# Single cell ATAC-Seq reveals cell type-specific transcriptional regulation and unique chromatin accessibility in human spermatogenesis

Xiaolong Wu[1, #], Mujun Lu[2, #], Damin Yun[1], Sheng Gao[1], Shitao Chen[3], Longfei Hu[4], Yunhao Wu[1], Xiaorong Wang[1], Enkui Duan[5, *], C. Yan Cheng[6, *], Fei Sun[1, *]

[1]Medical School, Institute of Reproductive Medicine, Nantong University, Nantong 226001, Jiangsu, China; [2]International Peace Maternity and Child Health Hospital, Shanghai Key Laboratory for Reproductive Medicine, School of Medicine, Shanghai Jiaotong University, Shanghai 200030, China; [3]Department of Urology and Andrology, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200001, China; [4]Singleron Biotechnologies Ltd., 211 Pubin Road, Nanjing, Jiangsu, China; [5]State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China; [6]The Mary M. Wohlford Laboratory for Male Contraceptive Research, Center for Biomedical Research, Population Council, 1230 York Ave, New York, NY10065

[#]These authors contributed equally to this work.

[*]Correspondence should be addressed to Fei Sun, Medical School, Institute of Reproductive Medicine, Nantong University, Nantong 226001, Jiangsu, China. E-mail: sunfei@ntu.edu.cn; Tel/fax: 0513-8505 1865; C. Yan Cheng, The Mary M. Wohlford Laboratory for Male Contraceptive Research, Center for Biomedical Research, Population Council, 1230 York Ave, New York, NY 10065; E-mail: yancheng01@aol.com; or Enkui Duan, State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. E-mail: duane@ioz.ac.cn.

## Abstract

During human spermatogenesis, germ cells undergo dynamic changes in chromatin organization/re-packaging and in transcriptomes. In order to better understand the underlying mechanism(s), scATAC-Seq of 5376 testicular cells from 3 normal men were performed. Data were analyzed in parallel with the scRNA-Seq data of human testicular cells. Ten germ cell types associated with spermatogenesis and 6 testicular somatic cell types were identified, along with 142,024 peaks located in promoter, genebody and CpG Island. We had examined chromatin accessibility of all chromosomes, with chromosomes 19 and 17 emerged as the leading chromosomes that displayed high chromatin accessibility. In accessible chromatin regions, transcription factor (TF)-binding sites were identified and specific motifs with high frequencies at different spermatogenesis stages were detected, including CTCF, BORIS, NFY, DMRT6, EN1, ISL1 and GLI3. Two most notable observations were noted. First, TLE3 was specifically expressed in differentiating spermatogonia. Second, PFN4 was found to be involved in actin cytoskeletal organization during meiosis. More important, unique regions upstream of *PFN4* and *TLE3* were shown to display high accessibility, illustrating their significance in supporting human spermatogenesis.

## Introduction

During spermatogenesis in men, germ cells undergo dynamic chromatin packing/repacking and organizational remodeling to facilitate the diverse expression of transcriptomes to support the production of functional haploid sperm in particular spermiogenesis (1). On the other hand, spermatogonial stem cells (SSCs) undergo active mitotic self-renewal to maintain the SSC population since a small pool of SSCs differentiate into meiotic cells periodically at puberty by ~12 years of age (2). Using RNA-Seq coupled with 5hmc (5-hydroxymethylcytosine) sequencing to examine stage-specific high-resolution profiles in adult germline stem cells (AGSC), including DNA methylation and histone modifications, AGSC signaling and chromatin/DNAme strategies emerge as the key regulatory mechanisms to maintain the crucial level of vital transcription factors to activate DNA-methylation in promoters during spermatogenesis (3). DNA methylation serves as an epigenetic marker, and following methylation establishment, spermatogonia begin a mitotic proliferative burst and differentiate into sperm (4). A recent study has demonstrated disordered genome methylation modification of germ cells in NOA patients, which was considerably different from reproductive process-mediated gene expression (5). In order to better understand how human adult SSC (hSSC) maintain the intriguing balance between self-renewal and differentiation, a study was performed to profile open chromatin (ATAC-Seq), DNA methylation, bulk, and single cell RNA transcriptomes (RNA-Seq) in hSSCs (6). This has identified four different developmental stages during hSSC differentiation and established the dynamic chromatin/transcription landscape operating in hSSCs (Guo, et al., 2017). During the mitotic-to-meiosis transition, mitotic-type open chromatin is closed, which is followed by the *de novo* formation of meiotic-type open chromatin.

Single cell RNA sequencing (scRNA-Seq) is a powerful tool to identify novel regulators of male germ cell development, which has revealed unexpected cell types in the adult testis that support spermatogenesis (7). scRNA-Seq also identifies unique molecular signatures in gene expression attributed to different development stages of germ cells (8). Single cell ATAC-Seq (assay for transposase-accessible chromatin using sequencing) is a new approach to study open chromatin in individual cells and to assess genome-wide chromatin accessibility, which provides insights into cell-to-cell variations in chromatin organization by gathering hundreds or thousands of single cells data (9). This approach has also been used to profile chromatin accessibility to identity 85 distinct patterns across 13 tissues in mice, mapping the regulatory landscape at single-cell resolution (10).

Relationship between dynamic gene expression, chromatin organization and accessibility in human spermatogenesis remains unknown. Hence scATAC-Seq was used to obtain detailed chromatin organization and accessibility data from *n*=3 normal men, and compared to scRNA-Seq data of normal men as recently reported (11). These analyses provided detailed assessment regarding the dynamic reorganization of open chromatin and transcriptomes during human spermatogenesis. We also investigated chromatin accessibility in all chromosomes, with chromosomes 19 and 17 emerged as the two chromosomes displaying considerable higher chromatin accessibility during spermatogenesis. We also identified the transcription factor (TF)-binding sites and uncovered some crucial motifs in high frequencies from accessible chromatin regions at different stages of spermatogenesis. *TLE3* and *PFN4* were shown to be actively involved in human spermatogenesis, with regions in their upstream to have high accessibility.

## Results

**scRNA-Seq data analysis recognize 10 germ cell types and 6 somatic cell types.** Seurat analysis was used to identify germ and somatic cell types. According to the clustering results, cells were classified to 16 cell types based on known markers. Ten of the cell types comprised of different germ cell classes distributed continuously and six somatic cell types distributed separately **(Fig. 1A)**. **Fig. 1B** displays the marker genes expression status. Among the continuous cell types, we confirmed SSCs, Diff.ingSPG and Diff.edSPG cells by GFRA1, KIT and STRA8, respectively **(Fig. 1C)**. The marker gene of meiotic-specific double-strand breaks, SPO11, was noted which expressed specifically in leptotene and zygotene clusters. TEX19 is a regulator of SPO11-dependent meiotic recombination in leptotene spermatocytes. Combined with the expression of these two genes, we identified leptotene spermatocytes. It has been reported that OVOL1 and OVOL2 are zinc finger transcription factors involved in pachytene progression and are localized in the XY body of spermatocytes at the pachytene and diplotene stages. NME8 expressed through the progress of pachytene to early stage spermatids. Using these three genes, the pachytene and diplotene spermatocytes were separated from other cells. After cell type determination, zygotene cells were identified by exclusion. Subsequently, round spermatids were identified by TXNDC2 and the sperm by TNP1 and PRM1. Through the expression pattern of TXNDC2, TNP1 and PRM1 genes, elongate spermatids, which were formed in-between round spermatids and sperm in time sequence, were identified. Besides germ cells, six testicular somatic cell types were identified. In these cell types, the marker genes used to identify macrophages were CD163 and CD14. Endothelial cell marker genes were PECAM1 and VWF. Leydig cell markers were VIT, IGF1, and DLK1. Myoid cell markers were GLI1, MYH11, and ACTA2. Sertoli cell

markers were PRND, WFDC2, AMH, and SOX9. Pericytes specifically expressed NOTCH3, SLIT3, PDGFRB, and MCAM.

**Preliminary analysis of scATAC-Seq.** Following scATAC-Seq, we checked the promoter ratio and UMI (unique molecular identifiers) number and set thresholds to select cells for downstream analysis. Fig. 2A shows the selected distribution of cells identified by scRNA-Seq *versus* the current scATAC-Seq data in this report, illustrating scATAC-Seq was a considerably powerful technique. According to the prediction scores, we isolated 4685 cells from a total of 5376 cells with scores greater than 0.3 (Fig. 2B). Human spermatogenesis is comprised of a series of complex cellular events, the prediction scores of 10 germ cell types were relatively low, with the exception of SSCs, when compared to somatic cells (Fig. 2C). For the cell types identification, SnapATAC was used to integrate scRNA-Seq and scATAC-Seq datasets. First, the software was used to perform preprocessing of sequencing data including demultiplexing, reads alignments and filtering, duplicate removal and barcode selection by SnapTools. Similarity of chromatin accessibility between cells were then assessed to resolve the heterogeneity of cell population by the normalize Jaccard similarity matrix, which were then used for the following dimensionality reduction. To combine with the annotated scRNA-Seq data, we used "snapToSeurat" function in SnapATAC to convert the snap object to Seurat object. The two datasets were merged by "FindTransferAnchors" and "TransferData" functions of Seurat. In the process of integration, prediction scores were generated. These 4685 cells and their corresponding types were shown in Fig. 2D. Cell types in the cell annotation result were analogous to the scRNA-Seq analysis as noted in Fig. 1A, including 10 cell types associated with human spermatogenesis and 6 testicular somatic cell types. Fig. 2E shows the percentage of different cell

types based scRNA-Seq vs. scATAC-Seq. As noted here, using scRNA-Seq, about 25% of the total cells were the six somatic cell types and only 75% were germ cells; however, with scATAC-Seq, only 10% of all cells were somatic cell types and 90% were germ cells (Fig. 2E). Except for sperm1, the relative percentage of the different germ cell types were comparable between the two sequence methods (Fig. 2E). Unlike scRNA-Seq, the Sertoli cell population was most predominant among the somatic cell types in human testes based on scATAC-Seq (Fig. 2E). Findings based on scATAC-Seq (Fig. 2E) are more consistent with earlier reports regarding the relative percentage of different somatic cell types (e.g., Sertoli cells) and germ cells types in human testes (12, 13), illustrating the scATAC-Seq method is a more reliable and robust technique to assess cell type distributions in human testes. Using the scATAC-Seq data, statistical analysis of peaks location to have 142,024 peaks (Fig. 2F). The number of peaks located in genebody were 69,718, about 3 times of the promoter (26,796), CpG Islands (21,014) and TEs (24,496). Part of the peaks were overlapping in the three regions. This finding is consistent with earlier findings regarding TEs wherein germline was silenced in several repressive pathways (14, 15). For example, small RNAs, in particular piRNA and siRNA, are produced from transposable elements (TEs), which, in turn, are used to silence TEs (16).

**Dynamic chromatin accessibility and its association with chromosomes 17 and 19.** As shown in Fig.3A, the total peak numbers in different germ cell types during spermatogenesis were varied considerably. From SSCs to difgSPG, and diSPG, the numbers of peaks progressively increased, suggesting multiple genes began to initiate transcription. This observation is consistent with earlier findings that multiple cell types, such as Leydig and Sertoli cells, that support the SSC niche to regulate the SSC-renewal and differentiation also become transcriptionally active during

postnatal testicular maturation when meiosis I/II begins (17). During meiosis, the number of peaks continued to rise, including zygotene and pachytene/diplotene spermatocytes (Fig. 3A), indicating that gene expression was robustly activated to support spermatogenesis. However, in the stages of RoundS.tids and Elongates.tids, the peak numbers began to decline since gene expression began to turn off following the formation of haploid spermatids which began the process of genetic material condensation (18). Using coaccessibility enrich analysis (Fig. 3B), the coaccessibility enrichment of chr19 was considerably higher than any other chromosomes. The coaccessibility might be due to the highest gene density of chr19 in all human chromosomes (19). Chromosome 17 was found to have the second highest coaccessibility enrichment (Fig. 3B). This unusual coaccessibility enrichment in chr19 and chr17 might be correlated with meiosis. Studies have shown that the DSBs (double-strand breaks) hotspots specified by PRDM9 protein since DSBs and PRDM9 were shown to be crucial to support recombination during meiosis, and PRDM9 evolved rapidly and with a large number of variants (20, 21). In the process of DSB generation, PRDM9 protein binds to DNA to promote the formation of recombination initial complex with the SPO11 in the complex to introduce meiotic DSBs, which in turn increased chromatin accessibility of chr 19 and 17. Because of their higher accessibility, chr 17 and 19 were delected to displayed peaks distribution in cell types of different spermatogenesis stages to extract additional information as noted in Fig. 3C, D. Different stages of spermatogenesis from early stage, differentiating spermatogonia, to the late stage, elongate spermatids were sorted by time sequence (Fig. 3C, D). In the differentiating spermatogonia of chromosome 19, cells had the highest chromatin accessibility than other stages. In contrast, the chr17, which displayed the second highest coaccessibility enrichment, had the lower chromatin accessibility in differentiating spermatogonia. Also, leptotene and zygotene spermatocytes had higher chromatin accessibility than

pachytene/diplotene spermatocytes (Fig. 3C), which indicated that at the early stage of meiosis, gene expression was more prevalent than the middle and later stage. These results suggest that more genes are working to support early and late spermatogenesis on chr19 and on chr17, respectively. Besides these two chromosomes, chr1, chr11, chr12, chr16, chr20 and chr22 also present co-accessibility enrichment versus other chromosomes.

**Motif analysis reveals proteins involved in human spermatogenesis.** Motif analysis was performed on the eight germ cell types (Fig. 1A). Since at late spermatogenesis, chromatin became inaccessible and specific mRNAs required for subsequent transcription were stored in cytoplasmic ribonucleoprotein particles, we thus selected only 8 germ cell types for analysis without sperm1 and 2 (22, 23). In Fig. 4A, some motifs were also found in high frequencies during spermatogenesis, including CTCF (zf), BORIS, NYF, DMRT6 (MD), EN1 (Homeobox), DMRT1 (DM), FOXL2 (Forkhead box L2), which were enriched in these 8 cell types and were readily identified (Fig. 4A). It was noted that most motifs were enriched in Pachytene/Diplotene (Fig. 4 B). The CTCF and BORIS motifs were found in seven germ cell types, except SSCs (Fig. 4B). CTCF is a DNA binding protein that regulates higher order chromatin structure and genome topology (24). Deletion of *Ctcf* led to defects of sperm head formation and chromatin compaction in mice (25). BORIS (also called CTCFL) is a paralogue of CTCF and specifically expressed in the testis, which shares the same exons encoding the 11 ZF domain as of CTCF genes. However, the pattern of expression of CTCF and BORIS are in sharp contrast to each other during male germ differentiation, correlating with re-setting of methylation marks. For instance, abnormal BORIS expression affects the normal function of CTCF including epigenetic dysregulation (26). As a sibling of the regulator of imprinted sites with CTCF, BORIS is considered an oncogene, and highly expressed in

primary spermatocytes (27). BORIS bound to the *CST* promoter to activate the regulation of *CST*, which played a critical role in spermatogenesis. NFY (also called CBF and CP1) is a trimeric protein comprised of NFYA, NFYB, and NFYC, which is known to bind specifically to CCAAT-box and found in 30% of eukaryotic promoters (28). Deletion of CBF-A (corresponding to NFY-B) was shown to influence PRM2 expression, causing abnormal DNA morphology in high percentage of sperm (29). DMRT are transcription factors of a highly conserved family whose members share a unique DNA binding motif. Dmrt6 coordinated an orderly programming in differentiating spermatogonia to allow mitotic/spermatogonial entry to the meiotic/spermatocyte during the epithelial cycle (30). The absence of *Dmrt6* disrupted spermatogonial differentiation, resulting in a disordered expression of spermatogonial differentiation factors, including DMRD1, SHOLH1, SOLHLH2 and STRA8 (a meiotic initiation factor) (31). En1 motif proteins were overexpressed in transcription dependencies of triple-negative breast cancer and associated with co-activators and transcriptional repressor factor TLE3, which highly expressed in the testis (32). However, the role of EN1 in human spermatogenesis is presently unknown. ISL1 and GLI3 were restrictively expressed in pachytene/diplotene (Fig.4B). ISL1 (insulin gene enhancer protein) can be antagonized by MEIS2 which is a regulator of the retinoic acid pathway (33). Retinoic acid is crucial to support sperm maturation in mammals, and it is essential to initiate meiosis in adult testes (34). Thus, ISL1 may modulate meiosis during spermatogenesis through retinoic acid (34). As a member of the Hedgehog (Hh/HH) signaling pathway, GLI3 regulates HH genes by targeting GLI1 promoter (35). GLI1 expression was only noted in spermatocytes and round spermatids, but GLI1 was not detected nor highly enriched in our data. Forkhead box (FOX) proteins, consisting of FOXL1 and FOXL2, are transcription factors involved in cell cycle regulation, including growth, proliferation, differentiation, and longevity. FOXL2 also regulates sex-determination by repressing SOX9 to inhibit testis formation (36), and the secretion of

LH and FSH to modulate spermatogenesis (37), and its disruption leads to male oligozoospermia (38).

We also explored genes with differential region within 10kb upstream or 1kb downstream of transcription initiation site. Overlapped genes for genes annotated by differential accessible regions and differential expressed genes in scRNA-Seq were selected, and the correlation between the chromatin accessibility and expression of genes was analyzed based on the scATAC-Seq logFC and scRNA-Seq avg_logFC. The R-value was 0.48, and the Pearson correlation was significant. The genes in quadrant I and quadrant III had the same change trend on scATAC-Seq logFC and scRNA-Seq avg_logFC (Fig.4C, D). The expression of these genes was regulated by the accessible regions. The expression of genes in the I quadrant was upregulated with higher chromatin accessibility. In the Difg.SPG, the genes were upregulated with higher chromatin accessibility, including *ZNF207*, *MORF4L1*, *COPZ1*, *CBX5*, *TAF1P*, *FAF2*, *TLE3* and others (Fig.4C). For example, *ZNF207* controls human embryonic stem cell self-renewal and differentiation (39). Also, in the Pachytene/Diplotene stage, we found that seven genes were upregulated with higher chromatin accessibility, including *AMZ2*, *TMEM38B*, *RTF1*, *OIP5*, *POMP*, *FAM228A* and *PFN4* (Fig.4D), illustrating their possible involvement in human spermatogenesis. Among these, *TLE3* and *PFN4* were further analyzed since these two genes displayed similar transcriptome expression level based on scRNA-Seq and accessibility region peak noted in scATAC-Seq.

**GO analysis in Difg.SPG and Pachytene/Diplotene.** GO analysis was performed for Difg. SPG (Fig. 5) and Pachytene/Diplotene spermatocytes (Fig. 6). For Diff.ed SPG GO analysis, several biological processes, cellular components and molecular function pertinent to spermatogenesis were noted (Fig. 5A). Based on results of the differentially accessible region, Difg.SPG and

Pachytene/Diplotene were selected for further analysis as noted in Fig. 4C, D. In these cells, TLE3 (Fig. 5B, C) and PFN4 (Fig. 6B, C) were found to be specifically and prominently expressed in Difg.SPG and Pachytene/Diplotene, respectively (Fig 5B, C and Fig 6B, C). Further analysis revealed some regions in the upstream of *TLE3* and *PFN4* that displayed high accessibility. TLE3 (transducin-like enhance of split 3) was highly expressed in difg.SPG (Fig 5B, C). In humans, TLE3 is a member of the Groucho/Grg/TLE family, which interacts with multiple transcription factors (40). Since these transcription factors lack the NDA binding regions, they regulate target gene expression through TLE3 by restricting DNA-binding proteins in particular histone acetylation (41). Based on scATAC-Seq, motifs were found with high accessibility enriched specifically in Difg.SPG in the upstream region close to *TLE3* (Fig. 5C). As such, proteins containing these motifs might influence the expression of TLE3, such as CCCTC-binding factor (CTCF), a highly conserved zinc-finger DNA-protein known to regulate chromatin organization to support spermatogenesis. Deletion of *Ctcf* in mice led to small testes and male infertility with defects in nuclear morphology and chromatin compaction (42).

For Pachytene/Diplotene GO analysis, several biological processes, pathways, cellular components, and molecular function (MF) were enriched (Fig. 6A). As shown in Fig. 6B and C, PFN4 was notably expressed in Pachytene/Diploptene and RoundS.tids. Profilin (*pfn*) is a small actin-binding protein consisting of four isoforms of Pfn1, Pfn2, Pfn3, and Pfn4, known to involve in actin polymerization (43, 44). Pfn1 and pfn2, pfn3 and pfn4 are highly expressed in the testis to support spermatogenesis (45). Besides serving as an actin binding protein to support actin polymerization, Pfn proteins create a hub composed of a complex network of proteins to support molecular interactions (46, 47). For instance, pfn interacts with Cdc42, Rac1, Ras, RhoA and Rap that are part of the actin cytoskeleton regulatory GTPase network (44). In our scATAC-Seq analysis, CTCT, BORIS (zf), GLI3, and

ISL1(homeobox) motif were enriched in the upstream of PFN4 (Fig. 5B), which were also enriched in the pachytene/diplotene spermatocytes (Fig 4.B). It is noted that the role of CTCT, BORIS (zf), GLI3, and ISL1 in human spermatogenesis are presently unknown. In mouse, BORIS is restrictively expressed in germ cells, exclusively detected in the nucleus of spermatocytes and cytoplasmic of spermatogonia in early spermatogenesis, and the BORIS promoter is hypomethylated in normal testis and completely unmethylated in sperm (27, 48). As a member of *Gli* proteins family, *Gli3* was detected in mouse germ cells, crucial to support germ cell development (49).

## Discussion

In this report, we have demonstrated unequivocally that scATAC-Seq is a better analytical tool than scRNA-Seq since considerably more information can be extracted from the former approach by identifying unique differentially accessible regions, including motifs and GO Enrich, such as in chromosomes of Difg.SPG and Pachytene/Diplotene spermatocytes. In this context, it is of interest to note that scRNA-Seq and scATAC-Seq were powerful and reliable bioinformatics tools used in different studies and have their advantages. Interestingly, the percentage of cell types obtained by scATAC-Seq and scRNA-Seq yielded a slightly different outcomes as noted in Figure 2E. In scRNA-Seq, some cells with a small number of transcripts were filtered out. For instance, using scRNA-Seq, which targeted the accessible region of chromatin to capture the DNA, fewer Sperm cells were detected when compared to scATAC-Seq. More important, the relative percentage of different somatic and germ cell types in human testes obtained by scATAC-Seq matched considerably better to the earlier published morphometric data from human testes (12, 13, 50, 51) when compared to scRNA-Seq as noted in Figure 2E. Thus, the influence generated by the small amounts of expressed

genes in Sperm1 based on scATAC-Seq was almost negligible when compared to scRNA-Seq. Thus, comparing to the traditional bulk ATAC-Seq and scRNA-Seq technique, scATAC-seq is a more powerful technique. Furthermore, based on scATAC-Seq analysis, we detected reorganization in accessible chromatin during human spermatogenesis, with considerable accessibility enrichment in chr19 and chr17. Studies have shown that chr19 has the highest gene density among all chromosomes (19). On the other hand, there are numerous DSBs hotpots in chr 17 (52). Interestingly, scATAC-Seq analysis has shown that the accessibility of chr19 and 17 are different at different stages of spermatogenesis. For instance, chr19 has the highest chromatin accessibility in Difg.SPG, but lowest in chr17, illustrating that genes on chr19 and chr17 function differently in supporting specific stages of spermatogenesis. Spermatogonia probably require core pluripotency transcription factors, such as NANOG and more open chromatin structure for successful transition to meiosis as earlier reported (53). As such, a higher accessibility of chr19 in Difg.SPG supports this earlier concept, yet chr17 does not present more chromatin accessibility in this germ cell stage. An earlier study that investigates chromatin organization during mitotic-to-meiosis transition has shown that chromatin structure is closed in mitotic state, but it is open to prepare for meiosis (54). The lower accessibility in Difg. SPG and higher accessibility in Diff.ed SPG of chr17 may suggest it plays a more prominent role in supporting mitosis to meiosis transition. On the other hand, the higher accessibility of chr 19 in Difg.SPG shows that it is necessary to support differentiating spermatogonia and the subsequent transformation during human spermatogenesis. Taken collectively, our findings are consistent with earlier reports, illustrating that chr 19 and chr 17 function differently to support human spermatogenesis. In brief, this change of chromatin accessibility suggests that these two chromosomes play different roles in different processes of spermatogenesis. During spermatogenesis, gene

expression was dynamic, and some crucial genes require unique expression at different stages of the epithelial cycle. However, the functional and structural relationship between open regions and gene expression remain unknown. Our results demonstrate that high open regions contribute to more gene expression, such as meiosis. Begin with round spermatids, the genetic material undergoes extreme chromatin condensation and a series of dramatic morphological changes to become mature sperm, resulting a considerable decline of overall gene expression (1).

According to the analysis reported here, TLE3 and PFN4, whose function in supporting human spermatogenesis remains unexplored, is considerably enriched in Difg. SPG and pachytene/diplotene, respectively. As described above, TLE3 is expressed in spermatogonia, spermatids, and Sertoli cell in the mouse testis. Earlier study has reported that TLE3 regulates the balance between osteoblast and adipocyte formation from bone marrow mesenchymal stem cell (55). More important, besides its high level of expression in Difg. SPG, there is a special region with high chromatin accessibility upstream of TLE3, which enriched the CTCF motif and it is essential for genome organization and cell differentiation (56). We speculate that TLE3 plays a key role to regulate spermatogonial stem cell differentiation to maintain the critical balance between SSCs and Diff.edSPG. Nonetheless, further studies are needed to confirm this possibility. In the mouse testis, TLE3 was highly expressed in spermatogonia, spermatids, and Sertoli cells, but not spermatocytes. In Sertoli cells, TLE3 was also found to play an important role in Sertoli cell differentiation and cell proliferation by regulating specific genes during testicular development (57). As a member of actin-binding proteins, our results have shown that PFN4 is enriched in the pachytene/diplotene and RoundS.tids. It is generally accepted that actin filaments are involved in spindle migration, protecting meiotic cells from

chromosome segregation errors that lead to aneuploidy. Also, in human oocytes, spindle assembly is regelated by actin-microtubule interplay coordination (58). During spermatogenesis, regulation of actin dynamics is a complex and sophisticated progress (59). Herein, the open regions upstream of the *PFN4* have identified four important motifs which are also enriched, including CTCT, BORIS(zf), GLI3, and ISL1. Collectively, these findings thus support the notion that PFN4 is a crucial factor in regulating meiosis, possibly through its effects on actin cytoskeletal dynamics during human spermatogenesis.

In summary, the use of scATAC-Seq has illustrated dynamic chromatin changes in human spermatogenesis, which include changes in chromatin accessibility and expression. Also, TLE3 is a marker gene of Difg. SPG and it is likely to be involved in stem cell differentiation. PFN4 plays a role in actin cytoskeletal organization in meiosis during human spermatogenesis.

## Materials and Methods

**scRNA-Seq analysis.** Dataset (GSE120508) (~6500 testicular cells from *n*=3 young human adults) quality control was performed using Seurat (V3.0.1). Cells with greater than 50% mitochondrial reads were excluded from the analysis. Number of detected genes per cell, and number of UMIs expressed per cell using following threshold parameters: nGene (200-5000), nUMI (0-30000). Resolution used to cluster is 1.5. Cell labeling was performed after clustering based on marker genes.

**Human testis samples for scATAC-seq library preparation.**

Adult human testicular biopsy samples for scATAC-Seq were obtained from three healthy men (donor #1: 31 years old; donor #2: 33 years old; donor #3: 36 years old) without notable underlying pathological conditions including signs of infertility and normal andrology parameters. These samples were obtained in the Male Clinic at the Nantong University Medical School (Nantong, China) with patients' consent for research (IRB Approved Protocol: TDLS-2020-36) entitled: Understanding the mechanism of modification of chromatin and RNA in male infertility. Testicular tissue was initially digested to suspension as earlier described (60). Cell nuclei were isolated and prepared for sequencing as outlined in the 10X Genomics Chromium single cell (sc) ATAC-Seq solution protocol (10). Nuclei were loaded with a capture target of 10,000 nuclei per sample. scATAC-Seq libraries were prepared for sequencing following the 10X Genomics single cell ATAC-Seq solution using protocol supplied by the manufacturer. scATAC-Seq libraries were sequenced using PE50 sequencing on an Illumina NovaSeq.

**Single cell ATAC-Seq analysis.** Data analysis was performed by Cell Ranger ATAC (1.2.0). Genome used in our analysis was GRCh38. For the downstream analysis, SnapATAC was used to convert data format based on fragments data. Thereafter, we picked out cells used in cell clustering step, utilizing diffusion map for data dimensionality reduction. The first 30 dimensionalities were selected to calculate Euclidean Distance to create KNN cluster (default k=15) of each cell. Cell clustering was carried out based on the graph that drawn by data calculated above. Uniform Manifold Approximation and Projection (umap) was subsequently applied to the results and to perform hierarchical clustering analysis between clusters. Cell annotation was added to the result of scATAC-Seq analysis, according to the annotation result of scRNA after identifying cell types.

Following this, FindTransferAnchors and TransferData (Seurat) were used to merge the two types of omics data, using the differential genes identified by scRNA-Seq analysis as medium. Then cells with prediction scores greater than 0.3 were selected for the following analyses. Also, the relative percentage of different germ cell types versus different somatic cell types obtained by scATAC-Seq and scRNA-Seq were compared.

**Peak calling and Motif analysis.** Peak calling and peak merge of predicted cell types were performed by MACS2 (cell types with cells more than 50). The default threshold of q value used by MACS2 was at 0.05. This thus generated a cell X peak matrix. Peaks located in gene functional region was counted by bedtools. The reads distribution of every cell type was displayed by IGV. We identified differentially accessible peaks by SnapATAC. SnapATAC selected closer cell clusters automatically to compare with each other. Through the exact test, differentially accessible peaks were identified. Peaks whose logFC > 0 were selected and sorted by p value from high to low. The top 200 of the selected peaks were then used in the following analysis. Motif analysis of differential accessible peaks selected before was carried out by HOMER. Motifs with p value < 0.05 of each cell type were used to draw the statistical graph.

## Coaccessibility analysis

To get the information of the relationship among peaks, we performed coaccessibility analysis to get the scores of peaks by Cicero. Select peaks with score > 0.5 to do chromosome enrichment statistic.

$$coaccessibility = \frac{\frac{chr\ CoAccessibility\ number}{total\ CoAccessibility}}{\frac{chrlength}{all\ length\ sum}}$$

**Differential accessible regions and differential expressed genes analysis.** Genes with differential region within 10kb upstream or 1kb downstream were regarded as those annotated by differential accessible regions. Next, the overlapped genes between genes annotated by differential accessible regions and differential expressed genes in scRNA-Seq were selected, and the correlation between the chromatin accessibility and expression of genes were analyzed based on the scATAC logFC and scRNA avg_logFC. The genes in quadrant I and quadrant III had the same change trend on scATAC logFC and scRNA avg_logFC. Expressions of these genes were regulated by the accessible regions. Expression of genes in I quadrant were upregulated with higher chromatin accessibility.

**GO pathway analysis.** After differential peaks were obtained, GO pathway analysis was performed by R package, rGREAT, to confirm peaks enriching pathway.

**Acknowledgements:** This work was supported by The National Key Research and Development Program of China (2018YFC1003500 to F.S) and National Natural Science Foundation of China (Grant No. 81901528 to X.W., 81971376 and 81670622 to M.L.). All authors declare no competing financial or other interests.

**Conflict of Interest Statement:** The Authors have nothing to declare.

## References

1 Hao, S.L., Ni, F.D. and Yang, W.X. (2019) The dynamics and regulation of chromatin remodeling during spermiogenesis. *Gene*, **706**, 201-210.

2 Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R. and Surani, M.A. (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature*, **452**, 877-881.

3 Hammoud, Saher S., Low, Diana H.P., Yi, C., Carrell, Douglas T., Guccione, E. and Cairns, Bradley R. (2014) Chromatin and Transcription Transitions of Mammalian Adult Germline Stem Cells and Spermatogenesis. *Cell Stem Cell*, **15**, 239-253.

4 Stewart, K.R., Veselovska, L. and Kelsey, G. (2016) Establishment and functions of DNA methylation in the germline. *Epigenomics*, **8**, 1399-1413.

5 Wu, X., Luo, C., Hu, L., Chen, X., Chen, Y., Fan, J., Cheng, C.Y. and Sun, F. (2020) Unraveling epigenomic abnormality in azoospermic human males by WGBS, RNA-Seq, and transcriptome profiling analyses. *J Assist Reprod Genet*, in press.

6 Guo, J., Grow, E.J., Yi, C., Mlcochova, H., Maher, G.J., Lindskog, C., Murphy, P.J., Wike, C.L., Carrell, D.T., Goriely, A. *et al.* (2017) Chromatin and Single-Cell RNA-Seq Profiling Reveal Dynamic Signaling and Metabolic Transitions during Human Spermatogonial Stem Cell Development. *Cell Stem Cell*, **21**, 533-546 e536.

7 Green, C.D., Ma, Q., Manske, G.L., Shami, A.N., Zheng, X., Marini, S., Moritz, L., Sultan, C., Gurczynski, S.J., Moore, B.B. *et al.* (2018) A Comprehensive Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. *Dev Cell*, **46**, 651-667 e610.

8 Chen, Y., Zheng, Y., Gao, Y., Lin, Z., Yang, S., Wang, T., Wang, Q., Xie, N., Hua, R., Liu, M. *et al.* (2018) Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. *Cell research*, **28**, 879-896.

9 Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y.

and Greenleaf, W.J. (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, **523**, 486-490.

10 Cusanovich, D.A., Hill, A.J., Aghamirzaie, D., Daza, R.M., Pliner, H.A., Berletch, J.B., Filippova, G.N., Huang, X., Christiansen, L., DeWitt, W.S. *et al.* (2018) A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. *Cell*, **174**, 1309-1324 e1318.

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

12 Amann, R.P. (2008) The cycle of the seminiferous epithelium in humans: a need to revisit? *J Androl*, **29**, 469-487.

13 Johnson, L., Staub, C., Neaves, W.B. and Yanagimachi, R. (2001) Live human germ cells in the context of their spermatogenic stages. *Human reproduction (Oxford, England)*, **16**, 1575-1582.

14 Nie, Y., Wilson, A.F., DeFalco, T., Meetei, A.R., Namekawa, S.H. and Pang, Q. (2020) FANCD2 is required for the repression of germline transposable elements. *Reproduction (Cambridge, England)*, **159**, 659-668.

15 Rojas-Ríos, P. and Simonelig, M. (2018) piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. *Development (Cambridge, England)*, **145**.

16 Friedli, M. and Trono, D. (2015) The developmental control of transposable elements and the evolution of higher species. *Annu Rev Cell Dev Biol*, **31**, 429-451.

17 Mäkelä, J.-A. and Hobbs, R.M. (2019) Molecular regulation of spermatogonial stem cell renewal and differentiation. *Reproduction (Cambridge, England)*, **158**, R169-R187.

18 Pandey, A., Yadav, S.K., Vishvkarma, R., Singh, B., Maikhuri, J.P., Rajender, S. and Gupta, G. (2019) The dynamics of gene expression during and post meiosis sets the sperm agenda. *Molecular*

*reproduction and development*, **86**, 1921-1939.

19 Grimwood, J., Gordon, L.A., Olsen, A., Terry, A., Schmutz, J., Lamerdin, J., Hellsten, U., Goodstein, D., Couronne, O., Tran-Gyamfi, M. *et al.* (2004) The DNA sequence and biology of human chromosome 19. *Nature*, **428**, 529-535.

20 Berg, I.L., Neumann, R., Lam, K.W., Sarbajna, S., Odenthal-Hesse, L., May, C.A. and Jeffreys, A.J. (2010) PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat Genet*, **42**, 859-863.

21 Oliver, P.L., Goodstadt, L., Bayes, J.J., Birtle, Z., Roach, K.C., Phadnis, N., Beatson, S.A., Lunter, G., Malik, H.S. and Ponting, C.P. (2009) Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. *PLoS genetics*, **5**, e1000753.

22 Bai, S., Fu, K., Yin, H., Cui, Y., Yue, Q., Li, W., Cheng, L., Tan, H., Liu, X., Guo, Y. *et al.* (2018) Sox30 initiates transcription of haploid genes during late meiosis and spermiogenesis in mouse testes. *Development*, **145**.

23 Ronfani, L. and Bianchi, M.E. (2004) Molecular mechanisms in male determination and germ cell differentiation. *Cell Mol Life Sci*, **61**, 1907-1925.

24 Ong, C.T. and Corces, V.G. (2014) CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet*, **15**, 234-246.

25 Hernandez-Hernandez, A., Lilienthal, I., Fukuda, N., Galjart, N. and Hoog, C. (2016) CTCF contributes in a critical way to spermatogenesis and male fertility. *Scientific reports*, **6**, 28355.

26 Klenova, E.M., Morse, H.C., Ohlsson, R. and Lobanenkov, V.V. (2002) The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Seminars in Cancer Biology*, **12**, 399-414.

27 Martin-Kleiner, I. (2012) BORIS in human cancers -- a review. *European journal of cancer (Oxford, England : 1990)*, **48**, 929-935.

28 Bucher, P. (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J Mol Biol*, **212**, 563-578.

29 Fukuda, N., Fukuda, T., Sinnamon, J., Hernandez-Hernandez, A., Izadi, M., Raju, C.S., Czaplinski, K. and Percipalle, P. (2013) The transacting factor CBF-A/Hnrnpab binds to the A2RE/RTS element of protamine 2 mRNA and contributes to its translational regulation during mouse spermatogenesis. *PLoS genetics*, **9**, e1003858.

30 Zhang, T. and Zarkower, D. (2017) DMRT proteins and coordination of mammalian spermatogenesis. *Stem Cell Res*, **24**, 195-202.

31 Zhang, T., Murphy, M.W., Gearhart, M.D., Bardwell, V.J. and Zarkower, D. (2014) The mammalian Doublesex homolog DMRT6 coordinates the transition between mitotic and meiotic developmental programs during spermatogenesis. *Development (Cambridge, England)*, **141**, 3662-3671.

32 Peluffo, G., Subedee, A., Harper, N.W., Kingston, N., Jovanović, B., Flores, F., Stevens, L.E., Beca, F., Trinh, A., Chilamakuri, C.S.R. *et al.* (2019) Is a Transcriptional Dependency in Triple-Negative Breast Cancer Associated with Brain Metastasis. *Cancer research*, **79**, 4173-4183.

33 Quaranta, R., Fell, J., Rühle, F., Rao, J., Piccini, I., Araúzo-Bravo, M.J., Verkerk, A.O., Stoll, M. and Greber, B. (2018) Revised roles of ISL1 in a hES cell-based model of human heart chamber specification. *eLife*, **7**.

34 Endo, T., Mikedis, M.M., Nicholls, P.K., Page, D.C. and de Rooij, D.G. (2019) Retinoic Acid and Germ Cell Development in the Ovary and Testis. *Biomolecules*, **9**.

35 Matissek, S.J. and Elsawa, S.F. (2020) GLI3: a mediator of genetic diseases, development and cancer. *Cell communication and signaling : CCS*, **18**, 54.

36 Veitia, R.A. (2010) FOXL2 versus SOX9: a lifelong "battle of the sexes". *Bioessays*, **32**, 375-380.

37 Justice, N.J., Blount, A.L., Pelosi, E., Schlessinger, D., Vale, W. and Bilezikjian, L.M. (2011) Impaired FSHbeta expression in the pituitaries of Foxl2 mutant animals. *Mol Endocrinol*, **25**, 1404-1415.

38 Li, Y., Schang, G., Wang, Y., Zhou, X., Levasseur, A., Boyer, A., Deng, C.-X., Treier, M., Boehm, U., Boerboom, D. *et al.* (2018) Conditional Deletion of FOXL2 and SMAD4 in Gonadotropes of Adult Mice Causes Isolated FSH Deficiency. *Endocrinology*, **159**, 2641-2655.

39 Fang, F., Xia, N., Angulo, B., Carey, J., Cady, Z., Durruthy-Durruthy, J., Bennett, T., Sebastiano, V. and Reijo Pera, R.A. (2018) A distinct isoform of ZNF207 controls self-renewal and pluripotency of human embryonic stem cells. *Nature communications*, **9**, 4384.

40 Gasperowicz, M. and Otto, F. (2005) Mammalian Groucho homologs: redundancy or specificity? *Journal of cellular biochemistry*, **95**, 670-687.

41 Jangal, M., Couture, J.-P., Bianco, S., Magnani, L., Mohammed, H. and Gévry, N. (2014) The transcriptional co-repressor TLE3 suppresses basal signaling on a subset of estrogen receptor α target genes. *Nucleic acids research*, **42**, 11339-11348.

42 Hernández-Hernández, A., Lilienthal, I., Fukuda, N., Galjart, N. and Höög, C. (2016) CTCF contributes in a critical way to spermatogenesis and male fertility. *Scientific reports*, **6**, 28355.

43 Tariq, N., Basharat, Z., Butt, S. and Baig, D.N. (2016) Distribution analysis of profilin isoforms at transcript resolution with mRNA-seq and secondary structure in various organs of Rattus norvegicus. *Gene*, **589**, 49-55.

44 Witke, W. (2004) The role of profilin complexes in cell motility and other cellular processes. *Trends Cell Biol*, **14**, 461-469.

45 Behnen, M., Murk, K., Kursula, P., Cappallo-Obermann, H., Rothkegel, M., Kierszenbaum, A.L. and Kirchhoff, C. (2009) Testis-expressed profilins 3 and 4 show distinct functional characteristics and localize in the acroplaxome-manchette complex in spermatids. *BMC Cell Biol*, **10**, 34.

46 Giesemann, T., Rathke-Hartlieb, S., Rothkegel, M., Bartsch, J.W., Buchmeier, S., Jockusch, B.M. and Jockusch, H. (1999) A role for polyproline motifs in the spinal muscular atrophy protein SMN. Profilins bind to and colocalize with smn in nuclear gems. *The Journal of biological chemistry*, **274**, 37908-37914.

47 Witke, W., Sutherland, J.D., Sharpe, A., Arai, M. and Kwiatkowski, D.J. (2001) Profilin I is essential for cell survival and cell division in early mouse development. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 3832-3836.

48 Hoffmann, M.J., Muller, M., Engers, R. and Schulz, W.A. (2006) Epigenetic control of CTCFL/BORIS and OCT4 expression in urogenital malignancies. *Biochem Pharmacol*, **72**, 1577-1588.

49 Persengiev, S.P., Kondova, I.I., Millette, C.F. and Kilpatrick, D.L. (1997) Gli family members are differentially expressed during the mitotic phase of spermatogenesis. *Oncogene*, **14**, 2259-2264.

50 Johnson, L., Zane, R.S., Petty, C.S. and Neaves, W.B. (1984) Quantification of the human Sertoli cell population: its distribution, relation to germ cell numbers, and age-related decline. *Biol Reprod*, **31**, 785-795.

51 Johnson, L., Petty, C.S. and Neaves, W.B. (1980) A comparative study of daily sperm production and testicular composition in humans and rats. *Biol Reprod*, **22**, 1233-1243.

52 Pratto, F., Brick, K., Khil, P., Smagulova, F., Petukhova, G.V. and Camerini-Otero, R.D. (2014)

DNA recombination. Recombination initiation maps of individual human genomes. *Science*, **346**, 1256442.

53 Bowles, J. and Koopman, P. (2010) Sex determination in mammalian germ cells: extrinsic versus intrinsic factors. *Reproduction*, **139**, 943-958.

54 Maezawa, S., Yukawa, M., Alavattam, K.G., Barski, A. and Namekawa, S.H. (2018) Dynamic reorganization of open chromatin underlies diverse transcriptomes during spermatogenesis. *Nucleic acids research*, **46**, 593-608.

55 Kokabu, S., Lowery, J.W. and Jimi, E. (2016) Cell Fate and Differentiation of Bone Marrow Mesenchymal Stem Cells. *Stem cells international*, **2016**, 3753581.

56 Zheng, H. and Xie, W. (2019) The role of 3D genome organization in development and cell differentiation. *Nature reviews. Molecular cell biology*, **20**, 535-550.

57 Lee, S., Jang, H., Moon, S., Lee, O.H., Lee, S., Lee, J., Park, C., Seol, D.W., Song, H., Hong, K. *et al.* (2019) Differential Regulation of TLE3 in Sertoli Cells of the Testes during Postnatal Development. *Cells*, **8**.

58 Roeles, J. and Tsiavaliaris, G. (2019) Actin-microtubule interplay coordinates spindle assembly in human oocytes. *Nat Commun*, **10**, 4651.

59 Su, W., Mruk, D.D. and Cheng, C.Y. (2013) Regulation of actin dynamics and protein trafficking during spermatogenesis--insights into a complex process. *Crit Rev Biochem Mol Biol*, **48**, 153-172.

60 Jingtao Guo, E.J.G., Chongil Yi,Hana Mlcochova, Geoffrey J. Maher,Cecilia Lindskog,Patrick J. Murphy,Candice L. Wike,Douglas T. Carrell, Anne Goriely,James M. Hotaling,and Bradley R. Cairns. (2017) Chromatin and Single-Cell RNA-Seq Profiling Reveal Dynamic Signaling and Metabolic Transitions during Human Spermatogonial Stem Cell Development. *cell stem cell*, in press.

**Figure Legends**

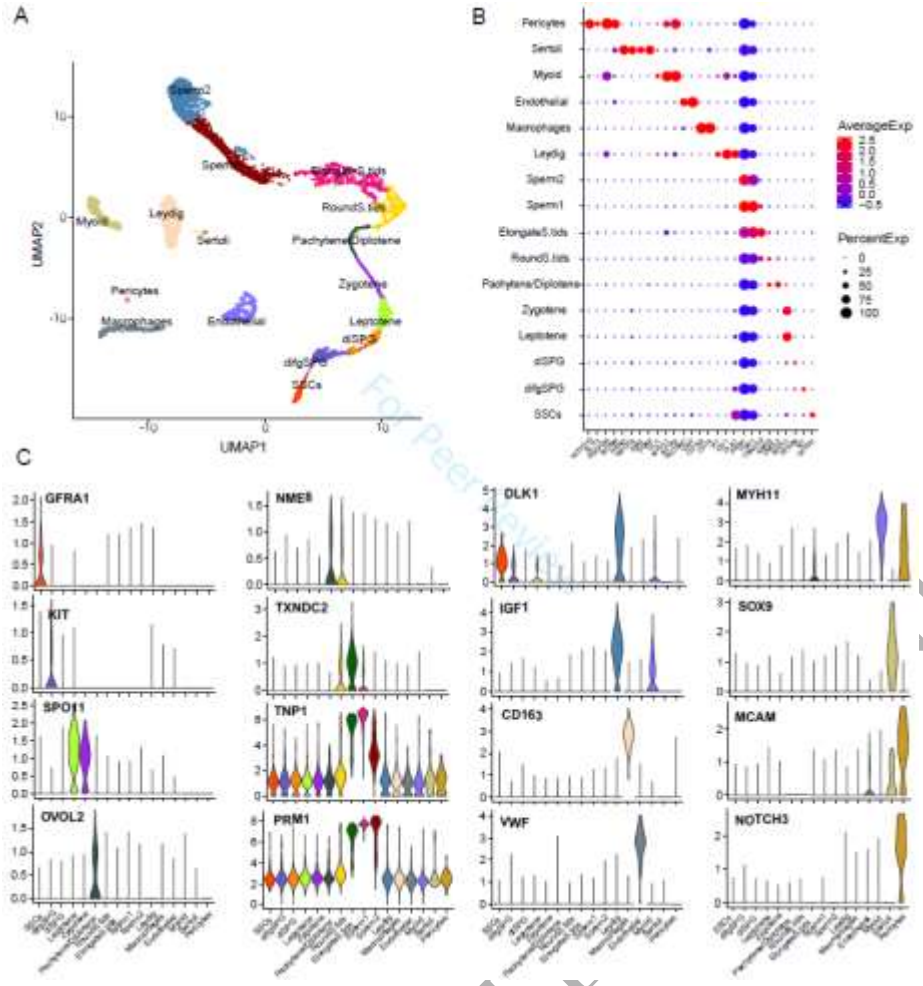


**Figure 1. Results of scRNA-Seq data analysis.** (A) Umap plot of adult human testicular cells. Colors based on cell types we identified are shown. (B) Expression of marker genes in different cell types (dot plot). The dot size and dot color represent expression percentage and expression average, correspondingly. (C) Expression patterns (violin plots) of partial marker genes.

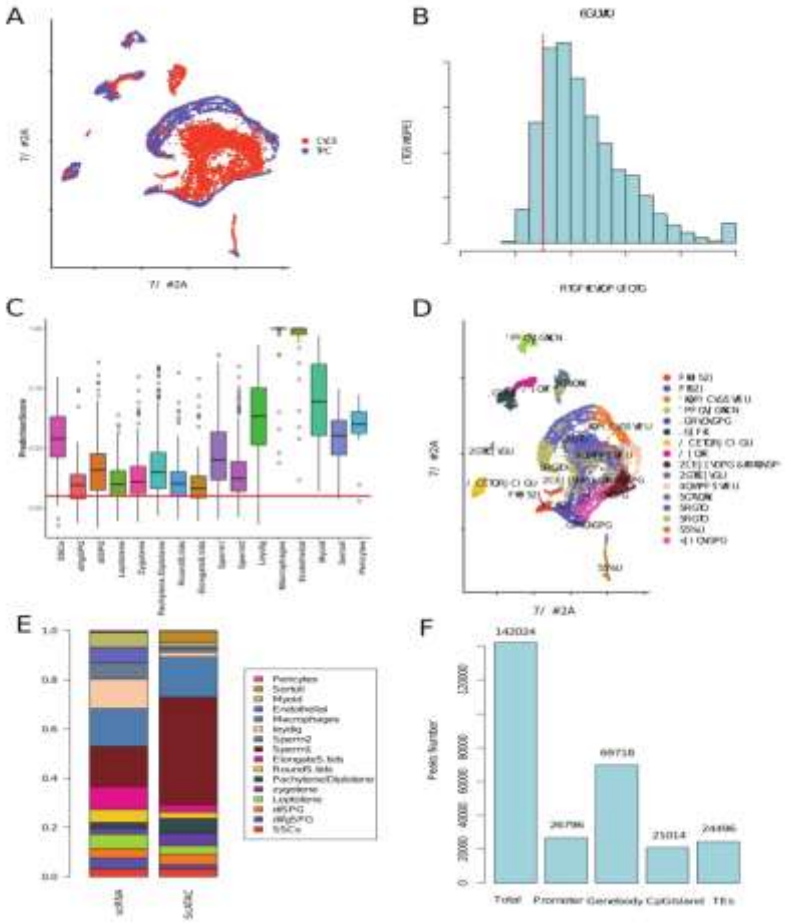


**Figure 2. Cell selection and preliminary analysis combining data of scATAC-Seq and scRNA-Seq.** (A) Umap plot of cells. Different colors represent cells from scATAC-Seq and scRNA-Seq. (B) Cell number of different scATAC-Seq cell prediction scores displayed by bar plot. We set 0.3 as threshold (red dashed line) to select cells for the following analysis. There are 4685 cells in whole 5376 cells satisfied our condition. (C) scATAC-Seq cell prediction score of different stage cell types from (B). (D) Umap plot of cells from (A). Cell types are shown by different colors. (E) Percentage of different cell types in cells selected frrm (B) classified by scRNA-Seq and scATAC-Seq. (F) Number of peaks in different locations displayed by bar plot.

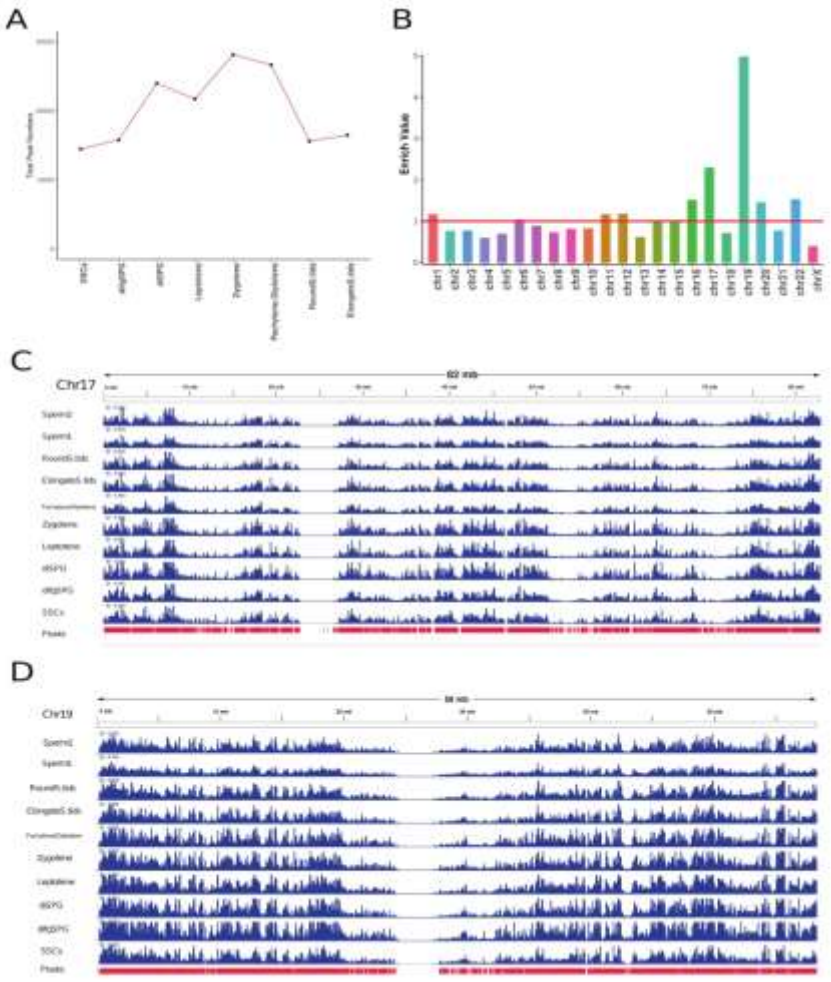


**Figure 3. Open regions in spermatogenesis and the result of accessibility analysis in different chromosomes.** (A) Peak number in different stage of spermatogenesis. (B) The dashed horizontal line represents 1.0 of the coaccessiblity enrich in the plot. (C) and (D) display accessibility peaks enrichment distribution of chr19 and chr17 for different germ cell types during spermatogenesis.

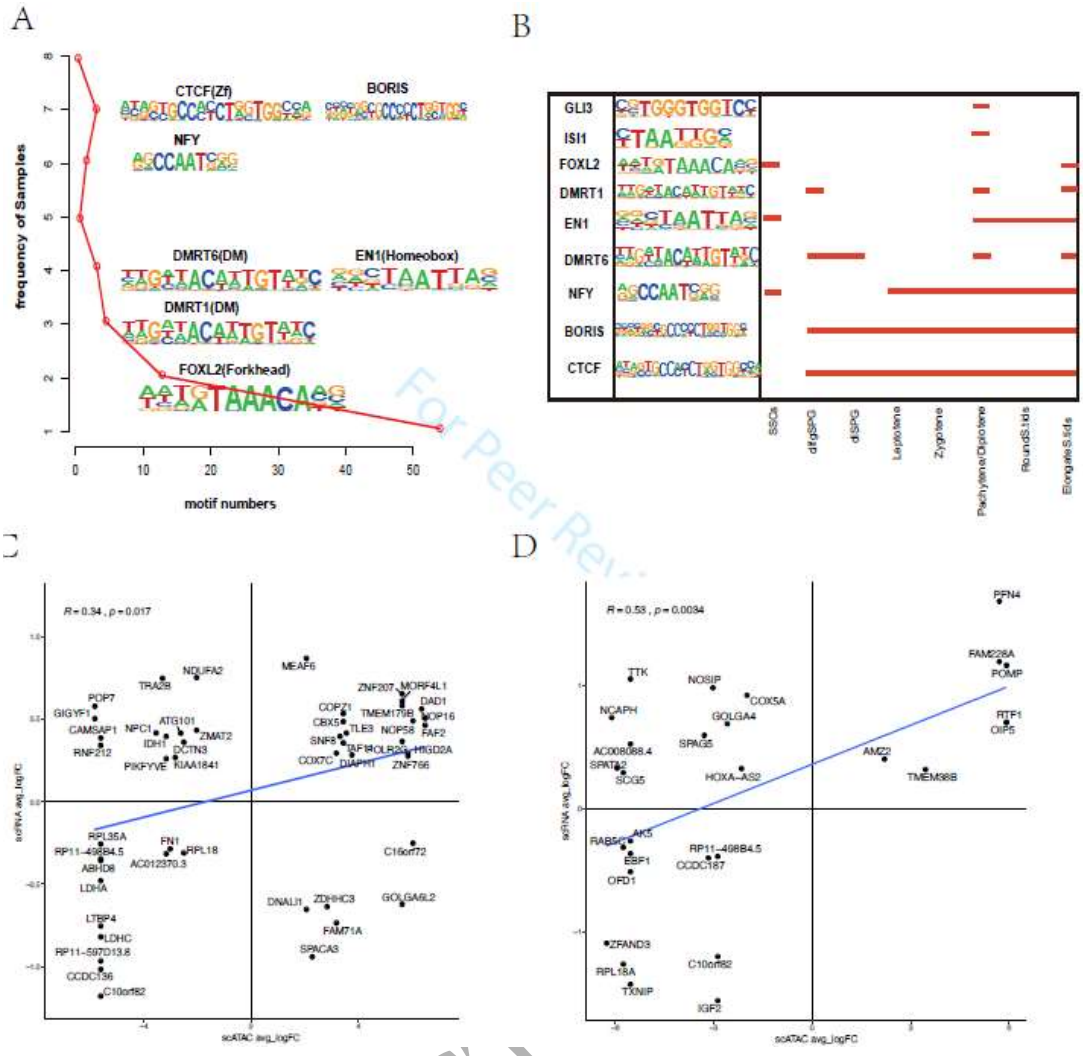


**Figure 4. Cell selection and motif analysis.** (A) Different motifs were enriched in 8 different germ cell types. The motif and sequence logo are corresponding to the point on their left. (B) Detailed cell types of motifs enriching. Motifs displayed on the left enrich in the germ cell types where the red lines indicate the expression of the corresponding genes. (C) and (D) show the correlation between chromatin accessibility and expression of genes which were analyzed based on the scATAC logFC and scRNA avg_logFC. X-axis, (scATAC logFC): >0 means open regions and <0 means closed regions. Y-axis, (scRNA avg logFC): >0 means expression and <0 means inhibit expression. Abbreviations used: Difg. SPG and Pachytene/Diplotene)

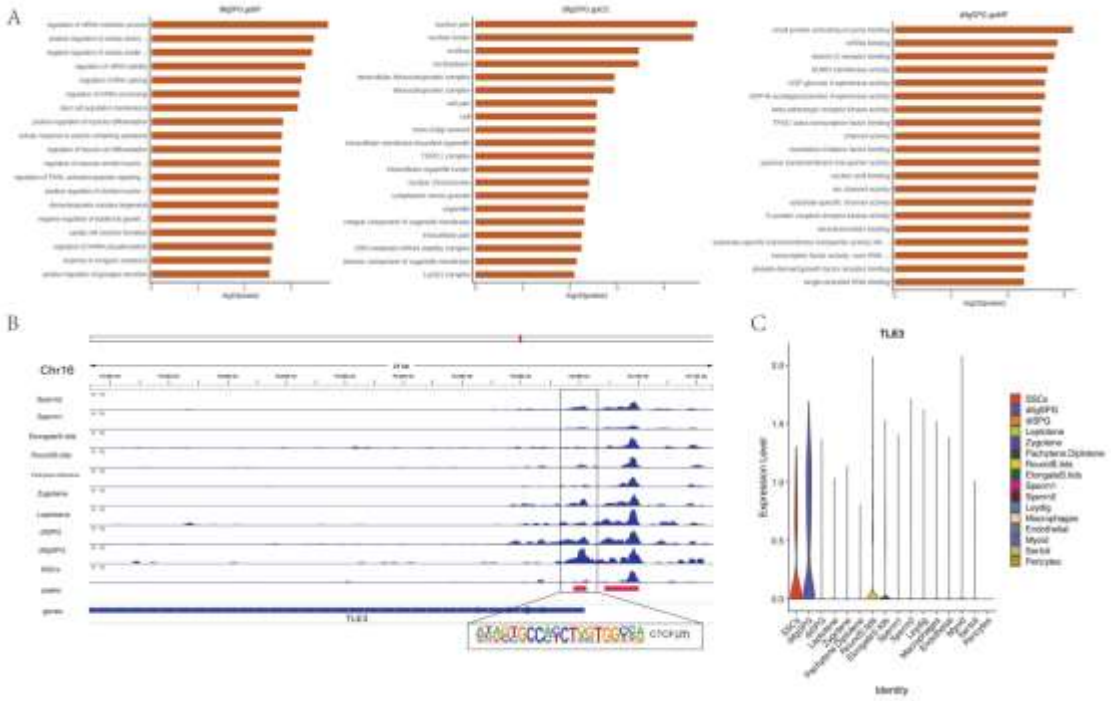


**Figure 5. GO analysis of Difg. SPG.** (A) Enriched GO terms including biological process, cellular component and molecular function in Difg. SPG. (B) Regions of enriching peaks in Difg. SPG on upstream of TLE3 (Chr16). Motifs located in this region are annotated. (C) Expression pattern of TLE3 in 8 germ cell types during spermatogenesis (violin plot).

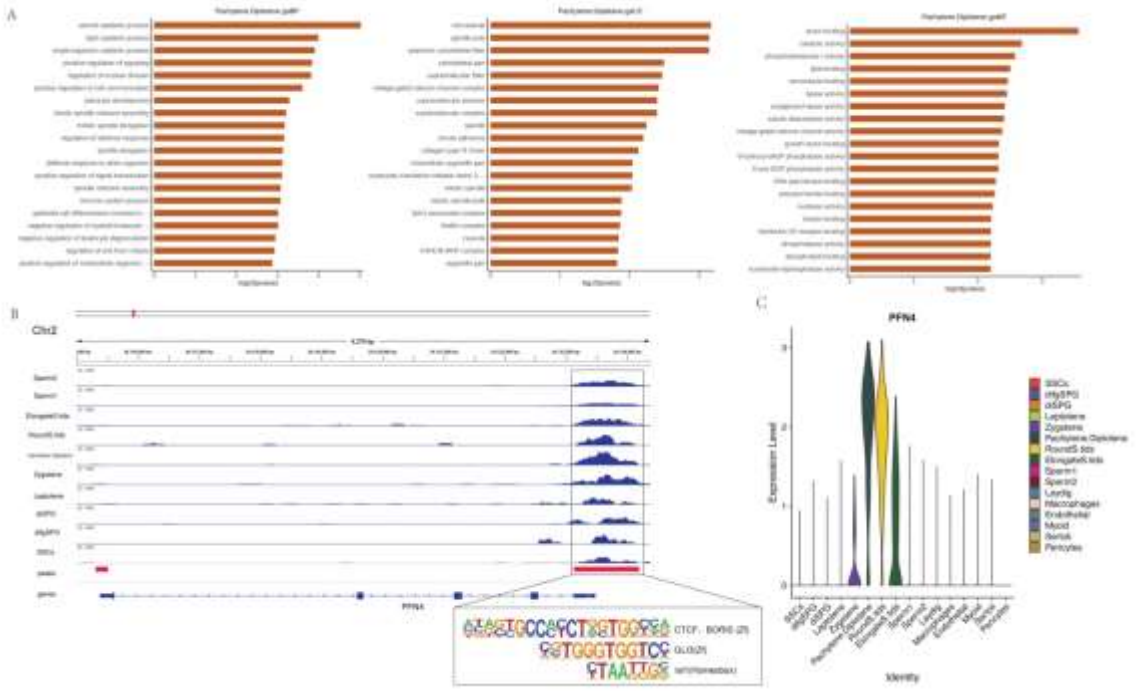


**Figure 6. GO analysis of Pachytene/Diplotene spermatocytes.** (A) Enriched GO terms including biological process, cellular component and molecular function in Pachytene/Diplotene spermatocytes. (B) Regions of enriching peaks in Pachytene/Diplotene on upstream of PFN4 (Chr2). Motifs located in this region are annotated. (C) Expression pattern of PFN4 in 8 germ cell types during spermatogenesis (violin plot).