



OPEN ACCESS

EDITED BY
Ruth Roberts,
Apconix, United Kingdom

REVIEWED BY
Yemin Zhang,
Wuhan University, China
Fei Yu,
Tongji University School of Medicine,
China

*CORRESPONDENCE
Guoqun Luo,
2751096131@qq.com
Huan Li,
leywoom@163.com

[†]These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

RECEIVED 06 July 2022
ACCEPTED 01 November 2022
PUBLISHED 24 November 2022

CITATION
Zhong Y, Chen X, Zhao J, Deng H, Li X, Xie Z, Zhou B, Xian Z, Li X, Luo G and Li H (2022), Integrative analyses of potential biomarkers and pathways for non-obstructive azoospermia. *Front. Genet.* 13:988047. doi: 10.3389/fgene.2022.988047

COPYRIGHT
© 2022 Zhong, Chen, Zhao, Deng, Li, Xie, Zhou, Xian, Li, Luo and Li. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrative analyses of potential biomarkers and pathways for non-obstructive azoospermia

Yucheng Zhong^{1†}, Xiaoqing Chen^{2†}, Jun Zhao^{1†}, Hao Deng^{1†}, Xiaohang Li¹, Zhongju Xie¹, Bingyu Zhou¹, Zhuojie Xian¹, Xiaoqin Li¹, Guoqun Luo^{1*} and Huan Li^{1*}

¹Assisted Reproductive Technology Center, Southern Medical University Affiliated Maternal & Child Health Hospital of Foshan, Foshan, China, ²Department of Breast Surgical Oncology, Southern Medical University Affiliated Maternal & Child Health Hospital of Foshan, Foshan, China

Background: Non-obstructive azoospermia (NOA) is the most severe form of male infertility. Currently, the molecular mechanisms underlying NOA pathology have not yet been elucidated. Hence, elucidation of the mechanisms of NOA and exploration of potential biomarkers are essential for accurate diagnosis and treatment of this disease. In the present study, we aimed to screen for biomarkers and pathways involved in NOA and reveal their potential molecular mechanisms using integrated bioinformatics.

Methods: We downloaded two gene expression datasets from the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) in NOA and matched the control group tissues were identified using the limma package in R software. Subsequently, Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), gene set enrichment analysis (GSEA), protein-protein interaction (PPI) network, gene-microRNAs network, and transcription factor (TF)-hub genes regulatory network analyses were performed to identify hub genes and associated pathways. Finally, we conducted immune infiltration analysis using CIBERSORT to evaluate the relationship between the hub genes and the NOA immune infiltration levels.

Results: We identified 698 common DEGs, including 87 commonly upregulated and 611 commonly downregulated genes in the two datasets. GO analysis indicated that the most significantly enriched gene was protein polyglycylation, and KEGG pathway analysis revealed that the DEGs were most significantly enriched in taste transduction and pancreatic secretion signaling pathways. GSEA showed that DEGs affected the biological functions of the ribosome, focaladhesion, and protein_expor. We further identified the top 31 hub genes from the PPI network, and friends analysis of hub genes in the PPI network showed that NR4A2 had the highest score. In addition, immune infiltration analysis found that CD8⁺ T cells and plasma cells were significantly correlated with ODF3 expression, whereas naive B cells, plasma cells, monocytes, M2 macrophages, and resting mast cells showed significant variation in the NR4A2 gene expression group, and there were differences in T cell regulatory immune cell infiltration in the FOS gene expression groups.

Conclusion: The present study successfully constructed a regulatory network of DEGs between NOA and normal controls and screened three hub genes using integrative bioinformatics analysis. In addition, our results suggest that functional changes in several immune cells in the immune microenvironment may play an important role in spermatogenesis. Our results provide a novel understanding of the molecular mechanisms of NOA and offer potential biomarkers for its diagnosis and treatment.

KEYWORDS

non-obstructive azoospermia, differentially expressed genes, hub genes, bioinformatics, biomarker

1 Introduction

Over the past few decades, the incidence of infertility has rapidly increased every year. Infertility affects approximately 15% of couples of childbearing age. Azoospermia is the most severe phenotype of male infertility, with approximately 10–15% of infertile men seeking medical attention (Lotti et al., 2014; Tournaye et al., 2017). Some congenital or acquired causes can be found in the vast majority of cases of obstructive azoospermia (OA) (Krausz, 2011; Pan et al., 2018). Non-obstructive azoospermia (NOA), the most severe form of male factor infertility, is characterized by the lack of sperm in the ejaculate and affects about 5–10% of infertile men (Jarow et al., 1989; Lotti et al., 2014; Tournaye et al., 2017). OA is typically characterized by normal spermatogenesis, whereas NOA represents a heterogeneous condition in which spermatogenesis is impaired, including insufficient spermatogenesis and maturation arrest in Sertoli cell-only syndrome (Krausz, 2011; Lotti et al., 2014; Tournaye et al., 2017). Klinefelter syndrome and Y chromosome microdeletions are the most common congenital causes of NOA (Forti et al., 2010; Krausz, 2011; Corona et al., 2017). The etiology of acquired NOA includes torsion, mumps, orchitis, cryptorchidism, and iatrogenic problems (Krausz, 2011; Lotti et al., 2014; Tournaye et al., 2017). For nearly half of the patients with NOA, the etiology remains unknown (Poongothai et al., 2009). Previously, NOA was considered an untreatable condition that required fertilization with donor sperm. With the advent of microdissection testicular sperm extraction (mTESE) and intracytoplasmic sperm injection (ICSI), these techniques have become the first-line treatments for NOA patients (Krausz, 2011; Lotti et al., 2014; Tournaye et al., 2017). Unfortunately, the probability of retrieving sperm is only about 50% in men with NOA owing to partial and heterogeneous preserved focal spermatogenesis (Krausz, 2011; Lotti et al., 2014; Tournaye et al., 2017). Currently, the molecular mechanisms underlying NOA pathology remain unclear. However, this remains an important challenge to solve. Therefore, the identification of genetic abnormalities in patients with NOA is critical.

Several studies have identified potential biomarkers involved in NOA using microarray analysis, weighted gene co-expression network analysis (WGCNA), whole-exome sequencing, and single-cell transcriptome sequencing (scRNA-seq) (Wang et al., 2018; Zheng et al., 2019; Chen et al., 2020; Zhao et al., 2020).

However, only a few of these biomarkers are currently used for the diagnosis of NOA. In addition, these studies did not integrate immune infiltration analysis, and the immune system plays an important role in testicular dysfunction and male infertility. Immune infiltration analysis can be used to study the infiltrated immune cells in the testes, to infer and discover the role of immune cells in NOA, and to develop diagnostic and therapeutic targets for the disease. This omission laid the foundation for the present study.

In the present study, we identified differentially expressed genes (DEGs) for NOA by analyzing two mRNA expression profiles downloaded from the Gene Expression Omnibus database (GEO, <http://www.ncbi.nlm.nih.gov/geo/>). Our results provide a novel understanding of the molecular mechanisms of NOA and offer potential biomarkers for its diagnosis and treatment.

2 Materials and methods

2.1 Microarray data

Two gene expression datasets (GSE45885 and GSE45887) (Zhao et al., 2021) related to male NOA were derived from GEO (GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]) (Chicco, 2022) using the R package GEOquery (Hunt et al., 2022) of R software program. The data type was expression profiling by array, and the species was *Homo sapiens*. Twenty-seven NOA and four normal control samples were obtained from GSE45885, and GSE45887 contained 16 NOA and four normal control samples. The characteristics of GSE45885 and GSE45887 patients can be seen in Supplementary Tables S1, S2. The raw data from the GSE45885 and GSE45887 datasets were normalized using the limma package (Liu et al., 2021). The data analysis process can be seen in Figure 1.

2.2 Identification of differentially expressed genes (DEGs)

DEGs with the threshold criterion of $|\log_2FC| > 1$ and $p < 0.05$ from NOA and normal control samples in GSE45885 and

GSE45887 datasets were screened with the limma package (Walker, 2009). Subsequently, heatmaps and volcano plots of DEGs from each dataset were plotted using the pheatmap package (Ning et al., 2022) and ggplot2 package (Gustavsson et al., 2022) in the R analysis platform.

2.3 Gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) enrichment analyses of DEGs

GO analysis is a common method for annotating and analyzing the biological processes of genes, including biological processes (BPs), molecular functions (MFs), and cellular components (CCs) (Ge et al., 2020). KEGG is a widely used database that stores information on genomes, biological pathways, diseases, and drugs (Kanehisa et al., 2021). We performed biological analyses using clusterProfiler (Wu et al., 2021) in the R software statistical analysis platform (significant at $p < 0.05$) to visualize the GO terms and KEGG pathway enrichment analysis of DEGs from NOA and normal control samples.

2.4 Gene set enrichment analysis (GSEA)

GSEA determines whether a set of predefined genes shows statistically significant, concordant differences between related phenotypes to screen for significant differential biological functions (Nguyen et al., 2021). We obtained the “C2. kegg.v7.4. symbols” gene set from MSigDB (Liberzon et al., 2015) for GSEA of the two datasets. Furthermore, GSEA was automatically completed and visualized using clusterProfiler (Wu et al., 2021) in R software, and statistical significance was set at $p < 0.05$.

2.5 Construction of protein-protein interaction (PPI) network

PPI (Das and Mitra, 2021) networks are composed of individual proteins that interact with each other, and are involved in various aspects of biological signal transmission, gene expression regulation, energy and substance metabolism, and cell cycle regulation. Systematic analysis of functional interactions between proteins is of great significance for comprehending the principle of proteins in biological systems, the reaction mechanism of biological signals, and the mechanisms of generation or development of diseases under special physiological states, as well as the functional relationships between proteins.

The STRING (Szklarczyk et al., 2021) database currently covers 9.6 million proteins and 13.8 million protein-protein interactions from 5,090 organisms and contains known and predicted protein-protein interactions. The results of the

STRING database were derived from experimental data, PubMed abstract texts, other database data, and bioinformatics methods. In this study, a PPI network of DEGs of NOA was constructed using the STRING database and visualized using the Cytoscape (Puig et al., 2020) software.

2.6 Construction of miRNA-hub gene and transcription factor-hub gene regulatory networks

Hub genes were identified using GOSemSim (Kamran and Naveed, 2022) in R software. In addition, the functional similarity between proteins was evaluated using the geometric mean of semantic acquaintance in CCs and MF using GOSemSim in R software. We also analyzed hub genes using the NetworkAnalyst database (<https://www.networkanalyst.ca/NetworkAnalyst/home.xhtml>) and constructed miRNA-hub gene interaction and transcription factor-hub gene networks.

2.7 Immune infiltration analysis by CIBERSORT

CIBERSORT (Le et al., 2021) is a deconvolution algorithm for the expression matrix of immune cell subtypes based on linear support vector regression to estimate the abundance of immune cell infiltration in a mixed-cell population using RNA-seq data. We analyzed immune cell infiltration and the proportion of NOA tissue using CIBERSORT and conducted a correlation analysis of related immune cells by obtaining the hub gene to evaluate the relationship between the hub gene and NOA immune infiltration levels.

2.8 Statistical analysis

R version 4.0.2 software (<https://cran.r-project.org/>) was used to conduct the statistical analyses. For the comparison of the two groups of continuous variables, the statistical significance of the normally distributed variables was estimated using the independent Student's *t*-test, and the differences between the non-normally distributed variables were analyzed using the Mann-Whitney *U* test. All statistical *p*-values were bilateral, and a *p*-value of <0.05 was considered statistically significant.

3 Results

3.1 Identification of DEGs in NOA

First, we obtained the corresponding data using the R package GEOquery package and standardized the original

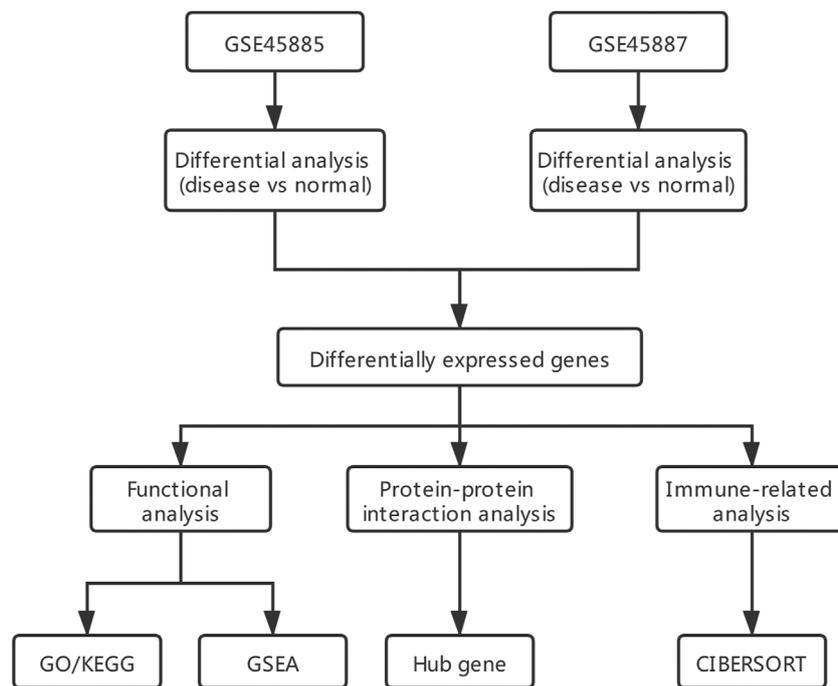


FIGURE 1 Analysis flow chart. GO/KEGG: Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG); GSEA: gene set enrichment analysis.

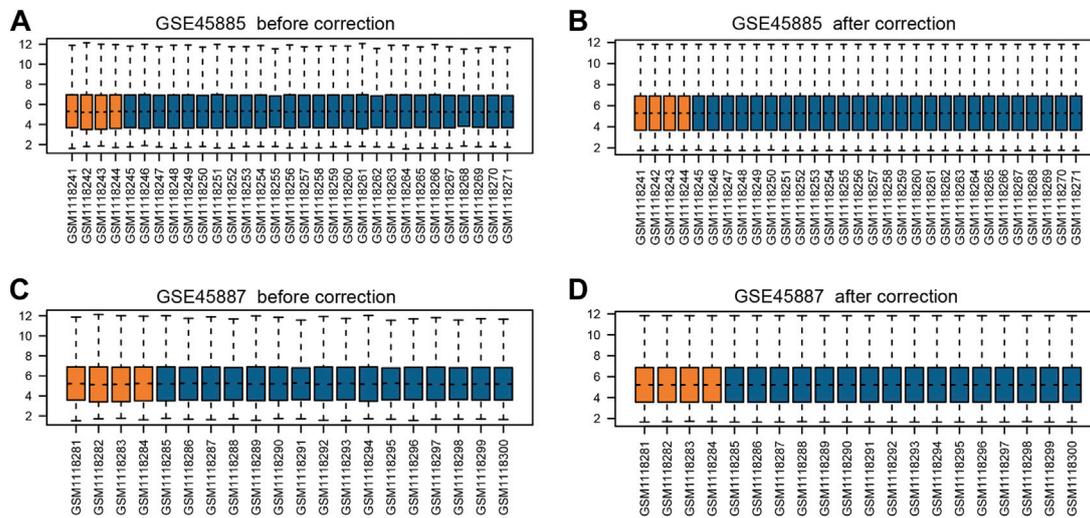
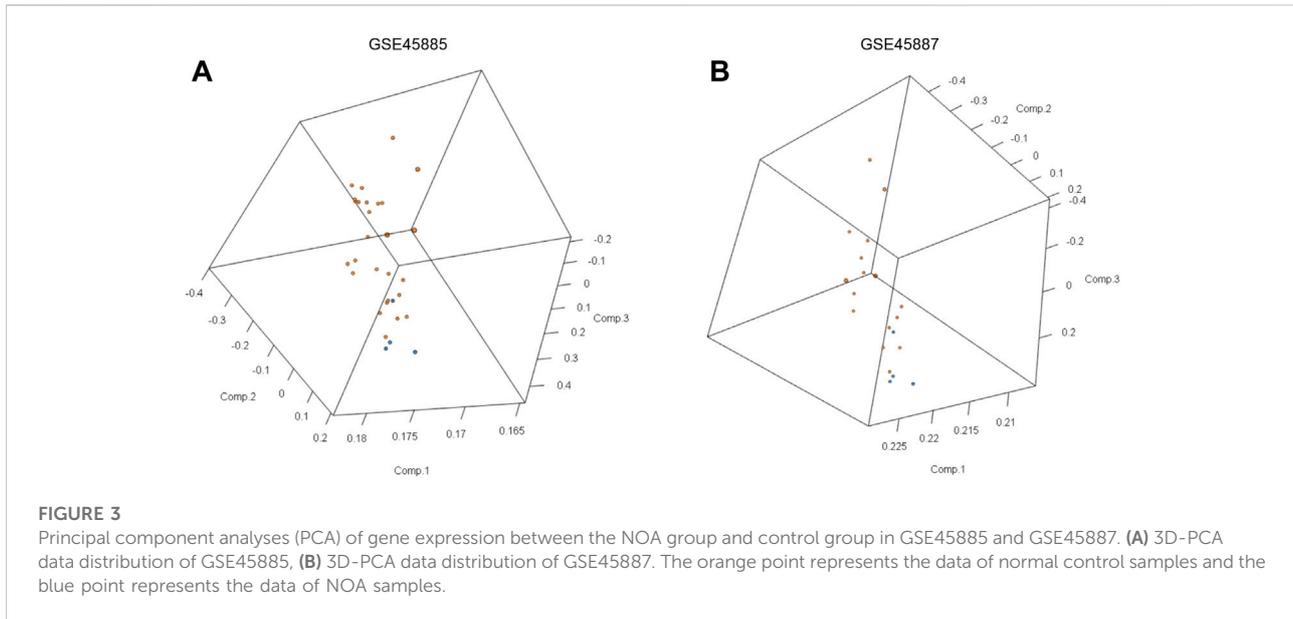


FIGURE 2 Box plots of gene expression data before and after normalization. (A) GSE45885 before correction, (B) GSE45885 after correction, (C) GSE45887 before correction, (D) GSE45887 after correction. The x-axis label represents the sample symbol and the y-axis label represents the gene expression values. The orange bar represents the data of normal control samples and the blue bar represents the data of NOA samples.

datasets using the limma package (Figures 2A–D). In order to further judge whether the data set samples have obvious overall differences in expression profiles, we

performed PCA analysis and found that the disease group and control group samples had obvious overall differences (Figures 3A,B).



Next, we used the limma package to perform differential analysis, and the results are as follows: the combined GEO dataset (Combined Datasets) has a total of 698 differentially expressed genes (DEGs) that satisfy the threshold of $|\log_{2}FC| > 0$ and $adj.p < 0.05$, under this threshold, 70 up-regulated genes ($\log_{2}FC > 0$ and $adj.p < 0.05$) and 507 down-regulated genes ($\log_{2}FC < 0$ and $adj.p < 0.05$) in the GSE45885 dataset (Figure 4A), and utilized the top 40 differentially expressed genes A classification heat map was drawn (Figure 4B); there were 17 up-regulated genes ($\log_{2}FC > 0$ and $adj.p < 0.05$) and 104 down-regulated genes ($\log_{2}FC < 0$ and $adj.p < 0.05$) in the GSE45887 dataset. The 40 differentially expressed genes were classified as heatmaps (Figure 4D). The top 40 differential genes obtained are: PRM1, PRM2, GTSF1L, TNP1, FSCN3, C19orf62, ACTL7A, AKAP4, ABHD1, GAPDHS, LELP1, FNDC8, BANF2, FAM205A, ACSBG2, ODF1, BPIFA3, GSG1, UBQLN3, SMCP, RN7SL648P, MGC24103, RN7SL751P, MIR145, MIR30E, MIR27B, MIR99A, MIR32, MIRLET7G, MIRLET7A2, MIR199A2, MT-TT, MT-TW, MT-TD, MIR509-1, RNU1-59P, MT-TK, MT-TL2, MT-TS2, MT-TH (Figures 4C,D).

3.2 PPI network construction and hub gene identification

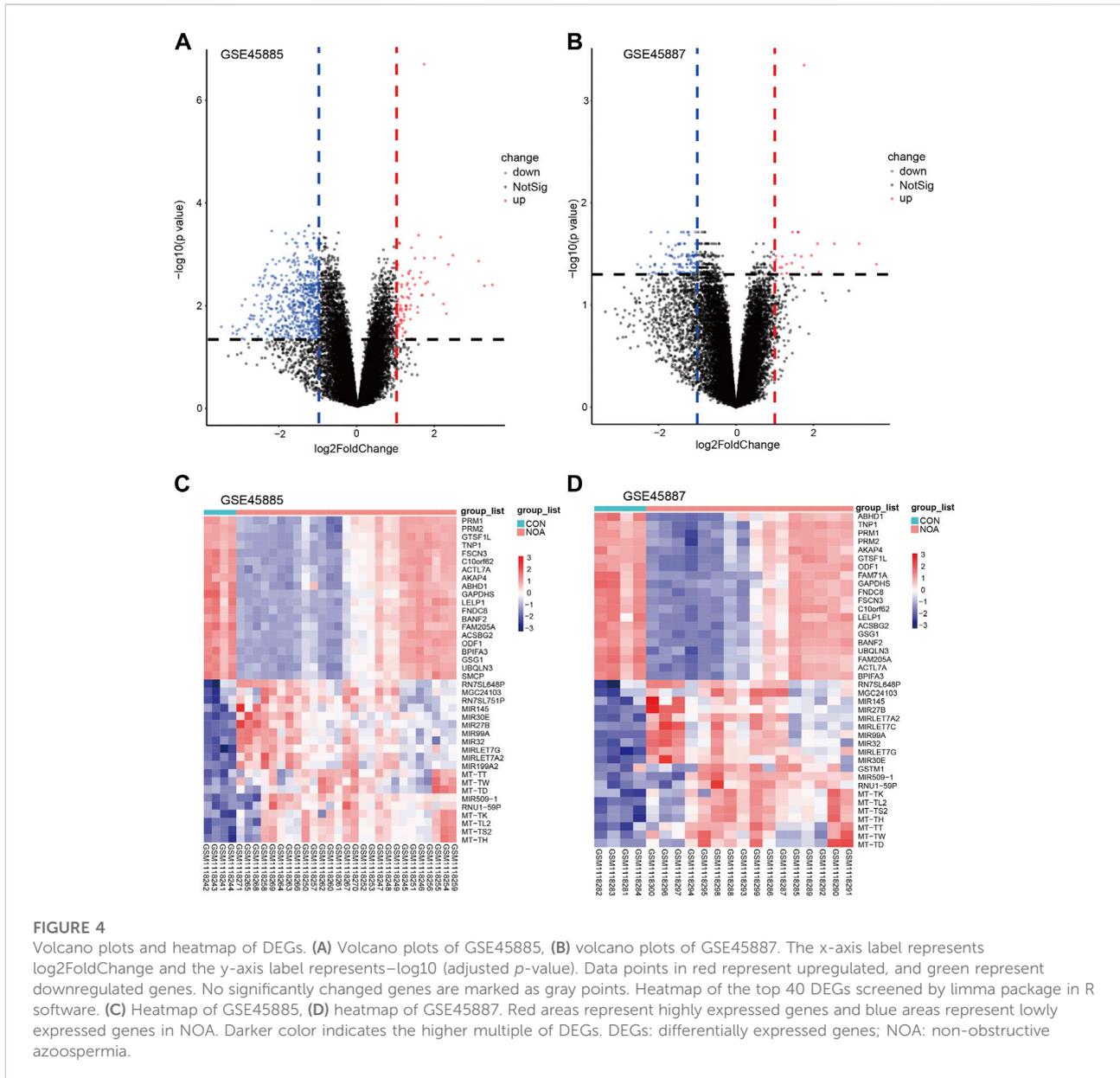
We intersected the differentially expressed genes from the analysis of the two datasets and visualized them using a Venn diagram, resulting in 119 differentially expressed genes (Figure 5A). Then, we conducted PPI (Protein-Protein Interaction Networks, PPI) analysis on 119 differential genes through the STRING database, and imported the

protein interaction results into Cytoscape for visualization and drawing, and found that 31 genes have strong similarities of biological functions (Figure 5B), are: TEK4, ODF3, ACTN3, DOT1L, MUC1, CCDC116, RSPH4A, ELL, GLT6D1, HSF1, SPATA31E1, NSUN4, TBL3, FAM187B, NR4A2, DUS1L, CCDC96, SNX2, FER1L5, DOC2A, FER1L6, STAC3, CLPB, FAM163A, SLFN1L, FOS, PLCD4, DUSP1, INPP1, CATSPER4, DNAH1.

Then, the networkAnalyst database was used to analyze these 119 genes, and jointly used the JASPAR database (Zheng et al., 2019) and miRTarBase v8.0 (Chen et al., 2020) to analyze the transcription factors of differential genes and their possible binding miRNAs. The results of the analysis were then imported into Cytoscape for visualization, displaying the transcription factor-differential gene network (Figure 5C) and the differential gene-miRNA network (Figure 5D). Figure 5E shows the friends analysis results of the hub gene. The top 8 genes with scores are NR4A2, FOS, COF3, MUC1, ELL, DUSP1, STAC3, TEK4, among which NR4A2 has the highest score, which may play an important role in the occurrence of NOA effect.

3.3 KEGG and GO enrichment analyses

To investigate the relationship between the DEGs of NOA and BCs, MFs, CCs, biological pathways, and diseases, functional and pathway enrichment analyses were performed. Regarding BPs of GO analysis, DEGs of NOA were significantly enriched in protein polyglycylation, negative regulation of cytokine-mediated signaling pathways, and negative regulation of



response to cytokine stimulus. For CCs, DEGs of NOA were significantly enriched in sperm flagellum, 9 + 2 motile cilium, axoneme, ciliary plasm, and motile cilium. Changes in MFs of hub genes were mainly enriched in RNA binding involved in posttranscriptional gene silencing, mRNA binding involved in posttranscriptional gene silencing, protein-glycine ligase activity, and protein-glycine ligase activity. Additionally, using the logFC value of genes, we visualized the changes in gene expression and the relationship between functions using GOpot software (Figure 6B). KEGG pathway enrichment demonstrates that the DEGs of NOA were primarily enriched in the taste transduction and pancreatic secretion signaling pathway Supplementary Table S3, and the most significant enrichment

signaling pathway, hSA04742: Taste transduction, is shown in Figure 6C.

3.4 GSEA and gene set variation analysis (GSVA)

GSEA was performed to determine the effect of gene expression on NOA, and the results show that the DEGs affected the biological functions of the ribosome, focaladhesion, and protein_expor (Figures 7A,B). To evaluate the enrichment of different metabolic pathways in different samples, we analyzed the expression levels of genes

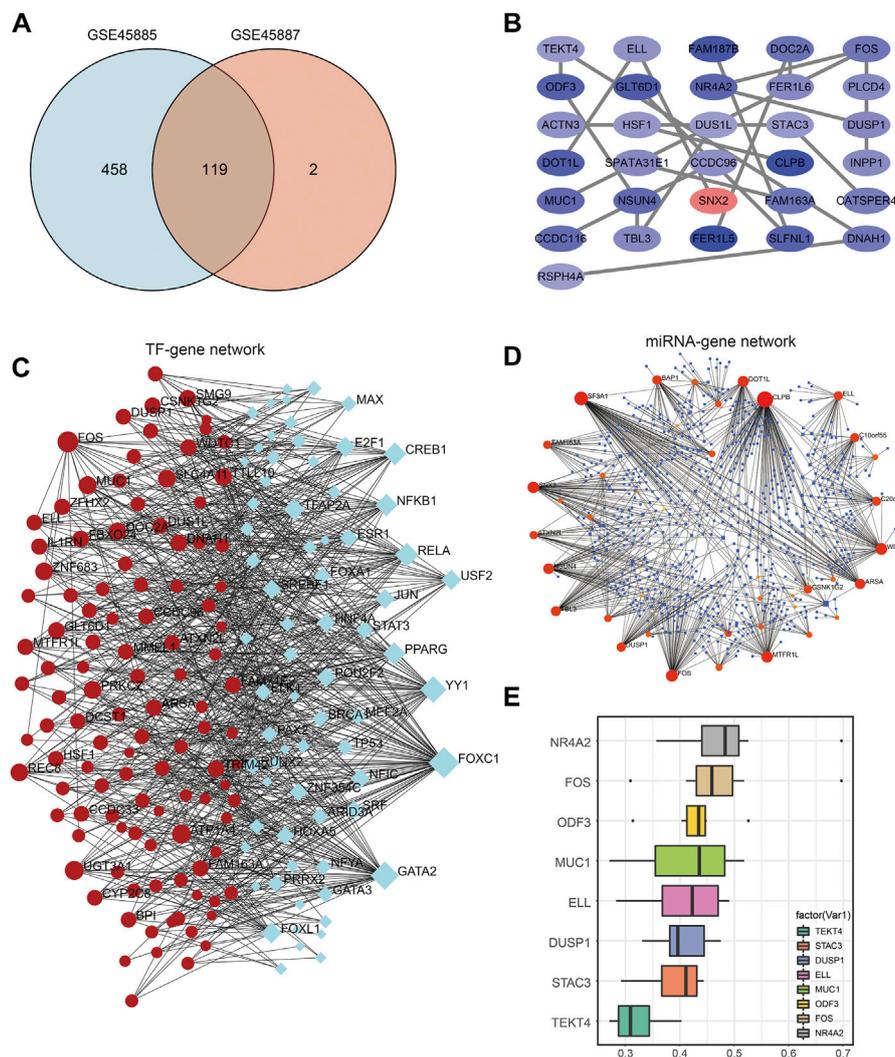


FIGURE 5 PPI network construction and hub gene regulatory network. **(A)** Venn diagram of co-expressed DEGs from GSE45885 and GSE45887. **(B)** DEG-related PPI networks of NOA. Red nodes represent highly expressed genes and blue nodes represent lowly expressed genes in NOA. **(C)** DEG-related TF-mRNA networks of NOA. **(D)** DEG-related miRNA-mRNA networks of NOA. **(E)** Friend analysis of hub gene in PPI network. DEGs: differentially expressed genes; NOA: non-obstructive azoospermia; PPI: protein-protein interaction.

in different samples that converted them into gene sets in different samples using GSVA, and the enrichment was visualized using the pheatmap package (Figure 7C). We found that sample grouping could distinguish the results of the GSEA.

3.5 Immune infiltration analysis

To assess the level of immune infiltration in male non-obstructive azoospermia tissue, we used the CIBERSORT algorithm to calculate the degree of infiltration of 22 types of immune cells in the tissue.

Using the wilcox.test algorithm to analyze and filter out immune cells with low expression abundance, a total of 15 immune cells were included, including B cells naive, B cells memory, Plasma cells, T cells CD8, T cells CD4 memory resting, T cells follicular helper, T cells regulatory (Tregs), NK cells resting, NK cells activated, Monocytes, Macrophages M0, Macrophages M2, Dendritic cells activated, Mast cells resting, Mast cells activated, and draw a panorama of immune cell entry in NOA (Figure 8A). Next, the correlation between individual immune cells in both datasets was assessed (Figure 8B). To assess the functional correlation between key genes and immune cells in male non-obstructive azoospermia, we selected the top three hub genes for analysis.

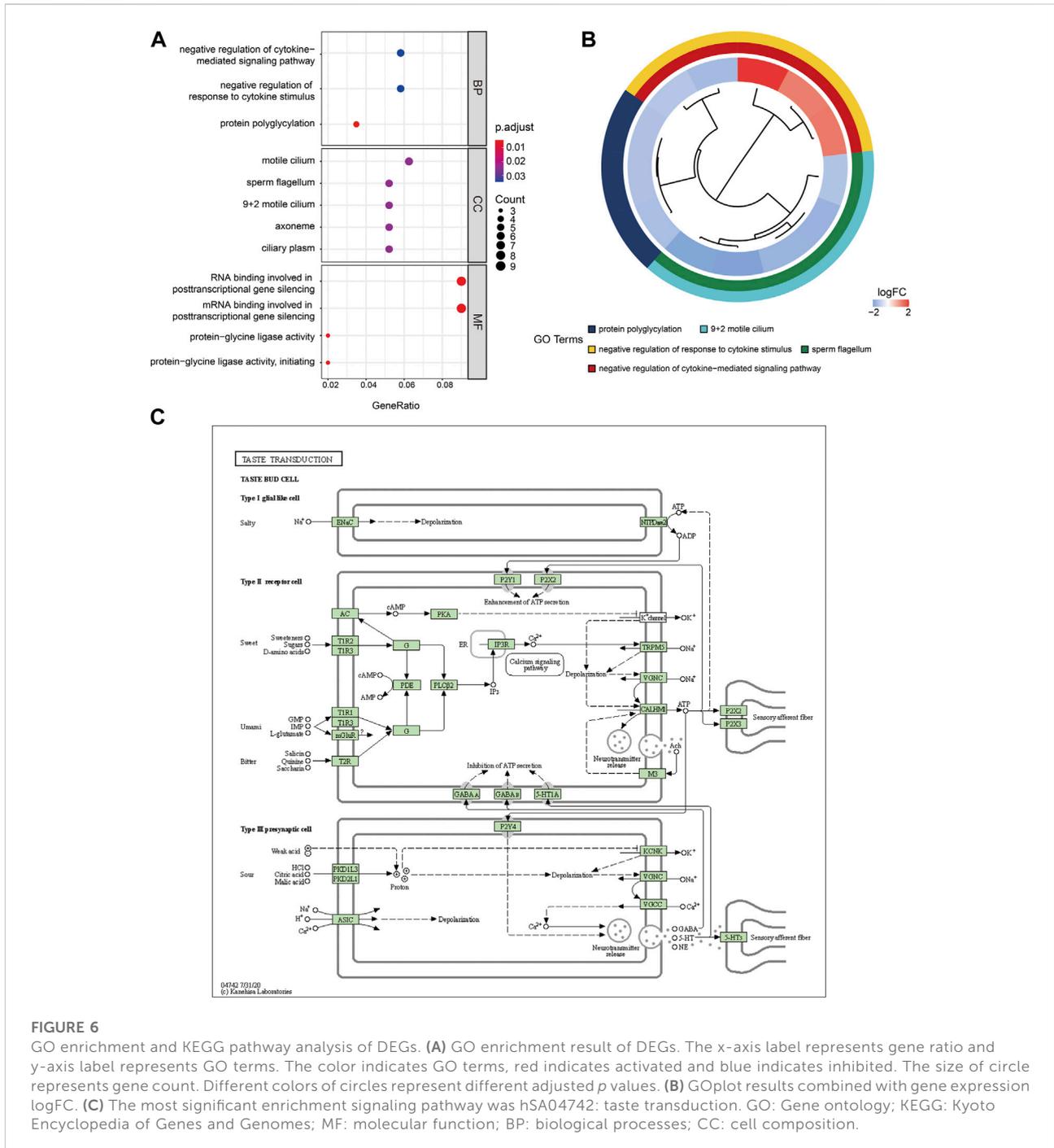
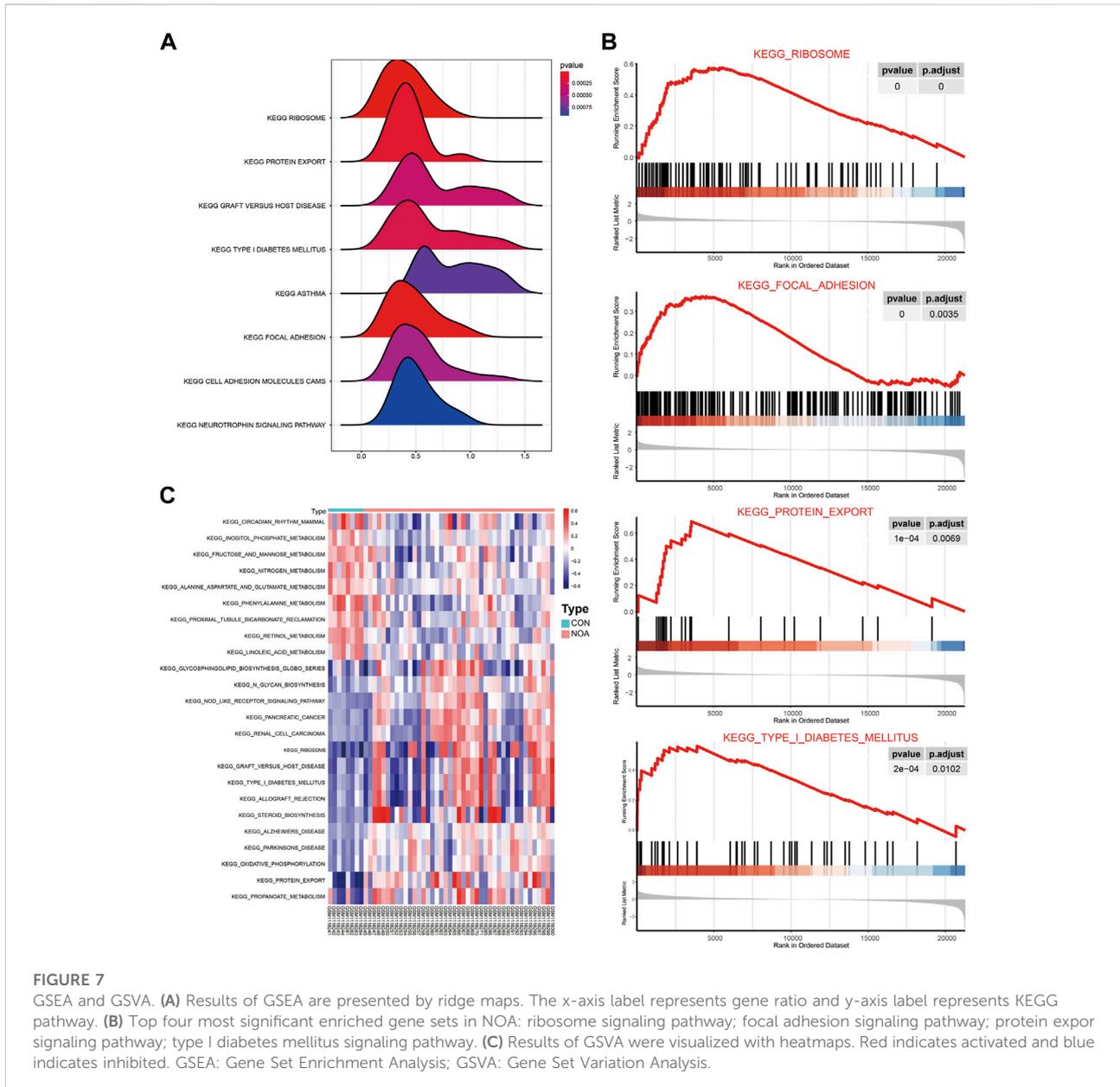


FIGURE 6

GO enrichment and KEGG pathway analysis of DEGs. (A) GO enrichment result of DEGs. The x-axis label represents gene ratio and y-axis label represents GO terms. The color indicates GO terms, red indicates activated and blue indicates inhibited. The size of circle represents gene count. Different colors of circles represent different adjusted *p* values. (B) GOplot results combined with gene expression logFC. (C) The most significant enrichment signaling pathway was hSA04742: taste transduction. GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; MF: molecular function; BP: biological processes; CC: cell composition.

According to the expression value of the hub gene, the samples were divided into high expression group and low expression group, and calculated the difference of immune cell infiltration levels between the high and low expression groups. As shown in Figures 8C–E, in the disease group and the ODF3 high and low expression group, there were differences in the immune cell infiltration of B cells naïve,

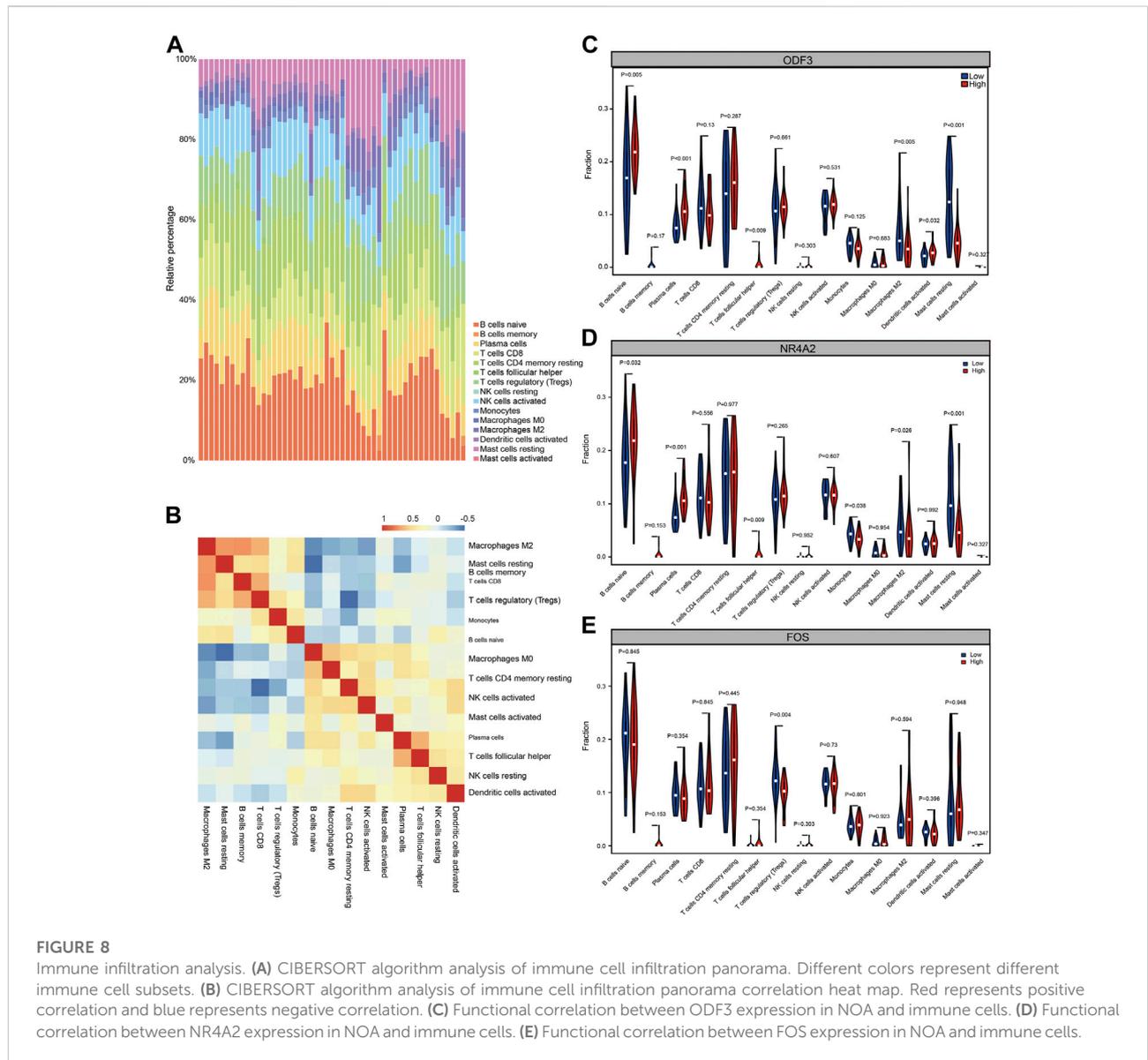
plasma cells, Macrophages M2, Dendritic cells activated, and Mast cells resting, among which Macrophages M2 and Mast cells resting were low. Expression genes, plasma cells, B cells naïve, and Dendritic cells activated are highly expressed genes; however, in the high and low expression groups of NR4A2 gene, there are differences in the immune cell infiltration of B cells naïve, Plasma cells, Monocytes,



Macrophages M2, and Mast cells resting. The expressed genes are B cells naïve, Plasma cells, and the low-expressed genes are Monocytes, Macrophages M2, Mast cells resting, and there are differences in the infiltration of T cells regulatory (Tregs) immune cells in the high and low expression groups of FOS genes, which are low-expressed genes. It was statistically significant ($p < 0.05$, statistically significant; $p < 0.01$, highly statistically significant; $p < 0.001$, extremely statistically significant). The above results indicate that the functional changes of several immune cells may play an important role in the immune microenvironment of NOA patients.

4 Discussion

Spermatogenesis is a complex and subtle process characterized by three specific functional phases: mitotic proliferation and expansion, meiosis, and spermiogenesis (Neto et al., 2016). Genetic mutations play an important role in NOA (Peña et al., 2020). Compared to that of OA, the pathogenesis of NOA is more complex and difficult to understand. However, there are currently no specific therapeutic targets for NOA. Thus, it is of great significance to construct a molecular regulatory network of NOA to search for therapeutic targets. In this study, we integrated the



GSE45885 and GSE45887 datasets and identified 527 and 121 common DEGs, respectively, using bioinformatics methods, among which 119 DEGs were at the intersection of the two datasets and 31 genes had similar biological functions. To further investigate the molecular mechanism of the pathogenesis of NOA, we analyzed the transcription factors of 119 DEGs and their possible binding miRNAs and successfully identified three hub genes (NR4A2, FOS, and ODF3). KEGG and GO enrichment analyses of DEGs showed that the DEGs of NOA were mainly enriched in taste transduction and pancreatic secretion signaling pathways. Moreover, the three most important hub genes (NR4A2, FOS, and ODF3) were analyzed for immune cell infiltration in NOA testicular tissue, suggesting that changes in the function of several immune cells in

the immune microenvironment may play an important role in spermatogenesis. These results suggest that NR4A2, FOS, and ODF3 are potential biomarkers for the diagnosis and treatment of NOA.

NR4A2 belongs to the nuclear receptor superfamily and is involved in multiple BPs including proliferation, metabolism, immunity, cellular stress, apoptosis, and DNA repair (Safe et al., 2016). In addition to Nr4a proteins, which are known transcription factors, they are also important receptors of hormones (Ke et al., 2004). Testosterone and dihydrotestosterone secreted from Leydig cells are crucial for the initiation and maintenance of spermatogenesis (Walker, 2009). Dysfunction of Leydig cells contributes to testicular spermatogenesis disorders in OA patients (Walker, 2009). A

previous study reported that NR4A2 plays an essential role in Leydig cells, and can regulate the transcription of steroidogenic acute regulatory protein (StAR) or 3 β -hydroxysteroid dehydrogenase (3 β -HSD) in Leydig cells (Martin and Tremblay, 2010; Hu et al., 2018). Dai et al. (2021), (Martin et al., 2008) investigated the expression patterns, distribution, and functions of NR4A2 in male Tianzhu white yaks and found that NR4A2 is involved in the regulation of male yak reproduction, especially steroid hormones and androgen metabolism. Our study further indicates that NR4A2 might play an essential role in spermatogenesis, and the detailed regulatory mechanism of NR4A2 in the spermatogenesis of patients with NOA needs to be further studied. FOS, also known as the c-fos gene, is a transcription factor. One study showed that the expression of c-fos mRNA is appreciably lower and ER β mRNA is higher in the testes of men with NOA than in those with OA, which suggests that c-fos transcriptional activity is associated with spermatogenesis. However, few studies have focused on c-Fos transcriptional activity during spermatogenesis, and the detailed mechanisms are unclear. The outer dense fiber (ODF) protein is a keratin that includes ODF1, ODF2, and ODF3 (Araújo et al., 2009). The tail is an important structure for proper sperm function. Studies have demonstrated that ODF proteins are the main cytoskeletal structures of the sperm tail and are preferentially expressed during mammalian spermiogenesis (Kierszenbaum, 2002; Petersen et al., 2002). Among them, ODF3 is transcribed in the testes, more specifically in spermatids, and is involved in sperm tail formation (Petersen et al., 2002; Sarkar et al., 2022). Unfortunately, there is little evidence that ODF3 is an important trigger for the pathophysiological process of NOA.

We conducted KEGG and GO enrichment analyses of DEGs and found that the DEGs of NOA were mainly enriched in taste transduction and pancreatic secretion signaling pathways. A recent study showed that olfactory receptor 2 is present in vascular macrophages (Orechioni et al., 2022). However, no studies have reported the involvement of these pathways in spermatogenesis. Whether these pathways are involved in spermatogenesis needs to be verified using larger sample sequencing and experiments.

A previous study has confirmed that inflammation in the reproductive tract contributes to testicular dysfunction and male infertility (Hedger, 2011). Therefore, we performed immune infiltration analysis and found that immune cells are involved in the pathophysiological process of NOA. One study reported that T cell subsets are essential for intact spermatogenesis and may be targets for the treatment of chronic orchitis and immune infertility (Gong et al., 2020). Suppressor T cells predominate in patients with obstructive azoospermia, whereas T cells of the helper phenotype predominate in patients with unilateral testicular obstruction (el-Demiry et al., 1987), but T cells and their subset distribution in the testis of NOA are unknown. In addition, several studies have suggested that B cells and macrophages contribute to spermatogenesis (Cabas et al.,

2011; de Oliveira et al., 2021), similar to our results. Dong et al. (2021) also found that macrophages are the most important immune cells in NOA by immune infiltration analysis. All these studies showed that immune cells play an important role in NOA, and the mechanism by which immune cells impact spermatogenesis should be further studied.

Although we obtained several key findings in the present study, there are still some limitations. First, although we analyzed two datasets with 51 patients, the effects of race and region on the findings were not observed. We look forward to verifying our results using a larger, real-world sample size. Second, our analysis was based on data from a public database and we did not conduct experiments to verify our results. Third, the hub genes, pathway enrichment, and results of immune infiltration analysis were not validated by external datasets or clinical samples. Fourth, we must admit that WGCNA is a better analytical method, while DEGs is the more common method to screen for differential genes. We believe that the analysis method combining WGCNA and DEGs will make the results more rigorous and credible, which is what we need to work on in the future. Finally, owing to the lack of patient prognosis information in the dataset, genes related to patient prognosis were not screened out.

In conclusion, the present study successfully constructed a regulatory network of DEGs between NOA and normal controls and screened three hub genes using integrative bioinformatics analysis. In addition, our results suggest that functional changes in several immune cells in the immune microenvironment may play an important role in spermatogenesis. Our results provide a novel understanding of the molecular mechanisms of NOA and offer potential biomarkers for its diagnosis and treatment.

Data availability statement

Publicly available datasets were analyzed in this study. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

All authors listed have made substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This study was supported by funding from Innovation Project of Women and Children medical research center affiliated to Foshan Institute of Fetal Medicine (FEYJZX-2020-008 and FEYJZX-2020-006) and Research Funds from Foshan Municipal Science and Technology Project (2020001005568). We thank all support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.988047/full#supplementary-material>

References

- Araújo, F. C., Oliveira, C. A., Reis, A. B., Del Puerto, H. L., Martins, A. S., and Reis, F. M. (2009). Expression of the proto-oncogene *c-fos* and the immunolocalization of *c-fos*, phosphorylated *c-fos* and estrogen receptor beta in the human testis. *Histol. Histopathol.* 24, 1515–1522. PMID: 19795350. doi:10.14670/HH-24.1515
- Cabas, I., Chaves-Pozo, E., Alcázar, A. G., Meseguer, J., Mulero, V., and García-Ayala, A. (2011). Dietary intake of 17 α -ethinylestradiol promotes leukocytes infiltration in the gonad of the hermaphrodite gilthead seabream. *Mol. Immunol.* 48, 2079–2086. doi:10.1016/j.molimm.2011.07.001
- Chen, S., Wang, G., Zheng, X., Ge, S., Dai, Y., Ping, P., et al. (2020). Whole-exome sequencing of a large Chinese azoospermia and severe oligospermia cohort identifies novel infertility causative variants and genes. *Hum. Mol. Genet.* 29, 2451–2459. doi:10.1093/hmg/ddaa101
- Chicco, D. (2022). geneExpressionFromGEO: An R package to facilitate data reading from gene expression Omnibus (GEO). *Methods Mol. Biol.* 2401, 187–194. doi:10.1007/978-1-0716-1839-4_12
- Corona, G., Pizzocaro, A., Lanfranco, F., Garolla, A., Pelliccione, F., Vignozzi, L., et al. (2017). Sperm recovery and ICSI outcomes in klinefelter syndrome: A systematic review and meta-analysis. *Hum. Reprod. Update* 23, 265–275. doi:10.1093/humupd/dmx008
- Dai, L., Zhang, Q., Shi, J., Bai, X., An, X., Zhang, B., et al. (2021). The distribution, expression patterns and functional analysis of NR1D1 and NR4A2 in the reproductive axis tissues of the male Tianzhu White Yak. *Animals.* 11, 3117. doi:10.3390/ani11113117
- Das, B., and Mitra, P. (2021). Protein interaction network-based deep learning framework for identifying disease-associated human proteins. *J. Mol. Biol.* 433 (19), 167149. doi:10.1016/j.jmb.2021.167149
- de Oliveira, S. A., Cerri, P. S., and Sasso-Cerri, E. (2021). Impaired macrophages and failure of steroidogenesis and spermatogenesis in rat testes with cytokines deficiency induced by diacerein. *Histochem. Cell Biol.* 156, 561–581. doi:10.1007/s00418-021-02023-7
- Dong, M., Li, H., Zhang, X., and Tan, J. (2021). Weighted correlation gene network analysis reveals new potential mechanisms and biomarkers in non-obstructive azoospermia. *Front. Genet.* 12, 617133. doi:10.3389/fgene.2021.617133
- el-Demiry, M. I., Hargreave, T. B., Busuttill, A., Elton, R., James, K., and Chisholm, G. D. (1987). Immunocompetent cells in human testis in health and disease. *Fertil. Steril.* 48, 470–479. doi:10.1016/s0015-0282(16)59421-7
- Forti, G., Corona, G., Vignozzi, L., Krausz, C., and Maggi, M. (2010). Klinefelter's syndrome: A clinical and therapeutical update. *Sex. Dev.* 4, 249–258. doi:10.1159/000316604
- Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36 (8), 2628–2629. doi:10.1093/bioinformatics/btz931
- Gong, J., Zeng, Q., Yu, D., and Duan, Y. G. (2020). T lymphocytes and testicular immunity: A new insight into immune regulation in testes. *Int. J. Mol. Sci.* 22, 57. doi:10.3390/ijms22010057
- Gustavsson, E. K., Zhang, D., Reynolds, R. H., Garcia-Ruiz, S., and Ryten, M. (2022). ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. *Bioinformatics* 38 (15), 3844–3846. doi:10.1093/bioinformatics/btac409
- Hedger, M. P. (2011). Toll-like receptors and signalling in spermatogenesis and testicular responses to inflammation—a perspective. *J. Reprod. Immunol.* 88, 130–141. doi:10.1016/j.jri.2011.01.010
- Hu, X., Ding, Z., Hong, Z., Zou, Z., Feng, Y., Zhu, R., et al. (2018). Spermatogenesis improved by suppressing the high level of endogenous gonadotropins in idiopathic non-obstructive azoospermia: A case control pilot study. *Reprod. Biol. Endocrinol.* 16, 91. doi:10.1186/s12958-018-0401-7
- Hunt, G. P., Grassi, L., Henkin, R., Smeraldi, F., Spargo, T. P., Kabiljo, R., et al. (2022). GEOexplorer: A webserver for gene expression analysis and visualisation. *Nucleic Acids Res.* 50 (W1), W367–W374. doi:10.1093/nar/gkac364
- Jarow, J. P., Espeland, M. A., and Lipshultz, L. I. (1989). Evaluation of the azoospermic patient. *J. Urol.* 142, 62–65. doi:10.1016/s0022-5347(17)38662-7
- Kamran, A. B., and Naveed, H. (2022). GONtoSim: A semantic similarity measure based on LCA and common descendants. *Sci. Rep.* 2 (1), 3818. doi:10.1038/s41598-022-07624-3
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., and Tanabe, M. (2021). Kegg: Integrating viruses and cellular organisms. *Nucleic Acids Res.* 49 (D1), D545–D551. doi:10.1093/nar/gkaa970
- Ke, N., Claassen, G., Yu, D. H., Albers, A., Fan, W., Tan, P., et al. (2004). Nuclear hormone receptor NR4A2 is involved in cell transformation and apoptosis. *Cancer Res.* 64, 8208–8212. doi:10.1158/0008-5472.CAN-04-2134
- Kierszenbaum, A. L. (2002). Keratins: Unraveling the coordinated construction of scaffolds in spermatogenic cells. *Mol. Reprod. Dev.* 61, 1–2. doi:10.1002/mrd.1124
- Krausz, C. (2011). Male infertility: Pathogenesis and clinical diagnosis. *Best. Pract. Res. Clin. Endocrinol. Metab.* 25, 271–285. doi:10.1016/j.beem.2010.08.006
- Le, T., Aronow, R. A., Kirshtein, A., and Shahriyari, L. (2021). A review of digital cytometry methods: Estimating the relative abundance of cell types in a bulk of cells. *Brief. Bioinform.* 22 (4), bbaa219. doi:10.1093/bib/bbaa219
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. PMID: 26771021; PMCID: PMC4707969. doi:10.1016/j.cels.2015.12.004
- Liu, S., Wang, Z., Zhu, R., Wang, F., Cheng, Y., and Liu, Y. (2021). Three differential expression analysis methods for RNA sequencing: Limma, EdgeR, DESeq2. *J. Vis. Exp.* 2021, 175. doi:10.3791/62528
- Lotti, F., Corona, G., Mondaini, N., Maseroli, E., Rossi, M., Filimberti, E., et al. (2014). Seminal, clinical and colour-Doppler ultrasound correlations of prostatitis-like symptoms in males of infertile couples. *Andrology* 2, 30–41. doi:10.1111/j.2047-2927.2013.00156.x
- Martin, L. J., Boucher, N., Brousseau, C., and Tremblay, J. J. (2008). The orphan nuclear receptor NUR77 regulates hormone-induced StAR transcription in Leydig cells through cooperation with Ca²⁺/calmodulin-dependent protein kinase I. *Mol. Endocrinol.* 22, 2021–2037. doi:10.1210/me.2007-0370
- Martin, L. J., and Tremblay, J. J. (2010). Nuclear receptors in Leydig cell gene expression and function. *Biol. Reprod.* 83, 3–14. doi:10.1095/biolreprod.110.083824
- Neto, F. T., Bach, P. V., Najari, B. B., Li, P. S., and Goldstein, M. (2016). Spermatogenesis in humans and its affecting factors. *Semin. Cell Dev. Biol.* 59, 10–26. doi:10.1016/j.semcdb.2016.04.009
- Nguyen, T. M., Bharti, S., Yue, Z., Willey, C. D., and Chen, J. Y. (2021). Statistical enrichment analysis of samples: A general-purpose tool to annotate metadata neighborhoods of biological samples. *Front. Big Data* 4, 725276. doi:10.3389/fdata.2021.725276
- Ning, W., Wei, Y., Gao, L., Han, C., Gou, Y., Fu, S., et al. (2022). HemI 2.0: An online service for heatmap illustration. *Nucleic Acids Res.* 50 (W1), W405–W411. doi:10.1093/nar/gkac480

- Orecchioni, M., Kobiyama, K., Winkels, H., Ghosheh, Y., McArdle, S., Mikulski, Z., et al. (2022). Olfactory receptor 2 in vascular macrophages drives atherosclerosis by NLRP3-dependent IL-1 production. *Science* 375, 214–221. doi:10.1126/science.abg3067
- Pan, M. M., Hockenberry, M. S., Kirby, E. W., and Lipshultz, L. I. (2018). Male infertility diagnosis and treatment in the era of *in vitro* fertilization and intracytoplasmic sperm injection. *Med. Clin. North Am.* 102, 337–347. doi:10.1016/j.mcna.2017.10.008
- Peña, V. N., Kohn, T. P., and Herati, A. S. (2020). Genetic mutations contributing to non-obstructive azoospermia. *Best. Pract. Res. Clin. Endocrinol. Metab.* 34, 101479. doi:10.1016/j.beem.2020.101479
- Petersen, C., Aumüller, G., Bahrami, M., and Hoyer-Fender, S. (2002). Molecular cloning of Odf3 encoding a novel coiled-coil protein of sperm tail outer dense fibers. *Mol. Reprod. Dev.* 61, 102–112. doi:10.1002/mrd.1136
- Poongothai, J., Gopenath, T. S., and Manonayaki, S. (2009). Genetics of human male infertility. *Singap. Med. J.* 50, 336–347.
- Puig, R. R., Holmås, S., Mironov, V., and Kuiper, M. (2020). Network building with the Cytoscape BioGateway app explained in five use cases. *Curr. Protoc. Bioinforma.* 72 (1), e106. doi:10.1002/cpbi.106
- Safe, S., Jin, U. H., Morpurgo, B., Abudayyeh, A., Singh, M., and Tjalkens, R. B. (2016). Nuclear receptor 4A (NR4A) family - orphans no more. *J. Steroid Biochem. Mol. Biol.* 157, 48–60. doi:10.1016/j.jsbmb.2015.04.016
- Sarkar, S., Yadav, S., Mehta, P., Gupta, G., and Rajender, S. (2022). Histone methylation regulates gene expression in the round spermatids to set the RNA payloads of sperm. *Reprod. Sci.* 29, 857–882. doi:10.1007/s43032-021-00837-3
- Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (D1), D605–D612. doi:10.1093/nar/gkaa1074
- Tournaye, H., Krausz, C., and Oates, R. D. (2017). Concepts in diagnosis and therapy for male reproductive impairment. *Lancet. Diabetes Endocrinol.* 5, 554–564. doi:10.1016/S2213-8587(16)30043-2
- Walker, W. H. (2009). Molecular mechanisms of testosterone action in spermatogenesis. *Steroids* 74, 602–607. doi:10.1016/j.steroids.2008.11.017
- Wang, M., Liu, X., Chang, G., Chen, Y., An, G., Yan, L., et al. (2018). Single-cell RNA sequencing analysis reveals sequential cell fate transition during human spermatogenesis. *Cell Stem Cell* 23, 599–614. e4. doi:10.1016/j.stem.2018.08.007
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation.* 2 (3), 100141. doi:10.1016/j.xinn.2021.100141
- Zhao, L., Yao, C., Xing, X., Jing, T., Li, P., Zhu, Z., et al. (2020). Single-cell analysis of developing and azoospermia human testicles reveals central role of Sertoli cells. *Nat. Commun.* 11, 5683. doi:10.1038/s41467-020-19414-4
- Zhao, M., Liu, Y., Ding, G., Qu, D., and Qu, H. (2021). Online database for brain cancer-implicated genes: Exploring the subtype-specific mechanisms of brain cancer. *BMC Genomics* 22 (1), 458. doi:10.1186/s12864-021-07793-x
- Zheng, W., Zou, Z., Lin, S., Chen, X., Wang, F., Li, X., et al. (2019). Identification and functional analysis of spermatogenesis-associated gene modules in azoospermia by weighted gene coexpression network analysis. *J. Cell. Biochem.* 120, 3934–3944. doi:10.1002/jcb.27677