Integration and Gene Co-Expression Network Analysis of scRNA-seq Transcriptomes Reveal Heterogeneity and Key Functional Genes in Human Spermatogenesis

Najmeh Salehi
Royan Institute for Reproductive Biomedicine
Mohammad Hossein Karimi-Jafari
University of Tehran
Mehdi Totonchi (m.totonchi@royaninstitute.org)
Royan Institute for Reproductive Biomedicine
Amir Amiri-Yekta
Royan Institute for Reproductive Biomedicine

Research Article

Keywords: single-cell RNA sequencing, spermatogenesis, weighted gene co-expression network analysis, heterogeneity, male infertility.

Posted Date: June 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-611284/v1

Abstract

Spermatogenesis is a complex process of cellular division and differentiation that begins with spermatogonia stem cells and leads to functional spermatozoa production. However, many of the molecular mechanisms underlying this process remain unclear. Single-cell RNA sequencing (scRNA-seq) is used to sequence the entire transcriptome at the single-cell level to assess cell-to-cell variability. Here, more than 33,000 testicular cells from five scRNA-seq datasets with normal spermatogenesis were integrated to identify single-cell heterogeneity on a more comprehensive scale. Clustering, cell type assignments, differential expressed genes and pseudotime analysis characterized 5 spermatogonia, 4 spermatocyte, and 4 spermatid cell types during the spermatogenesis process. The UTF1 and ID4 genes were introduced as most specific markers that can differentiate two undifferentiated spermatogonia stem cell sub-cellules, and C7orf61 and TNP, two round spermatid sub-cellules. The topological analysis of the weighted gene co-expression network along with the integrated scRNA-seq data revealed some bridge genes between spermatogenesis's main stages such as DNAJC5B, C1orf194, HSP90AB1, BST2, EEF1A1, CRISP2, PTMS, NFKBIA, CDKN3, and HLA-DRA. The importance of these key genes is confirmed by their role in male infertility in the available studies. It can be stated that, this integrated scRNA-seq of spermatogenic cells offers novel insights into cell-to-cell heterogeneity and suggests a list of key players with a pivotal role in male infertility from the fertile spermatogenesis datasets. These key functional genes can be introduced as candidates for filtering and prioritizing of genotype-to-phenotype association in male infertility.

Introduction

Spermatogenesis is a highly organized and complex process of differentiation events that produce sperm from the primordial germ cells[1]. Sperm production, occurring in the seminiferous tubules, is a continuous process that begins at puberty and continues throughout life[2]. This productivity depends on the activity of the spermatogonia stem cells (SSC), which are the stem cells of adult testicular tissue[3]. The SSCs are capable of perpetual self-renewal and differentiation divisions, which preserve the stem cell pool and spermatogenesis fuel, respectively[3,4]. Then, differentiating spermatogonia cells divide mitotically and produce two diploid spermatocytes, followed by two meiosis and the spermiogenesis process to produce haploid spermatids and sperm, respectively[1,4]. Between 1,500 and 2,000 genes are thought to play a role in controlling spermatogenesis, and genetic changes in these genes are expected to impair male fertility[5,6]. Currently, the genetic diagnostic for male infertility includes screening for a shortlist of candidate genes that should be expanded[7–9]. Therefore, a high-resolution profile of gene expression signatures in the process of spermatogenesis can be a starting point for solving male infertility[10].

Gene expression profiling assays, such as typical microarray or RNA-sequencing (RNA-seq) have been widely used to investigate the changes in testicular gene expression from birth to adulthood[11–14], and molecular mechanisms involved in male infertility[15,16]. These studies relied on the bulk RNA analysis of mixed aggregates of spermatogenic cells, that provide the average expression signal for a pool of

different cell types[17,18]. Therefore, they lose within and between cell types diversity or rare cell phenotypes[17]. To isolate spermatogenic cell types, some common approaches such as fluorescence-activated cell sorting (FACS), magnetic activated cell sorting (MACS), and STA-PUT are used[17,19]. However, these methods can only separate some types of spermatogenic cells and cannot isolate high-purity homogeneous spermatogenic cells from all types[10,20].

Single-cell RNA sequencing (scRNA-seq) provides the transcriptome profiles of individual cells that can investigate the variation within and between cell types and reveal rare cell types[17]. During the past few years, some studies have examined the transcriptome profiles of different cell types in human testicular tissue using scRNA-seq. Most of these studies have investigated spermatogenesis single-cell transcriptome in only fertile individuals or obstructive azoospermia (OA) patients[21–28]. And a limited number of studies in non-obstructive azoospermia (NOA) patients have been reported[29,30]. FACS, MACS, and STA-PUT were used to sort individual cell types before scRNA-seq in some studies[21–23,29]. However, scRNA-seq can examine thousands of individual cells in the steady-state of spermatogenesis without the need for prior sorting[22,23,25–27,29]. Also, single-cell transcriptomes of infants, juvenile and adult males were profiled to investigate the changes in the spermatogenesis cell types at birth, during puberty, and adulthood[23,25,27]. The common idea in all of these studies was the identification of cell types based on the expressions of key markers, finding of differentially expressed genes (DEGs) in each cell type with respect to the others, and their functional and biological enrichment which showed significant heterogeneity within and between spermatogenesis cell types.

In this study, we integrated the scRNA-seq data of human spermatogonia, spermatocyte, spermatid sorted cells[22] and steady-state spermatogenic cells[22,29]. The integrated analysis of these datasets provides a more comprehensive profile of spermatogenesis[31]. Then clustering, cell type assignments, DEGs, enrichment, and pseudotime trajectory analysis were performed to characterize cell heterogeneity. Furthermore, a related gene co-expression network was generated, and its topological analysis revealed bridge genes in this process. The role of these bridge genes in male infertility makes them candidates for infertility diagnostic tests.

Results

Clustering of integrated spermatogenesis dataset

The diverse human spermatogenesis scRNA-seq datasets, including spermatogonia, spermatocyte, spermatid sorted cells, and steady-state spermatogenic cells were collected from the GEO database. The cell types, sorting methods, scRNA-seq methods, GEO ID, and the initial number of genes and cells in each dataset were summarized in Figure 1A. After pre-processing, 33,185 spermatogenic cells were gathered. The integrated datasets in the UMAP low dimensional space showed that similar cells in different datasets were placed together in the UMAP space (Fig. 1B). Each dataset in the UMAP space of integrated data was presented in detail in Figure S1. The Spermatogenesis1 dataset which belongs to

steady-state spermatogenic cells[22], depicted the greatest similarity with the integrated data in the UMAP space (Fig. 1B, S1A). On the other hand, some of the Spermatocyte and Spermatid dataset cells, that were isolated using the STA-PUT method, are mixed in the UMAP space of integrated data (Fig. 1B, S1D, and S1E). The unsupervised, graph-based clustering revealed 16 clusters of testicular cells in this integrated data which is showed in the UMAP plot (Fig. 1C).

Cell type assignment shows heterogeneity among testicular cells.

Due to the importance of assigning cell type identity to the clusters, some of the known testicular germ and somatic cell markers such as NANOS2, PIWIL4, MAGEA4, SYCP3, OVOL2, DDX4, etc. expression during the human spermatogenesis were shown (Fig. 2A). The NANOS2 and PIWIL4 are major genes for the SSC maintenance and are expressed in self-renewing SSC[25,29,32–34]. These marker genes were specifically expressed in clusters-1 and -2 which were named Undiff. SPG1 and Undiff. SPG2, respectively (Fig. 2B, 2C, and S2). GFRA1 and SALL4 are well-known as markers of both undifferentiated and differentiating SSCs[35] which were expressed in clusters-1, -2, -10, and -13. So, clusters-10 and -13 were assigned to differentiating cells and termed as Diff.ing SPG1 and Diff.ing SPG2, respectively (Fig. 2B, 2C, and S2). Cluster-12 was identified as a differentiated spermatogonia cells cluster (Diff.ed SPG) due to the MAGEA4 and HMGA1 expression in cluster-1, -2, -10, -12, and -13 for all spermatogonia cells (Fig. 2B, 2C and S2)[29,35,36]. DMC1 and RAD51AP2 are mitotic genes expressed at the leptotene stage[37]. Accordingly, cluster-9 with the highest expression level of these genes belonged to leptotene cells, denoted as the Leptotene SPC cluster (Fig. 2B, 2C, and S3). PIWIL1 expression is initiated from spermatocyte to spermatid cells with the highest expression levels in zygotene and pachytene[38]. Also, SYCP3 upregulated from differentiated spermatogonia cells to the early round spermatids stage[39]. OVOL2 is expressed from zygotene to diplotene, relating to the presence of the sex body during mammalian male meiosis[40]. Based on these results, cluster-6, -5, -8, and -7 were recognized as the zygotene, pachytene, diplotene stages of spermatocytes and the early round spermatids, respectively, that were denoted as Zygotene SPC, Pachytene SPC, Diplotene SPC and Early round SPT (Fig. 2B, 2C, S3 and S4). TEX29 and SUN5 genes can be observed in the round spermatids[29], were expressed in cluster-3 and -4 (denoted as Round SPT1 and Round SPT2). Furthermore, ACR and PGK2 presented in spermatocytes and spermatids, from zygotene to round spermatids and elongating spermatids, respectively[22,41,42]. SPEM1 is expressed in the late stages of spermatid[43]. Thus, cluster-11 corresponded to the last stage of spermatid, which was named as Elongating SPT (Fig. 2B, 2C and S4). To detect somatic cells clusters, the expression pattern of CYP26B1 as Sertoli[44], INSL3 as Leydig[45], MYH11 as myoid[46], and ALDH1A1 as Sertoli, Leydig and myoid markers[47,48], were evaluated. On the other hand, CD68 and CD163 are known markers of macrophages. These investigations showed cluster-13, -14, -15 as somatic cell clusters. On the other hand, the DDX4 gene expression pattern, as germ cells marker, confirmed the somatic cell clusters assignment. All of these cell clustering analyses on datasets and cell-type assignments are summarized in Table 1.

The expression patterns of DEGs were compared among all cell type clusters (Table S1). The number of up and down-regulated genes (or positive and negative DEGs) in all germ cell types (13 clusters) were

measured and compared with each other. Among all spermatogenic cell clusters, Round SPT2 and Round SPT1 displayed the most up-regulated genes with 415 and 284 genes, respectively (Fig. 2D). On the other hand, Pachytene SPC and Zygotene SPC presented the most down-regulated genes with 370 and 345 genes (Fig. 2E).

The cell assignment results demonstrated five spermatogonia cells. Among them, the Undiff. SPG1, Undiff. SPG2 and Diff.ing SPG1 positive DEGs were enriched especially for BPs related to translation (Fig. 2E). Translation in undifferentiated stem cells is usually kept low and must be strictly regulated[49]. Nevertheless, stem cells need to maintain the proper expression level of the main stem factors to keep their specific properties and characteristics[49]. Also, a higher RNA production in mouse spermatogonia cells was reported in earlier studies[50]. The Diff.ing SPG2 and Diff.ed SPG cells were enriched with terms of the cell cycle, chromosome organization, DNA metabolic process, and cellular macromolecular complex subunit organization (Fig. 2E). The cell cycle or cell-division cycle is started in differentiating spermatogonia cells with mitotic division and continued in spermatocyte cells with meiosis division[51]. During mitosis, extensive chromosome organization is needed to transport genetic material to daughter cells[52]. In the Leptotene SPC cells, BPs of spermatogenesis and meiosis were enriched in addition to Diff.ed SPG BPs (Fig. 2E). The meiosis process was the main BP in the spermatocyte cells. The cell wall macromolecule catabolic process genes were highly expressed in Round SPT1 and Round SPT2 (Fig. 2E). Furthermore, spermatid development and sperm motility were up-regulated in Round SPT2 and Elongating SPT. Finally, in Elongating SPT cells, spermatogenesis, spermatid development, sperm motility, nucleus organization, and spermatid nucleus differentiation were enriched (Fig. 2E). The BP enrichment seems reasonable since the closer cells in the differentiation process, the more similar BPs are enriched.

Developmental ordering of spermatogenesis cells

The developmental order of these cells and clusters on the UMAP space were in agreement with spermatogenesis cells order (Fig. 3A). The PTGDS and ZNF428 were the top up-regulated genes in somatic and Undiff. SPG1, which were expressed at the same time (Fig. 3B). Then ID4, TKTL1, HIST1H4C, HIST1H4C, TEX101 CETN3, PPP3R2, GLIPR1L1, LINC00643, LINC00919, GOLGA6L2, PRM2 were expressed sequentially which were top up-regulated genes in Undiff. SPG2, Diff.ing SPG1, Diff.ing SPG2, Diff.ed SPG, Leptotene SPC, Zygotene SPC, Pachytene SPC, Diplotene SPC, Early round SPT, Round SPT1, Round SPT2, and Elongating SPT cell clusters, respectively (Fig. 3B).

Weighted gene co-expression network indicates bridge genes between testicular cells

The clustering dendrogram of genes in the WGCN resulted in 6 modules (Fig. 4A). The eigengene dendrogram and eigengene adjacency heatmap displayed the inter-modular relationships which revealed a high correlation between turquoise and yellow modules (Fig. 4B). Also, there is a correlation between the red and the green modules and between these modules with the brown one. The brown module eigengenes location on the UMAP space and its higher values in cluster-1, -2, -10, -12, and -13 indicated

that this module related to the co-expression genes in spermatogonia cells (Fig. 4C and 4D). The blue module eigengenes fitted to the location of the spermatocyte cells on the UMAP and cluster-5, -6, -8, and -9 (Fig. 4C and 4D). These results for turquoise and yellow modules displayed that these modules were related to co-expression genes in spermatid cells. The co-expression genes in the somatic cells were presented in red and green modules which revealed higher expression in the cluster-14, -15, and -16 (Fig. 4C and 4D).

The WGCN of the integrated data was constructed and shown with Cytoscape (Fig. 5A). The co-expression network is colored based on the BC value for each node (Table S2) and its top ten nodes, DNAJC5B, C1orf194, HSP90AB1, BST2, EEF1A1, CRISP2, PTMS, NFKBIA, CDKN3, and HLA-DRA were highlighted (Fig. 5B). The results demonstrated all these genes were expressed in all cell type clusters with different levels. BST2, EEF1A1, PTMS, NFKBIA, and HLA-DRA revealed higher expression at the beginning of the pseudotime trajectory in somatic cells (Fig. 5C and 5D). HSP90AB1 was one other bridge gene in this network that was particularly expressed in spermatogonia cells. C1orf194 and CDKN3 were specially expressed in the middle of the pseudotime trajectory and spermatocyte cell clusters (Fig. 5C and 5D). DNAJC5B as the gene with the highest BC value and CRISP2 were other bridge genes that were expressed in the spermatid cell clusters especially, the Elongating SPT cluster (Fig. 5C and 5D). Then these analyses were performed between brown and blue modules in the WGCN to find bridge genes between the spermatogonia and spermatocyte cells as sequential cell types in spermatogenesis. The mentioned BCs were presented in Table S3. C1orf194, HSP90AB1, MFSD6L, TPD52L3, PTMA, PHF7, BOLL, TEX40, C6orf48, and NDUFAF3 were detected as the bridge genes between the brown and blue modules in the WGCN (Fig. S6A and S6B). The gene expression along the time trajectory and clusters (Fig. S6C-D) showed most of these genes expressed in the middle of time and spermatocyte cells. Then, the bridge genes between spermatocyte and spermatid cells were evaluated, using BC between related modules (Fig. S7A, Table S4). The centrality analysis identified DNAJC5B, C1orf194, CDKN3, CRISP2, MFSD6L, CCDC89, CALM2, TPD52L3, SPACA7, and RCN2 as bridge genes (Fig. S7B). These genes expressions were well-distributed between both cell type clusters and along the pseudotime trajectory (Fig. S7C-D).

Discussion

In this study, we integrated five scRNA-seq datasets of more than 33,000 testicular cells, to identify pure and comprehensive cell profiles for spermatogenesis. Although four of these datasets were retrieved from *Hermann et al.,*[22] containing the steady-state of spermatogenesis and three sorted spermatogenic major cell types, they were not integrated. The value of integrating and re-analyzing these datasets is due to genetic diversity and different developmental timing between different individuals. On the other hand, only a few samples were evaluated in each study, that those tissues were available due to a disease or trauma [10]. Our data integration led to the coverage of similar cell types in different datasets. However, sorted spermatocyte and spermatid datasets overlapped which can be a drawback of the STA-PUT method to isolate pure cells[10,53]. The integration, in our study, led to 16 clusters within the

spermatogenesis complex process. The goal of scRNA-seq datasets integration is to improve cell classifications and identification of cell type-dependent gene expression differences[31].

The evaluation of marker gene expression identified two, two, and one cluster for undifferentiating, differentiating, and differentiated spermatogonia cells, respectively. While, the number of spermatogonia clusters in the Spermatogenesis1 and Spermatogenesis2 dataset alone were four and three, respectively[22,29]. The spermatogonia cells presented fewer up-regulated genes than down-regulated ones which were involved in the translational process and started the cell cycle through the spermatogenesis. The UTF1 and ID4 genes are known marker genes for SSC[54,55] that were differentially expressed in Undiff. SPG1 and Undiff. SPG2 clusters, though ID4 is also expressed in Undiff. SPG1 with lower averaged log fold change value. A similar result was found for these genes that marked distinctly with partially overlapping in undifferentiated spermatogonia cells, which proved the heterogeneity in these cells[55]. ASB9 gene was detected as a top DEG in Diff.ing SPG1 cluster is inconsistent with its identified expression in early differentiating spermatogonia cells[25]. Diff.ing SPG2 belongs to the late differentiating spermatogonia cells, due to the similarities in top DEGs with Diff.ed SPG cells. All results insist on heterogeneity within the spermatogonia cell population which was declared in some previous studies[56–59]. The four cell types of spermatocytes (leptotene, zygotene, pachytene, and diplotene) were identified distinctly which their DEGs significantly enrich meiosis BP. These differential stages of meiotic prophase I associated with genes down-regulation that is consistent with low RNA production during early meiosis in humans[60] and mice[50]. Whereas the Spermatogenesis1 and Spermatogenesis2 datasets alone revealed four and seven spermatocyte clusters, respectively[22,29]. Three, one, one and one of seven spermatocyte clusters in Spermatogenesis2 were fitted to leptotene, zygotene, pachytene, and diplotene cells, respectively, and the other one of seven clusters was a mixture of spermatocyte cell types[29]. Five spermatid clusters demonstrated the heterogeneity in spermatid cells with one, two, and one cluster for early-round, round, and elongating spermatid cells. In addition to spermatocytes, the Early round SPT cluster also enriched meiosis BP which produces round spermatids [61]. The C7orf61 and TNP1 are two known round spermatids that belonged to top DEGs of Round SPT1 and Round SPT2, respectively. These results indicate the presence of heterogeneous spermatid cells during the spermatogenesis process which presented many up- and down-regulated genes compared to other spermatogenesis cells. On the other hand, the Spermatogenesis1 and Spermatogenesis2 datasets lonely presented 7 and 4 spermatid cell types, respectively[22,29]. The expression of the top DEG of each cluster in pseudotime proved another confirmation on the cell type assignment and ordering. Based on these results the clustering of the integrated scRNAseq of the spermatogenic cells led to more comprehensive clustering than each of those datasets separately.

The "Guilt by Association" is one of the concepts that provide the use of gene co-expression networks to identify gene functions and molecular mechanisms in biological processes[62]. Gene co-expression network on scRNA-seq data can find functional modules related to a specific state[63,64]. In this regard, the WGCN analysis detected six modules. Adaptation of the expression pattern of these six modules with cell

clusters and eigengene dendrograms led to the attribution of these co-expression gene modules to the main stages of testicular cells, including somatic, spermatogonia, spermatocyte, and spermatid cells. Topological analysis of a cell-type-specific gene co-expression network can be useful to find the main functional genes between modules[63]. Among the network topological analysis, BC represents the influence of a node on its neighbors and the spread of information, in other words, a node with a high value of BC can be the bridge point between network modules[65,66]. The investigation of the BC of WGCN of these integrated testicular scRNA-seq datasets showed DNAJC5B, C1orf194, HSP90AB1, BST2, EEF1A1, CRISP2, PTMS, NFKBIA, CDKN3, and HLA-DRA were the top ten genes with the highest BC values. Interestingly, studies have shown that most of these genes have played a role in infertility disorders. C1orf194 was differentially expressed in the asthenozoospermic infertile group in comparison to the normozoospermic infertile group[67]. HSP90AB1 interacted with the catalytic domain of Kdm3a, that mutant Kdm3a can cause male infertility in mice[68]. Furthermore, the Hsp90ab1 gene lacking was reported to cause embryo death during implantation in mice [69]. The EEF1A1 heterozygous mutation led to spermatogenesis arrest phenotype and male infertility in tilapia[70]. The low CRISP2 expressions in asthenozoospermic[71,72] and teratoasthenozoospermic[73] patients were reported. An association was identified between NFKBIA gene polymorphisms and idiopathic male infertility risk[74]. The expression of the CDKN3 gene was reduced in teratozoospermic men[75]. GWAS studies showed HLA-DRA gene-related SNPs were significantly related to Nonobstructive Azoospermia[76,77]. Interestingly, five of these genes highly expressed in the somatic cells which is consistent with the high effects of somatic cells on the different stages of spermatogenesis[78]. Then to find specific bridge genes between the main stages of spermatogenesis, we zoomed in sequential stages of testicular cell genes in the WGCN. The top ten BC genes between spermatogonia and spermatocyte modules were C1orf194, HSP90AB1, MFSD6L, TPD52L3, PTMA, PHF7, BOLL, TEX40, C6orf48, and NDUFAF3. The top two BC genes between these modules, C1orf194, and HSP90AB1, were similar to the top BC genes of the global WGCN. A down-regulation of TPD52L3 was reported in oligozoospermia [79]. Disruption of Phf7 caused infertility in male mice by decreasing sperm count and increasing abnormal sperm ratio[80]. The relation of BOLL deletion or mutation with unfunctional sperm production that led to infertility has been reported in different studies [81–84]. The expression of TEX40, a calcium entry protein, is reduced in asthenozoospermic males [85] and targeted disruption of TEX40 led to severe male subfertility in mice [86]. In the next step, the top ten BC genes between spermatocyte and spermatid modules were examined as two sequential modules to find the bridge genes between them. The DNAJC5B, C1orf194, CDKN3, CRISP2, MFSD6L, CCDC89, CALM2, TPD52L3, SPACA7, and RCN2 genes were identified as the top ten BC genes. The four (DNAJC5B, C1orf194, CDKN3, and CRISP2) and three (C1orf194, MFSD6L and TPD52L3) genes between these modules were similar to the top BC genes of the global, and spermatogonial-spermatocyte part of the WGCN, respectively. C1orf194 was detected as the top BC gene in all global, spermatogonial-spermatocyte, and spermatocyte-spermatid parts of the WGCN.

In summary, five different testicular scRNA-seq datasets were integrated to construct comprehensive spermatogenesis transcriptome-wide data. The clustering, cell type assignments, DEGs, and pseudotime

analysis revealed heterogeneity in spermatogenesis's main stages. Then, the WGCN along with the integrated scRNA-seq data identified functional modules associated with spermatogenesis main stages. The BC analysis on this cell-type-specific WGCN discovered some bridge genes between the spermatogenesis main stages such as DNAJC5B, C1orf194, HSP90AB1, BST2, EEF1A1, CRISP2, PTMS, NFKBIA, CDKN3, and HLA-DRA. Some of these bridge genes are highly expressed in the somatic cells, emphasizing the role of somatic cells in spermatogenesis. Available studies about these genes showed that perturbation of these genes led to male infertility disorders, which confirms the functional role of top betweenness genes in this cell-type-specific WGCN. These functional bridge genes can be use as candidates for filtering and prioritizing genetic variants and gene expression alterations with the goal of introducing a male infertility panel. So, our study not only offers knowledge about cell-to-cell heterogeneity in spermatogenesis but also introduces key genes between the functional modules of normal spermatogenesis that may play important roles in male infertility disorders. These results can be a starting point for experimental research to investigate the function of these genes in male infertility.

Methods

The scRNA-seq datasets and preprocessing

The scRNA-seq datasets related to human spermatogonia, spermatocyte, spermatid sorted cells (GEO: GSE109037) and steady-state spermatogenic cells (GEO: GSE109037 and GSE106487) with normal spermatogenesis were retrieved from the gene expression omnibus (GEO) repository[87]. These datasets were produced across 10x Genomics and SMART-seq2 technologies. The Seurat3.2 R package [88] was used for data analysis. To filter out low-quality cells, at first, cells with less than 2,000 expressed genes and genes expressed in less than 3 cells were removed. Then, cells with a very low or high number of genes and cells with a high percentage of mitochondrial genes were filtered. Standard preprocessing, normalizing, and identifying 2,000 highly variable features were performed individually for each dataset. Totally, 33,185 cells were collected for integration.

Data integration and analysis

Anchor strategy[89] was used to integrate these datasets, which were produced across multiple technologies. Finding an accurate set of anchors is the basis for subsequent integration analyses. Thus, these datasets were integrated with 2000 anchors, resulting in a batch corrected expression matrix for all cells. The new integrated matrix was used for scaling and the principal component analysis (PCA). The first 35 principal components (PCs) were selected based on the variance percentage of each PC to perform UMAP non-linear dimension reduction[90] to visualize, explore and separate cells. The graph-based clustering approach of the Seurat3.2 R package was used to find clusters with a dimensionality of 35 and a resolution of 0.2. The cell type of each cluster was assigned based on the expression of specific markers of spermatogenic cells obtained from the literature.

Differentially expressed genes and enrichment analysis

To find differentially expressed genes (DEG), the non-parametric Wilcoxon rank sum test[91] was used. The minimum percentage in both cell groups (min.pct) and the log fold-change of the average expression between the two cell groups (logfc.threshold) were set to 0.25 and 0.5, respectively. The up and down-regulated genes in each cluster in comparison to all other clusters were quantified based on positive and negative averaged log fold-change values, respectively. Up-regulated genes with averaged log fold-change higher than 0.7 and adjusted p-value (based on Bonferroni correction) less than 0.05 were selected for enrichment analysis. The Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.7[92] was used for gene enrichment analysis. The biological processes (BPs) terms with the lowest Benjamini correction score (adjusted p-value) were used to plot the heat map.

Pseudotime Analysis

For pseudotime analysis, the Monocle3 R package was used[93]. The integrated data, dimension reduction and clustering information were imported from Seurat to the Monocle3 package. To order the cells in pseudotime, Monocle3 learns a trajectory that reconstructs the progress of a cell in a cell differentiation process. After the graph learning, the cells were ordered according to their progress.

Co-expression network construction and analysis

To reveal correlations between gene expression of these integrated cells, a weighted gene co-expression network (WGCN) was created by the WGCNA R package[94]. To construct the WGCN with scale-free topology, different values of soft thresholding power β were assessed for the network topology analysis, and the value of 6 was selected. The Pearson correlation coefficient and the signed network options were used to measure the correlation between the expression of each pair of genes and to maintain only positive correlations, respectively. The topological overlap measure (TOM), which investigates the similarities between gene pairs based on the number of shared neighbors in the resulting co-expression network, was used to identify modules. Modules in the WGCN were depicted in different colors. Genes that lacked similar co-expression to other genes in the network, were assigned to the gray module. So, the gray module was removed from more analysis. The relationships between the detected modules were depicted by module eigengenes that are the first principal component of the expressions in modules. Constructed WGCN was exported to Cytoscape[95]. To find essential genes in this network, the betweenness centrality (BC) of each node was measured. A node with the highest BC value indicates the bridge node in that network[66]. Cytoscape and its plugin CytoNCA[65], were used for network visualization and centralities measurements, respectively.

Codes are available at https://github.com/nasalehi/scRNAseq_spermatogenesis

Declarations

Acknowledgements: This study was supported by Iran Science Elites Federation, the Royan Institute, Iran National Science Foundation (INSF) (grant number: 99013926), and Institute for Research in

Fundamental Sciences (IPM).

Author contributions: N.S. designed the study, analyzed the data, and wrote the manuscript; M.H.K. validated the analysis and revised the manuscript; A.A.Y. and M.T. designed the study, revised, and proofread the manuscript. All authors read and approved the final manuscript.

Competing interests: The authors declare no competing interests.

References

1. De Kretser, D. M., Loveland, K. L., Meinhardt, A., Simorangkir, D. & Wreford, N. Spermatogenesis. in *Human Reproduction* **13**, 1–8 (Oxford University Press, 1998).
2. Watson, R. R. *Handbook of Fertility: Nutrition, Diet, Lifestyle and Reproductive Health*. *Handbook of Fertility: Nutrition, Diet, Lifestyle and Reproductive Health* (Elsevier Inc., 2015). doi:10.1016/C2013-0-19077-0
3. De Rooij, D. G. & Grootegoed, J. A. Spermatogonial stem cells. *Curr. Opin. Cell Biol.* **10**, 694–701 (1998).
4. Valli, H., Phillips, B. T., Orwig, K. E., Gassei, K. & Nagano, M. C. Spermatogonial Stem Cells and Spermatogenesis. in *Knobil and Neill's Physiology of Reproduction: Two-Volume Set* **1**, 595–635 (Elsevier Inc., 2015).
5. De Braekeleer, M., Nguyen, M. H., Morel, F. & Perrin, A. Genetic aspects of monomorphic teratozoospermia: a review. *J. Assist. Reprod. Genet.* **32**, 615–623 (2015).
6. Matzuk, M. M. & Lamb, D. J. The biology of infertility: Research advances and clinical challenges. *Nature Medicine* **14**, 1197–1213 (2008).
7. Asero, P. *et al.* Relevance of genetic investigation in male infertility. *Journal of Endocrinological Investigation* **37**, 415–427 (2014).
8. Xavier, M. J., Salas-Huetos, A., Oud, M. S., Aston, K. I. & Veltman, J. A. Disease gene discovery in male infertility: past, present and future. *Human Genetics* 1–13 (2020). doi:10.1007/s00439-020-02202-x
9. Oud, M. S. *et al.* A systematic review and standardized clinical validity assessment of male infertility genes. *Hum. Reprod.* **34**, 932–941 (2019).
10. Suzuki, S., Diaz, V. D. & Hermann, B. P. What has single-cell RNA-seq taught us about mammalian spermatogenesis? *Biol. Reprod.* **0**, 1–18 (2019).
11. Shima, J. E., McLean, D. J., McCarrey, J. R. & Griswold, M. D. The Murine Testicular Transcriptome: Characterizing Gene Expression in the Testis During the Progression of Spermatogenesis1. *Biol. Reprod.* **71**, 319–330 (2004).
12. Laiho, A., Kotaja, N., Gyenesei, A. & Sironen, A. Transcriptome Profiling of the Murine Testis during the First Wave of Spermatogenesis. *PLoS One* **8**, e61558 (2013).
13. Chalmel, F. *et al.* High-resolution profiling of novel transcribed regions during rat spermatogenesis. *Biol. Reprod.* **91**, 1–13 (2014).

14. Rolland, A. D. *et al.* RNA profiling of human testicular cells identifies syntenic lncRNAs associated with spermatogenesis. *Hum. Reprod.* **34**, 1278–1290 (2019).
15. Gatta, V. *et al.* Testis transcriptome analysis in male infertility: New insight on the pathogenesis of oligo-azoospermia in cases with and without AZFc microdeletion. *BMC Genomics* **11**, 401–410 (2010).
16. Razavi, S. M. *et al.* Comprehensive functional enrichment analysis of male infertility. *Sci. Rep.* **7**, 1–14 (2017).
17. Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental and Molecular Medicine* **50**, 1–14 (2018).
18. Shalek, A. K. *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363–369 (2014).
19. Bellve, A. R. *et al.* Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. *J. Cell Biol.* **74**, 68–85 (1977).
20. Chen, Y. *et al.* Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. *Cell Res.* **28**, 879–896 (2018).
21. Guo, J. *et al.* 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.e6 (2017).
22. Hermann, B. P. *et al.* The Mammalian Spermatogenesis Single-Cell Transcriptome, from Spermatogonial Stem Cells to Spermatids. *Cell Rep.* **25**, 1650-1667.e8 (2018).
23. Guo, J. *et al.* The adult human testis transcriptional cell atlas. *Cell Res.* **28**, 1141–1157 (2018).
24. Neuhaus, N. *et al.* Single-cell gene expression analysis reveals diversity among human spermatogonia. *Mol. Hum. Reprod.* **23**, 79–90 (2017).
25. Sohni, A. *et al.* The Neonatal and Adult Human Testis Defined at the Single-Cell Level. *Cell Rep.* **26**, 1501-1517.e4 (2019).
26. Xia, B. *et al.* Widespread Transcriptional Scanning in the Testis Modulates Gene Evolution Rates. *Cell* **180**, 248-262.e21 (2020).
27. Guo, J. *et al.* The Dynamic Transcriptional Cell Atlas of Testis Development during Human Puberty. *Cell Stem Cell* **26**, 262-276.e4 (2020).
28. Shami, A. N. *et al.* Single-Cell RNA Sequencing of Human, Macaque, and Mouse Testes Uncovers Conserved and Divergent Features of Mammalian Spermatogenesis. *Dev. Cell* **54**, 529-547.e12 (2020).
29. Wang, M. *et al.* Single-Cell RNA Sequencing Analysis Reveals Sequential Cell Fate Transition during Human Spermatogenesis. *Cell Stem Cell* **23**, 599-614.e4 (2018).
30. Zhao, L. Y. *et al.* Single-cell analysis of developing and azoospermia human testicles reveals central role of Sertoli cells. *Nat. Commun.* **11**, 5683 (2020).
31. Adey, A. C. Integration of Single-Cell Genomics Datasets. *Cell* **177**, 1677–1679 (2019).

32. Sada, A., Suzuki, A., Suzuki, H. & Saga, Y. The RNA-binding protein NANOS2 is required to maintain murine spermatogonia! Stem Cells. *Science (80-.).* **325**, 1394–1398 (2009).
33. Sada, A., Hasegawa, K., Pin, P. H. & Saga, Y. NANOS2 acts downstream of glial cell line-derived neurotrophic factor signaling to suppress differentiation of spermatogonial stem cells. *Stem Cells* **30**, 280–291 (2012).
34. Sasaki, T., Shiohama, A., Minoshima, S. & Shimizu, N. Identification of eight members of the Argonaute family in the human genome. *Genomics* **82**, 323–330 (2003).
35. von Kopylow, K. & Spiess, A. N. Human spermatogonial markers. *Stem Cell Res.* **25**, 300–309 (2017).
36. He, Z., Kokkinaki, M., Jiang, J., Dobrinski, I. & Dym, M. Isolation, characterization, and culture of human spermatogonia. *Biol. Reprod.* **82**, 363–372 (2010).
37. Dai, J., Voloshin, O., Potapova, S. & Camerini-Otero, R. D. Meiotic Knockdown and Complementation Reveals Essential Role of RAD51 in Mouse Spermatogenesis. *Cell Rep.* **18**, 1383–1394 (2017).
38. Fernandes, M. G. *et al.* Human-specific subcellular compartmentalization of P-element induced wimpy testis-like (PIWIL) granules during germ cell development and spermatogenesis. *Hum. Reprod.* **33**, 258–269 (2018).
39. Bisig, C. G. *et al.* Synaptonemal complex components persist at centromeres and are required for homologous centromere pairing in mouse spermatocytes. *PLoS Genet.* **8**, 1002701 (2012).
40. Chizaki, R., Yao, I., Katano, T., Matsuda, T. & Ito, S. Restricted Expression of Ovol2/MOVO in XY Body of Mouse Spermatocytes at the Pachytene Stage. *J. Androl.* **33**, 277–286 (2012).
41. Sun, M. *et al.* Efficient generation of functional haploid spermatids from human germline stem cells by three-dimensional-induced system. *Cell Death Differ.* **25**, 747–764 (2018).
42. Kashiwabara, S. ichi, Arai, Y., Kodaira, K. & Baba, T. Acrosin biosynthesis in meiotic and postmelotic spermatogenic cells. *Biochem. Biophys. Res. Commun.* **173**, 240–245 (1990).
43. Zheng, H. *et al.* Lack of Spem1 causes aberrant cytoplasm removal, sperm deformation, and male infertility. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 6852–6857 (2007).
44. Li, H., MacLean, G., Cameron, D., Clagett-Dame, M. & Petkovich, M. Cyp26b1 expression in murine sertoli cells is required to maintain male germ cells in an undifferentiated state during embryogenesis. *PLoS One* **4**, e7501 (2009).
45. Rossato, M. *et al.* The novel hormone INSL3 is expressed in human testicular Leydig cell tumors: A clinical and immunohistochemical study. *Urol. Oncol. Semin. Orig. Investig.* **29**, 33–37 (2011).
46. Chen, S. R. & Liu, Y. X. Myh11-Cre is not limited to peritubularmyoid cells and interaction between Sertoli and peritubular myoid cells needs investigation. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E2352 (2016).
47. Arnold, S. L. *et al.* Importance of ALDH1A enzymes in determining human testicular Retinoic acid concentrations. *J. Lipid Res.* **56**, 342–357 (2015).
48. Vernet, N. *et al.* Retinoic acid metabolism and signaling pathways in the adult and developing mouse testis. *Endocrinology* **147**, 96–110 (2006).

49. Gabut, M., Bourdelais, F. & Durand, S. Ribosome and Translational Control in Stem Cells. *Cells* **9**, 497 (2020).
50. Monesi, V. Ribonucleic acid synthesis during mitosis and meiosis in the mouse. *J. Cell Biol.* **22**, 521–532 (1964).
51. Creasy, D. M. & Chapin, R. E. Male Reproductive System. in *Fundamentals of Toxicologic Pathology: Third Edition* 459–516 (Elsevier Inc., 2018). doi:10.1016/B978-0-12-809841-7.00017-4
52. Batty, P. & Gerlich, D. W. Mitotic Chromosome Mechanics: How Cells Segregate Their Genome. *Trends in Cell Biology* **29**, 717–726 (2019).
53. Bryant, J. M., Meyer-Ficca, M. L., Dang, V. M., Berger, S. L. & Meyer, R. G. Separation of Spermatogenic cell types using STA-PUT velocity sedimentation. *J. Vis. Exp.* e50648 (2013). doi:10.3791/50648
54. Jung, H., Roser, J. F. & Yoon, M. UTF1, a putative marker for spermatogonial stem cells in stallions. *PLoS One* **9**, e108825 (2014).
55. Sachs, C. *et al.* Evaluation of candidate spermatogonial markers ID4 and GPR125 in testes of adult human cadaveric organ donors. *Andrology* **2**, 607–614 (2014).
56. Niedenberger, B. A., Busada, J. T. & Geyer, C. B. Marker expression reveals heterogeneity of spermatogonia in the neonatal mouse testis. *Reproduction* **149**, 329–338 (2015).
57. Grisanti, L. *et al.* Identification of Spermatogonial Stem Cell Subsets by Morphological Analysis and Prospective Isolation. *Stem Cells* **27**, 3043–3052 (2009).
58. Liao, J. *et al.* Revealing cellular and molecular transitions in neonatal germ cell differentiation using single cell RNA sequencing. *Dev.* **146**, 1–15 (2019).
59. Kubota, H. Heterogeneity of Spermatogonial Stem Cells. in *Advances in Experimental Medicine and Biology* **1169**, 225–242 (Springer New York LLC, 2019).
60. Jan, S. Z. *et al.* Unraveling transcriptome dynamics in human spermatogenesis. *Dev.* **144**, 3659–3673 (2017).
61. Mays-Hoopes, L. L., Bolen, J., Riggs, A. D. & Singer-Sam, J. Preparation of spermatogonia, spermatocytes, and round spermatids for analysis of gene expression using fluorescence-activated cell sorting. *Biol. Reprod.* **53**, 1003–1011 (1995).
62. Lamere, A. T. & Li, J. Inference of gene co-expression networks from single-cell RNA-sequencing data. in *Methods in Molecular Biology* **1935**, 141–153 (Humana Press Inc., 2019).
63. Cha, J. & Lee, I. Single-cell network biology for resolving cellular heterogeneity in human diseases. *Experimental and Molecular Medicine* **52**, 1798–1808 (2020).
64. Li, Y. *et al.* Elucidation of Biological Networks across Complex Diseases Using Single-Cell Omics. *Trends in Genetics* **36**, 951–966 (2020).
65. Tang, Y., Li, M., Wang, J., Pan, Y. & Wu, F. X. CytoNCA: A cytoscape plugin for centrality analysis and evaluation of protein interaction networks. *BioSystems* **127**, 67–72 (2015).
66. Newman, M. E. J. A measure of betweenness centrality based on random walks. *Soc. Networks* **27**, 39–54 (2005).

67. Bansal, S. K., Gupta, N., Sankhwar, S. N. & Rajender, S. Differential genes expression between fertile and infertile spermatozoa revealed by transcriptome analysis. *PLoS One* **10**, e0127007 (2015).
68. Kasioulis, I. *et al.* Kdm3a lysine demethylase is an Hsp90 client required for cytoskeletal rearrangements during spermatogenesis. *Mol. Biol. Cell* **25**, 1216–1233 (2014).
69. Dun, M. D., Aitken, R. J. & Nixon, B. The role of molecular chaperones in spermatogenesis and the post-testicular maturation of mammalian spermatozoa. *Human Reproduction Update* **18**, 420–435 (2012).
70. Chen, J. *et al.* Heterozygous mutation of eEF1A1b resulted in spermatogenesis arrest and infertility in male tilapia, Oreochromis niloticus. *Sci. Rep.* **7**, e43733 (2017).
71. Zhou, J.-H. *et al.* The Expression of Cysteine-Rich Secretory Protein 2 (CRISP2) and Its Specific Regulator miR-27b in the Spermatozoa of Patients with Asthenozoospermia 1. *Biol. Reprod.* **92**, 28–29 (2015).
72. Heidary, Z., Zaki-Dizaji, M., Saliminejad, K. & Khorramkhorshid, H. R. Expression analysis of the CRISP2, CATSPER1, PATE1 and SEMG1 in the sperm of men with idiopathic asthenozoospermia. *J. Reprod. Infertil.* **20**, 70–75 (2019).
73. Gholami, D. *et al.* The expression of Cysteine-Rich Secretory Protein 2 (CRISP2) and miR-582-5p in seminal plasma fluid and spermatozoa of infertile men. *Gene* **730**, 144261 (2020).
74. Wang, T., Hu, T., Zhen, J., Zhang, L. & Zhang, Z. Association of MTHFR, NFKB1, NFKBIA, DAZL and CYP1A1 gene polymorphisms with risk of idiopathic male infertility in a Han Chinese population. *Int. J. Clin. Exp. Pathol.* **10**, 7640–7649 (2017).
75. Platts, A. E. *et al.* Success and failure in human spermatogenesis as revealed by teratozoospermic RNAs. *Hum. Mol. Genet.* **16**, 763–773 (2007).
76. Zhao, H. *et al.* A genome-wide association study reveals that variants within the HLA region are associated with risk for nonobstructive azoospermia. *Am. J. Hum. Genet.* **90**, 900–906 (2012).
77. Hu, Z. *et al.* Association analysis identifies new risk loci for non-obstructive azoospermia in Chinese men. *Nat. Commun.* **5**, e3857 (2014).
78. Fritz, I. B. Somatic cell-germ cell relationships in mammalian testes during development and spermatogenesis. *Ciba Foundation symposium* **182**, (1994).
79. Montjean, D. *et al.* Sperm transcriptome profiling in oligozoospermia. *J. Assist. Reprod. Genet.* **29**, 3–10 (2012).
80. Wang, X. *et al.* PHF7 is a novel histone H2A E3 ligase prior to histone-toprotamine exchange during spermiogenesis. *Dev.* **146**, 191445 (2019).
81. Li, T. *et al.* Histomorphological comparisons and expression patterns of BOLL gene in sheep testes at different development stages. *Animals* **9**, 105–116 (2019).
82. Luetjens, C. M. *et al.* Association of meiotic arrest with lack of BOULE protein expression in infertile men. *J. Clin. Endocrinol. Metab.* **89**, 1926–1933 (2004).

83. Kee, K., Angeles, V. T., Flores, M., Nguyen, H. N. & Reijo Pera, R. A. Human DAZL, DAZ and BOULE genes modulate primordial germ-cell and haploid gamete formation. *Nature* **462**, 222–225 (2009).
84. Yung, M. L., Chia, L. C. & Yu, S. C. Posttranscriptional regulation of CDC25A by BOLL is a conserved fertility mechanism essential for human spermatogenesis. *J. Clin. Endocrinol. Metab.* **94**, 2650–2657 (2009).
85. Sinha, A., Singh, V., Singh, S. & Yadav, S. Proteomic analyses reveal lower expression of TEX40 and ATP6V0A2 proteins related to calcium ion entry and acrosomal acidification in asthenozoospermic males. *Life Sci.* **218**, 81–88 (2019).
86. Chung, J. J. *et al.* Catsperς regulates the structural continuity of sperm ca2+ signaling domains and is required for normal fertility. *Elife* **6**, e23082 (2017).
87. Barrett, T. *et al.* NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res.* **41**, 991–995 (2013).
88. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
89. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
90. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (2018).
91. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biometrics Bull.* **1**, 80 (1945).
92. Dennis, G. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**, (2003).
93. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
94. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559–572 (2008).
95. Shannon, P. *et al.* Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

Tables

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Figures

A

Datasets Names	Cell Types (Sorting Method)	scRNA-seq Method	GEO ID	Genes * Cells	Total Genes * Total Cells (After Pre-processing)
Spermatogenesis1	Steady-state spermatogenic cells	10x Genomics	GSE109037	33,694 * 7,134	Union of genes: 33,551 * 33,185
Spermatogenesis2	Steady-state spermatogenic cells	SMART-seq2	GSE106487	24,153 * 3,059	
Spermatogonia	Sorted spermatogonia cells (FACS)	10x Genomics	GSE109037	33,694 * 11,104	
Spermatocyte	Sorted spermatocyte cells (STA-PUT)	10x Genomics	GSE109037	33,694 * 4,884	
Spermatid	Sorted spermatid cells (STA_PUT)	10x Genomics	GSE109037	33,694 * 7,434	

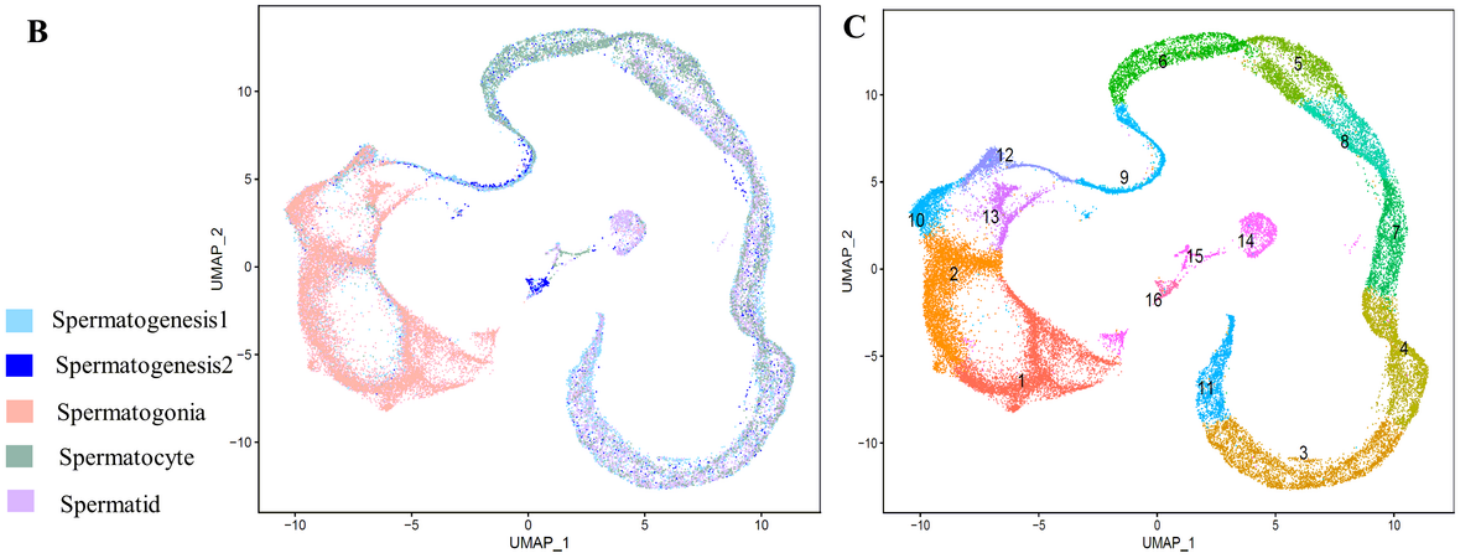

Figure 1

Profiling and integrating testicular datasets. (A) Datasets information of adult human testicular cells that were integrated and analyzed, (B, C) UMAP plot of integrated human testicular cells. Cells are colored based on (B) the original datasets, (C) clustering results.

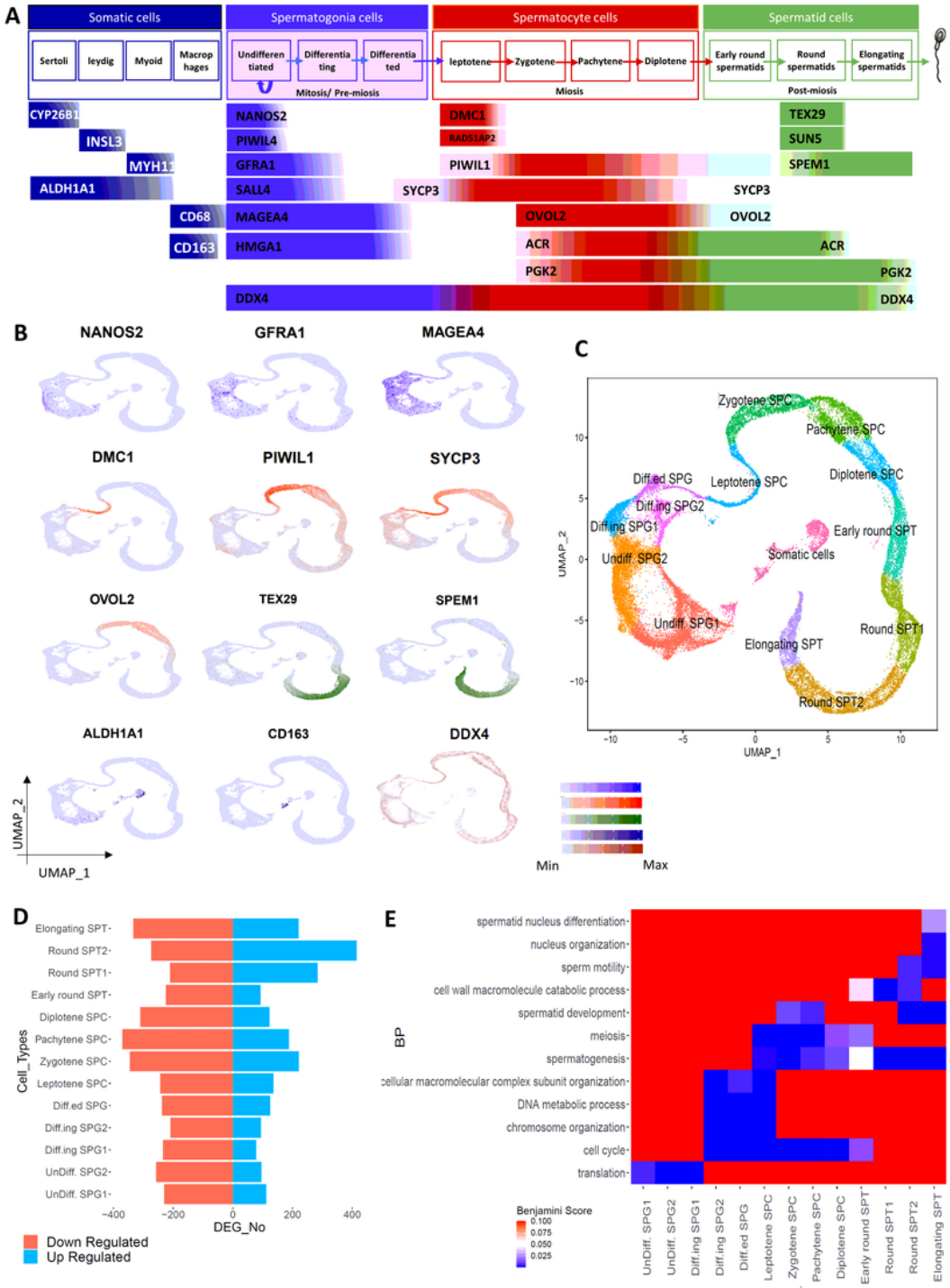

Figure 2

Cell type assignment of clusters. (A) Gene markers of testicular cells were categorized based on different somatic and spermatogenesis cells, (B) Gene expression patterns of these markers on the UMAP space, (C) Cell type assignment of clusters based on gene markers expression patterns, (A) The number of up- and down-regulated genes in germ cell types, (B) The biological processes enrichment for up-regulated genes of each cluster

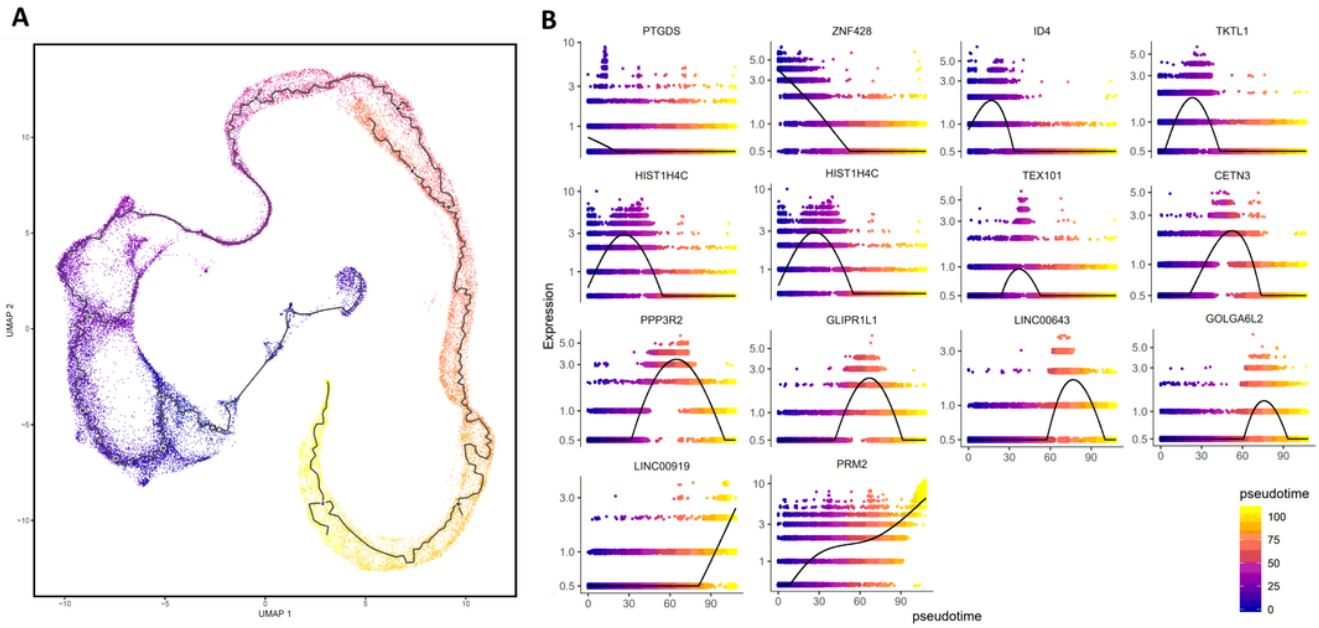

Figure 3

Developmental ordering of spermatogenesis cells. (A) The pseudotime analysis of testicular cells on the UMAP space, (B) The expression of top positive DEGs in each cluster along the pseudotime.

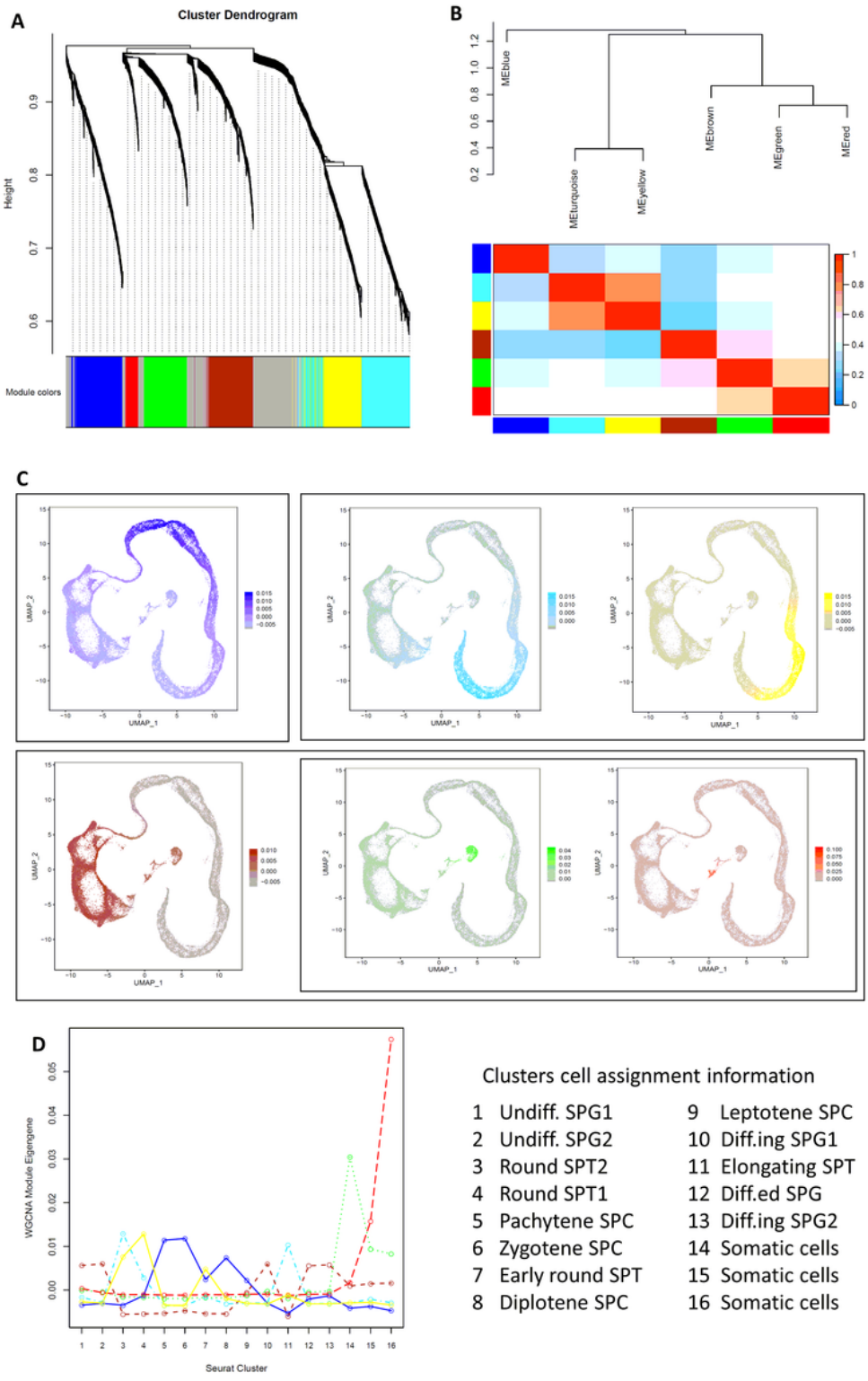

Figure 4

Weighted gene co-expression network analysis. (A) The clustering dendrogram of the weighted gene co-expression network. The resulted modules are depicted in different colors. (B) The eigengene dendrogram and eigengene adjacency heatmap of modules, (C) The gene expression patterns on the UMAP space for each module with their corresponding colors, (D) The eigengene of each module in each cluster.

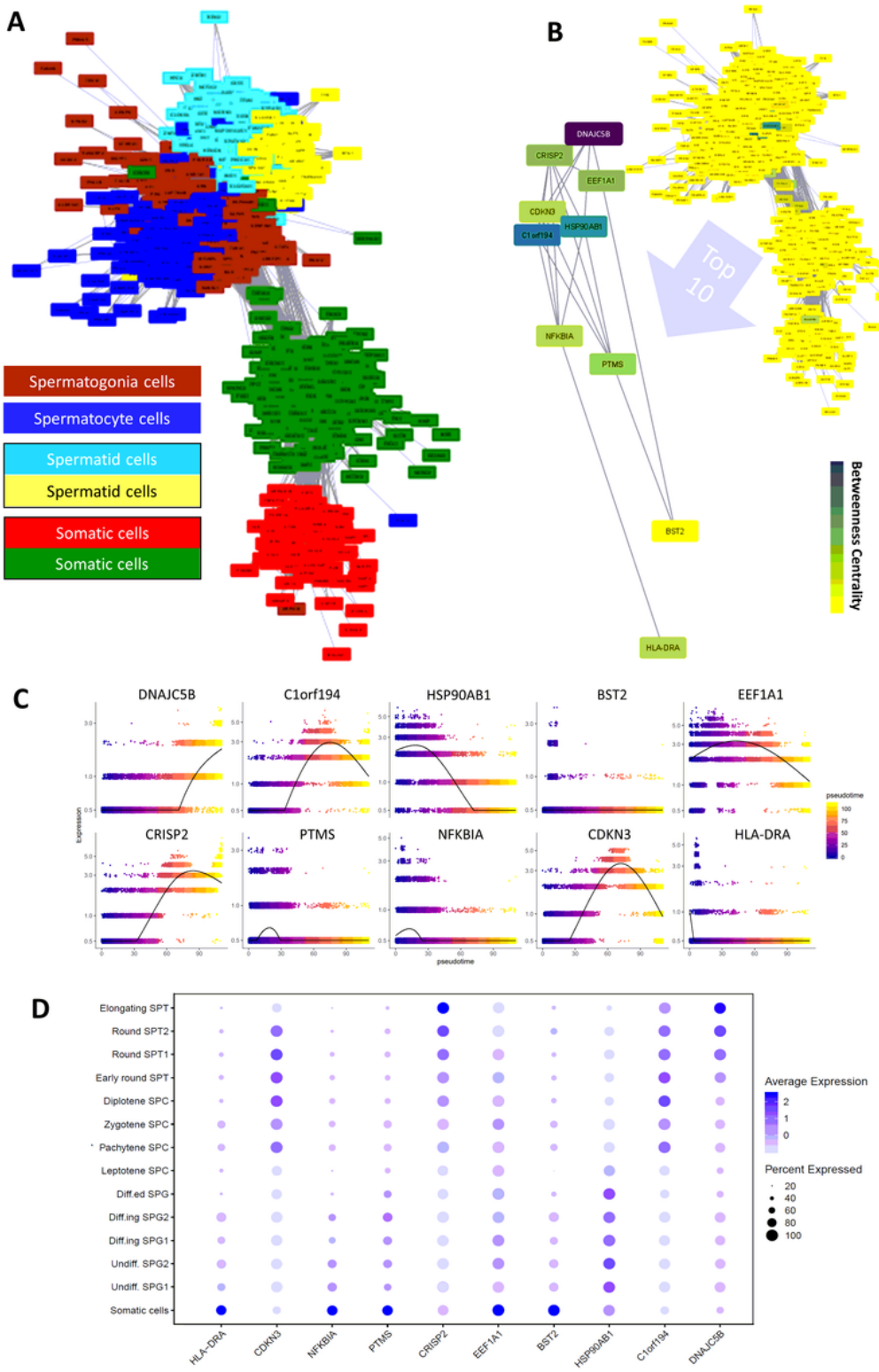

Figure 5

Betweenness centrality analysis of the weighted gene co-expression network. (A) The presentation of the weighted gene co-expression network, (B) The co-expression network colored based on the betweenness centralities from yellow to purple. The top ten genes with the highest betweenness centralities are highlighted. (C, D) The expressions of these top betweenness centrality genes along (C) the pseudotime and (D) the cell-types clusters.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.docx
- TableS1.csv
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- GraphicalAbstract.tif
- Table1.jpg