



OPEN ACCESS

EDITED BY

Katja Teerds,
Wageningen University and
Research, Netherlands

REVIEWED BY

Daniela Fietz,
University of Giessen, Germany
Taotao Sun,
First Affiliated Hospital of Zhengzhou
University, China

*CORRESPONDENCE

Li Liu

✉ uro_ly@163.com

Xiaoqiang Liu

✉ liuxiaoqiang2308@163.com

†These authors have contributed equally to
this work

RECEIVED 18 December 2023

ACCEPTED 22 March 2024

PUBLISHED 26 September 2024

CITATION

Pan Y, Chen X, Zhou H, Xu M, Li Y, Wang Q,
Xu Z, Ren C, Liu L and Liu X (2024)
Identification and validation of SHC1
and FGFR1 as novel immune-related
oxidative stress biomarkers of
non-obstructive azoospermia.
Front. Endocrinol. 15:1356959.
doi: 10.3389/fendo.2024.1356959

COPYRIGHT

© 2024 Pan, Chen, Zhou, Xu, Li, Wang, Xu,
Ren, Liu and Liu. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). 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.

Identification and validation of SHC1 and FGFR1 as novel immune-related oxidative stress biomarkers of non-obstructive azoospermia

Yang Pan[†], Xiangyu Chen[†], Hang Zhou[†], Mingming Xu[†],
Yuezheng Li[†], Qihua Wang, Zhunan Xu, Congzhe Ren,
Li Liu* and Xiaoqiang Liu*

Department of Urology, Tianjin Medical University General Hospital, Tianjin, China

Background: Non-obstructive azoospermia (NOA) is a major contributor of male infertility. Herein, we used existing datasets to identify novel biomarkers for the diagnosis and prognosis of NOA, which could have great significance in the field of male infertility.

Methods: NOA datasets were obtained from the Gene Expression Omnibus (GEO) database. CIBERSORT was utilized to analyze the distributions of 22 immune cell populations. Hub genes were identified by applying weighted gene co-expression network analysis (WGCNA), machine learning methods, and protein–protein interaction (PPI) network analysis. The expression of hub genes was verified in external datasets and was assessed by receiver operating characteristic (ROC) curve analysis. Gene set enrichment analysis (GSEA) was applied to explore the important functions and pathways of hub genes. The mRNA–microRNA (miRNA)–transcription factors (TFs) regulatory network and potential drugs were predicted based on hub genes. Single-cell RNA sequencing data from the testes of patients with NOA were applied for analyzing the distribution of hub genes in single-cell clusters. Furthermore, testis tissue samples were obtained from patients with NOA and obstructive azoospermia (OA) who underwent testicular biopsy. RT-PCR and Western blot were used to validate hub gene expression.

Results: Two immune-related oxidative stress hub genes (*SHC1* and *FGFR1*) were identified. Both hub genes were highly expressed in NOA samples compared to control samples. ROC curve analysis showed a remarkable prediction ability (AUCs > 0.8). GSEA revealed that hub genes were predominantly enriched in toll-like receptor and Wnt signaling pathways. A total of 24 TFs, 82 miRNAs, and 111 potential drugs were predicted based on two hub genes. Single-cell RNA sequencing data in NOA patients indicated that *SHC1* and *FGFR1* were highly expressed in endothelial cells and Leydig cells, respectively. RT-PCR and Western blot results showed that mRNA and protein levels of both hub genes were significantly upregulated in NOA testis tissue samples, which agree with the findings from analysis of the microarray data.

Conclusion: It appears that *SHC1* and *FGFR1* could be significant immune-related oxidative stress biomarkers for detecting and managing patients with NOA. Our findings provide a novel viewpoint for illustrating potential pathogenesis in men suffering from infertility.

KEYWORDS

non-obstructive azoospermia, immune cell, oxidative stress, WGCNA, machine learning, single-cell RNA-seq

Introduction

Infertility is defined as unsuccessful pregnancy after engaging in normal, unprotected sexual intercourse for more than 1 year (1). Approximately 30% of couples worldwide are infertile, and almost half of the cases are attributed to male factors (2). There is significant evidence to indicate that the incidence of male infertility has increased in recent years. Azoospermia is diagnosed as the inability to identify a single sperm cell in three consecutive ejaculations (3). As such, azoospermia is defined as an absolute spermatozoa deficiency in the ejaculate, even on microscopic examination of a cell pellet obtained by centrifugation of the semen sample (4). Azoospermia is a key cause of infertility in men; this condition can be defined as either obstructive azoospermia (OA) or non-obstructive azoospermia (NOA). NOA is found in approximately 10% of infertile men and 1% of the general male population (5), making it a primary cause of male infertility (6). The pathogenesis of NOA includes genetic, developmental, hormonal, environmental, and other reasons, and its management remains challenging (7). Following the development of testicular sperm extraction (TESE), the combination of TESE and intracytoplasmic sperm injection (ICSI) has been regarded as a primary management strategy for male patients with NOA (5, 8). However, retrieval of spermatozoa by TESE is only successful in approximately 50% of NOA cases, due largely to the heterogeneity of NOA. As a result, exploration of the exact pathogenesis and molecular markers for managing NOA cases is important.

The rapid advancement of high-throughput sequence techniques has greatly facilitated research on the genetic characteristics of spermatogenesis. EAU guidelines also illustrate the importance of karyotype abnormalities and Y chromosome microdeletions in NOA cases (4). While numerous reports have shown that a variety of genetic mutations could be associated with NOA, the failure of spermatogenesis may also be linked to various unknown factors (9). Overall, the pathogenesis of NOA is relatively complex and may result from the alternation of multiple genes rather than a single monogenetic factor. Identifying novel biomarkers with high statistical efficiency could hold significant value for assessing the clinical outcomes of patients with NOA.

Many previous studies have revealed that NOA was associated with immune (10, 11) and oxidative stress in testis tissues and cells

(12, 13). Gene microarray assessment of testis tissues could aid in identifying novel biomarkers for diagnosing and predicting prognosis in patients with NOA. In the present study, a weighted gene co-expression network analysis (WGCNA) method was applied to accurately explore the complicated molecular mechanisms and identify potential immune- and oxidative stress-associated gene biomarkers for NOA. WGCNA is commonly utilized to identify highly correlated module genes and confirm hub genes associated with phenotype characteristics (14). This novel method has been widely applied by a diverse range of biological researchers in the field of molecular genetics, cancer, and others. WGCNA differs from traditional analyses of microarray-generated differentially expressed genes (DEGs). This novel method can create various co-expression clusters to contextualize high-dimensional microarray data into fewer variables, presenting an enhanced view of associating gene clusters with phenotypes.

In the present study, we combined WGCNA and three machine learning methods to identify novel immune- and oxidative stress-related biomarkers for NOA based on testis tissue expression microarray. Subsequently, single-cell RNA sequencing data from the testes of patients with NOA were used to validate hub gene distribution in single-cell clusters. Furthermore, testis tissue samples were obtained from patients who underwent testicular biopsy. Hub gene expression levels were further validated by real-time polymerase chain reaction (RT-PCR) and Western blot. Overall, this study provides insight into the underlying pathogenesis and can contribute to the identification of crucial medicinal targets for patients with NOA suffering from infertility.

Materials and methods

Datasets

Figure 1 presents a flowchart of this study. NOA microarray data were obtained from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>). We selected the dataset GSE9210 to further identify hub genes, as it contained the largest sample size, including data from 47 patients with NOA and 11 OA controls.

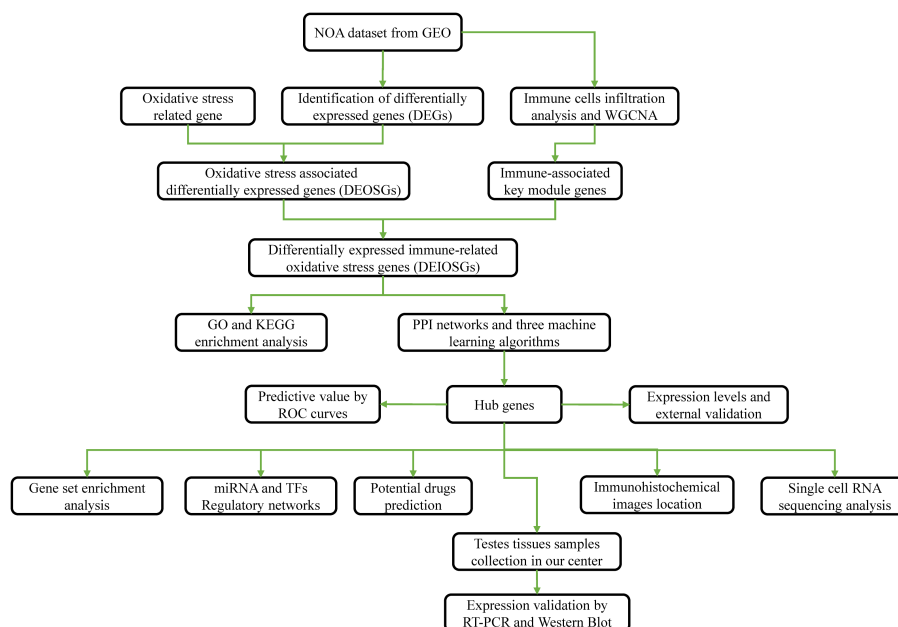


FIGURE 1
Flowchart for this study.

Identification of differentially expressed genes

The “limma” package (15) with the standard of $|\log_2$ (Fold Change) $| > 1$ and p -value < 0.05 was utilized to identify DEGs. The “pheatmap” and “ggplot2” packages were utilized to plot heatmap and volcanic map of DEGs, separately.

Moreover, 1,792 genes related to oxidative stress were obtained from the GeneCards database (<https://www.genecards.org/>) with a criterion of >5 scores. The 1,792 genes were intersected with DEGs to generate the differentially expressed oxidative stress-associated genes (DEOSGs).

Immune infiltration analysis and WGCNA constructio

CIBERSORT uses a deconvolution algorithm to assess the composition and abundance of immune cells according to microarray data (16). An online website, CIBERSORTx (<https://cibersortx.stanford.edu/>), was utilized to determine the distribution matrix of different immune cell populations between NOAs and controls in the GSE9210 dataset. The WGCNA method was used to identify candidate disease biomarkers, as described in our previous study (17). The “WGCNA” package was further used to perform WGCNA and identify the modules most strongly related to immune cells in patients with NOA. The sample data were preprocessed and outliers were removed. Subsequently, the “WGCNA” package was applied to create the correlation matrix. The optimal soft threshold was selected to convert the correlation matrix into an adjacency matrix, and a topological overlap matrix (TOM) was created from the adjacency matrix. The TOM-based phase dissimilarity metric was

utilized to categorize genes with similar expression patterns into gene modules using average linkage hierarchical clustering. The module with the significantly highest correlation to immune cells was regarded as the core module. Eventually, the genes in the core module were intersected with DEOSGs. The intersected genes were identified as differentially expressed immune-related oxidative stress genes (DEIOSGs) for subsequent analysis.

GO and KEGG enrichment analysis

To evaluate the functional and pathway outcomes of DEIOSGs, GO and KEGG enrichment were analyzed using the “clusterProfiler” package (18, 19). The “ggplot2” package was utilized to visualize enrichment analysis outcomes.

Identification of hub genes

Three machine learning algorithms—least absolute shrinkage and selection operator (LASSO) regression (20), support vector machine–recursive feature elimination (SVM-RFE) (21), and random forest (RF) (22)—were applied to identify hub genes based on DEIOSGs. The genes from each machine learning method were intersected to identify the common hub genes.

Furthermore, protein–protein interaction (PPI) networks based on DEIOSGs were generated using the STRING database (<https://cn.string-db.org/>). Subsequently, PPI networks were analyzed using the cytoHubba plugin in the Cytoscape software (v3.7.2). Twelve algorithms in the cytoHubba plugin were performed, after which the top 10 genes from every algorithm were intersected to identify common genes, as another hub gene.

Eventually, the genes identified by these two methods were regarded as hub genes.

Gene set enrichment analysis

Single-gene gene set enrichment analysis (GSEA) was conducted to explore the potential function of each hub gene using the “clusterProfiler” and “org.Hs.eg.db” packages (23).

Constructing the mRNA–miRNA–TFs network and predicting drugs

The JASPAR and TarBase databases extracted from the NetworkAnalyst (<https://www.networkanalyst.ca/>) were used to forecast hub genes encoding transcription factors (TF) and microRNAs (miRNA), respectively. In addition, the DSigDB database through the Enrichr platform (<https://amp.pharm.mssm.edu/Enrichr/>) was applied to forecast potential drugs.

Hub gene expression analysis and validation

Hub gene expression comparisons between the NOA and control groups were explored using data extracted from the GSE9210 database. To validate hub gene expression trends, we obtained another dataset (GSE145467), which included 10 NOA samples and 10 OA control samples. The Mann–Whitney test was applied to compare hub gene expressions.

Predictive value for NOA using hub genes

To evaluate the performance of hub genes, we assessed their ability to predict clinical phenotype by constructing receiver operating characteristic (ROC) curves using the “timeROC” package. Furthermore, samples from the validation set (GSE145467) were utilized to verify predictive accuracy. The predictive value of hub genes in each dataset was assessed using the areas under ROC curves (AUCs).

Immunohistochemical analysis

Immunohistochemical data and images were obtained from the Human Protein Atlas (HPA) database (<https://www.proteinatlas.org/>) to explore hub gene distribution in testis tissues.

Single-cell RNA sequencing analysis

A single-cell RNA sequencing (scRNA-seq) database for testis diseases, named the Male Health Atlas database (<http://malehealthatlas.cn/>), was utilized to assess hub gene distribution in single-cell clusters of testis tissues from patients with NOA.

malehealthatlas.cn/), was utilized to assess hub gene distribution in single-cell clusters of testis tissues from patients with NOA.

Testis tissue sample collection

Testis tissue samples were obtained from patients with NOA and OA controls who underwent testicular biopsy. Samples were immediately flash-frozen in liquid nitrogen for preservation. All patient procedures were performed in accordance with the Declaration of Helsinki and were approved by the Ethics Committee. Written informed consent was obtained from all cases in our study.

Expression validation by RT-PCR

Total RNA from testes was extracted using TRIzol reagent (Invitrogen) as per the manufacturer’s instructions. cDNA was synthesized using commercial kits. A qPCR commercial kit was also utilized to amplify transcript and analyze the mRNA levels of genes in testes. The reaction was performed in a 20- μ L volume comprising 10 μ L of mix, 10 μ M forward and reverse primers, 1 μ g of diluted cDNA sample, and RNase-free water. The optimal PCR conditions were 95°C for 120 s, followed by 40 cycles of 95°C for 15 s and 60°C for 30 s. After checking for reference gene suitability, beta-actin was utilized as an internal control gene. The relative mRNA expression levels of target genes were normalized to beta-actin, and results were analyzed using the $2^{-\Delta\Delta CT}$ method. The primers were purchased from Sangon Biotech (Shanghai). All primers were tested and quality controlled by this company and met the standards. The primers are shown in Additional file 1 (Supplementary Table S1).

Expression validation by Western blot

Total protein was extracted as per the manufacturer’s instructions (cat. no. R0020; Solarbio, China). Protein concentration was measured using BCA Protein Quantitation Assay (cat. no. PC0020; Solarbio, China). Protein samples were separated using SDS-PAGE and transferred onto polyvinylidene fluoride (PVDF) membranes (cat. no. FFP26, Beyotime, China). The membranes were then blocked in non-fat powdered milk (cat. no. D8340; Solarbio, China) and incubated overnight in primary antibodies at 4°C. Primary antibodies included rabbit anti-beta-actin (1:5,000; cat. no. bs-0061R; Bioss, China), rabbit anti-FGFR1 (1:1,000; cat. no. 60325-1-Ig; Proteintech, USA), and rabbit anti-SHC1 (1:1,000; cat. no. NBP3-21850; Novus Biologicals, USA). Subsequently, the PVDF membranes were incubated in the secondary antibody (1:5,000; cat. no. bs-0295G-HRP; Bioss, China) for 1 h, then washed three times in TBS-T. Finally, the ECL Plus kit (cat. no. PE0010; Solarbio, China) was applied to expose the immunoreactive blots. ImageJ software was used to scan the blot density. Hub gene protein expression was normalized to the expression of beta-actin.

Statistical analysis

R Project software (version 4.2.2) was used to analyze and visualize the data. Graphics were created using GraphPad Prism 8.0 software. p -values < 0.05 were considered as statistically significant.

Results

Oxidative stress-associated differentially expressed genes

In total, 1,194 DEGs were identified from the GSE9210 dataset. The volcano map of all DEGs (Figure 2A) and the heatmap of the top 40 DEGs (Figure 2B) were constructed. Moreover, 1,792 oxidative stress-associated genes were obtained from the GeneCards database with a criterion of >5 scores. After the intersection, 109 DEOSGs were identified (Figure 2C).

Immune cell infiltration analysis and WGCNA results

Application of the CIBERSORT algorithm revealed that 8/22 immune cells (plasma cells, CD8⁺ T cells, CD4⁺ naïve T cells, CD4⁺ memory-activated T cells, follicular helper T cells, monocytes, M2 macrophages, and resting mast cells) were found to be significantly different between the NOA and OA control groups (Figures 3A–C).

WGCNA revealed that the soft-threshold power was calibrated to 9 ($R^2 = 0.86$) (Figure 4A), and a total of nine modules were identified (Figures 4B, C). Among them, the yellow module had the most strongly positive correlations with CD8⁺ T cells (Figure 4C). Owing to its significance in relation to immune infiltrating cells, the yellow module including 189 genes was chosen for further investigation. Nineteen DEOSGs were identified as the intersection between the 109 DEOSGs and the yellow module genes (Figure 5A).

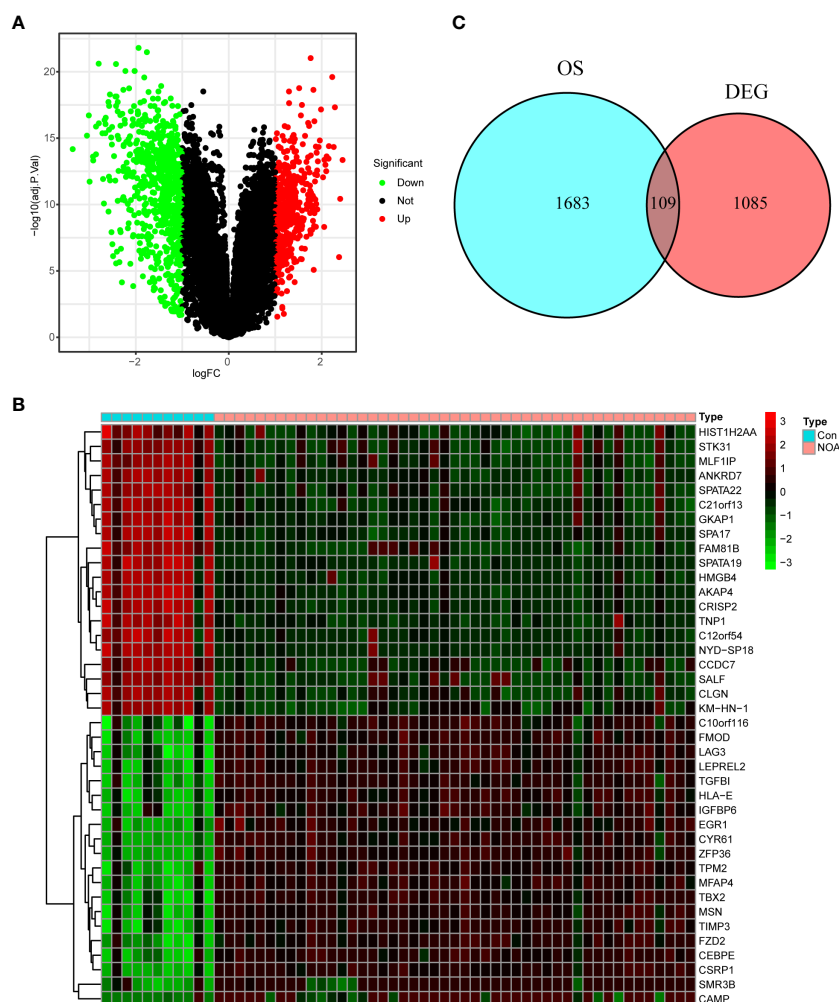


FIGURE 2
Identification of oxidative stress-associated differentially expressed genes (DEOSGs). (A) Volcano plot of differentially expressed genes (DEGs). (B) Heatmap of the top 40 DEGs. (C) Venn diagrams of generating DEOSGs.

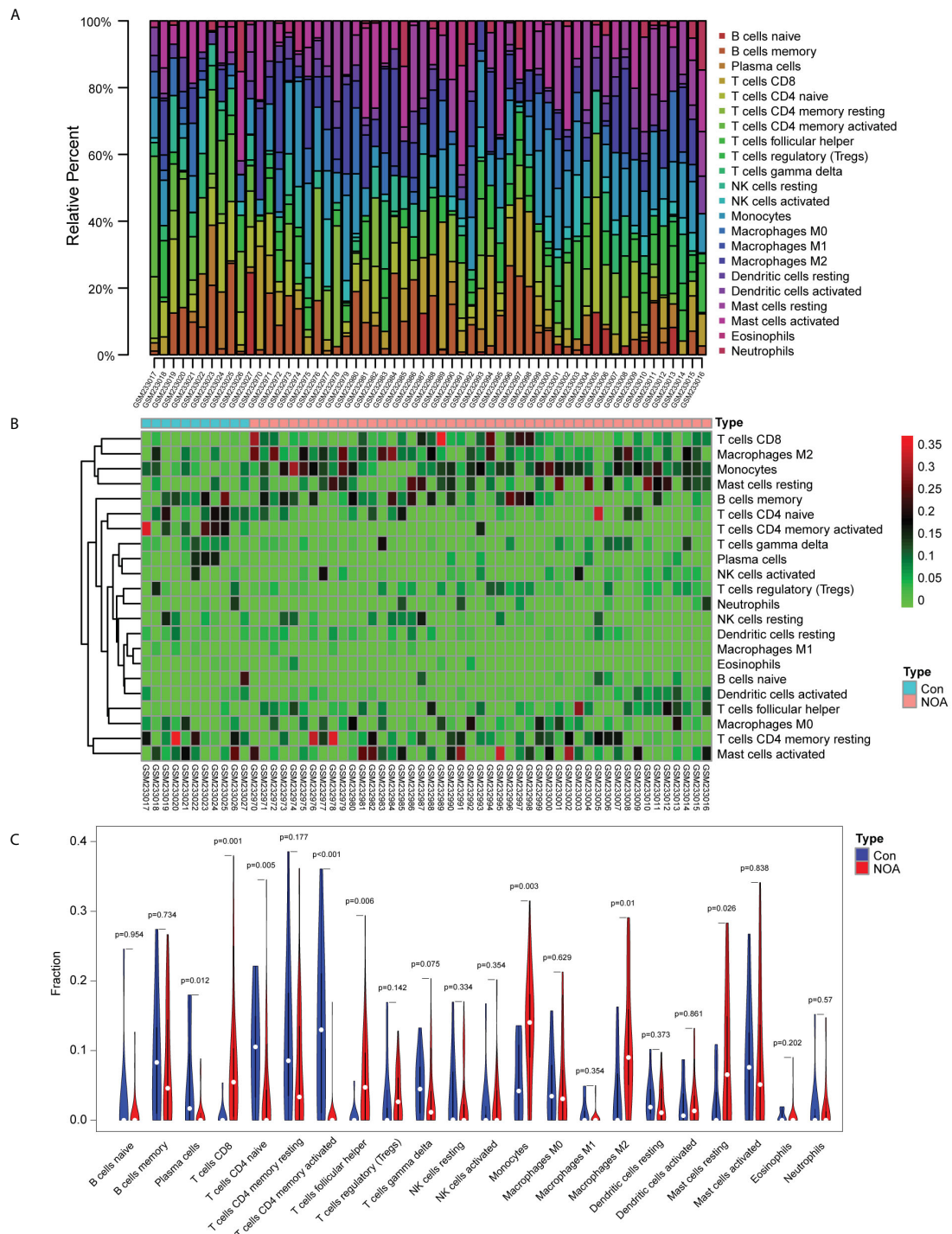


FIGURE 3 Immune infiltration analysis based on the CIBERSORT algorithm and non-obstructive azoospermia (NOA) dataset. **(A)** Relative percentage of 22 immune cells in each sample. **(B)** Heatmap of 22 immune cells in each sample. **(C)** Comparison of 22 immune cells between non-obstructive azoospermia (NOA) samples and control samples.

GO and KEGG

GO and KEGG enrichment analysis based on 19 DEIOSGs were conducted. The biological processes of GO analysis were engaged in the regulation of reactive oxygen species metabolic process, mesenchymal

cell differentiation, and other functions. The cellular components of the GO analysis included endoplasmic reticulum lumen, the external side of the plasma membrane, and collagen-containing extracellular matrix. The molecular functions of GO analysis predominantly included growth factor binding, proteoglycan binding, and testosterone

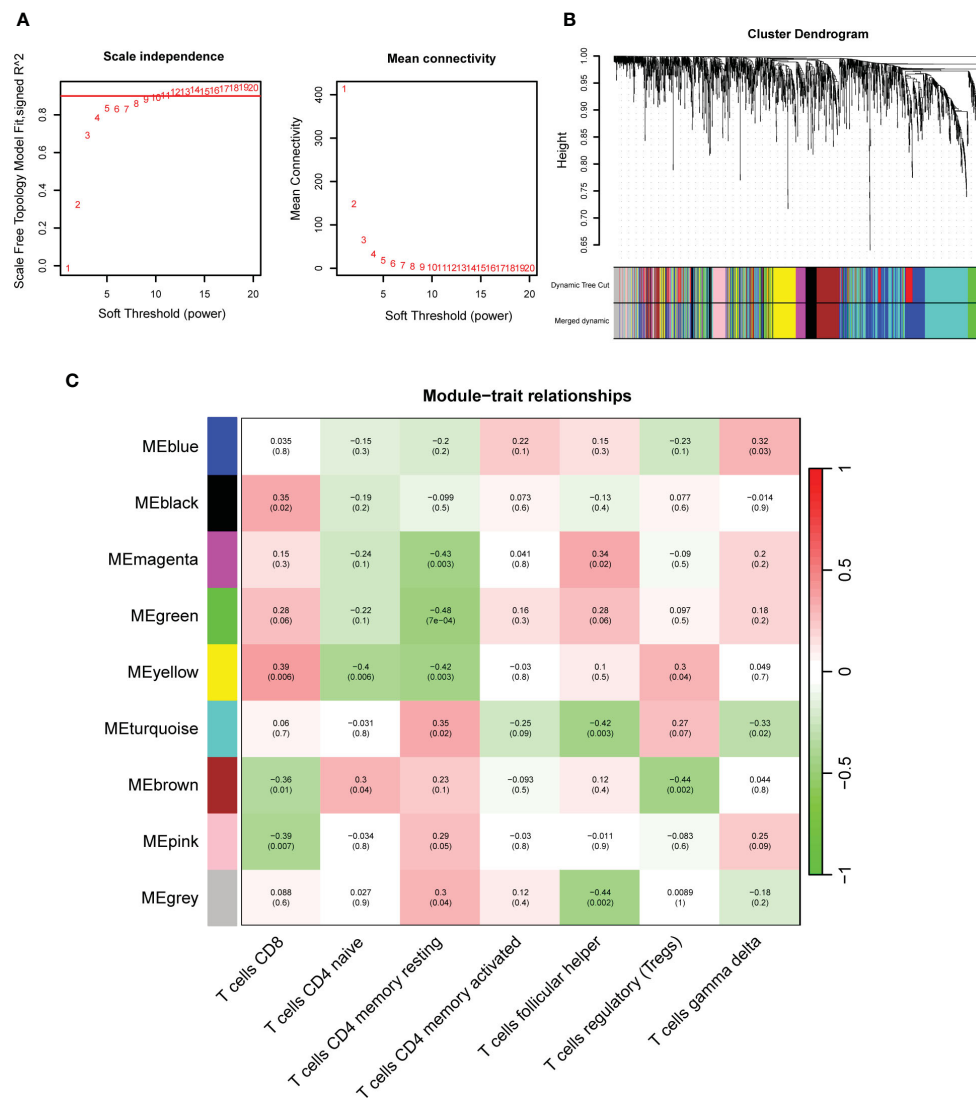


FIGURE 4

Construction of weighted gene co-expression networks analysis (WGCNA). (A) Choosing the best soft-threshold power. (B) Dynamic tree cut and merged dynamic in WGCNA. (C) Nine immune-related gene modules revealed by the WGCNA.

dehydrogenase activity (Figures 5B, C). KEGG analysis of 19 DEIOSGs was predominantly engaged in human papillomavirus infection, MAPK signaling pathway, steroid hormone biosynthesis, and other pathways (Figure 5D).

Identification of hub genes

From the 19 DEIOSGs, 2, 2, and 3 genes were identified using LASSO (Figure 6A), SVM-RFE (Figure 6B), and RF (Figure 6C) analyses, respectively. Intersection of these gene lists identified one common gene, *SHC1* (Figure 7C). Furthermore, we identified another hub gene, *FGFR1*, through the intersection of genes from 12 cytoHubba algorithms in PPI networks (Figures 7A, B, D). Eventually, these two hub genes (*SHC1* and *FGFR1*) were identified as novel immune-related oxidative stress biomarkers for NOA.

Expression validation

In NOA samples, the expression of the two hub genes was found to be significantly higher than those in the OA control samples (Figures 8A, B). This observation was further confirmed by analyzing an external NOA dataset (GSE145467), which further revealed that two hub genes were more highly expressed in NOA samples compared to OA control samples (Figures 8C, D). All of these comparisons were statistically significant.

ROC curve for predicting the NOA phenotype

ROC curve analysis was performed and the AUCs were calculated to detect the diagnostic accuracy of the two hub genes.

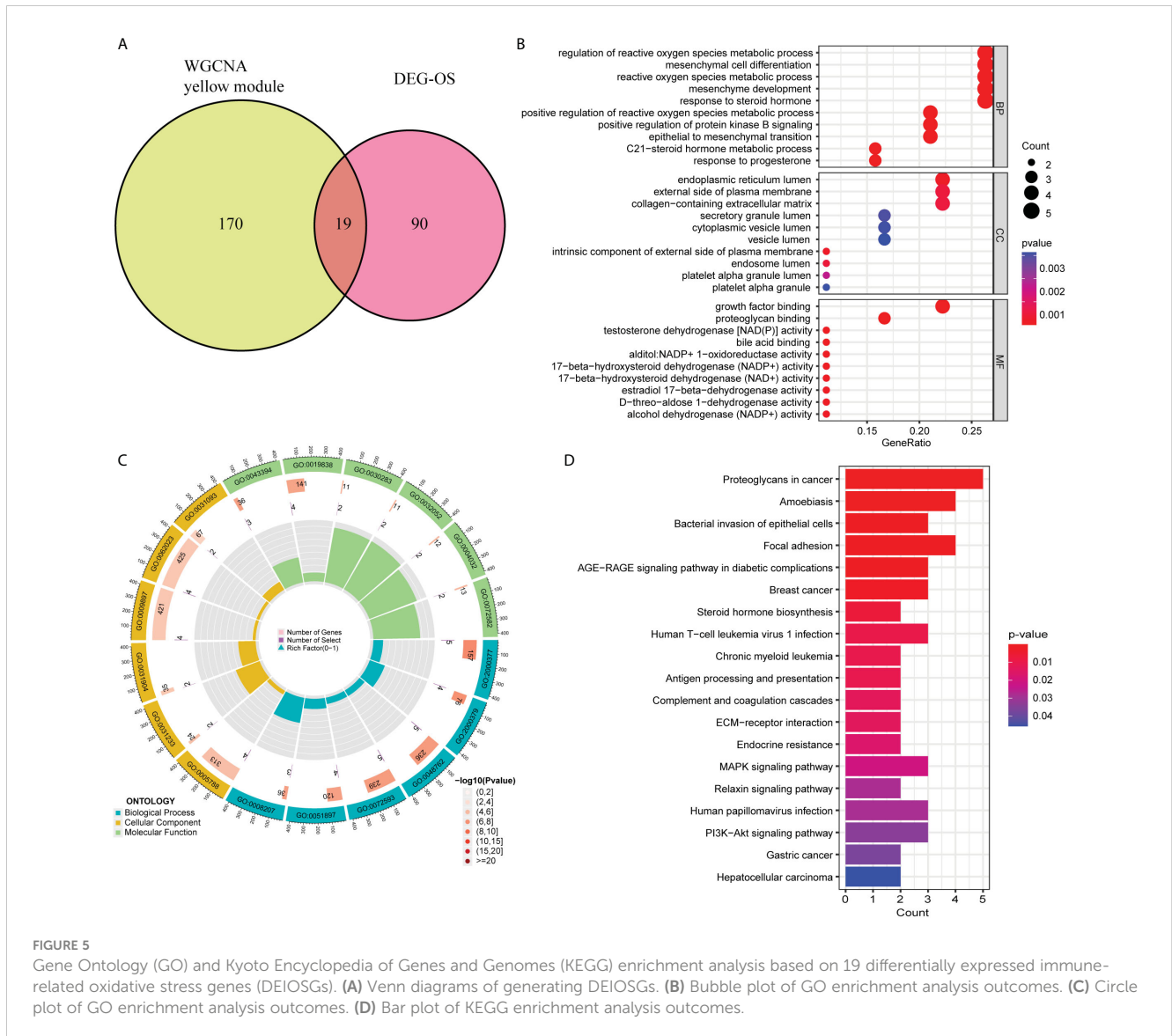


FIGURE 5

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis based on 19 differentially expressed immune-related oxidative stress genes (DEIOSGs). (A) Venn diagrams of generating DEIOSGs. (B) Bubble plot of GO enrichment analysis outcomes. (C) Circle plot of GO enrichment analysis outcomes. (D) Bar plot of KEGG enrichment analysis outcomes.

To predict the NOA phenotype, AUCs were 0.99 (95% CI: 0.965–1.0) for *SHC1* and 0.954 (95% CI: 0.861–1.0) for *FGFR1* (Figures 9A, B) in the integrated training group (GSE9210). This observation was further confirmed through analysis of an external NOA dataset (GSE145467), which revealed that the AUCs of *SHC1* and *FGFR1* were 0.84 (95% CI: 0.64–0.98) and 0.87 (95% CI: 0.66–1.000), respectively (Figures 9C, D). All AUCs in these datasets for predicting the NOA phenotype were relatively high.

Single-gene GSEA assessment

GSEA of the hub gene *SHC1* was predominantly enriched in allograft rejection, autoimmune thyroid disease, leishmania infection, toll-like receptor (TLR) signaling pathway, and type I diabetes mellitus (Figure 10A). GSEA of the hub gene *FGFR1* revealed predominant enrichment in cardiac muscle contraction, drug metabolism cytochrome P450, leishmania infection, leukocyte trans-endothelial migration, and the WNT signaling pathway (Figure 10B).

Regulatory network construction and potential drug prediction

Evaluation of the two hub genes in the JASPAR database identified 24 TFs (Figure 10C). Among them, two (*FOXC1* and *PPARG*) had a degree of ≥ 2 . Using the TarBase database, 82 miRNAs were predicted (Figure 10C). Among them, 11, including has-mir-200b-3p and has-mir-124-3p, had a degree of ≥ 2 . An overall regulatory network comprising hub genes, TFs, and miRNAs was further constructed (Figure 10C). Moreover, a total of 111 potential drugs were predicted from the DSigDB database based on the two hub genes (Additional file 1 Supplementary Table S2).

Immunohistochemical location

The localization distribution of hub genes in human testis tissues was explored in the HPA database. *SHC1* was highly positive in spermatogonia (Figure 11A). *SHC1* was moderately

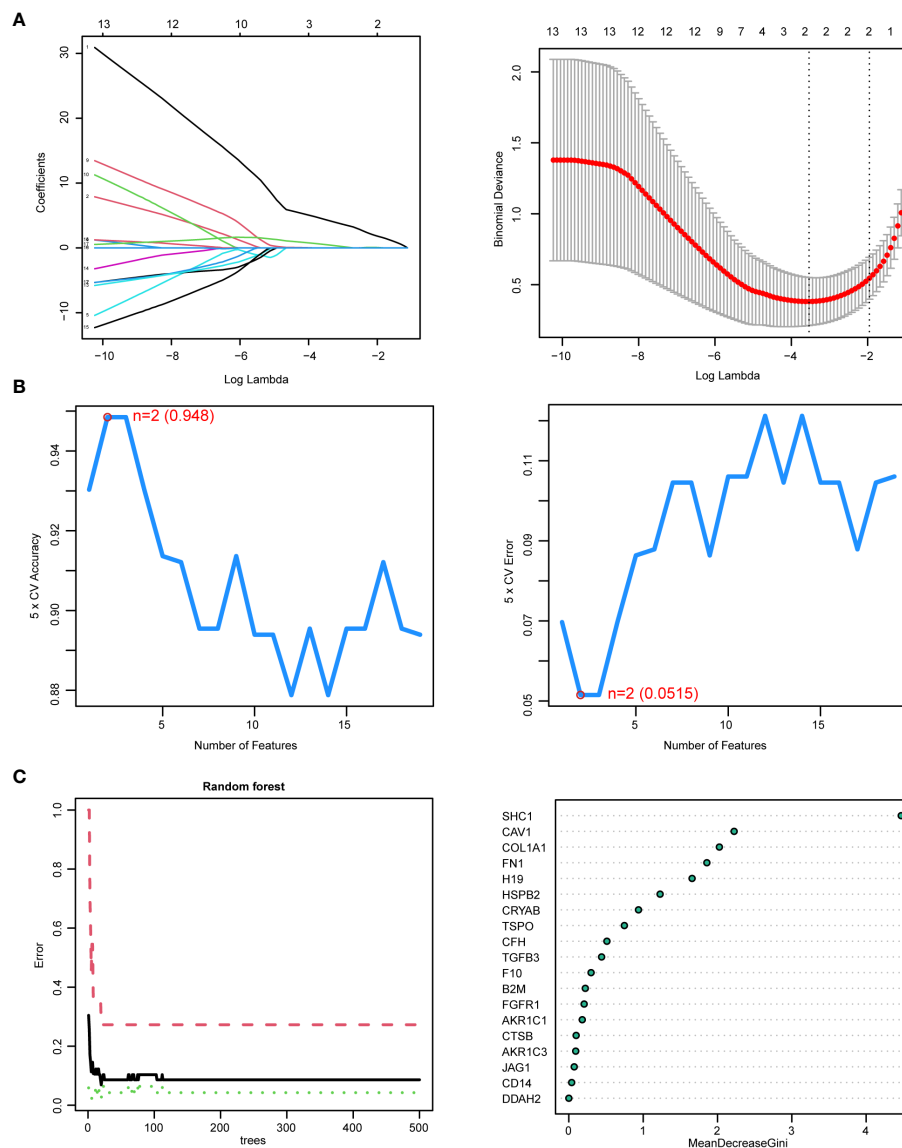


FIGURE 6

Hub genes identified by three machine learning algorithms based on 19 differentially expressed immune-related oxidative stress genes (DEIOSGs). **(A)** Outcomes of least absolute shrinkage and selection operator (LASSO) regression algorithm. (Left) LASSO plot showed that the variations in the size of coefficients for parameters shrank as the value of k penalty increased. (Right) LASSO logic coefficient penalty diagram. **(B)** Outcomes of the support vector machine–recursive feature elimination (SVM-RFE) algorithm. (Left) The relationship between the prediction accuracy of SVM-RFE and the number of features. (Right) The relationship between the prediction error rate of SVM-RFE and the number of features. **(C)** Outcomes of random forest (RF) algorithm. (Left) The error rate confidence intervals for random forest mode. (Right) The dot graph illustrating the relative importance of genes in the random forest model.

expressed in Leydig cells, Sertoli cells, pachytene spermatocytes, preleptotene spermatocytes, and round or early spermatids. However, *SHC1* expression was negative in peritubular cells and elongated or late spermatids. Moreover, *FGFR1* was lowly expressed in the seminiferous tubules and Leydig cells (Figure 11B).

scRNA-seq data analysis

The scRNA-seq data were analyzed to assess the distribution of *SHC1* and *FGFR1* in 11 cell clusters from the human testis NOA atlas in the Male Health Atlas database (Figures 12A, B). The results

revealed that *SHC1* was the most highly expressed in endothelial cells (Figures 12C, E), while *FGFR1* was the most highly expressed in Leydig, peritubular myoid, endothelial, and vascular smooth muscle cells (Figures 12D, F).

RT-PCR and Western blot validation in clinical samples

We collected testis tissue samples from patients with NOA and patients with OA who underwent tissue sampling. The gene expression levels of the two hub genes were validated using RT-

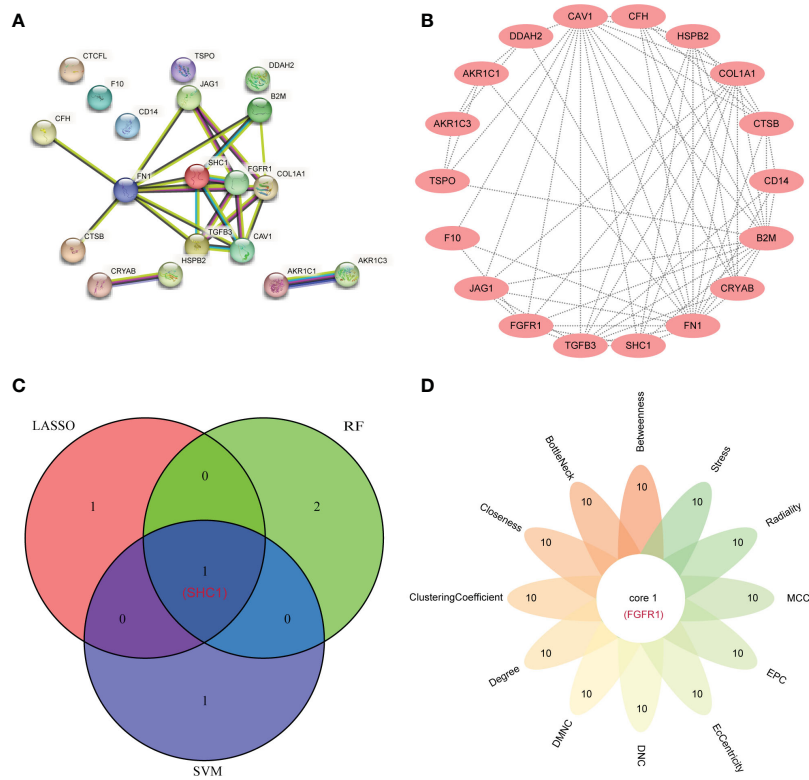


FIGURE 7 Hub gene identification by the protein-protein interaction (PPI) network. **(A)** The PPI network from the STRING database based on 19 differentially expressed immune-related oxidative stress genes (DEIOSGs). **(B)** PPI network visualization using the Cytoscape software. **(C)** Venn diagrams for intersecting the genes from three machine learning methods. **(D)** Flower Venn diagrams for intersecting the genes from 12 algorithms using the cytoHubba plugin in the Cytoscape software.

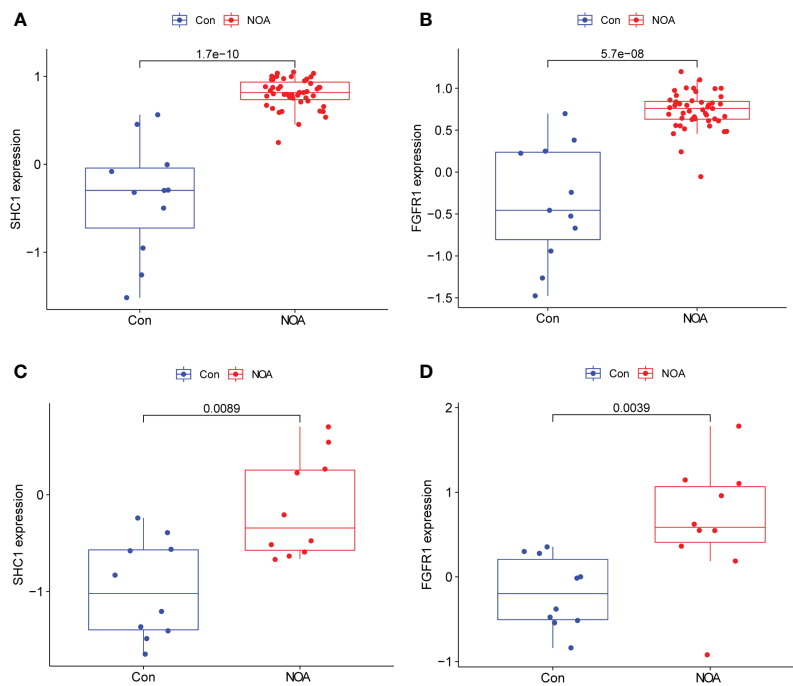


FIGURE 8 Hub gene expression and external validation. **(A, B)** Expression of two hub genes in the GSE9210 dataset. **(C, D)** Expression validation of two hub genes in an external dataset, GSE145467.

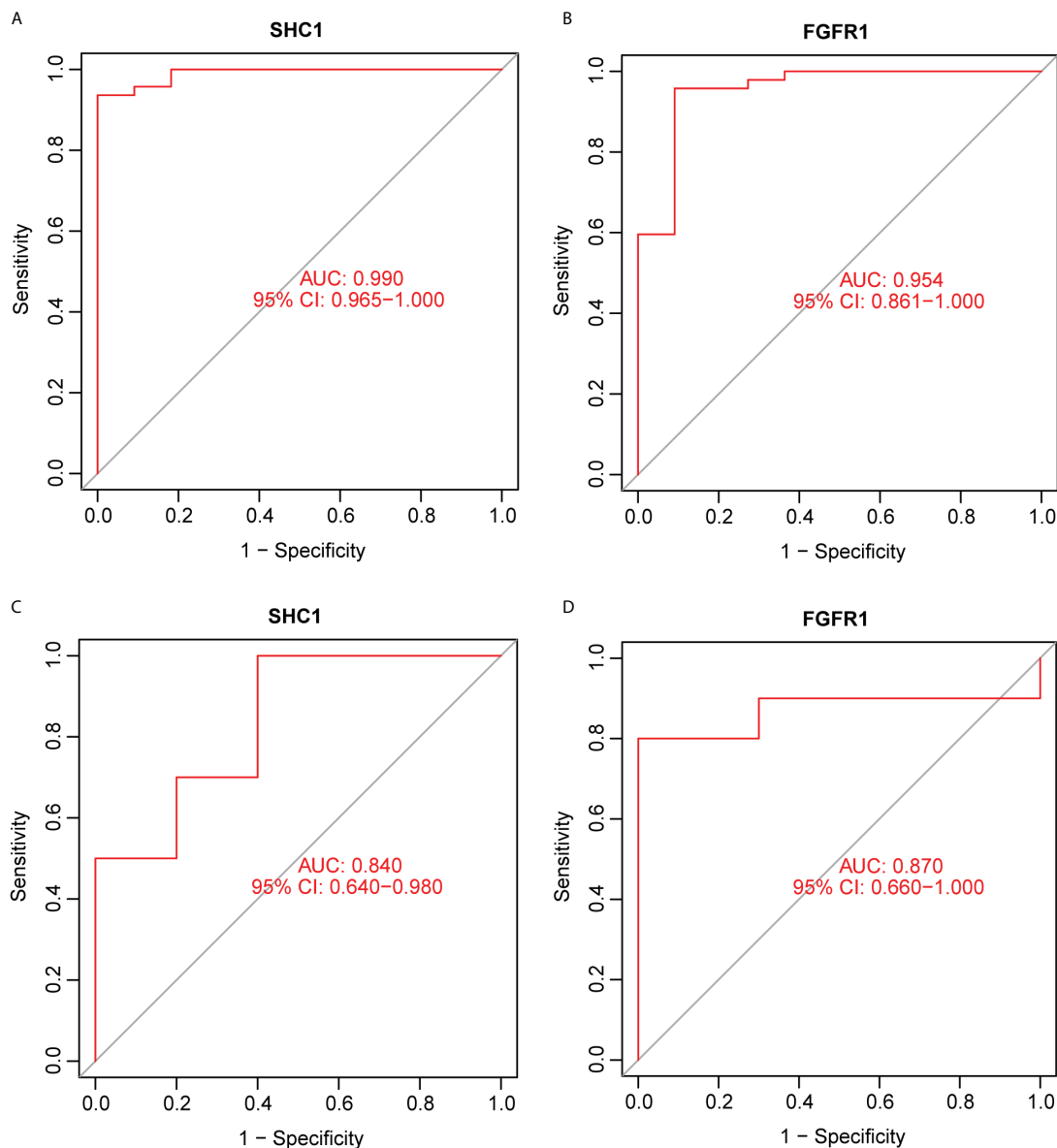


FIGURE 9

Receiver operating characteristic (ROC) curve for predicting the non-obstructive azoospermia (NOA) phenotype using hub genes. (A, B) Two hub genes were analyzed using ROC curves for predicting the NOA phenotype in the GSE9210 dataset. (C, D) Two hub genes were analyzed using ROC curves for predicting the NOA phenotype based on an external dataset, GSE145467.

PCR. RT-PCR results revealed that two hub genes were statistically different between NOA and OA testis tissue samples (Figure 13A). Compared to that in OA control samples, mRNA expression of *SHC1* and *FGFR1* were significantly upregulated in NOA testis tissue samples, which is consistent with the findings from the microarray data. Furthermore, we used testis tissue samples of patients with NOA and OA to validate the protein expression levels of the two hub genes using Western blot (Figure 13B). Similarly, the protein expression profiles of *SHC1* and *FGFR1* were both significantly upregulated in the NOA group than in the OA control group (Figure 13C), which is consistent with the RT-PCR results.

Discussion

Men with NOA generally present with varying severities and etiologies of abnormalities in spermatogenesis, while the exact pathogenesis remains unclear. Some previous reports have mentioned that genetic alternations could be associated with spermatogenesis abnormalities (24–26). The current management strategies for NOA are limited due to individual differences, necessitating the development of novel molecular markers that can aid in the pathogenesis of NOA. Thus, identifying novel biomarker genes is vital to improve the management of patients with NOA. Immune infiltration and oxidative stress have been

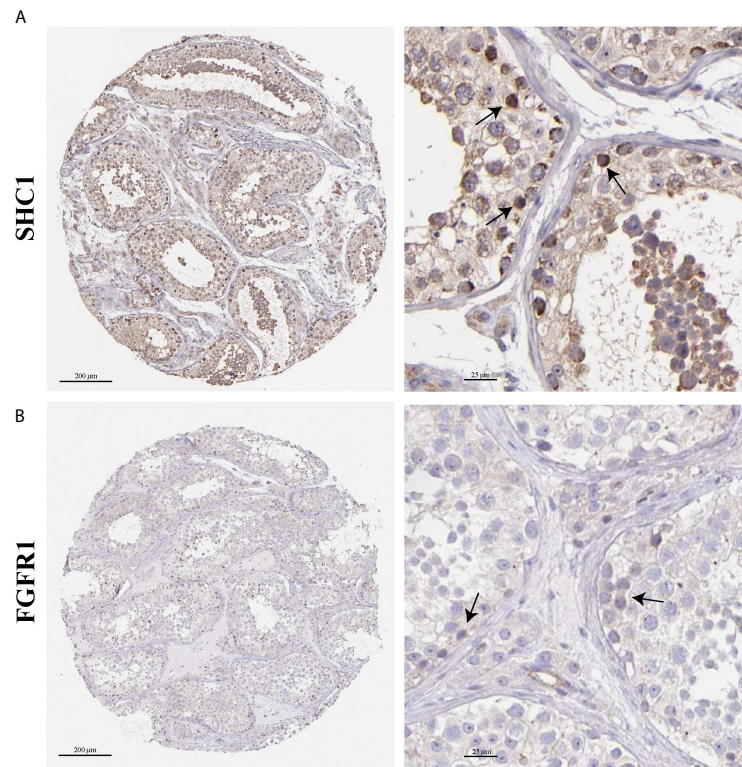


FIGURE 11

Immunohistochemistry images on hub gene expression based on testis tissues from the Human Protein Atlas (HPA) database. (A) Representative immunohistochemistry results of SHC1. (B) Representative immunohistochemistry results of FGFR1. Black arrow indicates positive cells.

Acting as the adapter protein, *SHC1* possessed SH2 domains, which were shared among signal proteins of the cytoplasm (31, 32). This was related to numerous receptor-related signal processes including antigen, hormone, and growth factor (33). In contrast to their non-transformed counterparts, numerous types of transformed cells contained *SHC1*, which was also highly phosphorylated (34). *SHC1* regulates cancer growth owing to its tumor-specific activation, suggesting that it served as a useful predictive marker and a target for therapeutic intervention (35, 36). However, the role of *SHC1* in NOA has not received much attention in the literature. The present study showed that *SHC1* was significantly increased in the testis tissues of NOA cases, indicating an extremely high diagnostic accuracy. Furthermore, functional enrichment analysis indicated that oxidative stress and immune processes were related to *SHC1*, which may be the primary mechanism responsible for *SHC1* participation in NOA.

FGFR1, a member of the fibroblast growth factor receptor (FGFR) family, plays a crucial role in biological process (37). Fibroblast growth factor (FGF) binds to the extracellular domain of FGFR, resulting in its activation. Subsequently, receptor dimerization occurs, leading to the phosphorylation of the C-terminal tyrosine (38). The key kinase and pathways were phosphorylated and activated as a result of the activation of the FGF tyrosine kinase family by various FGF ligands, thus regulating multiple physiological responses such as embryogenesis and angiogenesis (39). In various human malignancies, *FGFR1* gene translocation, mutation, and amplification can result in the

abnormal activation of the FGFR signaling system, thereby promoting carcinogenesis and tumor progression (40–42). FGF can also promote angiogenesis by activating *FGFR1* in endothelial cells (43). For example, *FGFR1* activation contributes to the epithelial–mesenchymal transition (EMT) and metastasis in breast cancer (44), as well as the carcinogenesis and EMT of prostate cancer (45). Consequently, *FGFR1* was associated with the emergence of many disease and function disorders. However, studies on the involvement of *FGFR1* in NOA development are lacking. The present study showed that *FGFR1* was highly expressed in the testes of patients with NOA, and it could be related to the pathogenesis of this condition. This finding opens the door for *FGFR1* to be used as a novel NOA diagnostic marker.

Single-gene GSEA of the hub gene *SHC1* revealed enrichment of the TLR signaling pathway. TLR has the ability to detect the molecular mode associated with causative agents. This detection triggers a series of gene expression alternations, which work in concert to eliminate harmful bacteria (46). The TLR signaling pathway plays a direct role in the activation, growth, differentiation, development, and functioning of T cells in various physiological activities (47). TLR is of utmost importance in the protection against infections, immunodeficiency, and tumor growth. Based on our findings, the TLR signaling pathway may be activated by the high expression of *SHC1* in testis tissues and result in inflammation and activation of the immune response, thus mediating the occurrence of NOA. However, this finding

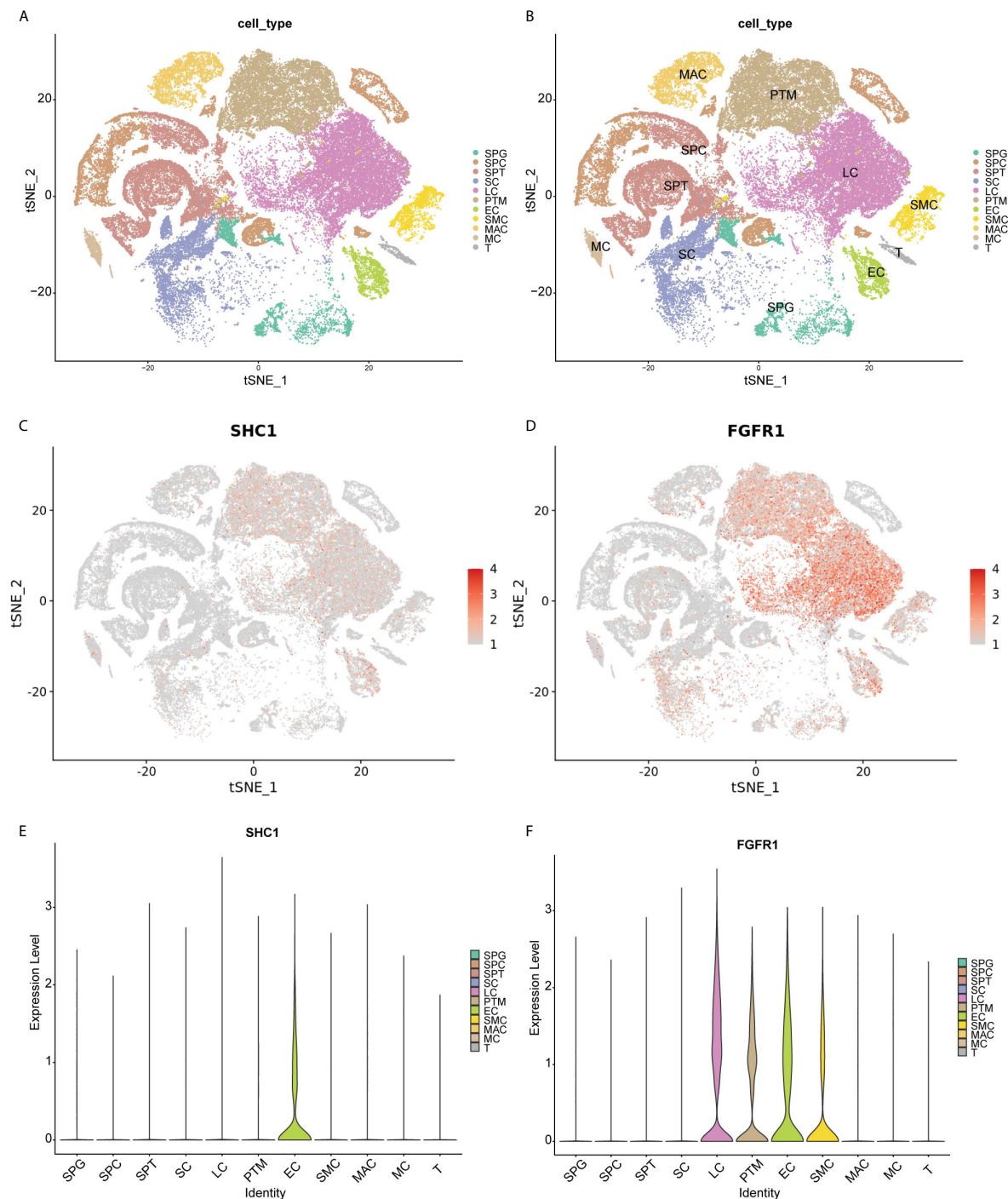


FIGURE 12

Single-cell RNA sequencing data analysis. (A, B) T-distribution stochastic neighbor embedding (TSNE) analysis outcomes of 11 cell clusters based on human testis non-obstructive azoospermia atlas in the Male Health Atlas database. (C, D) The expression distribution of two hub genes (SHC1 and FGFR1) in 11 cell clusters. (E, F) The violin plot showing the expression distribution of two hub genes (SHC1 and FGFR1) in 11 cell clusters. SPG, spermatogonia; SPC, spermatocyte; SPT, spermatids/sperm; SC, Sertoli cell; LC, Leydig cell; PTM, peritubular myoid cell; EC, endothelial cell; SMC, vascular smooth muscle cell; MAC, macrophage; MC, mast cell; T, T cell.

necessitates additional confirmation through *in vivo* and *in vitro* experimentation.

Single-gene GSEA of the hub gene FGFR1 revealed enrichment of the Wnt signaling pathway. The Wnt pathway plays a pivotal role in various evolutionary and illness-related processes. As such, it is a

crucial factor in the development and maintenance of organic tissue functions by regulating their native stem cells (48). Recent studies have further shed light on the involvement of the Wnt signaling pathway in the differentiation of human primordial germ cells and the maintenance of mouse spermatogonial

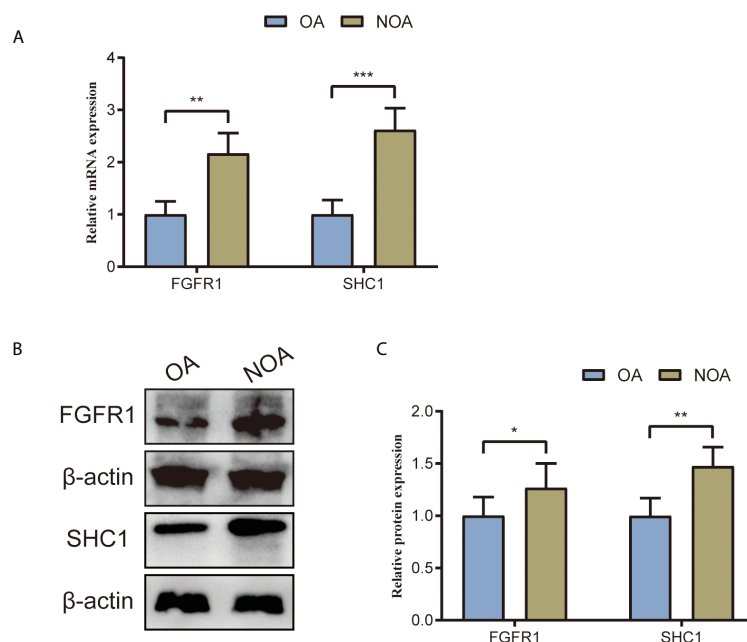


FIGURE 13

Two hub genes' (SHC1 and FGFR1) mRNA and protein expression levels validated by RT-PCR and Western blot using testes tissue samples collected from patients with NOA and OA in our center. (A) Relative mRNA expression of SHC1 and FGFR1. (B) Western blot outcomes of two hub genes (SHC1 and FGFR1). (C) Relative protein expression of SHC1 and FGFR1. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

stem cells (49–51). Moreover, several previous studies have reported that the downregulation of Wnt signaling in spermatogonia was strongly related to the development of NOA (52). These studies further proposed the significance of the Wnt signaling pathway in inducing human spermatogonial stem cells. Additionally, they indicated that the absence of spermatogonial stem cells in testis tissues of azoospermia could be linked to the inactivation of the Wnt signaling pathway. Similarly, our study also showed that the Wnt signaling pathway significantly decreased in the testis tissues of patients with NOA. Furthermore, our results indicated that FGFR1-mediated Wnt signaling pathway alternations could be a significant aspect of the pathogenesis of NOA.

Our study has some limitations that should be considered. First, our findings are predominantly based on publicly available data. Despite performing expression validation using testicular samples, further experimental validation with larger and multi-center samples is required to validate the results. Second, although the HPA is an extremely helpful tool to verify the protein expression levels in normal tissues, the formalin-fixed material depicts a minor cellular preservation and the spermatogenesis does not seem to be intact. It will be more encouraging to perform immunohistochemistry on the testis samples in the future. Our study revealed that oxidative stress- and immune-related hub genes, *SHC1* and *FGFR1*, could be primary factors in the development of NOA, and these genes are closely interconnected. The findings of immune analysis indicate that mast cells, T cells, monocytes, and macrophages were crucial to the pathogenesis of NOA. Moreover, enrichment analysis indicated that hub genes were primarily concentrated in TLR and Wnt

signaling pathways, and were associated with the occurrence of NOA. Thus, focusing on these hub genes and important signaling pathways might be crucial and hopeful. Moreover, in this study, we identified two key genes using machine learning and WGCNA, with validation of their importance performed using single-cell RNA-seq data, RT-PCR, and Western blot. As such, our study has the potential to contribute to a deeper understanding of the pathogenesis of NOA and could further aid in the identification of crucial targets for pharmaceutical exploration.

Conclusions

This study identified *SHC1* and *FGFR1* as oxidative stress- and immune-related hub genes. These hub genes were significantly upregulated in the testes of patients with NOA and had good predictive values for the NOA phenotype. The TLR and Wnt signaling pathways might be related to the development of NOA. *SHC1* and *FGFR1* could be novel immune-related oxidative stress biomarkers and crucial targets for the pathogenesis of patients with NOA.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Tianjin Medical University General Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YP: Investigation, Methodology, Writing – original draft. XC: Conceptualization, Validation, Visualization, Writing – review & editing. HZ: Formal Analysis, Visualization, Writing – review & editing. MX: Data curation, Formal Analysis, Software, Writing – review & editing. YL: Visualization, Data curation, Formal Analysis, Writing – review & editing. QW: Validation, Formal Analysis, Writing – review & editing. ZX: Validation, Writing – review & editing. CR: Validation, Visualization, Writing – review & editing. LL: Project administration, Resources, Supervision, Writing – review & editing. XL: Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Our study

was supported by the National Natural Science Foundation of China (No. 82171594).

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/fendo.2024.1356959/full#supplementary-material>

References

- Carson SA, Kallen AN. Diagnosis and management of infertility: A review. *Jama*. (2021) 326:65–76. doi: 10.1001/jama.2021.4788
- Nagirnaja L, Mørup N, Nielsen JE, Stakaitis R, Golubickaite I, Oud MS, et al. Variant PNLDC1, defective piRNA processing, and azoospermia. *N Engl J Med*. (2021) 385:707–19. doi: 10.1056/NEJMoa2028973
- Aziz N. The importance of semen analysis in the context of azoospermia. *Clinics (Sao Paulo)*. (2013) 68 Suppl 1:35–8. doi: 10.6061/clinics/2013(Sup01)05
- 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) 80:603–20. doi: 10.1016/j.eururo.2021.08.014
- Kherraf ZE, Cazin C, Bouker A, Fourati Ben Mustapha S, Hennebicq S, Septier A, et al. Whole-exome sequencing improves the diagnosis and care of men with non-obstructive azoospermia. *Am J Hum Genet*. (2022) 109:508–17. doi: 10.1016/j.ajhg.2022.01.011
- Tournaye H, Krausz C, Oates RD. Novel concepts in the aetiology of male reproductive impairment. *Lancet Diabetes Endocrinol*. (2017) 5:544–53. doi: 10.1016/S2213-8587(16)30040-7
- Chiba K, Enatsu N, Fujisawa M. Management of non-obstructive azoospermia. *Reprod Med Biol*. (2016) 15:165–73. doi: 10.1007/s12522-016-0234-z
- Corona G, Minhas S, Giwercman A, Bettocchi C, Dinkelman-Smit M, Dohle G, et al. Sperm recovery and ICSI outcomes in men with non-obstructive azoospermia: a systematic review and meta-analysis. *Hum Reprod Update*. (2019) 25:733–57. doi: 10.1093/humupd/dmz028
- Krausz C, Riera-Escamilla A. Genetics of male infertility. *Nat Rev Urol*. (2018) 15:369–84. doi: 10.1038/s41585-018-0003-3
- Menzies FM, Shepherd MC, Nibbs RJ, Nelson SM. The role of mast cells and their mediators in reproduction, pregnancy and labour. *Hum Reprod Update*. (2011) 17:383–96. doi: 10.1093/humupd/dmq053
- Salzmann M, Tosev G, Heck M, Schadendorf D, Maatouk I, Enk AH, et al. Male fertility during and after immune checkpoint inhibitor therapy: A cross-sectional pilot study. *Eur J Cancer*. (2021) 152:41–8. doi: 10.1016/j.ejca.2021.04.031
- Aitken RJ, Baker MA. The role of genetics and oxidative stress in the etiology of male infertility—A unifying hypothesis? *Front Endocrinol (Lausanne)*. (2020) 11:581838. doi: 10.3389/fendo.2020.581838
- Moubasher AE, Taha EA, Younis A, Fakhry ME, Morsy H. Testicular tissue oxidative stress in azoospermic patients: Effect of cryopreservation. *Andrologia*. (2020) 52:e13817. doi: 10.1111/and.13817
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf*. (2008) 9:559. doi: 10.1186/1471-2105-9-559
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res*. (2015) 43:e47. doi: 10.1093/nar/gkv007
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. (2015) 12:453–7. doi: 10.1038/nmeth.3337
- Xu M, Zhou H, Hu P, Pan Y, Wang S, Liu L, et al. Identification and validation of immune and oxidative stress-related diagnostic markers for diabetic nephropathy by WGCNA and machine learning. *Front Immunol*. (2023) 14:1084531. doi: 10.3389/fimmu.2023.1084531
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. (2021) 2:100141. doi: 10.1016/j.xinn.2021.100141
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118
- Li Z, Sillanpää MJ. Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. *Theor Appl Genet*. (2012) 125:419–35. doi: 10.1007/s00122-012-1892-9
- Lin X, Li C, Zhang Y, Su B, Fan M, Wei H. Selecting feature subsets based on SVM-RFE and the overlapping ratio with applications in bioinformatics. *Molecules*. (2017) 23. doi: 10.3390/molecules23010052
- Ghosh D, Cabrera J. Enriched random forest for high dimensional genomic data. *IEEE/ACM Trans Comput Biol Bioinform*. (2022) 19:2817–28. doi: 10.1109/TCBB.2021.3089417

23. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U.S.A.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
24. Cioppi F, Rosta V, Krausz C. Genetics of azoospermia. *Int J Mol Sci.* (2021) 22. doi: 10.3390/ijms22063264
25. Du L, Chen W, Cheng Z, Wu S, He J, Han L, et al. Novel gene regulation in normal and abnormal spermatogenesis. *Cells.* (2021) 10. doi: 10.3390/cells10030666
26. Eddy EM. Regulation of gene expression during spermatogenesis. *Semin Cell Dev Biol.* (1998) 9:451–7. doi: 10.1006/scdb.1998.0201
27. Zheng W, Zhang S, Jiang S, Huang Z, Chen X, Guo H, et al. Evaluation of immune status in testis and macrophage polarization associated with testicular damage in patients with nonobstructive azoospermia. *Am J Reprod Immunol.* (2021) 86:e13481. doi: 10.1111/aji.13481
28. Wright KD, Miller BS, El-Meanawy S, Tsaih SW, Banerjee A, Geurts AM, et al. The p52 isoform of SHC1 is a key driver of breast cancer initiation. *Breast Cancer Res.* (2019) 21:74. doi: 10.1186/s13058-019-1155-7
29. Ma S, Rubin BP. Apoptosis-associated tyrosine kinase 1 inhibits growth and migration and promotes apoptosis in melanoma. *Lab Invest.* (2014) 94:430–8. doi: 10.1038/labinvest.2014.13
30. Jackson JG, Yoneda T, Clark GM, Yee D. Elevated levels of p66 Shc are found in breast cancer cell lines and primary tumors with high metastatic potential. *Clin Cancer Res.* (2000) 6:1135–9.
31. Pelicci G, Lanfrancone L, Grignani F, McGlade J, Cavallo F, Forni G, et al. A novel transforming protein (SHC) with an SH2 domain is implicated in mitogenic signal transduction. *Cell.* (1992) 70:93–104. doi: 10.1016/0092-8674(92)90536-L
32. Ravichandran KS. Signaling via Shc family adapter proteins. *Oncogene.* (2001) 20:6322–30. doi: 10.1038/sj.onc.1204776
33. Mood K, Saucier C, Ishimura A, Bong YS, Lee HS, Park M, et al. Oncogenic Met receptor induces cell-cycle progression in *Xenopus* oocytes independent of direct Grb2 and Shc binding or Mos synthesis, but requires phosphatidylinositol 3-kinase and Raf signaling. *J Cell Physiol.* (2006) 207:271–85. doi: 10.1002/jcp.20564
34. Furlan T, Khalid S, Nguyen AV, Günther J, Troppmair J. The oxidoreductase p66Shc acts as tumor suppressor in BRAFV600E-transformed cells. *Mol Oncol.* (2018) 12:869–82. doi: 10.1002/1878-0261.12199
35. Li X, Xu Z, Du W, Zhang Z, Wei Y, Wang H, et al. Aiolos promotes anchorage independence by silencing p66Shc transcription in cancer cells. *Cancer Cell.* (2014) 25:575–89. doi: 10.1016/j.ccr.2014.03.020
36. Pawlicki JM, Cookmeyer DL, Maseda D, Everett JK, Wei F, Kong H, et al. NPM-ALK-induced reprogramming of mature TCR-stimulated T cells results in dedifferentiation and Malignant transformation. *Cancer Res.* (2021) 81:3241–54. doi: 10.1158/0008-5472.CAN-20-2297
37. Turner N, Grose R. Fibroblast growth factor signalling: from development to cancer. *Nat Rev Cancer.* (2010) 10:116–29. doi: 10.1038/nrc2780
38. Schlessinger J, Plotnikov AN, Ibrahim OA, Eliseenkova AV, Yeh BK, Yayon A, et al. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol Cell.* (2000) 6:743–50. doi: 10.1016/S1097-2765(00)00073-3
39. Wang K, Ji W, Yu Y, Li Z, Niu X, Xia W, et al. FGFR1-ERK1/2-SOX2 axis promotes cell proliferation, epithelial-mesenchymal transition, and metastasis in FGFR1-amplified lung cancer. *Oncogene.* (2018) 37:5340–54. doi: 10.1038/s41388-018-0311-3
40. Carstens JL, Shahi P, Van Tsang S, Smith B, Creighton CJ, Zhang Y, et al. FGFR1-WNT-TGF- β signaling in prostate cancer mouse models recapitulates human reactive stroma. *Cancer Res.* (2014) 74:609–20. doi: 10.1158/0008-5472.CAN-13-1093
41. Fontugne J, Wong J, Cabel L, Neyret-Kahn H, Karboul N, Maillé P, et al. Progression-associated molecular changes in basal/squamous and sarcomatoid bladder carcinogenesis. *J Pathol.* (2023) 259:455–67. doi: 10.1002/path.6060
42. Jones DT, Hutter B, Jäger N, Korshunov A, Kool M, Warnatz HJ, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat Genet.* (2013) 45:927–32. doi: 10.1038/ng.2682
43. Zhu X, Qiu C, Wang Y, Jiang Y, Chen Y, Fan L, et al. FGFR1 SUMOylation coordinates endothelial angiogenic signaling in angiogenesis. *Proc Natl Acad Sci U.S.A.* (2022) 119:e2202631119. doi: 10.1073/pnas.2202631119
44. Servetto A, Kolipara R, Formisano L, Lin CC, Lee KM, Sudhan DR, et al. Nuclear FGFR1 regulates gene transcription and promotes antiestrogen resistance in ER(+) breast cancer. *Clin Cancer Res.* (2021) 27:4379–96. doi: 10.1158/1078-0432.CCR-20-3905
45. Cheng C, Wang J, Xu P, Zhang K, Xin Z, Zhao H, et al. Gremlin1 is a therapeutically targetable FGFR1 ligand that regulates lineage plasticity and castration resistance in prostate cancer. *Nat Cancer.* (2022) 3:565–80. doi: 10.1038/s43018-022-00380-3
46. Lim KH, Staudt LM. Toll-like receptor signaling. *Cold Spring Harb Perspect Biol.* (2013) 5:a011247. doi: 10.1101/cshperspect.a011247
47. Duan T, Du Y, Xing C, Wang HY, Wang RF. Toll-like receptor signaling and its role in cell-mediated immunity. *Front Immunol.* (2022) 13:812774. doi: 10.3389/fimmu.2022.812774
48. Duchartre Y, Kim YM, Kahn M. The Wnt signaling pathway in cancer. *Crit Rev Oncol Hematol.* (2016) 99:141–9. doi: 10.1016/j.critrevonc.2015.12.005
49. Kojima Y, Sasaki K, Yokobayashi S, Sakai Y, Nakamura T, Yabuta Y, et al. Evolutionarily distinctive transcriptional and signaling programs drive human germ cell lineage specification from pluripotent stem cells. *Cell Stem Cell.* (2017) 21:517–32.e5. doi: 10.1016/j.stem.2017.09.005
50. Yeh JR, Zhang X, Nagano MC. Indirect effects of Wnt3a/ β -catenin signalling support mouse spermatogonial stem cells *in vitro*. *PLoS One.* (2012) 7:e40002. doi: 10.1371/journal.pone.0040002
51. Yamaji M, Seki Y, Kurimoto K, Yabuta Y, Yuasa M, Shigeta M, et al. Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet.* (2008) 40:1016–22. doi: 10.1038/ng.186
52. Wang X, Qu M, Li Z, Long Y, Hong K, Li H. Valproic acid promotes the *in vitro* differentiation of human pluripotent stem cells into spermatogonial stem cell-like cells. *Stem Cell Res Ther.* (2021) 12:553. doi: 10.1186/s13287-021-02621-1