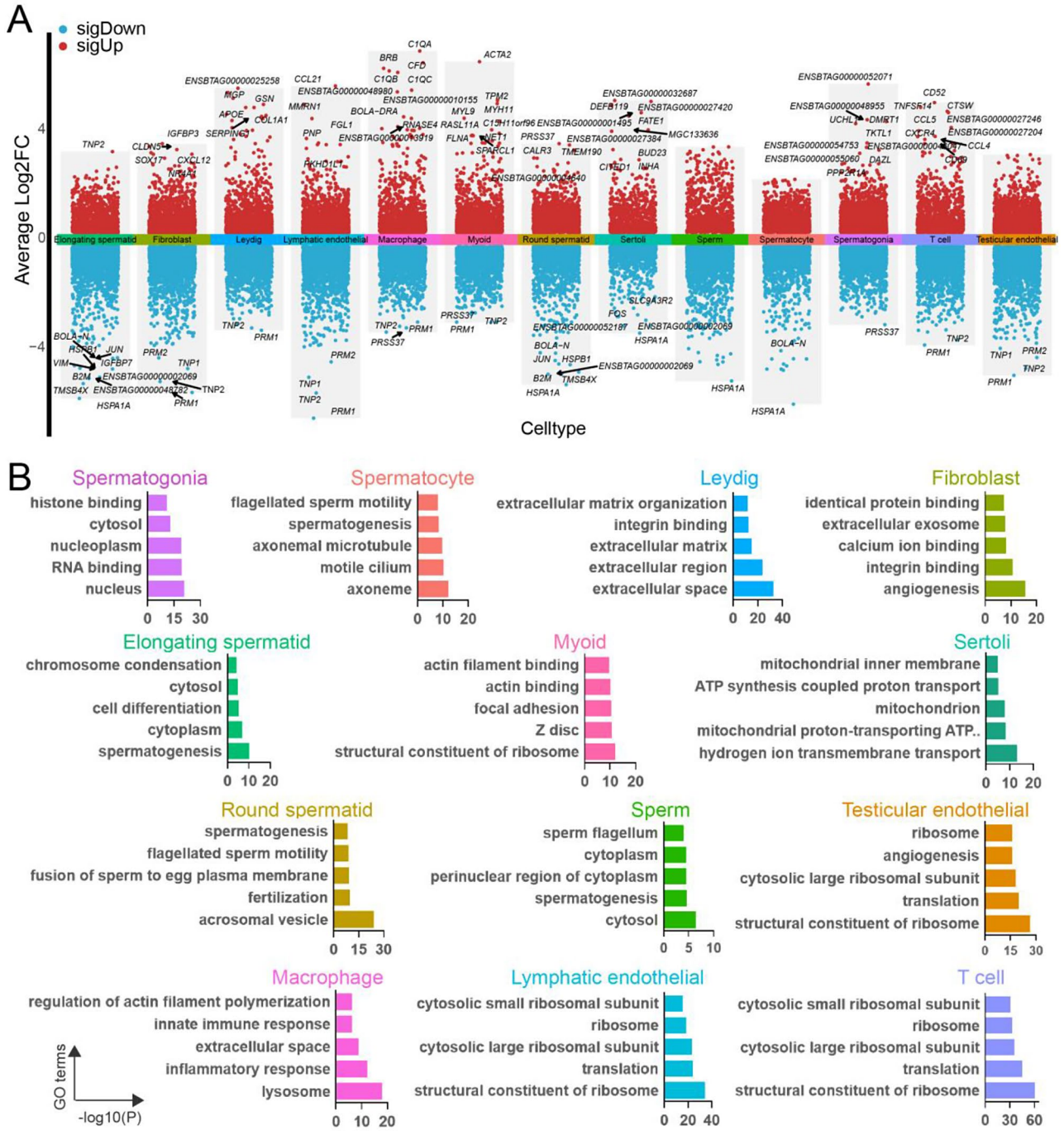

Fig. 3 Cell clustering analysis of scRNA-seq from cattle testes. (**A**) Significantly differentially-expressed genes for the testicular cell types. (**B**) The enriched GO terms shown for each cell type. The top 5 up-regulated GO terms are given on differentially expressed genes corresponding cell types by DAVID

testicular cell types holds significance in the context of the biological processes underlying spermatogenesis [5, 13].

The development trajectory analysis in cattle testis germ cells

To investigate the differentiation pathways of germ cells in the testis, trajectory analysis was performed utilizing the Monocle 2 package [40]. A subset of 3,000 germ cells from the single-cell RNA sequencing dataset was chosen, with the top 1000 genes exhibiting significant differential expression selected for further investigation. Initially, the developmental trajectory of germ cells was reconstructed by arranging them along a pseudo time axis according to their gene expression profiles. Following this, the Monocle 2 algorithm was utilized to identify branching points

and infer differentiation pathways within germ cells of the cattle testis. Our pseudo time analysis revealed a sequential progression from spermatogonia to spermatocytes, round spermatids, elongating spermatids, and ultimately sperm, providing insight into the spatial distribution of various germ cell types within the testis (Fig. 4A-B). During the developmental trajectory of germ cells, we observed multiple gene expression changes, such as heightened expression of *HSPA8* and *HNRNPA2B1* in spermatogonia, alongside a well-documented gradual upregulation of protamine genes (*PRM1*, *PRM2*) crucial for sperm chromatin condensation (Fig. 4C) [32].

Additionally, a heatmap was generated to visually represent the hierarchical relationships among clusters of genes displaying differential expression throughout pseudo time in cattle testis germ cells (Fig. 4D). The top 20 GO terms for clusters 1, 2, and 3 are emphasized in bold on the right side of the heatmap (Table S4). Cluster 1, consisting of pre-branch spermatogonia and spermatocytes, displayed enrichment in ribosome and RNA processing pathways. Cluster 2 exhibited enrichment in genes related to sperm biology, while cluster 3 showed enrichment in genes associated with acrosome generation.

Regulatory activity and transcription factors in cattle testis germ cells

In the context of testicular single-cell sequencing regulatory activity pertains to the modulation of particular gene regulatory networks within individual cells by transcription factors (TFs). These networks consist of genes regulated by the same TF or combinations thereof. TFs are proteins that interact with specific DNA sequences, thereby modulating gene expression. To identify key TFs and associated regulatory activities in our testicular single-cell sequencing dataset, we utilized the SCENIC pipeline followed by employing the AUCell algorithm to quantify and score regulon activity across individual cells.

Our study identified the top 100 transcription factor regulates that displayed differential activity at various stages of germ cell development in cattle testes (Fig. 5A; Figure S5) (Table S5). Specifically, we observed consistent high activity of the HMGA1 regulation throughout the entire process of germ cell development [25]. Conversely, the *ZFP14* and *RFX6* regulation exhibited heightened activity in elongating spermatids compared to other germ cell types. Additionally, TFs such as *TAL1*, *TCF12*, and *SREBF2* displayed specific regulated activities in sperm cells, although their roles in fertilization and embryonic development require further investigation. Our study has uncovered heightened and distinct expression patterns of YY1 in spermatogonia, indicating significant functions of this transcription factor in this particular cell type. Previous research has emphasized YY1's participation in the maintenance of spermatogonial stem cells and meiosis [21, 52]. The YY1 regulatory network of target genes in spermatogonia was established, with a particular emphasis on YY1's involvement in the cattle testis (Fig. 5B). These identified genes are either known or predicted to be involved in encoding splicing factors or proteins that govern splicing processes (Fig. 5C). Our observations revealed a heightened activity of the SOX9 regulation in various somatic cell types, which has been recognized for its pivotal role in Sertoli cell differentiation and function [7, 18]. Additionally, the expression of the RXRB regulation in Sertoli cells was found to regulate cholesterol homeostasis and spermiation [48].

Functional role of alternative splicing in cattle testis spermatogonia

Our study demonstrates the important functional roles of dynamic alterations in mRNA processing factors during spermatogonia development in bovine testes. Specifically, heterogeneous nuclear ribonucleoproteins (hnRNPs) like RBM39, HNRNPA2B1, and HNRNPH3, as well as serine/argenine rich proteins (SRs), cleavage and polyadnylation specificity factor (CPSF) and splicing-related accessory proteins which were implicated in spliceosomal complex assembly and mRNA splicing, were found to be significantly enriched (Figs. 4D and 5B-C; Figure S6) (https://www.wikipathways.org/instance/WP1023) [6]. Moreover, via KEGG enrichment analysis, it was also revealed that the DEGs in spermatogonia were remarkably enriched in the spliceosome signaling pathway (Table S3). Subsequently, an enrichment analysis was performed utilizing upregulated DEGs in spermatogonia. This analysis revealed a significant enrichment of GO terms associated with ribonucleoprotein complexes, spliceosomes, mRNA binding, and splicing (Fig. 6A) (Table S6), highlighting the importance of alternative splicing in the maintenance of spermatogonial developmental trajectory and functional competence.

Our prior investigation revealed numerous mRNA splicing-enriched genes as downstream targets of YY1. In order to confirm YY1-mediated regulation of mRNA splicing in cattle testis, we utilized ChIP-Seq to identify YY1 binding sites and their corresponding target genes. Through de novo motif discovery and comparison with established motif databases using Homer software, we determined that the highest-ranking motif aligned with YY1 transcription factor binding sites (Fig. 6B), validating the relevance of our experimental findings for subsequent analysis. This analysis revealed an enrichment of genes targeted by YY1 that are implicated in spliceosome assembly, RNA processing, mRNA splicing, and associated biological processes (Fig. 6C) (Table S7). Additionally, peak binding sequences were identified in the upstream regulatory regions of genes previously

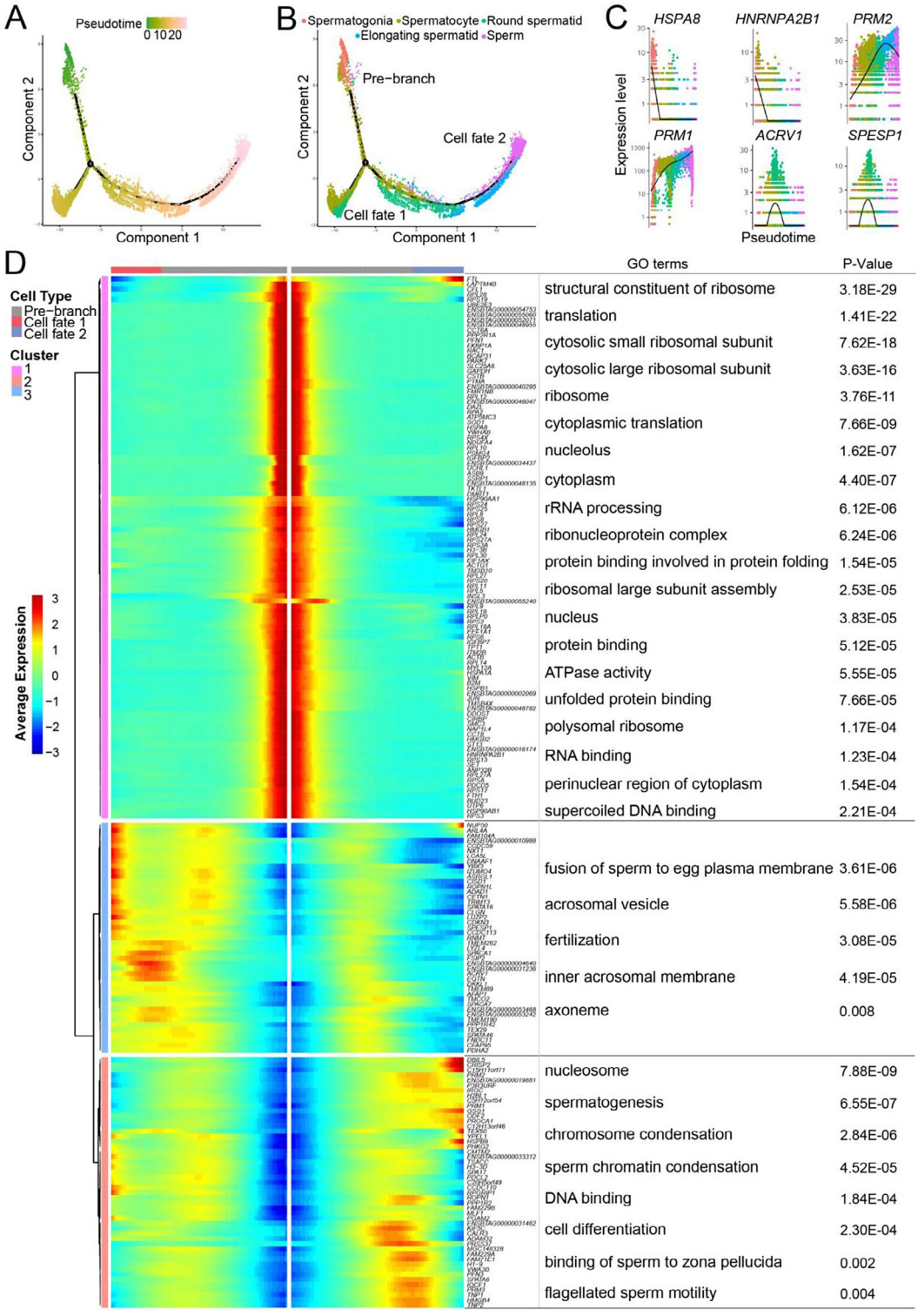

Fig. 4 The developmental trajectory of cattle testicular germ cells. (**A**) Pseudotime trajectory order of cattle testicular germ cells. Each dot represents an individual cell. The color bar from dark to light indicates the start and end of the trajectory. (**B**) The successive development trajectory of testicular germ cells. Cells are colored to their predicted locations along pseudotime trajectory. Three discrete cellular states (pre-branch, cell fate 1, cell fate 2) during germ cell development. (**C**) Expression profile of the selected genes involved in cellular state along the pseudotime trajectory. Black line indicates expression tendency across germ cells. (**D**) Heatmap and GO terms enrichment analysis of the significantly differentially-expressed genes. The significantly DEGs were enriched on the three gene clusters showed by the heatmap analysis. Associated GO terms with each gene cluster is shown on the corresponding right

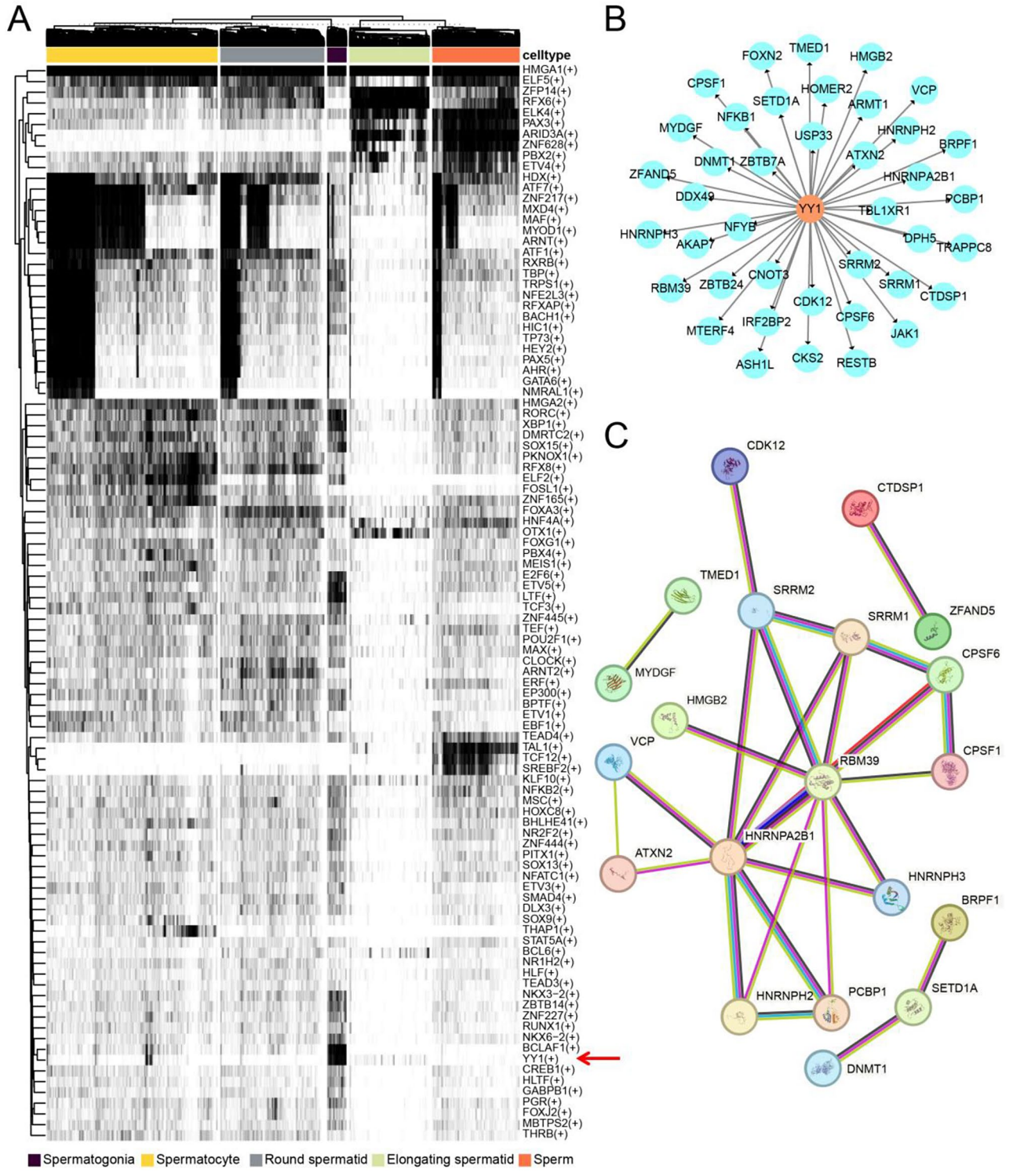

Fig. 5 Transcription factor regulatory network of cattle testis in germ cell development. (**A**) Binary heatmap depicting regulon activity in germ cell-type. 100 regulon activity between transcription factors and target genes in single cell were conducted by SCENIC. Consensus DNA motifs regulated by representative TFs are depicted. Color legends refer to cell cluster. (**B**) The target genes of YY1 regulon in spermatogonia. (**C**) The enriched analysis of YY1 target genes (**B**) predicted by STRING (Version 12.0)

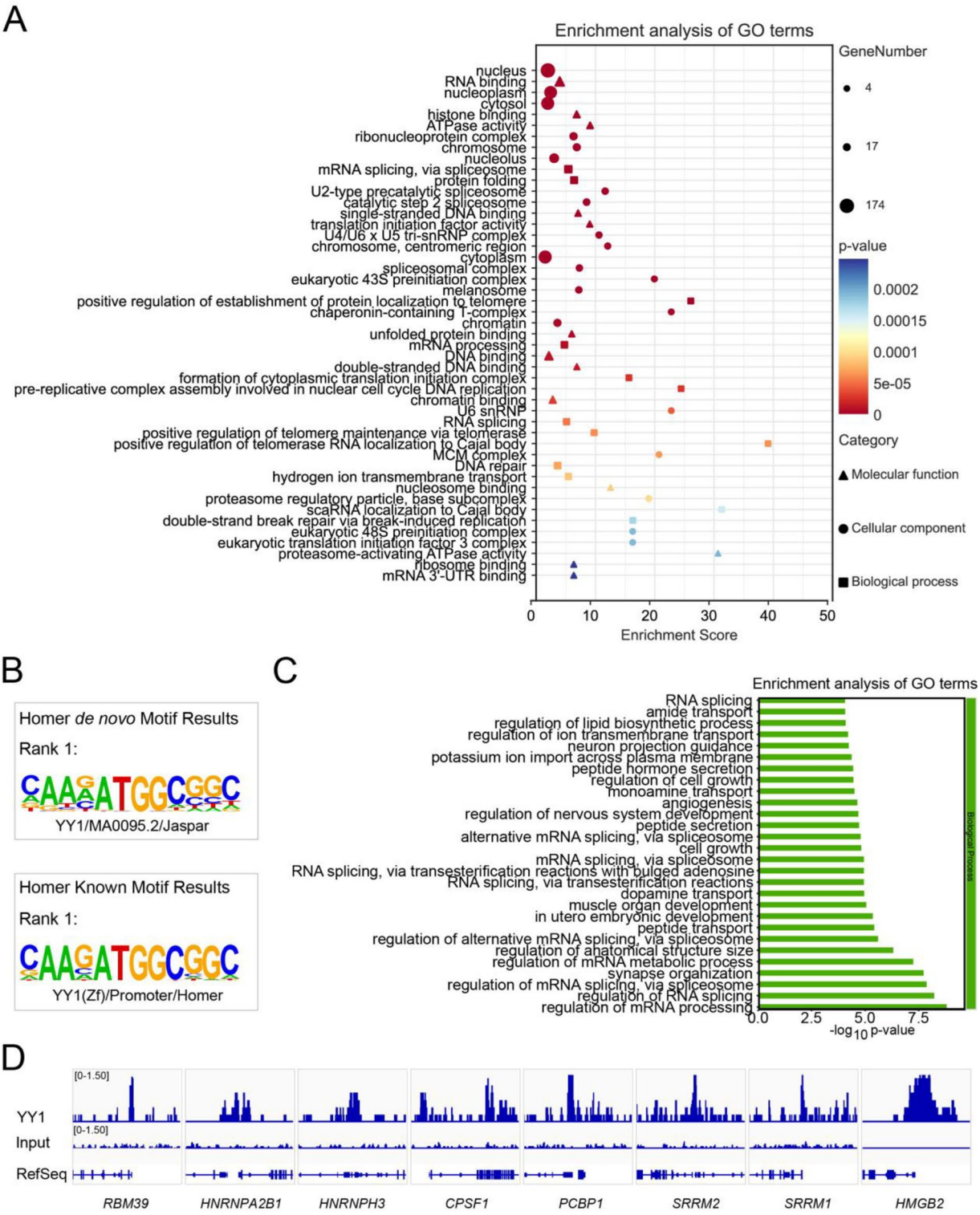

Fig. 6 Functional role of mRNA splicing in cattle testicular spermatogonia. (**A**) The enriched GO terms of upregulated DEGs (FDR ≤ 0.01) in spermatogonia. (**B**) Logo showing the rank 1 of de novo motif results and known binding motif results using Homer. Homer analysis results showed the logo of the de novo motif (upper) high coincident with the known binding motif (lower). (**C**) Enrichment analysis of the biological processes of GO terms for YY1-targeted genes ($P \leq 0.0001$). (**D**) Peak binding sequences in the upstream regulatory regions of YY1-regulated target genes

predicted to be regulated by YY1, including hnRNPs (*RBM39*, *HNRNPA2B1*, *HNRNPH3*), SRs (*SRRM2*, *SRRM1*), CPSF (*CPSF1*), accessory proteins (*PCBP1*) and *HMGB2* (Fig. 6D). These findings reinforce YY1's regulatory influence on mRNA splicing.

Discussion

The genetic influence of breeding bulls on cow populations surpasses 75%, underscoring the pivotal role of bull semen in animal reproduction. The quality of bull semen significantly affects reproductive efficacy and genetic enhancement in livestock. Despite the importance of this aspect, our comprehension of spermatogenesis in bull testes remains restricted. This study utilized scRNA-seq to investigate the transcriptomic profile of bull testes at cellular resolution. Our analysis has revealed distinct cell populations present in testicular tissue, providing a comprehensive understanding of spermatogenesis and its regulatory mechanisms in cattle, which is consistent with similar findings in other animal species. The presence of this cellular diversity highlights the dynamic and tightly controlled nature of spermatogenesis in cattle.

The gene expression profiles of individual cell types exhibited distinct functional enrichment, underscoring their distinctive biological functions. Extensive documentation of major testicular cell types and their associated biological markers has been conducted in mammalian research [13, 16, 17]. This study identified a multitude of uniquely expressed genes that have not been previously characterized within each cell type. Additionally, we screened and identified genes co-expressed in two cattle cell types: *TOP2A* in spermatogonia and spermatocytes, and *EDNRA* and *TCF21* in Leydig and Myoid cells. These genes are believed to have significant functions in particular stages or cell types in the process of cattle spermatogenesis. Our results emphasize the importance of multiple specific regulatory genes in this process, highlighting the need for additional annotation and functional studies. Additionally, our research indicates that enriched GO terms associated with spermatogenesis in cattle share similarities with those identified in yak and porcine species, particularly in spermatogonia, spermatids, and Sertoli cells [49, 55].

Trajectory analysis was utilized in this study to explore the differentiation pathways of testicular germ cells. The analysis revealed notable changes in gene expression during different stages of germ cell development, such as the downregulation of genes related to mRNA processing, the upregulation of genes associated with chromatin condensation, and increased expression involved in acrosome formation. Distinct clusters representing germ cells at different stages of spermatogenesis were identified, including spermatogonia, spermatocytes, round spermatids, elongating spermatids, and ultimately, spermatids. For instance, pathways such as ribosome and RNA processing likely play pivotal roles in the developmental transition from spermatogonia to spermatocytes. Each cell type is characterized by unique gene expression profiles indicative of their developmental stage and functional roles. Additionally, our research identified crucial transcriptional regulators and signaling pathways that govern cattle spermatogenesis. Notably, transcription factors YY1 and BCLAF1 displayed unique expression patterns in spermatogonia, as validated by SCENIC analysis of regulon activity in human and mouse models [31]. Pathway analysis further revealed enrichment of the YY1 regulatory network linked to mRNA splicing, emphasizing its role in coordinating spermatozoa production. Furthermore, we described the activities of transcription factor regulons not previously reported during spermatogenesis (e.g., *ZFP14*, *RFX6*, *TAL1*, *TCF12*, *SREBF2*). These findings contribute to unraveling the complex regulatory networks driving testicular cell differentiation, development, spermatogenesis, and other biological processes.

The AS process is predominantly controlled by the spliceosome, which orchestrates the assembly and identification of pre-mRNA splicing sites through the regulation of RNA-binding proteins (RBPs) [29, 39]. RBPs associated with spliceosome proteins engage with RNA and other proteins to establish hnRNPs that are integral in the modulation of various cellular functions, such as transcription, pre-mRNA splicing, mRNA transportation, localization, degradation, storage, and translation [19, 38, 50]. Despite numerous discoveries, the precise mechanisms through which AS influences spermatogenesis remain predominantly obscure. Particularly, the factors driving the extensive mRNA splicing and processing in spermatogonia, the regulatory mechanisms governing alternative splicing mRNA during ontogenesis, and the functions of various RNA-binding proteins in this complex process remain inadequately elucidated. The equilibrium between self-renewal and differentiation of spermatogonial cells within the testicular environment is imperative for the sustenance of spermatogenesis [9, 35]. However, many questions regarding the regulatory mechanisms governing this intricate process's homeostasis remain unanswered. In our study, we identified upregulated RBPs significantly enriched in GO terms related to mRNA splicing (e.g., hnRNPs, spliceosomal complex, mRNA 3'-UTR binding) in spermatogonia, highlighting their critical role in cattle spermatogenesis.

Multiple transcription factors, such as FOXO1, BCL6B, ETV5, TAF4B, and SOHLH1, have been identified as key regulators in the processes of self-renewal and differentiation of spermatogonial stem cells [3, 27, 45, 46]. Our study revealed that YY1 displays distinct expression patterns in spermatogonia compared to other germ cells,

with a notable enrichment of RNA-binding proteins involved in mRNA splicing among its target genes. Using ChIP-seq analysis, a multitude of YY1 binding sites have been discovered within the upstream regulatory regions of spliceosome proteins crucial for mRNA splicing in cattle testes. Additionally, functional RBPs associated with spermatogenesis have been identified [12]. Previous studies have demonstrated elevated expression of HNRNPA2B1 in mouse testicular cells, with knockout of *Hnrnpa2b1* leading to infertility in mice [51]. Additionally, HNRNPH1 regulates alternative splicing of developmental-related target genes in reproductive cells by recruiting splicing factors PTBP2 and SRSF3, thereby maintaining normal development and function of germ cells [11]. Together, these findings underscore YY1's regulatory role in RBP spliceosome proteins (hnRNPs, SRs, CPSF, and other accessory proteins) involved in the self-renewal and differentiation of spermatogonia in cattle (Figure S7).

This research provides a thorough analysis of single-cell RNA sequencing data obtained from cattle testes, offering a comprehensive overview of gene expression patterns among different cell types. The study identifies potential biomarkers and emphasizes their importance in understanding the processes involved in spermatogenesis, thus establishing a basis for further investigation into the molecular mechanisms influencing cattle fertility and infertility. Furthermore, our research results offer strong evidence supporting the essential functional significance of mRNA processing in spermatogonia within cattle testes, laying the groundwork for further investigation into their functional roles in spermatogonial development. These findings enhance our comprehension of cattle reproductive biology, not only adding to fundamental knowledge but also showing potential for advancements in reproductive health within livestock and potentially in human fertility research.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13062-024-00579-7.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Acknowledgements
Not applicable.

Author contributions
J.M.H. conceived and designed the experiments. X.G.W. and C.H.Y. conducted the computational analyses. Y.X., X.C.W., Q.J. and Z.H.J. performed experiments. Y.R.Z, J.P.W., and Y.P.G. collected and processed the testis tissues, collected samples and prepared materials. The manuscript was written by X.G.W. and J.M.H. All the authors read and approved the manuscript.

Funding
This study was supported by the National Natural Science Foundation of China (32272854, 32202677, 32202648), National Key Research and Development Program of China (2021YFF1000700), Shandong Agricultural Elite Variety Project (2022LZGC012, 2021LZG011), and Agricultural Scientific and Technological Innovation Project of Shandong Academy of Agricultural Sciences (CXGC2024A03 and CXGC2024B03).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
All experimental procedures involving animals were ethically approved by the Institutional Animal Care and Use Committee of the Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences (IACC20060101).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 31 July 2024 / Accepted: 5 December 2024
Published online: 26 December 2024

References

1 Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017;14(11):1083–86.
2 Butler ML, Bormann JM, Weaber RL, Grieger DM, Rolf MM. Selection for bull fertility: a review. Transl Anim Sci. 2019;4(1):423–41.
3 Ballow D, Meistrich ML, Matzuk M, Rajkovic A. Sohlh1 is essential for spermatogonial differentiation. Dev Biol. 2006;294(1):161–7.
4 Cheng C, Chen W, Jin H, Chen X. A review of single-cell RNA-Seq annotation, integration, and cell-cell communication. Cells. 2023;12(15):1970.
5 Chen Y, Zheng Y, Gao Y, Lin Z, Yang S, Wang T, Wang Q, Xie N, Hua R, Liu M, Sha J, Griswold MD, Li J, Tang F, Tong MH. Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. Cell Res. 2018;28(9):879–96.
6 Carmody SR, Wente SR. mRNA nuclear export at a glance. J Cell Sci. 2009;122(Pt 12):1933–7.
7 Chen M, Wang J, Liu N, et al. Pig SOX9: expression profiles of sertoli cell (SCs) and a functional 18 bp indel affecting testis weight. Theriogenology. 2019;138:94–101.
8 Dennis G Jr, Sherman BT, Hosack DA, et al. DAVID: database for annotation, visualization, and Integrated Discovery. Genome Biol. 2003;4(5):P3.
9 De Rooij DG, Griswold MD. Questions about spermatogonia posed and answered since 2000. J Androl. 2001;22(5):746–57.
10 Fawcett DW. The mammalian spermatozoon. Dev Biol. 1975;44(2):394–436.
11 Feng S, Li J, Wen H, Liu K, Gui Y, Wen Y, Wang X, Yuan S. hnRNPH1 recruits PTBP2 and SRSF3 to modulate alternative splicing in germ cells. Nat Commun. 2022;13(1):3588.
12 Fu K, Tian S, Tan H, Wang C, Wang H, Wang M, Wang Y, Chen Z, Wang Y, Yue Q, Xu Q, Zhang S, Li H, Xie J, Lin M, Luo M, Chen F, Ye L, Zheng K. Biological and RNA regulatory function of MOV10 in mammalian germ cells. BMC Biol. 2019;17(1):39.
13 Guo J, Grow EJ, Mlcochova H, Maher GJ, Lindskog C, Nie X, Guo Y, Takei Y, Yun J, Cai L, Kim R, Carrell DT, Goriely A, Hotaling JM, Cairns BR. The adult human testis transcriptional cell atlas. Cell Res. 2018;28(12):1141–57.

14 Gan H, Cai T, Lin X, et al. Integrative proteomic and transcriptomic analyses reveal multiple post-transcriptional regulatory mechanisms of mouse spermatogenesis. Mol Cell Proteom. 2013;12(5):1144–57.
15 Guo J, Grow EJ, Yi 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.
16 Green CD, Ma Q, Manske GL, Shami AN, Zheng X, Marini S, Moritz L, Sultan C, Gurczynski SJ, Moore BB, et al. A comprehensive roadmap of murine spermatogenesis defined by single-cell RNA-seq. Dev Cell. 2018;46:651–e66710.
17 Grive KJ, Hu Y, Shu E, Grimson A, Elemento O, Grenier JK, Cohen PE. Dynamic transcriptome profiles within spermatogonial and spermatocyte populations during postnatal testis maturation revealed by single-cell sequencing. PLoS Genet. 2019;15:e1007810.
18 Hemendinger RA, Gores P, Blacksten L, Harley V, Halberstadt C. Identification of a specific sertoli cell marker, Sox9, for use in transplantation. Cell Transpl. 2002;11(6):499–505.
19 Hall KB. RNA and proteins: mutual respect. F1000Res. 2017;6:345.
20 Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: a brief overview. Clin Transl Med. 2022;12(3):e694.
21 Kim JS, Chae JH, Cheon YP, Kim CG. Reciprocal localization of transcription factors YY1 and CP2c in spermatogonial stem cells and their putative roles during spermatogenesis. Acta Histochem. 2016;118(7):685–92.
22 Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. Mol Cell. 2015;58(4):610–20.
23 Kubota H, Brinster RL. Spermatogonial stem cells. Biol Reprod. 2018;99(1):52–74.
24 Kim JY, Jung HJ, Yoon MJ. VASA (DDX4) is a putative marker for Spermatogonia, Spermatocytes and Round spermatids in stallions. Reprod Domest Anim. 2015;50(6):1032–8.
25 Liu J, Schiltz JF, Ashar HR, Chada KK. Hmga1 is required for normal sperm development. Mol Reprod Dev. 2003;66(1):81–9.
26 Legrand JMD, Hobbs RM. RNA processing in the male germline: mechanisms and implications for fertility. Semin Cell Dev Biol. 2018;79:80–91.
27 Lovasco LA, Gustafson EA, Seymour KA, de Rooij DG, Freiman RN. TAF4b is required for mouse spermatogonial stem cell development. Stem Cells. 2015;33(4):1267–76.
28 Liu W, Lu X, Zhao ZH, et al. SRSF10 is essential for progenitor spermatogonia expansion by regulating alternative splicing. Elife. 2022;11:e78211.
29 Lei WL, Du Z, Meng TG, Su R, Li YY, Liu W, Sun SM, Liu MY, Hou Y, Zhang CH, Gui Y, Schatten H, Han Z, Liu C, Sun F, Wang ZB, Qian WP, Sun QY. SRSF2 is required for mRNA splicing during spermatogenesis. BMC Biol. 2023;21(1):231.
30 Lukassen S, Bosch E, Ekici AB, Winterpacht A. Single-cell RNA sequencing of adult mouse testes. Sci Data. 2018;5:180192.
31 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.
32 Merges GE, Meier J, Schneider S, et al. Loss of Prm1 leads to defective chromatin protamination, impaired PRM2 processing, reduced sperm motility and subfertility in male mice. Development. 2022;149(12):dev200330.
33 Maniatis T, Tasic B. Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature. 2002;418(6894):236–43.
34 Merkin J, Russell C, Chen P, Burge CB. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. Science. 2012;338(6114):1593–9.
35 Nakagawa T, Nabeshima Y, Yoshida S. Functional identification of the actual and potential stem cell compartments in mouse spermatogenesis. Dev Cell. 2007;12(2):195–206.
36 Qin J, Huang T, Wang Z, Zhang X, Wang J, Dang Q, Cui D, Wang X, Zhai Y, Zhao L, Lu G, Shao C, Li S, Liu H, Liu Z. Bud31-mediated alternative splicing is required for spermatogonial stem cell self-renewal and differentiation. Cell Death Differ. 2023;30(1):184–94.
37 Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-cell mRNA quantification and Differential Analysis with Census. Nat Methods. 2017;14(3):309–15.
38 Ripin N, Parker R. Formation, function, and pathology of RNP granules. Cell. 2023;186(22):4737–56.
39 Sun L, Lv Z, Chen X, et al. Splicing factor SRSF1 is essential for homing of precursor spermatogonial stem cells in mice. Elife. 2024;12:RP89316.
40 Song H, Wang L, Chen D, Li F. The function of Pre-mRNA Alternative Splicing in Mammal Spermatogenesis. Int J Biol Sci. 2020;16(1):38–48.
41 Sharma S, Schlatt S, Van Pelt A, Neuhaus N. Characterization and population dynamics of germ cells in adult macaque testicular cultures. PLoS ONE. 2019;14(6):e0218194.
42 Schmid R, Grellscheid SN, Ehrmann I, et al. The splicing landscape is globally reprogrammed during male meiosis. Nucleic Acids Res. 2013;41(22):10170–84.
43 Staub C, Johnson L, Review. Spermatogenesis in the bull. Animal. 2018;12(s1):s27–35.
44 Shami AN, Zheng X, Munyoki SK, et al. Single-cell RNA sequencing of Human, Macaque, and mouse testes uncovers conserved and divergent features of mammalian spermatogenesis. Dev Cell. 2020;54(4):529–47.
45 Song HW, Wilkinson MF. Transcriptional control of spermatogonial maintenance and differentiation. Semin Cell Dev Biol. 2014;30:14–26.
46 Simon L, Ekman GC, Garcia T, et al. ETV5 regulates sertoli cell chemokines involved in mouse stem/progenitor spermatogonia maintenance. Stem Cells. 2010;28(10):1882–92.
47 Souza FM, Lopes FB, Rosa GJM, Magnabosco CU. Economic values of reproductive, growth, feed efficiency and carcass traits in Nellore cattle. J Anim Breed Genet. 2022;139(2):170–80.
48 Vernet N, Dennefeld C, Klopfenstein M, et al. Retinoid X receptor beta (RXRB) expression in sertoli cells controls cholesterol homeostasis and spermiation. Reproduction. 2008;136(5):619–26.
49 Wang X, Pei J, Xiong L, et al. Single-cell RNA sequencing reveals Atlas of Yak Testis cells. Int J Mol Sci. 2023;24(9):7982.
50 Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009;136(4):701–18.
51 Wang X, Fan X, Zhang J, et al. hnRNPA2B1 represses the disassembly of arsenite-induced stress granules and is essential for male fertility. Cell Rep. 2024;43(2):113769.
52 Wu S, Hu YC, Liu H, Shi Y. Loss of YY1 impacts the heterochromatic state and meiotic double-strand breaks during mouse spermatogenesis. Mol Cell Biol. 2009;29(23):6245–56.
53 Yasuzumi G. Electron microscope studies on spermiogenesis in various animal species. Int Rev Cytol. 1974;37(0):53–119.
54 Yu XW, Li TT, Du XM, et al. Single-cell RNA sequencing reveals atlas of dairy goat testis cells. Zool Res. 2021;42(4):401–5.
55 Zhang L, Li F, Lei P, et al. Single-cell RNA-sequencing reveals the dynamic process and novel markers in porcine spermatogenesis. J Anim Sci Biotechnol. 2021;12(1):122.
56 Zhou S, Dong J, Xiong M, et al. UHRF1 interacts with snRNAs and regulates alternative splicing in mouse spermatogonial stem cells. Stem Cell Rep. 2022;17(8):1859–73.

Publisher's note