Received: 16 July 2024 | Accepted: 7 November 2024

DOI: 10.1002/rmb2.12616

Reproductive Medicine and Biology

WILEY

ORIGINAL ARTICLE

Identification of differentially expressed genes in human testis biopsies with defective spermatogenesis

Shashika D. Kothalawala[1,2] | Stefan Günther[3] | Hans-Christian Schuppe[4,5] | Adrian Pilatz[4,5] | Florian Wagenlehner[4,5] | Sabine Kliesch[6] | Liza O'Donnell[2] | Daniela Fietz[1,5]

[1]Institute for Veterinary Anatomy, Histology and Embryology, Justus-Liebig University of Giessen, Giessen, Germany

[2]Centre for Reproductive Health, Hudson Institute of Medical Research, Clayton, Victoria, Australia

[3]Max-Planck Institute for Heart and Lung Research, Bad Nauheim, Germany

[4]Clinic of Urology, Pediatric Urology and Andrology, Justus-Liebig University of Giessen, Giessen, Germany

[5]Hessian Centre of Reproductive Medicine, Justus-Liebig University of Giessen, Giessen, Germany

[6]Centre of Reproductive Medicine and Andrology, University of Muenster, Muenster, Germany

Correspondence
Daniela Fietz, Institute for Veterinary Anatomy, Histology and Embryology, Frankfurter Straße 98, 35392, Giessen, Germany.
Email: daniela.fietz@vetmed.uni-giessen.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: IRTG DFG GRK 1871/1

Abstract

Purpose: Sperm morphology and motility are major contributors to male-factor infertility, with many genes predicted to be involved. This study aimed to elucidate differentially expressed transcripts in human testis tissues of normal and abnormal spermatogenesis that could reveal new genes that may regulate sperm morphology and function.

Methods: Human testis biopsies were collected from men with well-characterized phenotypes of normal spermatogenesis, spermatid arrest, and Sertoli cell-only phenotype, and transcriptional differences were quantified by RNA-sequencing (RNA-Seq). Differentially expressed genes (DEGs) were filtered based on predominant expression in spermatids and gene functional annotations relevant to sperm morphology and motility. Selected 10 DEGs were validated by qRT-PCR and the localization of two proteins was determined in testis biopsies.

Results: The analysis revealed 6 genes (*SPATA31E1*, *TEKT3*, *SLC9C1*, *PDE4A*, *CFAP47*, and *TNC*) that are excellent candidates for novel genes enriched in developing human sperm. The immunohistochemical localization of two proteins, ORAI1 and SPATA31E1, in testis biopsies, verified that both are expressed in developing human germ cells, with SPATA31E1 enriched in late spermatocytes and spermatids.

Conclusion: This study identified human germ cell-enriched genes that could play functional roles in spermiogenesis and could thus be important in the development of morphologically normal, motile sperm.

KEYWORDS
RNA-sequencing, sperm morphology, sperm motility, spermatogenesis, spermiogenesis

1 | INTRODUCTION

Infertility, that is, the inability to conceive within 1 year,[1,2] has become a global issue with an increase in prevalence in recent years, with 10%–15% of couples suffering from infertility problems.[3] Male factor infertility is prevalent in approximately half of infertile couples[4] and can be caused by various factors such as congenital or acquired urogenital abnormalities, malignancies, urogenital tract

infections, endocrine disturbances, and genetic abnormalities.[2] In around 30%–40% of cases, the cause of infertility remains unclear and is thus referred to as idiopathic infertility.[5,6]

Ejaculate analysis as standardized by the World Health Organization[7] is the gold standard method to assess male fertility. Abnormalities in sperm number are defined as oligozoospermia (low sperm count <15 million/mL) or azoospermia (no sperm in the ejaculate). Impaired motility defects are defined as asthenozoospermia, that is, low progressive motility, and sperm morphology defects are defined as teratozoospermia, where sperm exhibit a wide range of defined structural abnormalities. The most common diagnosis is a combination of all three abnormalities and thus classified as oligoasthenoteratozoospermia (OAT), where both qualitative and quantitative abnormalities of sperm number and function are detected.[3] Morphological and functional abnormalities can be broadly divided into sperm head and flagellum defects.[8] Sperm abnormalities can arise during sperm development in the testis (spermatogenesis and/or spermiogenesis) or during passage through the epididymis.

Germ cell development, particularly of spermatocytes and spermatids, involves remarkable transcriptional complexity and diversity.[9] Studies have shown that spermatogenesis is regulated by nearly 2000 genes, and thus normal sperm motility and morphology too rely on a large number of genes.[3,10,11] The causes of defects in sperm number and/or function are complex, with potentially 30%–40% of idiopathic infertility cases caused by genetic abnormalities such as chromosomal defects or gene mutations.[3,10–12] So far, relatively few clinically relevant genes have been identified.[3,13,14]

Identifying specific markers of the different infertility phenotypes could benefit the diagnosis of infertility. We have previously identified that many sperm-specific proteins are released by the seminiferous tubules into the surrounding interstitial fluid, and from here they can enter the circulation.[15,16] Thus, the identification of novel proteins that are expressed specifically in sperm, or that differ with infertility phenotype, could be applied to the development of non-invasive diagnostics, such as blood tests, that could be used to inform clinicians about the likelihood of sperm being present in the testes and thus whether the patient is likely to have sperm available for retrieval via surgery.[15] Identifying genes that are essential for human male fertility will facilitate genetic counseling of patients, infertility diagnosis, and selection of the best treatment.[14,17,18] Importantly, elucidating genes that are specific to male germ cells and are essential for fertility also provides new opportunities for the development of non-hormonal male contraceptive approaches. Sperm production involves many genes that are highly restricted to germ cells and are not found elsewhere in the body[9,19] thus contraceptive strategies that target genes restricted to spermatid development could be highly specific and effective, with limited off-target effects. Recent approaches to male contraceptive development focus on proteins that are essential and specific for sperm production and/or function to limit the likelihood of off-target effects.[20] For example, a testis-specific bromodomain protein, BRDT has been identified as a potential target for contraception, and research has shown that mice with *Brdt* homozygous null mutation are sterile.[21] Such non-hormonal contraceptives are in the earliest phases of preclinical testing.[20] Identifying novel genes and proteins that are involved in human sperm production will have application for the design of non-hormone-based male contraceptives that could specifically inhibit sperm production or function.

RNA-Seq is a powerful tool that is used not only to discover RNA transcripts but also to compare mRNA expression[22] and identify DEGs between different phenotypes.[23] Even though next-generation sequencing has evolved from whole tissue (bulk) RNA-Seq to single-cell RNA-Seq (scRNA-Seq), bulk RNA-Seq remains one of the most used tools due to its high reproducibility compared to scRNA-Seq, high sequence coverage, and the usage of reliable statistical power and bioinformatical tools.[24] The cost-effective nature and the wide availability in many labs, and the ease of data interpretation, make it more affordable and feasible than scRNA-Seq. However, recently scRNA-Seq has been used as a new, popular application in identifying genes in human testicular samples.[25] The current study focused on identifying novel genes with a potential role in the development of human spermatozoa using RNA-Seq. We used histologically well-defined testis samples from men with azoospermia and different levels of spermatogenic failure and analyzed transcriptional differences using a custom bioinformatics strategy to identify novel genes with potential roles in impaired sperm morphology and motility.

2 | MATERIALS AND METHODS

2.1 | Sample collection and RNA extraction

Human testis biopsies were collected at the Clinic of Urology, Pediatric Urology and Andrology, Justus-Liebig University of Giessen, Germany, and the Centre for Reproductive Medicine and Andrology, University Clinic Muenster, Germany. Patients were classified as having either obstructive (OA) or non-obstructive azoospermia (NOA) after diagnosis by clinical parameters (sex hormones, ultrasound, and ejaculate analysis) (Appendix Table A1).[7,26] The mean age of the patients in the study group was 33 (29–38, SD = 3.41). Surgeries were performed after written informed consent and after approval by the local Ethics Committee (AZ 152/16 and 26/11). Biopsies for histological evaluation were fixed with Bouin's solution overnight, stained with hematoxylin and eosin following standard protocols, and evaluated following Bergmann and Kliesch score count analysis. Another piece of the biopsy was stored in liquid nitrogen for RNA analysis. The testis biopsies used for RNA-Seq were further classified into phenotypes based on histological evaluation of fixed tissues. The following samples were selected: normal/intact spermatogenesis from OA patients (normal spermatogenesis; NSP, $n = 3$; clinical score 10), biopsies with a phenotype of spermatid arrest from NOA patients (SDA, $n = 4$, four biopsies from 3 patients were used; contain somatic cells and germ cells up to round spermatids, clinical score 4–5), and biopsies with a phenotype of Sertoli cell only (SCO, $n = 3$; complete absence of germ cells, clinical score 0). Total RNA was extracted using the PeqGOLD total

RNA kit following the manufacturer's protocol (PeqGOLD total RNA kit, VWR Life Sciences). The quality and quantity of RNA were determined by Nanodrop (Eppendorf photometer). The sample quality and purity were further assessed by Labchip GX Touch (Perkin Elmer). Three samples that were outliers in PCA plots (Appendix Figure A1) were excluded from further downstream bioinformatical analysis. Thus, the RNA-Seq data analysis was limited to $n=3$ NSP, $n=2$ SDA, and $n=2$ SCO.

2.2 | RNA-seq and bioinformatical analysis

Libraries were prepared for RNA (barcoded separately and pooled) and run on Illumina NextSeq 500 sequencer at Max-Planck Institute, Bad Nauheim, Germany. Data generated were subjected to appropriate aligner platforms and software packages for analysis. RNA and library preparation integrity were verified with LabChip Gx Touch 24 (Perkin Elmer). Ten nanograms of total RNA were used as input for SMARTer® Stranded Total RNA-Seq Kit—Pico Input Mammalian (Takara Clontech). Sequencing was performed on the NextSeq500 instrument (Illumina) using v2 chemistry, resulting in an average of 37 M reads per library with a 1×75 bp single-end setup. The resulting raw reads were assessed for quality, adapter content, and duplication rates with FastQC.[28] Trimmomatic version 0.39 was employed to trim reads after a quality drop below a mean of Q20 in a window of 10 nucleotides.[29] Only reads between 30 and 150 nucleotides were cleared for further analyses. Trimmed and filtered reads were aligned versus the Ensembl human genome version hg38 (GRCh38.27) using STAR 2.6.1d with the parameter "outFilterMismatchNoverLmax 0.1" to increase the maximum ratio of mismatches to mapped length to 10%.[30] The number of reads aligning to genes was counted with the feature Counts 1.6.5 tool from the Subread package.[31] Only reads mapping at least partially inside exons were admitted and aggregated per gene. Reads overlapping multiple genes or aligning to multiple regions were excluded. The Ensemble annotation was enriched with UniProt data (release 06.06.2014) based on Ensembl gene identifiers (Activities at the Universal Protein Resource [UniProt]). Differentially expressed genes (DEGs) between groups were defined as follows: base mean ≥5, Log_2foldchange (FC) ≤−0.585 or ≥0.585, FDR ≤0.05. DEGs were filtered for specific terms relevant to sperm function based on Gene Ontology (GO), UniProt protein, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway terms using a bioinformatics strategy explained in the results.

All downstream analyses were based on the DESeq normalized gene count matrix. Deseq2: *p* values were based on the Wald test that was corrected for multiple testing using the Benjamini-Hochberg method. The statistical significance was determined by FDR and *p* values of <0.05. The methodology for calculating the FC was as follows: DESeq2 = log2 of (DESeq-norm mean (CondA) + 1)/(DESeq-norm mean (CondB) + 1). Volcano and MA plots were produced to visualize DEG expression. A global clustering heatmap of samples was created based on the Euclidean distance of regularized log-transformed gene counts. Dimension reduction analyses (principal component analyses = PCA) were performed on regularized log-transformed counts using the R packages FactoMineR (FactoMineR: A Package for Multivariate Analysis). DEGs were analyzed for gene set overrepresentation analyses with KOBAS (KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases).

2.3 | Complementary DNA (cDNA) synthesis, reverse transcription polymerase chain reaction (RT-PCR)

Primers were designed for the selected genes using NCBI and OligoExplorer software (Appendix Figure A1). Reagents used for cDNA synthesis and RT-PCR (if not otherwise stated) are from Applied Biosystems by Thermo Fisher Scientific. cDNA was synthesized from all available RNA samples, and RT-PCR was performed for the 10 selected genes, as described earlier.[32] Briefly, extracted total RNA (the same sample cohort used for RNA-seq +2 added samples to maintain $n=3$ per group) was adjusted to the final concentration of 200 µg/mL, treated with 1 U/µL DNAse 1 (Roche), 1× DNAse buffer (Roche), 1 U/µL RNAse inhibitor (Invitrogen), and incubated at 37°C for 30 min to degrade genomic DNA. cDNA synthesis was performed using 1× Gold PCR buffer, 1 mM dNTPs, 5 mM $MgCl_2$, 25 µM random hexamer primers (Invitrogen), 1 U/µL multiscript reverse transcriptase, 1 U/µL RNAse inhibitor, 18 µL distilled water, and 9 µL of treated mRNA in a 60 µL reaction. The mix was incubated at 21°C for 8 min, 42°C for 15 min, heated at 99°C for 5 min, and stored at 4°C.

RT-PCR was optimized for each primer with 1× Gold PCR buffer, 1 mM $MgCl_2$, 200 µM nucleotide mix, 10 pmol of forward and reverse primers (Appendix Figure A1), 0.65 U of Gold Amplitaq enzyme in a 25 µL reaction under following conditions: 95°C for 10 min, 45 cycles of 94°C for 45 s (denaturation), 60°C for 45 s (annealing), 72°C for 45 s (extension), and final extension 72°C for 10 min. For *PDE4A* and *SLC8B1*, the annealing temperatures were 58°C. Distilled water was used as the negative control for each reaction. PCR products were visualized using agarose gel electrophoresis (2% agarose gels, VWR Life Sciences) with Gel Green nucleic acid stain (Biotium Inc., Fremont, USA). To check primer and product specificity, PCR products were purified and sequenced (Eurofins Genomics, Ebersberg, Germany).

2.4 | Quantification of gene expression by qRT-PCR

TaqMan probes were selected according to the NCBI gene accession numbers and purchased from Thermo Fisher (Appendix Figure A1). Three samples per group were used from previously synthesized cDNA. Briefly, qRT-PCR assay mix was prepared with 10 µL of 2X TaqMan gene expression master mix (Applied Biosystems by Thermo Fisher Scientific), 1 µL of 20× TaqMan gene expression assay, 8 µL of

sterile distilled water, and 1 μL of cDNA to a final volume of 20 μL. Triplicates were run for each gene and *RPS27* and *RPL19* were used as reference genes and distilled water as the negative control. Plates were run at 95°C for10 min enzyme activation and 40 cycles of 95°C for 15 s (denaturation) and 60°C for 1 min (annealing). Relative gene expressions as fold change were manually calculated in Excel using the $\Delta\Delta C_q$ method,[33] statistical significance analysis (One-way ANOVA), and graphs were plotted using the software GraphPad Prism (version 9.4.1).

2.5 | Immunohistochemistry (IHC)

IHC was performed to localize proteins of two genes of interest (*ORAI1* and *SPATA31E1*) in archived, Bouin's fixed-paraffin-embedded tissue sections (NSP $n=5$, SDA $n=5$, SCO $n=5$) This sample cohort comprised of samples from (a) the patients' samples used for RNA-seq and (b) other patients' archive materials to increase the sample number. IHC was performed using previously described methods.[32] Each slide contained a negative control tissue which excluded the primary antibody treatment and placenta tissue slides were included for each protein as a negative control. Briefly, 5-μm-thick paraffin sections were mounted on coated slides, de-waxed, and rehydrated using xylol and a decreasing alcohol series, respectively. Heat-mediated antigen retrieval was performed in Tris-EDTA buffer (pH 9, SPATA31E1) or citrate buffer (pH 6, ORAI1) for 20 min in a microwave oven. Sections were then incubated in 3% hydrogen peroxide (H_2O_2) in Tris-EDTA buffer to inhibit endogenous peroxidase activity, followed by incubation in 1.5% bovine serum albumin (ROTH GmbH) in Tris-EDTA buffer for 30 min to block nonspecific binding. Sections were incubated overnight at 4°C with primary antibodies (ORAI1 1:100, SPATA31E1 1:5000, Appendix Figure A1). The next day, sections were washed and incubated with secondary antibodies (Appendix Figure A1) at room temperature for 1 h. The staining reaction was visualized using NovaRed (VECTOR Laboratories). After counterstaining with hematoxylin (MORPHISTO Laborchemikalien & Histologie Service GmbH), sections were dehydrated and mounted using Kaiser's gelatin (Merck, Darmstadt, Germany). A Leica DM750 microscope was used to observe the samples.

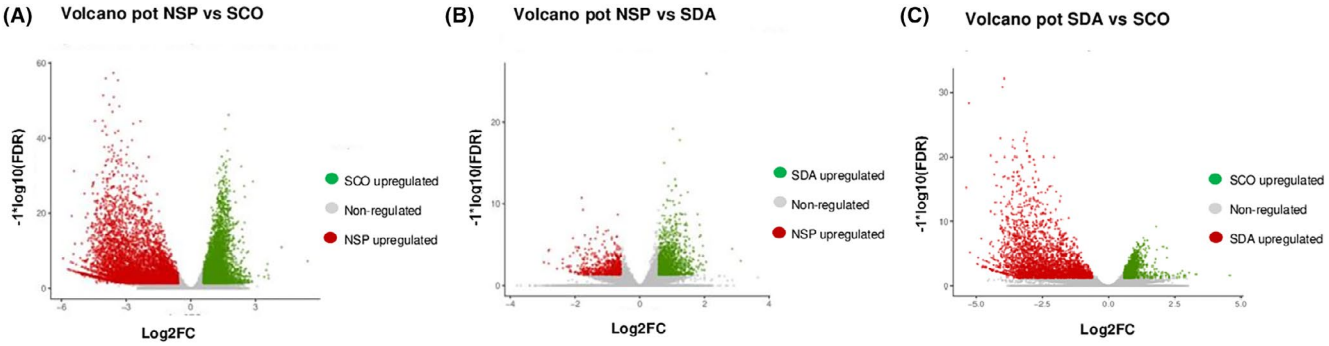
3 | RESULTS

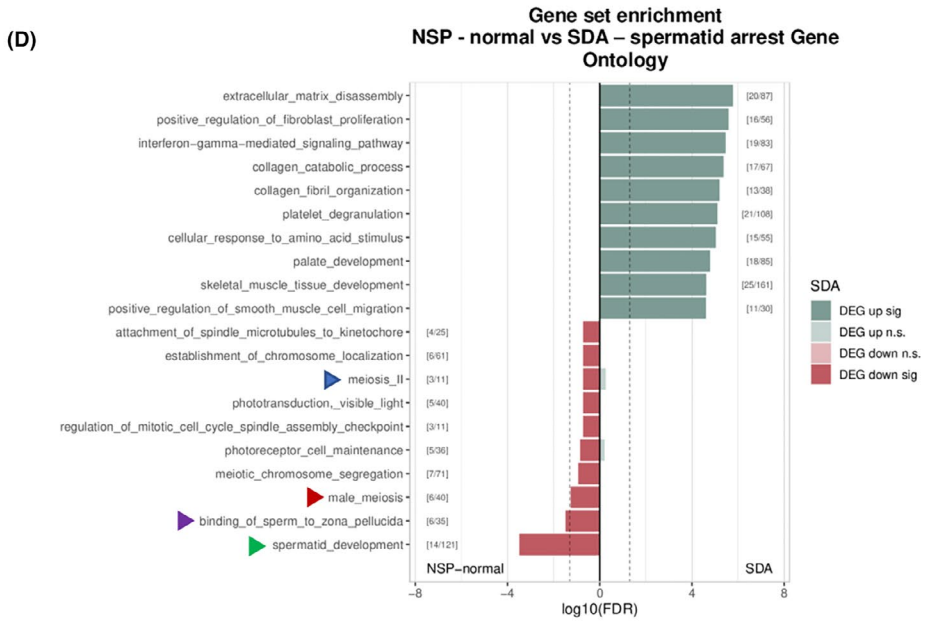
3.1 | RNA-sequencing and selection of candidate genes

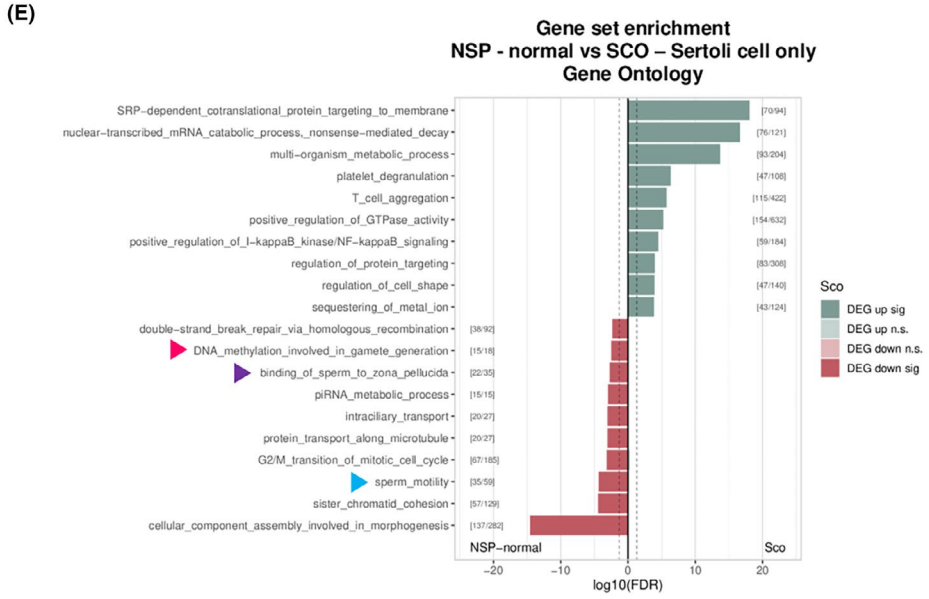
A total of 260,736814 reads were processed from all the selected samples subjected to RNA-sequencing and further analysis. After further filtering, 88,148401 reads were used. A total of over 49 million reads were generated in NSP samples, over 24 million reads in SDA samples, and over 13 million reads in SCO samples. DEGs were identified based on a mean minimum count ≥5, FDR ≤0.05, and $\log_2$FC of >0.585 or <−0.585. The number of significantly DEG between groups were as follows: NSP vs. SDA $n=1873$, SDA versus SCO $n=4017$, and NSP versus SCO $n=10\,253$ (Figures 1A–C and 2A, Appendix S1). The complete RNA-Seq raw data set has been submitted to GEO database (GEO accession GSE224929) (Appendix S1). The filtered RNA-Seq data set and the full list of DEGs between groups are available in Appendix Figure A1.

An important consideration regarding the analysis of these samples is that DEGs will be altered between the testis biopsies with different phenotypes because the cellular content differs. This fact allowed the design of a bioinformatic strategy to select genes that were likely to be predominantly expressed in germ cells compared to testicular somatic cells. The SDA samples lack elongated spermatids and therefore genes predominantly expressed during spermiogenesis should be downregulated compared to NSP biopsies. Consistent with this, gene set enrichment analysis of Gene Ontology (GO) terms comparing SDA to NSP showed that downregulated genes in SDA were significantly associated with terms relevant to spermiogenesis including fertilization and spermatid development (Figure 1D). Conversely, SCO biopsies will show a relative increase in the expression of somatic cell genes compared to NSP and SDA because they lack germ cells and thus RNA from somatic cells will be overrepresented and RNA from germ cells will be highly downregulated. Consistent with these expected changes, we noted the expression of the well-known spermatid-specific genes such as *TNP2* were downregulated in SDA versus NSP (*TNP2*: SDA vs. NSP $\text{Log}_2\text{FC}=-1.12$) and was further downregulated in SCO versus NSP (*TNP2*: SCO vs. NSP

FIGURE 1 Representation of RNA-Seq data. (A) The volcano plot showis upregulated genes of NSP and SCO. Each dot corresponds to a gene. Red dots are genes upregulated in the NSP group, and green dots are upregulated genes in the SCO group. (B) The volcano plot shows upregulated genes of NSP and SDA, where red dots correspond to genes upregulated in the NSP group, and green dots are genes upregulated in the SDA group. (C) The volcano plot shows upregulated genes of SDA and SCO, where red dots are genes upregulated in the SDA group, and green dots are genes upregulated in the SCO group. Volcano plots were made against Log2FC versus −1×log10(FDR) with statistical regulation: Counts >5, log2FC <> + −0.585, FDR ≤0.05. (D) The gene set enrichment plot of NSP vs. SDA shows the top 10 up- or downregulated genes clustered under relative gene ontology (GO) terms. A number of genes were downregulated in the spermatid arrest group compared to normal spermatogenesis which is categorized under meiosis II (3/11) highlighted with a blue arrow, male meiosis (6/40, highlighted with a red arrow), binding of sperm to zona pellucida (6/35, highlighted with a purple arrow), and spermatid development (14/121, highlighted with a green arrow). (E) The gene set enrichment plot of NSP versus SCO showing the top 10 up- or downregulated genes clustered under relative gene ontology (GO) terms. A number of genes were downregulated in the Sertoli cell-only group compared to normal spermatogenesis which is categorized under DNA methylation in gamete formation (15/18, highlighted with a pink arrow), binding of sperm to zona pellucida (22/35, highlighted with a purple arrow) and sperm motility (35/59, highlighted with light blue arrow). FDR, false discovery rate; log2FC, the fold change; NSP, normal spermatogenesis; SCO, Sertoli cell only; SDA, spermatid arrest.







log2FC = −2.48). Also, as expected, the Sertoli cell-specific gene *SOX9* was upregulated in SCO vs. NSP biopsies (*SOX9*:SCO vs. NSP – log2FC = +1.51). These changes were apparent in all SDA and SCO samples, confirming that the cellular changes observed in the histological sections were also apparent in the frozen samples used for mRNA analysis.

Gene set enrichment analysis of Gene Ontology comparing SDA versus NSP showed that there were a significant number of genes downregulated in SDA, which were associated with the terms related to male germ cell development, including "meiosis II" (3/11 genes), "male meiosis" (6/40 genes), "binding of sperm to zona pellucida" (6/35 genes), and "spermatid development" (14/121 genes)

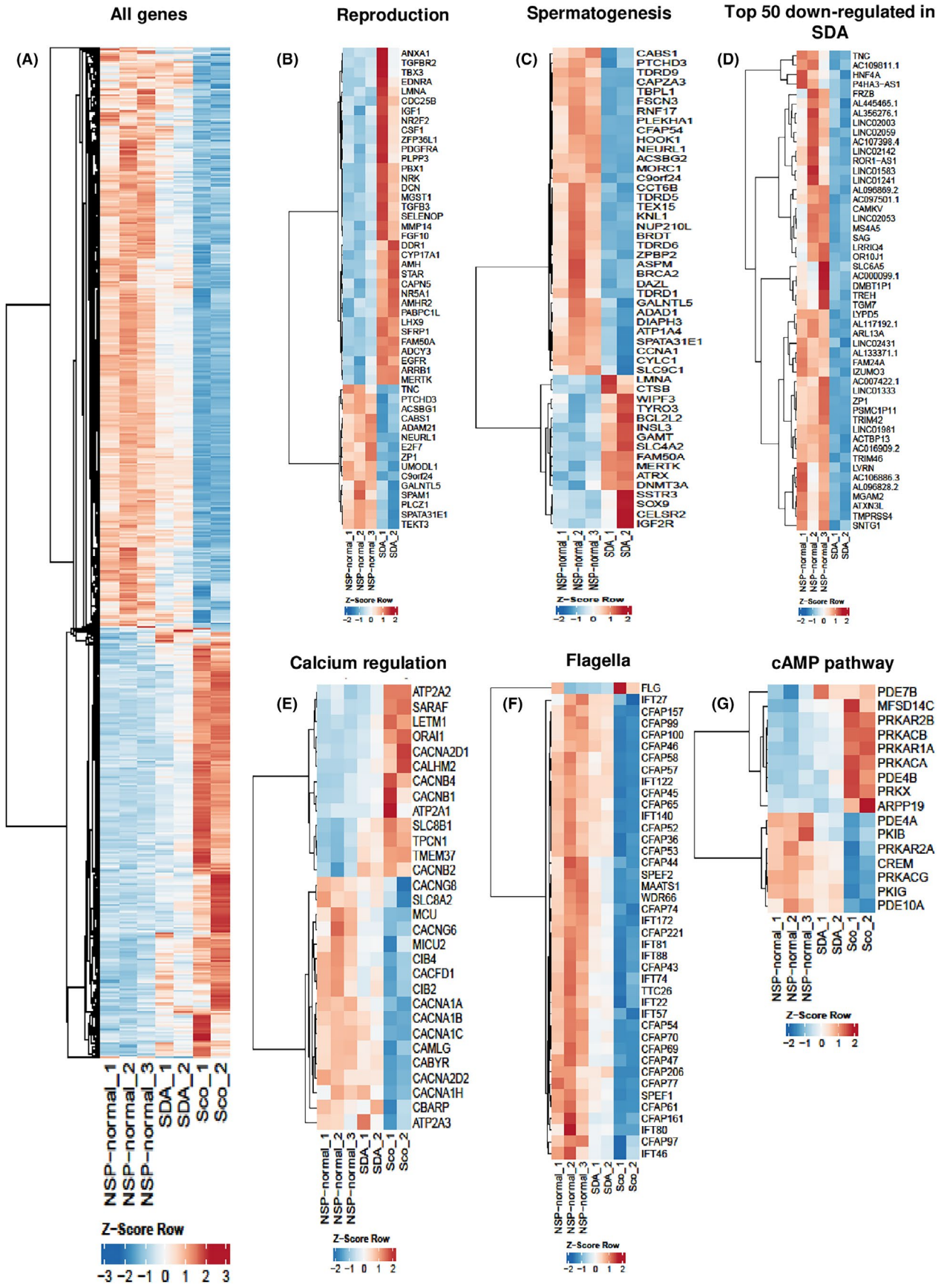


FIGURE 2 Heat maps of DEGs in the RNA-Seq. (A) Heat map of all DEGs. (B) DEGs annotated as involved in reproduction according to GO/KEGG terms. (C) DEGs annotated as involved in spermatogenesis according to GO/KEGG terms. (D) The top 50 most downregulated genes in SDA compared to NSP. (E) DEGs annotated as involved in calcium pathway regulation according to GO/KEGG terms. (F) DEGs annotated as associated with flagella according to GO/KEGG terms. (G) DEGs annotated as involved in cAMP pathway regulation according to GO/KEGG terms. Colors on the heat maps correspond to a *z*-score which is a measure of relative expression. Blue indicates low relative expression, and red color indicates high relative expression. DEG, differentially expressed gene; GO, gene ontology; NSP, normal spermatogenesis; SCO, Sertoli cell only; SDA, spermatid arrest.

TABLE 1 Genes selected and its foldchange values according to different groups.

	Gene	Filter condition—related GO /KEGG pathway/UniProt protein terms	Log2FC SDA/NSP	Log2FC SCO/SDA	Log2FC SCO/NSP	Conditions checked
01	*CFAP47*	Flagella proteins	−0.83	−2.54	−3.46	• Protein coding • UniProt gene/protein • Base mean and *p* values • KEGG pathway terms • GO terms • Gene information on NCBI and PubMed • Ensembl protein profiles • Information on the human protein atlas, GermOnline
02	*ORAI1*	Calcium regulation	0.79	1.05	1.78	
03	*CACNB2*		1.12	Not DE	1.09	
04	*SLC8B1*		1.14	Not DE	1.47	
05	*TMEM37*		1.53	Not DE	1.92	
06	*PDE4A*	cAMP pathway	−0.79	−0.1.41	−2.26	
07	*SPATA31E1*	Spermatogenesis	−1.65	−2.46	−4.19	
08	*SLC9C1*		−0.79	−2.34	−3.21	
09	*TEKT3*	Reproduction	−0.81	−3.39	−4.29	
10	*TNC*	Highly downregulated in SDA	−1.75	Not DE	−2.62	

Abbreviations: DE, differentially expressed; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; NCBI, National Center for Biotechnology Information; NSP, normal spermatogenesis; SCO, Sertoli cell only; SDA, Spermatid arrest.

(Figure 1D). The analysis of NSP vs. SCO also showed a significant number of genes downregulated in SCO were associated with GO terms associated with germ cell development including "DNA methylation in gamete formation" (15/18), "binding of sperm to zona pellucida" (22/35), and "sperm motility" (35/59) (Figure 1E).

A list of genes was selected under particular annotated terms related to the study objective and further refined after gathering more information from published data sets. As the aim of our study was to identify genes involved in human sperm development, our subsequent bioinformatic analyses primarily focused on analyzing genes that were predominantly expressed during spermiogenesis. To identify genes associated with sperm development and function, we also filtered DEGs for relevant terms in GO, UniProt, and KEGG pathways, particularly those annotated as being involved in sperm motility and morphology. First, we selected DEGs with functions known to be related to sperm morphology and motility, focusing on reproduction (Figure 2B), spermatogenesis (Figure 2C), calcium (Ca2+) pathway/regulation (Figure 2E), flagella (Figure 2F), and cyclic adenosine monophosphate (cAMP)-dependent protein kinase A (PKA) pathway (Figure 2G). The latter two pathways were chosen based on the fact that they are known to be essential for sperm motility.[34–36] Second, we focused on highly downregulated DEGs between NSP versus SDA (Figure 2D) because this identifies genes highly expressed in elongating and elongated spermatids.

We refined the list of DEGs using various published data sets to focus on genes that (a) are likely to be involved in human spermatogenesis due to their expression pattern, (b) have functions relevant to sperm development, morphology, or motility, and (c) are highly enriched or specifically expressed in the human testis. DEGs identified in functional categories were analyzed in Ensembl, Human Protein Atlas, GermOnline, UniProt, and PubMed. Genes were further refined based on limited information on a potential role in male fertility in PubMed, to facilitate the selection of genes with novel roles in spermatogenesis (Appendix Figure A1).

Using this approach, we selected six genes with strong evidence to support a novel role in human sperm development, morphology, and/or motility. These genes were downregulated in SDA vs. NSP (Table 1). *CFAP47* (Cilia- and Flagella-Associated Protein 4; log2FC −0.83 in SDA) was selected from the gene cluster annotated with Uniport term, flagella proteins. *PDE4A* (Phosphodiesterase 4A; log2FC −0.79 in SDA) was selected from the genes clustered under cAMP pathway, *SPATA31E1* (SPATA31 subfamily E member 1; log2FC −1.65), *SLC9C1* (Solute Carrier Family 9 Member C1; log2FC −0.79), and *TEKT3* (Tektin 3; log2FC −0.81) were selected from the spermatogenesis gene cluster and *TNC* (Tenascin C; log2Fc −1.75) was selected from the highly downregulated gene cluster in SDA versus NSP. Four genes, *ORAI1* (ORAI Calcium Release-Activated Calcium Modulator 1), *CACNB2* (Calcium Voltage-Gated Channel Auxiliary Subunit Beta 2), *SLC8B1* (Solute Carrier Family 8 Member B1), and *TMEM37* (Transmembrane Protein 37) were selected based on their roles in calcium regulation, which is known to be important for multiple aspects of spermatogenesis as well as for sperm function.

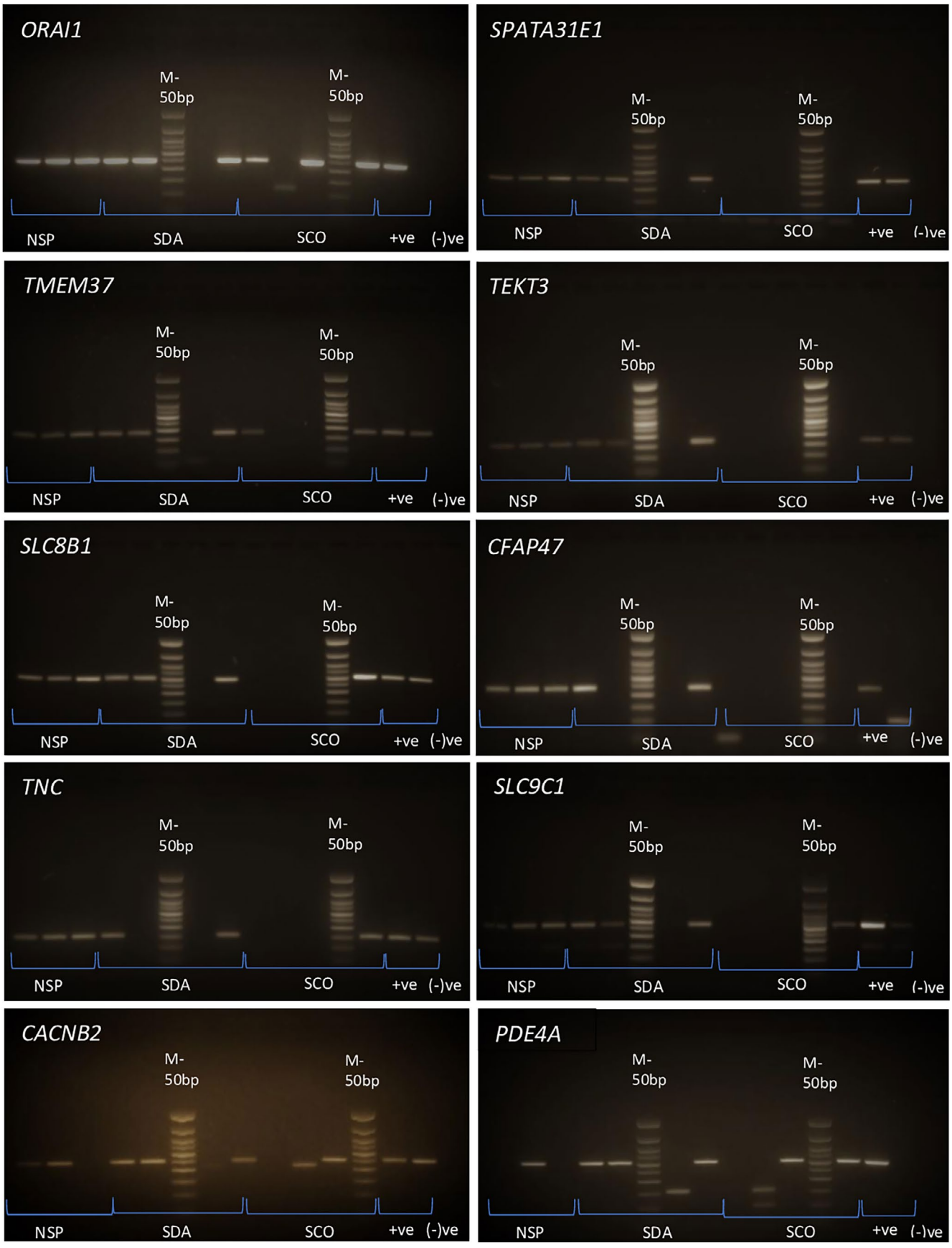


FIGURE 3 RT-PCR results of selected 10 genes. RT-PCR was conducted for 10 genes selected and agarose gel electrophoresis was performed to visualize DNA bands produced. NSP ($n=3$), SDA ($n=4$), and SCO ($n=4$) samples were used with two known testis RNA samples as positive controls and distilled water as the negative control. All the negative controls were free from DNA (last well of each gel) and at least one positive control produced the respective DNA band for genes: *ORAI1*–230bp, *SPATA31E1*—239bp, *TMEM37*—174bp, *TEKT3*—161bp, *SLC8B1*—269bp, *CFAP47*—244bp, *TNC*—151bp, *SLC9C1*—259bp, *CACNB2*—241bp, *PDE4A*—281bp. (+)ve, positive control; (−)ve, negative control; M, molecular marker (50bp); NSP, normal spermatogenesis; SCO, Sertoli cell only; SDA, spermatid arrest.

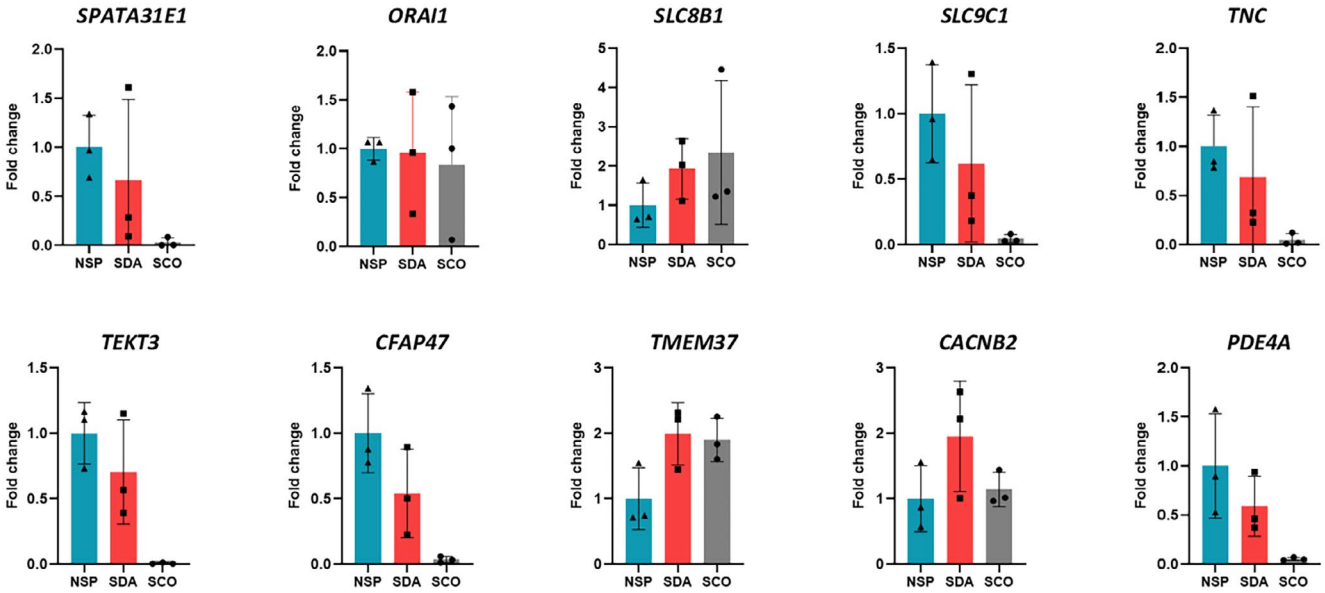


FIGURE 4 Relative gene expression of each candidate gene selected. qRT-PCR was performed to determine the changes in mRNA expression in selected genes in biopsies with normal vs. impaired spermatogenesis. RNA originated from testis samples of normal spermatogenesis (NSP, $n=3$), spermatid arrest (SDA, $n=3$), and Sertoli cell only (SCO, $n=3$) (each sample was assessed in triplicate). Two housekeeping genes were used for each sample. All the negative controls showed no signal. Each dot represents the average of triplicate measures for a single patient sample. The average value for NSP is set at 1, whereas SDA and SCO values represent foldchange from this. Data are shown as mean ± SD, $n=3$ samples per group. NSP, normal spermatogenesis; SCO, Sertoli cell only; SD, standard deviation; SDA, spermatid arrest.

3.2 | Analysis of candidate gene expression in biopsies from men with intact vs. impaired spermatogenesis

mRNA transcripts for selected 10 genes were detected in all NSP samples by RT-PCR. Genes *SPATA31E1*, *TEKT3*, and *CFAP47* were detected in NSP and SDA patients but not in SCO, confirming expression in germ cells but not in somatic cells (Figure 3). We further analyzed quantitative gene expression changes by qRT-PCR. *SPATA31E1*, *ORAI1*, *SLC9C*, *TNC*, *TEKT3*, *CFAP47*, and *PDE4A* were reduced in SDA compared to NSP biopsies, and the levels were further reduced in SCO biopsies confirming the enrichment of these genes in germ cells. *SLC8B1*, *TMEM37*, and *CACNB2* were selected based on likely functional roles in calcium regulation which is important for sperm function, however, their expression was increased in SDA and SCO compared to NSP, suggesting that they are not enriched in germ cells (Figure 4). The coefficient of variation for the two housekeeping genes were as follows: NSP, 3.47%, 4.03%; SDA, 2.20%, 3.27%; SCO, 8.44%, 12.08%.

3.3 | Protein localization of SPATA31E1 and ORAI1 by IHC

Due to practical considerations, we chose only two genes for further analysis by immunohistochemistry to verify protein localization in human germ cells. The gene *SPATA31E1* was selected to determine its protein localization because this is a novel gene that is potentially germ cell-specific as predicted by RNA-Seq, and there was limited information on its molecular function. The gene *ORAI1* was selected as it is related to the calcium pathway and had previously been shown to be related to sperm motility in mice, however, its localization in the human testis has not previously been determined. In NSP, SPATA31E1 was not detected in spermatogonia and early spermatocytes but in nuclei of pachytene spermatocytes from approximately stage IV of spermatogenesis.[37] This nuclear staining remained detectable in spermatids until step 4–5. Cytoplasmic staining of round spermatids was apparent in some sections where the staining intensity was high. Elongating spermatid cytoplasm was generally negative (Figure 5A). Both lightly stained and unstained

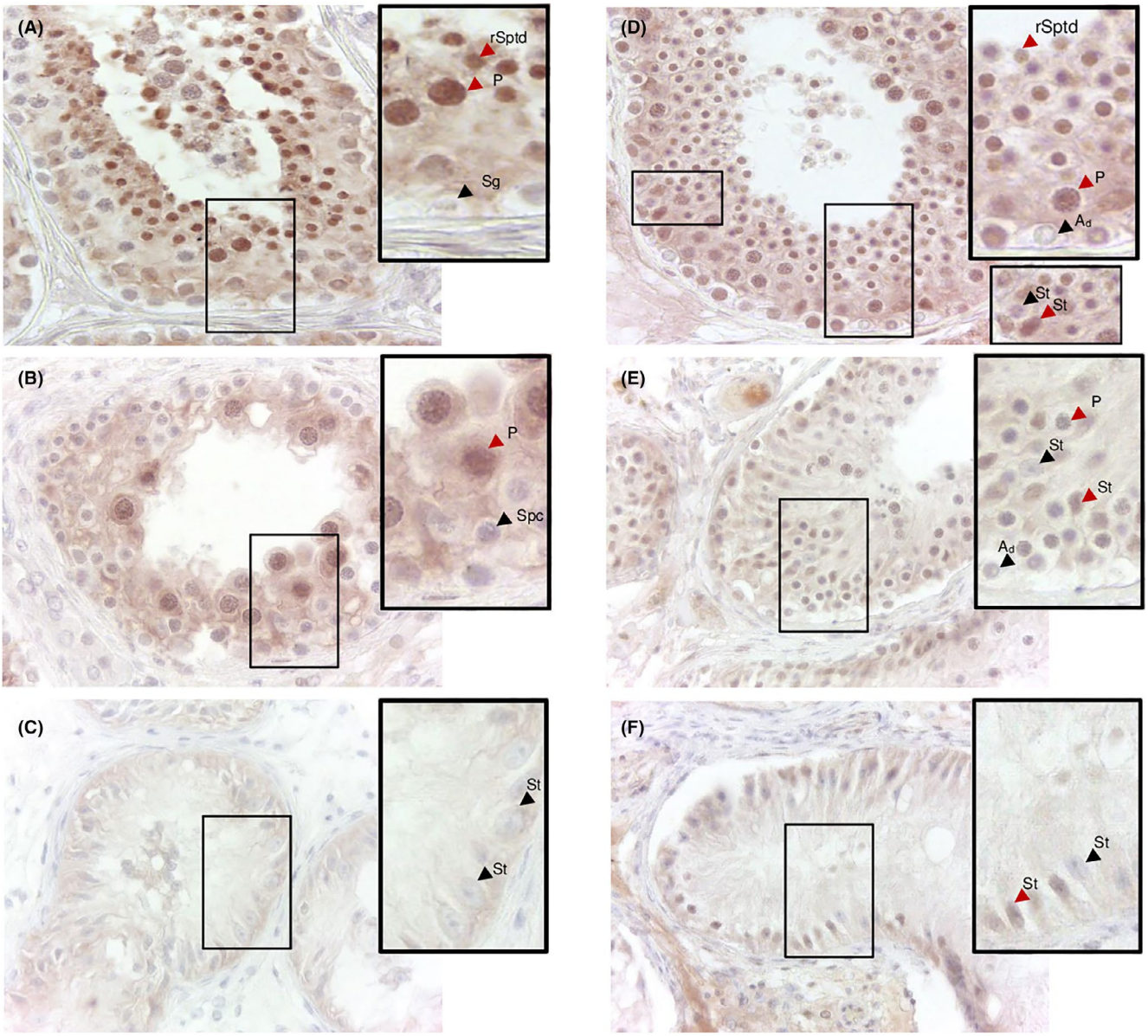


FIGURE 5 Immunohistochemistry of SPATA31E1 and ORAI1 in human testis biopsies. (A–C) SPATA31E1, (D–F) ORAI1. Immunostaining was detected using NovaRed, and sections were counterstained with hematoxylin. (A) SPATA31E1 staining in NSP. *Inset*: Black arrowhead indicates immuno-negative spermatogonia (Sg) and red arrowheads indicate nuclear staining in pachytene spermatocytes (P) and round spermatids (rSptd). (B) SPATA31E1 staining in SDA. *Inset*: Black arrowhead indicates immuno-negative early spermatocytes (Spc) and red arrowhead nuclear staining in a late pachytene spermatocyte (P). (C) SPATA31E1 staining in SCO. *Inset*: Black arrows indicate immuno-negative Sertoli cell nuclei (St). (D) ORAI1 staining in NSP. *Inset*: Red arrows indicate nuclear staining of pachytene spermatocytes (P), round spermatids (rSptd), and Sertoli cell (St), and black arrows indicate immuno-negative A_{dark} spermatogonia (A_d) and some immuno-negative Sertoli cell nuclei (St). (E) ORAI1 staining in SDA. *Inset*: Red arrows indicate stained pachytenes (P) and Sertoli cell nuclei (St), and black arrows indicate immuno-negative A_{dark} spermatogonia (A_d) and Sertoli cells nuclei (St). (F) ORAI1 staining in SCO. *Inset*: Red arrow indicates immunostaining of Sertoli cell nucleus (St), and black arrow indicates immuno-negative Sertoli cell nucleus (St).

Sertoli cell nuclei were observed in NSP; this seems likely due to cyclic variation in Sertoli cells (Figure 5A and Appendix Figure A1). Basal Sertoli cell cytoplasm staining could also be observed in tubules with the greatest staining intensity in each patient (Appendix Figure A1). Within these tubules, the staining seemed to vary according to the stage of the spermatogenic cycle. In the tubules/tissues with lower staining intensity, Sertoli cell cytoplasmic staining was not evident (Appendix Figure A1).

As in NSP, SDA biopsies showed SPATA31E1 staining in nuclei of late pachytene spermatocytes and early round spermatids with mature spermatids absent in these biopsies. Both stained and unstained Sertoli cell nuclei were evident as observed in NSP biopsies (Figure 5B). Some tubules also showed evidence of Sertoli cell cytoplasmic staining (Appendix Figure A.2). In SCO, we noted greater heterogeneity between patients, with some patients showing very little immunostaining ($n = 2$) and others showing clear Sertoli cell

nuclear and light cytoplasmic immunostaining ($n = 3$) (Figure 5C, Appendix Figure A.3). Interestingly, both stained and unstained Sertoli cell nuclei were observed in the same tubules of some patients, suggesting that the cyclic variation in Sertoli cell nuclear SPATA31E1 immunostaining observed in NSP and SDA biopsies is preserved in SCO patients (Figure 5C, Appendix Figure A.3). The specificity of immunostaining was confirmed by obtaining unstained negative control testis tissues in each slide and unstained placenta tissue slides.

For ORAI1, staining in NSP samples was detected in the nuclei of all germ cells except for A_{dark} spermatogonia. Nuclear staining was observed in Sertoli cells, however, some Sertoli cell nuclei were unstained suggesting nuclear localization could be stage-specific (Figure 5D), similar to SPATA31E1. In SDA samples, ORAI1 was also observed in nuclei of all germ cells present (up to round spermatids) except A_{dark} spermatogonia. Sertoli cell nuclei also showed immunostaining, with some immuno-negative nuclei (Figure 5E) suggesting that stage-specificity of Sertoli cell staining was preserved in SDA biopsies. In SCO samples, both stained and unstained Sertoli cell nuclei were observed (Figure 5F). No germ cell cytoplasm staining was observed for ORAI1 in all three phenotypes analyzed. The specificity of immunostaining was confirmed by obtaining unstained negative control testis tissues in each slide and unstained placenta tissue slides.

4 | DISCUSSION

Next-generation sequencing and RNA-Seq have been applied for the identification of altered gene expression in abnormal testicular histology and function and to identify novel genes involved in disease development.[38,39] The application of RNA-Seq technology in male reproductive research has contributed to the understanding of RNA complexity and molecular mechanisms of impaired male fertility.[40] Spermatogenesis is believed to be regulated by up to 2000 genes, of which approximately 700–900 genes are exclusively expressed in the male germ line.[3,41,42] Since there are so many genes likely involved in sperm production, many genes causing male infertility are probably yet to be identified.[18]

The current study focused on discovering genes that could be linked to human sperm development, morphology, and/or motility. In an untargeted approach, we used bulk RNA-Seq analysis to quantify gene expression levels in testis biopsies from men with obstructive azoospermia, with intact spermatogenesis (normal spermatogenesis, NSP), and non-obstructive azoospermia due to spermatid arrest (SDA) or germ cell arrest (Sertoli cell-only, SCO). By definitive histological characterization of tissue specimens, the cellular components of the starting material are precisely known in contrast to tissue samples used without prior histological assessment. Additionally, knowledge on cellular composition is essential for the interpretation of DEGs between the sample groups. Many genes will be differentially expressed simply due to the different cell compositions of biopsies; for example, in SDA, where spermatids are lost, a marked downregulation of genes enriched in spermatids will result. Biopsies with SCO compared to NSP will show an upregulation of somatic cell genes and a downregulation of germ cell-specific genes. Careful selection of genes with particular expression patterns and functional annotations can be used to identify novel genes and corresponding proteins that could play important roles in human spermatogenesis. By using bulk RNA-Seq on such well-defined samples with known cell populations and by comparing gene expressions between such sample groups, we were able to understand gene expression unique to the respective phenotype, including not only germ cells but also somatic cells present. By this, we were able to understand the total gene expressional changes in the testicular milieu. We particularly focused on the comparison of NSP (i.e., intact spermatogenesis with all germ cell stages, including elongated spermatids or "testicular sperm" present) and SDA (i.e., spermatid arrest with only round spermatids present suggesting a defective spermiogenesis) samples.

To further refine target gene selection, we classified the list of differentially expressed genes (DEGs) based on their functions, expression patterns, and the published literature. We particularly focused on genes associated with calcium and cAMP pathways since these are important for sperm motility.[34–36,43] Abnormalities in the flagella result in asthenozoospermia which impairs male fertility[18] and thus we also selected genes likely associated with flagella function based on gene annotations. We also selected genes annotated as having an involvement in spermatogenesis and/or spermiogenesis and focused on genes that were highly downregulated in biopsies showing spermatid arrest biopsies compared to NSP, as these were likely to be predominantly expressed in elongating or elongated spermatids. Finally, we selected genes that had little published information but were enriched in germ cells. Using this approach, we narrowed down our selection to 10 genes with putative roles in sperm development and/or function. The data suggested that *SPATA31E1*, *TEKT3*, *SLC9C1*, *PDE4A*, *CFAP47*, and *TNC* are predominantly expressed in human germ cells. Thus, we have identified several genes that may play previously unrecognized roles in human sperm production and function. Immunohistochemical analyses of two proteins, that is, SPATA31E1 and ORAI1, confirmed the localization to human germ cells, which have not previously been described.

We selected *CFAP47* based on annotations and expression patterns suggesting a role in sperm flagella function. *CFAP47* mRNA was reduced in SDA and SCO testis samples by qRT-PCR, suggesting it is expressed in postmeiotic germ cells. Analysis of *CFAP47* expression patterns in the Human Protein Atlas (HPA) confirmed *CFAP47* is expressed in early and late spermatids[44] and annotations in UniProt suggest that the protein is highly expressed in spermatozoa.[45] A recent study that was published while these experiments were being performed, revealed that deleterious X-linked variants of *CFAP47* were associated with asthenoteratozoospermia in a whole-exome sequencing study,[45] thus providing confidence in our approach to the selection of genes that are likely to have a functional role in spermiogenesis.

PDE4A was selected because of its likely involvement in the regulation of cAMP, which is known to be involved in sperm motility,[35] yet PDE4A's role in human spermatozoa is not clear. Our RNA-Seq and qRT-PCR data suggest that *PDE4A* is enriched in human spermatids, and this proposition is confirmed by *PDE4A* tissue expression data in HPA. This protein has been shown to be enriched in rat spermatids[46] and thus it is likely to be conserved in spermatids in both rodents and humans.

TEKT3, *TNC*, and *SLC9C1* were selected based on their annotations suggesting an involvement in reproduction and spermatogenesis. Our results show that these genes are detected in normal spermatogenesis but are reduced in biopsies with spermatid arrest and SCO, suggesting they are predominantly expressed in human germ cells, particularly in early/late spermatids. Consistent with our findings, *TEKT3* mRNA is enriched in human spermatids according to the HPA,[44] and its protein is detected in sperm flagella according to UniProt. *TEKT3* has shown an association with sperm motility in null mice by showing reduced progressive motility and structural defects in the sperm.[47] *TNC* mRNA is also enriched in human spermatids compared to other germ cells,[44] yet little is known of its role in these germ cells. According to the HPA, *SLC9C1* mRNA is highly enriched in human testis compared to other tissues, with high expression in spermatids,[44] suggesting its expression is mainly restricted to developing spermatids in humans. *SLC9C1* was recently published as a causative gene for asthenozoospermia,[48] again confirming that our approach to gene selection is able to reveal genes with functional significance in human spermatids. In summary, our RNA-Seq and bioinformatics approach proved to be useful in identifying proteins that are expressed in developing sperm and thus potentially play a role in sperm development or function.

We focused on *SPATA31E1* because little is known about this gene and the molecular function has not been identified. According to HPA,[49] SPATA31E1 is highly testis-specific, making it an attractive target for contraception. The role of *SPATA31E1* in male fertility is not well studied, other than a publication in stallions suggesting that SNPs in *SPATA31E1* may have a deleterious effect on fertility.[50] *SPATA31E1* was recently mentioned as a human spermatid-specific gene and a potential drug target.[39] We showed *SPATA31E1* is expressed in intact human spermatogenesis and is reduced in SDA and SCO biopsies. Tissue expression analyses in HPA suggest that *SPATA31E1* mRNA and protein are highly enriched in the testis, restricted to early and late spermatids, with some mRNA (but not protein) in the brain.[44] This expression pattern is consistent with our data showing it is reduced in testis biopsies from patients with spermatid arrest and shows very low expression in SCO. We used immunohistochemistry to show that SPATA31E1 is localized to germ cell nuclei from late pachytene spermatocytes (stages IV, V) suggesting this protein could play a role in the later stages of meiosis. It was also detected in early round spermatids and elongating spermatids suggesting a role during spermiogenesis. Further studies are needed to investigate the expression of SPATA31E1 in human spermatozoa.

SPATA31E1 was also detectable in human Sertoli cell nuclei, and we noted both immuno-positive and immuno-negative Sertoli cell nuclei in NSP and SDA, suggesting the protein expression in Sertoli cell nuclei depends on the stage of the spermatogenic cycle. We detected *SPATA31E1* mRNA in very low but detectable levels in SCO biopsies by RNA-Seq and qRT-PCR, and we observed some specific immunostaining in Sertoli cell nuclei that varied between cells, tubules, and between patients. Because of the helical nature of human spermatogenesis and the fact that multiple stages are observed in a single tubule cross-section, it was difficult for us to distinguish the precise stages at which Sertoli cell nuclear expression of SPATA31E1 varied. Intriguingly, we also observed apparent stage-specificity of SPATA31E1 nuclear (and to some degree, cytoplasmic) Sertoli cell localization in seminiferous tubules lacking germ cells in SCO biopsies. Stage-specific Sertoli cell protein expression can be observed in the mouse embryonic testis, even in the absence of germ cells, leading to speculation that fetal Sertoli cells, in the presence or absence of germ cells, show innate cyclic variation in their function that does not rely on the presence of postnatal germ cells.[51] Thus, it is tempting to speculate that the variation in SPATA31E1 nuclear localization in SCO reflects the preservation of human Sertoli cell cyclic activity in the absence of germ cells. We also noted heterogeneity between SCO patients in terms of whether SPATA31E1 was expressed in the nuclei or not. This could potentially reflect Sertoli cell maturity or function that is influenced by the genetic, epigenetic, and/or hormonal status of the patient. Further studies should investigate whether SPATA31E1 Sertoli cell nuclear staining is related to Sertoli stage, maturity, and functional status.

We analyzed the protein expression of ORAI1 in the testis because its localization has not been described previously. ORAI1 is a calcium-selective ion channel-calcium transporter class and is a membrane calcium subunit that is activated when calcium stores are depleted.[52–54] ORAI1 was expressed in various organs including the testis.[49] Our immunohistochemical data demonstrate that ORAI1 was expressed in germ cells and Sertoli cells, suggesting multiple roles in spermatogenesis. It was also stage-specific in Sertoli cell nuclei in all samples, including SCO. *Orai1*−/− mice were sterile showing defects in elongating spermatid development with markedly reduced motility and low sperm count, suggesting a functional role in germ cell development.[55] Orai1 has also been suggested to be involved in the turnover of Sertoli cell intercellular junctions in the rat testis.[56]

In conclusion, we analyzed transcriptional changes in human testis biopsies with well-characterized phenotypes, comparing gene expression in biopsies from testes with normal spermatogenesis, spermatid arrest, and those lacking germ cells. We used various bioinformatic approaches to select 10 differentially expressed genes that were likely to be testis and/or spermatid specific, and thus potentially play a functional role, in developing sperm. Of these genes, *SPATA31E1*, *TEKT3*, *SLC9C1*, *PDE4A*, *CFAP47*, and *TNC* showed excellent evidence of enrichment in human germ cells in the testis. We also identified ORAI1 as a protein with potential roles in Sertoli and germ cell function, and SPATA31E1 as a testis-specific protein expressed in spermatocytes, spermatids, and Sertoli cells that could be further explored for its utility as a contraceptive target due to its restricted expression. Our data provide proteins that could be explored

for their utility as targets for contraception, predictive markers for successful surgical sperm retrieval, or involvement in human male infertility. Further investigations are required with larger sample sizes from independent sample groups to further validate these findings.

ACKNOWLEDGMENTS

We thank all the patients for the consent given to utilize their samples for the research. We thank all the technicians especially Ms. Kerstin Wilhelm and Ms. Tania Bloch of Clinic of Urology, Pediatric Urology and Andrology, Justus-Liebig University of Giessen, Germany. Ms. Alexandra Hax of the Institute for Veterinary Anatomy, Histology and Embryology, Justus-Liebig University of Giessen, Germany, and Ms. Sameena Mahmood from Max Planck Institute, Bad Nauheim, Germany.

FUNDING INFORMATION

This study was funded by the Deutsche Forschungsgemeinschaft (DFG)—IRTG *DFG GRK 1871/1*.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ETHICS STATEMENT

Ethical approval was obtained from the Ethic's Committee of the Medical Faculty, Justus Liebig University of Giessen (reference numbers 26/11 and 152/16).

HUMAN RIGHTS STATEMENT AND INFORMED CONSENT

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and its later amendments. Informed consent was obtained from all patients to be included in the study.

ORCID

Shashika D. Kothalawala https://orcid.org/0000-0001-9783-3647
Daniela Fietz https://orcid.org/0000-0002-7945-1810

REFERENCES

1. Jungwirth A, Giwercman A, Tournaye H, Diemer T, Kopa Z, Dohle G, et al. European Association of Urology guidelines on male infertility: the 2012 update. Eur Urol. 2012;62(2):324–32.
2. Minhas S, Bettocchi C, Boeri L, Capogrosso P, Carvalho J, Cilesiz NC, et al. European Association of Urology guidelines on male sexual and reproductive health: 2021 update on male infertility. Eur Urol. 2021;80(5):603–20.
3. Tüttelmann F, Ruckert C, Ropke A. Disorders of spermatogenesis: perspectives for novel genetic diagnostics after 20 years of unchanged routine. Med Genet. 2018;30(1):12–20.
4. Agarwal A, Mulgund A, Hamada A, Chyatte MR. A unique view on male infertility around the globe. Reprod Biol Endocrinol. 2015;13:37.
5. Lin YH, Lin YM, Teng YN, Hsieh TY, Lin YS, Kuo PL. Identification of ten novel genes involved in human spermatogenesis by microarray analysis of testicular tissue. Fertil Steril. 2006;86(6):1650–8.
6. Panner Selvam MK, Agarwal A, Peter PN, Baskaran S, Bendou H. Sperm proteome analysis and identification of fertility-associated biomarkers in unexplained male infertility. Gen. 2019;10(7):522.
7. World Health Organization. WHO laboratory manual for the examination and processing of human semen. 6th ed. Geneva, Switzerland: WHO; 2021 https://www.who.int/publications/i/item/9789240030787
8. Ray PF, Toure A, Metzler-Guillemain C, Mitchell MJ, Arnoult C, Coutton C. Genetic abnormalities leading to qualitative defects of sperm morphology or function. Clin Genet. 2017;91(2):217–32.
9. Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, et al. Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. Cell Rep. 2013;3(6):2179–90.
10. Tüttelmann F, Rajpert-De Meyts E, Eberhard Nieschlag E, Simoni M. Gene polymorphisms and male infertility—a meta analysis and mini review. Reprod Biomed Online. 2007;15:643–58.
11. Houston BJ, Riera-Escamilla A, Wyrwoll MJ, Salas-Huetos A, Xavier MJ, Nagirnaja L, et al. A systematic review of the validated monogenic causes of human male infertility: 2020 update and a discussion of emerging gene-disease relationships. Hum Reprod Update. 2021;28(1):15–29.
12. Vogt PH. Molecular genetic of human male infertility: from genes to new therapeutic perspectives. Curr Pharm Des. 2004;10:471–500.
13. McLachlan RI, O'Bryan MK. Clinical review: State of the art for genetic testing of infertile men. J Clin Endocrinol Metab. 2010;95(3):1013–24.
14. Krausz C, Riera-Escamilla A, Moreno-Mendoza D, Holleman K, Cioppi F, Algaba F, et al. Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. Genet Med. 2020;22(12):1956–66.
15. O'Donnell L, Rebourcet D, Dagley LF, Sgaier R, Infusini G, O'Shaughnessy PJ, et al. Sperm proteins and cancer-testis antigens are released by the seminiferous tubules in mice and men. FASEB J. 2021;35(3):e21397.
16. Fietz D, Sgaier R, O'Donnell L, Stanton PG, Dagley LF, Webb AI, et al. Proteomic biomarkers in seminal plasma as predictors of reproductive potential in azoospermic men. Front Endocrinol (Lausanne). 2024;15:1327800.
17. O'Flynn O'Brien KL, Varghese AC, Agarwal A. The genetic causes of male factor infertility: a review. Fertil Steril. 2010;93(1):1–12.
18. Coutton C, Escoffier J, Martinez G, Arnoult C, Ray PF. Teratozoospermia: spotlight on the main genetic actors in the human. Hum Reprod Update. 2015;21(4):455–85.
19. Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, Chung SS, Demougin P, Gattiker A, et al. The conserved transcriptome in human and rodentmale gametogenesis. PNAS. 2007;104:8346–51.
20. Long JE, Lee MS, Blithe DL. Update on novel hormonal and nonhormonal male contraceptive development. J Clin Endocrinol Metab. 2021;106(6):e2381–e2392.
21. Shang E, Nickerson HD, Wen D, Wang X, Wolgemuth DJ. The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. Development. 2007;134(19):3507–15.
22. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. Genome Biol. 2016;17:13.
23. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. Nat Rev Genet. 2019;20(11):631–56.

24. Jeon H, Xie J, Jeon Y, Jung KJ, Gupta A, Chang W, et al. Statistical power analysis for designing bulk, single-cell, and spatial transcriptomics experiments: Review, tutorial, and perspectives. Biomol Ther. 2023;13(2):221.
25. Dong F, Ping P, Ma Y, Chen XF. Application of single-cell RNA sequencing on human testicular samples: a comprehensive review. Int J Biol Sci. 2023;19(7):2167–97.
26. Jungwirth A, Diemer T, Kopa Z, Krausz C, Minhas S, Tournaye H, et al. EAU Guidelines on Male Infertility. 2018. http://uroweb.org/guidelines/compilations-of-all-guidelines/.
27. Fietz D, Kliesch S. Testicular biopsy and histology. In: Nieschlag E, Behre HM, Nieschlag S, editors. Andrology: Male Reproductive Health and Dysfunction. 4th ed. Berlin: Springer Verlag; 2023. p. 181–96.
28. Andrews S. A quality control tool for high throughput sequence data Babraham Bioinformatics. 2010 http://www.bioinformatics.babraham.ac.uk/projects/fastqc
29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
30. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.
31. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
32. Pleuger C, Fietz D, Hartmann K, Schuppe HC, Weidner W, Kliesch S, et al. Expression of ciliated bronchial epithelium 1 during human spermatogenesis. Fertil Steril. 2017;108(1):47–54.
33. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001;29:2002–7.
34. Suarez SS, Varosi SM, Dai X. Intracellular calcium increases with hyperactivation in intact. Proc Natl Acad Sci USA. 1993;90:4660–4.
35. Pereira R, Sa R, Barros A, Sousa M. Major regulatory mechanisms involved in sperm motility. Asian J Androl. 2017;19(1):5–14.
36. Mata-Martinez E, Sanchez-Cardenas C, Chavez JC, Guerrero A, Trevino CL, Corkidi G, et al. Role of calcium oscillations in sperm physiology. Biosystems. 2021;209:104524.
37. Clermont Y. The cycle of the seminiferous epithelium in man. Am J Anat. 1963;112:35–51.
38. Dado-Senn B, Skibiel AL, Fabris TF, Zhang Y, Dahl GE, Penagaricano F, et al. RNA-seq reveals novel genes and pathways involved in bovine mammary involution during the dry period and under environmental heat stress. Sci Rep. 2018;8(1):11096.
39. Robertson MJ, Kent K, Tharp N, Nozawa K, Dean L, Mathew M, et al. Large-scale discovery of male reproductive tract-specific genes through analysis of RNA-seq datasets. BMC Biol. 2020;18(1):103.
40. Bianchi E, Boekelheide K, Sigman M, Braun JM, Eliot M, Hall SJ, et al. Spermatozoal large RNA content is associated with semen characteristics, sociodemographic and lifestyle factors. PLoS One. 2019;14(5):e0216584.
41. Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. Nat Med. 2008;14(11):1197–213.
42. Chalmel F, Lardenois A, Evrard B, Mathieu R, Feig C, Demougin P, et al. Global human tissue profiling and protein network analysis reveals distinct levels of transcriptional germline-specificity and identifies target genes for male infertility. Hum Reprod. 2012;27(11):3233–48.
43. Suarez SS. Control of hyperactivation in sperm. Hum Reprod Update. 2008;14(6):647–57.
44. Uhlén M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.
45. Liu C, Tu C, Wang L, Wu H, Houston BJ, Mastrorosa FK, et al. Deleterious variants in X-linked CFAP47 induce asthenoteratozoospermia and primary male infertility. Am J Hum Genet. 2021;108(2):309–23.
46. Salanova M, Chun SY, Iona S, Puri C, Stefanini M, Conti M. Type 4 cyclic adenosine monophosphate-specific phosphodiesterases are expressed in discrete subcellular compartments during rat spermiogenesis. Endocrinology. 1999;140(5):2297–306.
47. Roy A, Lin YN, Agno JE, DeMayo FJ, Matzuk MM. Tektin 3 is required for progressive sperm motility in mice. Mol Reprod Dev. 2009;76(5):453–9.
48. Cavarocchi E, Whitfield M, Chargui A, Stouvenel L, Lores P, Coutton C, et al. The sodium/proton exchanger SLC9C1 (sNHE) is essential for human sperm motility and fertility. Clin Genet. 2021;99(5):684–93.
49. Pontén F, Gry M, Fagerberg L, Lundberg E, Asplund A, Berglund L, et al. A global view of protein expression in human cells, tissues, and organs. Mol Syst Biol. 2009;5:337.
50. Schrimpf R, Gottschalk M, Metzger J, Martinsson G, Sieme H, Distl O. Screening of whole genome sequences identified high-impact variants for stallion fertility. BMC Genomics. 2016;17:288.
51. Timmons PM, Rigby PWJ, Poirier F. The murine seminiferous epithelial cycle is pre-figured in the Sertoli cells of embyonic testis. Development. 2002;129:635–47.
52. Mercer JC, DeHaven WI, Smyth JT, Wedel B, Boyles RR, Bird GS, et al. Large store-operated calcium-selective currents due to Co-expression of Orai1 or Orai2 with the intracellular calcium sensor, Stim1. J Biol Chem. 2006;281:24979–90.
53. Frischauf I, Zayats V, Deix M, Hochreiter A, Polo IJ, Muik M, et al. A calcium-accumulating region, CAR, in the channel Orai1 enhances Ca2+ permeation and SOCE-induced gene transcription. Sci Signal. 2015;8(408):ra131.
54. Lunz V, Romanin C, Frischauf I. STIM1 activation of Orai1. Cell Calcium. 2019;77:29–38.
55. Davis FM, Goulding EH, D'Agostin DM, Janardhan KS, Cummings CA, Bird GS, et al. Male infertility in mice lacking the store-operated Ca(2+) channel Orai1. Cell Calcium. 2016;59(4):189–97.
56. Lyon K, Adams A, Piva M, Asghari P, Moore ED, Vogl AW. Ca2+ signaling machinery is present at intercellular junctions and structures associated with junction turnover in rat Sertoli cells. Biol Reprod. 2017;96(6):1288–302.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kothalawala SD, Günther S, Schuppe H-C, Pilatz A, Wagenlehner F, Kliesch S, et al. Identification of differentially expressed genes in human testis biopsies with defective spermatogenesis. Reprod Med Biol. 2024;23:e12616. https://doi.org/10.1002/rmb2.12616