Original Article

Identification of biomarkers associated with macrophage infiltration in non-obstructive azoospermia using single-cell transcriptomic and microarray data

Xi Luo[1,2,3,4,5#^], Haishan Zheng[1,2,3#], Zhen Nai[1,2,3], Mingying Li[1,2,3], Yonggang Li[1,2,3], Na Lin[1,2,3], Yunxiu Li[1,2,3], Ze Wu[1,2,3^]

[1]Department of Reproductive Medicine, The First People's Hospital of Yunnan Province, Kunming, China; [2]Reproductive Medical Center of Yunnan Province, The Affiliated Hospital of Kunming University of Science and Technology, Kunming, China; [3]NHC Key Laboratory of Periconception Health Birth in Western China, Kunming, China; [4]Faculty of Life science and Technology, Kunming University of Science and Technology, Kunming, China; [5]Medical school, Kunming University of Science and Technology, Kunming, China

Contributions: (I) Conception and design: X Luo; (II) Administrative support: Z Wu, Yunxiu Li, N Lin; (III) Provision of study materials or patients: X Luo, H Zheng; (IV) Collection and assembly of data: X Luo, H Zheng; (V) Data analysis and interpretation: X Luo, H Zheng, N Lin, Yonggang Li, Z Nai, M Li, Yunxiu Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Yunxiu Li; Na Lin; Xi Luo. Department of Reproductive Medicine, The First People's Hospital of Yunnan Province, 157 Jinbi Road, Kunming, China. Email: lyrarshow@qq.com; kunhualinna@163.com; luoxi_ypfph@163.com.

Background: Non-obstructive azoospermia (NOA) is a common clinical cause of male infertility. Research suggests that macrophages are linked to testicular function; however, their involvement in NOA remains unknown.

Methods: To evaluate the importance of macrophages infiltration in NOA and identify the macrophage-related biomarkers, the gene-expression microarray data GSE45885 and the single-cell transcriptomic data GSE149512 were utilized from the Gene Expression Omnibus (GEO). A single-sample gene set enrichment analysis (ssGSEA) was conducted to investigate immune cell proliferation. The Seurat package was used for the single-cell data analysis, and the limma package was used to identify the differentially expressed genes between the NOA and normal samples. Moreover, we conducted a weighted gene co-expression network analysis (WGCNA) to identify the macrophage-related key modules and genes, and conducted Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses for the functional exploration. To identify the macrophage-related biomarkers, we conducted least absolute shrinkage and selection operator (LASSO) and support vector machine-recursive feature elimination (SVM-RFE) analyses. Real-time quantitative polymerase chain reaction (RT-qPCR) was used to verify the marker genes present in NOA.

Results: We confirmed that open reading frame 72 gene on chromosome 9 (*C9orf72*) [area under the curve (AUC) =0.861] and cartilage-associated protein (*CRTAP*) (AUC =0.917) were the hub genes of NOA, and the RT-qPCR analysis revealed the critical expression of both genes in NOA.

Conclusions: Through the combination of tissue transcriptomic and single-cell RNA-sequencing analyses, we concluded that macrophage infiltration is significant in different subtypes of NOA, and we hypothesized that *C9orf72* and *CRTAP* play critical roles in NOA due to their high expression in macrophages.

Keywords: Non-obstructive azoospermia (NOA); immune cell infiltration; macrophages; Gene Expression Omnibus (GEO); bioinformatics

^ ORCID: Xi Luo, 0000-0002-9779-3697; Ze Wu, 0000-0002-8899-0445.

Submitted Oct 21, 2022. Accepted for publication Dec 29, 2022. Published online Jan 31, 2023.
doi: 10.21037/atm-22-5601
View this article at: https://dx.doi.org/10.21037/atm-22-5601

Introduction

The prevalence of infertility in men has been reported to be almost equal to that in women (1), and there is no doubt that male infertility represent an obstacle to couples seeking to achieve pregnancy (2). Non-obstructive azoospermia (NOA) is defined as the absence of sperm production and has been associated with infertility in men (3); however, but unlike obstructive azoospermia (OA), the etiology underlying NOA is complex. The intrinsic cause of NOA is spermatogenic failure; however, many other aspects may also be involved. In a European survey of male-infertility reports, the pathogenies of azoospermia were comprehensively summarized as varicocele, undescended testis, a malignant testicular tumor, hypogonadism, and genetic abnormalities (3). The causes of NOA are complex; however, some treatments can be beneficial and can result in healthy biological offspring.

Intracytoplasmic sperm injection (ICSI) combined with testicular sperm extraction (TESE) and microdissection-TESE have been applied in clinical practice for many years to great success (4,5). However, the first step in treatment is to accurately diagnose the defect(s) and appropriately classify its(their) pathology. In relation to the diagnosis of NOA, testicular histopathology has shown distinct patterns [e.g., hypo-spermatogenesis, germ-cell arrest, Sertoli cell-only syndrome (SCOS), and seminiferous tubule hyalinization have been recorded in testicular histopathologic reports concerning NOA] (6). However, it has not yet been confirmed whether there are commonalities among these subtypes, and this issue requires further elucidation. Additionally, there continue to be many unknowns regarding the molecular diagnosis of NOA (7,8).

The human testis principally consists of seminiferous tubules and Leydig cells, and the tubules are the locus of spermatogenesis and comprise the spermatogenic epithelium and Sertoli cells. There are millions of germ cells at different phases within the tubules that develop synchronously. Moreover, the Leydig cells, which are located in the interstitium of the testis and are attached to the seminiferous tubules, are the cells primarily responsible for the production of testosterone. Approximately 95% of testosterone is synthesized in the testis, and the remaining 5% comes from the zona reticularis of the adrenal gland (9). The testicular interstitium also contains immune cells, vascular cells, and peritubular cells that are reported to be essential for spermatogenesis (10). It has been suggested that the testis is immunoprivileged and can resist autoimmunity (11,12).

The main role of macrophages in testis is maintenance of organ homeostasis. However, numerous studies have indicated that macrophages are involved in the process of testicular spermatogenesis (13-16). A unique population of testicular macrophages resides in close proximity to undifferentiated spermatogonia and expresses spermatogonial proliferation and differentiation-inducing proteins, including colony-stimulating factor 1 (CSF1) and enzymes involved in retinoic acid production (17). Moreover, investigators have established that macrophages not only directly affect spermatogenesis, but also affect male fertility by regulating the Leydig cells (10). There is also evidence of mast cell involvement in male infertility (18), and researchers recently hypothesized that mast cells further compromise spermatogenesis by provoking tubular hyalinization and sclerosis (19). Although many biomarkers related to NOA were found already (20). However, the molecular mechanism(s) underlying immune cell action in NOA have not yet been fully elucidated, and thus, more molecular researches focusing on the immune cells in NOA is needed.

In the present study, we adopted both data sets of single-

Highlight box

Key findings

- *C9orf72* and *CRTAP* may play critical roles in macrophage infiltration within NOA.

What is known and what is new?

- The main role of macrophages in testis is maintenance of organ homeostasis.
- Macrophages infiltration and polarization are involved in the process of testicular spermatogenesis.
- High expression of *C9orf72* and *CRTAP* in the macrophage infiltration of NOA were demonstrated.

What is the implication, and what should change now?

- A different perspective may be taken on the pathogenesis of NOA, particularly in terms of immune infiltration.

cell ribonucleic acid (RNA)-sequencing (scRNA-seq) and tissue transcriptomic sequencing approaches to examine macrophage infiltration in NOA. The scRNA-seq data showed that each NOA subtype had its own distinctive feature in the macrophage profile. Using bioinformatics technologies, we also identified the target macrophage differentially expressed genes (mDEGs) in NOA and identified appropriate biomarkers. Our findings may shed light on the molecular underlying of macrophage action in NOA, and thus lead to improvements in clinical treatments. We present the following article in accordance with the STREGA reporting checklist (available at https://atm.amegroups.com/article/view/10.21037/atm-22-5601/rc).

Methods

Data acquisition and sample collection

Herein study conformed to the provisions of the Declaration of Helsinki (as revised in 2013). We extracted gene-expression microarray data from the GSE45885 data set and single-cell transcriptomic data from the GSE149512 data set from the Gene Expression Omnibus (GEO) database. The GSE45885 (21) data set contained the gene-expression profile data of 31 testicular tissues as follows: 4 normal and 11 manifesting post-meiotic arrest in NOA, 7 at NOA meiotic arrest, 2 at NOA pre-meiotic arrest, and 7 from NOA Sertoli cell-only syndrome samples. The GSE149512 (22) data set included single-cell transcriptomic data from 17 testicular tissues from 10 healthy samples, 3 from Klinefelter syndrome (KS) patients, 1 chromosome Yq azoospermia factor a (AZFa) microdeletion (AZFa_DEL), and 3 idiopathic NOA (iNOA) samples. To further validate our findings, NOA and normal samples were collected from The First People's Hospital of Yunnan Province with the approval of the Hospital's Ethics Committee (Approval ID: KHLL2020-KY012) and evaluated by real-time quantitative polymerase chain reaction (RT-qPCR). The samples were collected under the condition of fully informed consent before the patient took part in the testicular sperm aspiration (TESA) procedure and were confirmed by pathological evaluation. We collected a total of 26 samples (18 from men with NOA and 8 from OA).

Single-cell analysis for cell clustering and macrophage identification

After conducting a quality control check of the GSE149512 data set and ensuring that it met the requirements for a single-cell analysis, we conducted a single-cell analysis to determine the presence of macrophages in NOA. The data were normalized, and the JackStraw function was employed to perform the principle component analysis (PCA) to reduce dimensionality. Unsupervised clustering was then performed using the FindNeighbors and FindClusters functions in the Seurat package, and the resolution parameters ranged from 0.01–0.15. The principal clusters were visualized via t-distributed stochastic neighbor embedding (t-SNE) and uniform manifold approximation and projection (UMAP) using the shared nearest neighbor modularity-optimization algorithm. We then applied the FindAllMarkers function, which was configured with a min.pct of 0.2, a log fold change (FC) threshold of 0.25, and a P value of 0.05, and searched for positive marker genes in distinct cell clusters and assessed variance using the Wilcoxon method. Marker genes from the cell clusters were matched with the CellMarker database (23) to identify cell types (particularly macrophages), and the SingleR algorithm was then used to validate the identified cell types.

Estimation of immune infiltration

To evaluate the abundance of immune cells in the GSE45885 data set, we conducted a single-sample gene set enrichment analysis (ssGSEA) (24), which computed a separate enrichment score for each sample, and then employed the Wilcoxon test to compare immune cell infiltration between the NOA and control samples.

WGCNA

To identify the macrophage-related key module genes in NOA, we conducted a weighted gene co-expression network analysis (WCGNA) (25) in the GSE45885 data set. Soft-thresholding was then applied by enhancing the difference between strong and weak correlations with a power of 23; this soft-thresholding power enabled us to achieve a network topology that approximated scale-free status. The modules were then segmented using a dynamic tree-cutting algorithm to construct the module-clustering graph, and the modules with highly correlated macrophages were merged. We determined the correlation between the gene expression and sample trait (macrophage cell score) using the following criteria: a gene significance value >0.2; and a module membership value >0.6.

Screening DEGs and the development of a characteristic model

The limma package (26) was used to screen the DEGs between the normal and NOA samples in the GSE45885 data set. The selection criteria were as follows: a |log2 FC| value >0.5 and a P value <0.05. The FindAllMarker function was then used to screen the DEGs of the macrophage cluster in the GSE149512 data set, and the following settings were used: a min.pct of 0.2, a log FC threshold of 0.25, and a P value of 0.05. The intersection of the 2 DEG data sets was then intersected with the key module genes from the WGCNA, and we ultimately obtained the macrophage DEGs for all the NOA subtypes.

To further validate the ability of the candidate genes to discriminate between NOA patients and healthy men, we used the glmnet and caret packages to execute the least absolute shrinkage and selection operator (LASSO) (27) and support vector machine (SVM)-recursive feature elimination (RFE) (28) algorithms for the candidate genes, respectively. The characteristic genes in the NOA samples were found by merging the candidate genes obtained using the 2 methods.

RTq-PCR

To further validate the findings from the bioinformatics computations, we implemented RT-qPCR to compare the expression levels of open reading frame (ORF) 72 gene on chromosome 9 (*C9orf72*) and cartilage-associated protein (*CRTAP*) in the normal and NOA samples. RNA from the testicular tissues was collected in TRIzol (Invitrogen, USA), and complementary deoxyribonucleic acid (cDNA) was synthesized using a First-Strand cDNA Synthesis Kit (Servicebio, China). Finally, RT-qPCR was performed using the Bio-Rad CFX Connect real-time system (Bio-Rad, USA). The following gene primers were used: (I) *C9orf72* forward, ATGAGTCAGGGCTCTTTGTA and reverse, TCTATGTGTGTGGTGGGATA; (II) *CRTAP* forward, GCTGCTCACACCTTTCTACT and reverse, GTTCCTCTTCATCATTTCGT; and (III) internal reference *GAPDH* forward, CCCATCACCATCTTCCAGG and reverse, CATCACGCCACAGTTTCCC. The PCR program cycle was as follows: pre-denaturation for 1 min at 95 ℃, denaturation for 20 s at 95 ℃, annealing for 20 s at 55 ℃, and synthesis for 30 s at 72 ℃ for a total of 40 cycles.

Functional-enrichment analysis

To explore the potential functions and pathways of the DEGs in NOA, a gene set enrichment analysis (GSEA) was performed using ClusterProfiler (29). We conducted a Gene Ontology (GO) analysis to reveal the biological functions of the genes, including the biological processes (BPs), cellular components (CCs), and molecular functions (MFs). Additionally, we conducted a Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis to determine the signal-transduction pathways involved.

Statistical analysis

The statistical analysis was conducted using R-Studio (R-Studio, Inc., version 2022.02.0-443), and significant differences were designated as *P<0.05, **P<0.01, and ***P<0.001. A volcano plot was generated using ggplot2 software (version 3.3.2), and the *heatmap* software package (version 3.16.0) was used to plot expression patterns within heatmaps. We employed the pROC package to plot receiver-operating characteristic (ROC) curves (30), and calculated the areas under the ROC curves (AUCs) to assess the sensitivity and specificity of the characteristic genes. In general, the larger the AUC value, the more robust the predictive power of diagnosed NOA.

Results

Single-cell analysis confirms macrophage infiltration in NOA

To determine the macrophage landscape in NOA, we employed the GSE149512 data set that comprised 17 single-cell gene-expression profiles. We first identified the top 2,000 highly variable genes (*Figure 1A*), and then normalized the data set, conducted the PCA dimensionality reduction analysis, and identified 16 principal components (P<0.05) (*Figure 1B*). An unsupervised clustering analysis was performed using t-SNE and UMAP, and our results revealed that the number of cell clusters obtained by the 2 methods was consistent and showed 18 cell clusters (*Figure 1C,1D*).

Based upon the aforementioned analysis, we observed a superior clustering effect with a resolution of 0.1 and with the cells clustered into 18 classes. We subsequently applied FindAllMarkers to discern a positive marker gene for each cluster, and the Wilcoxon method was used to

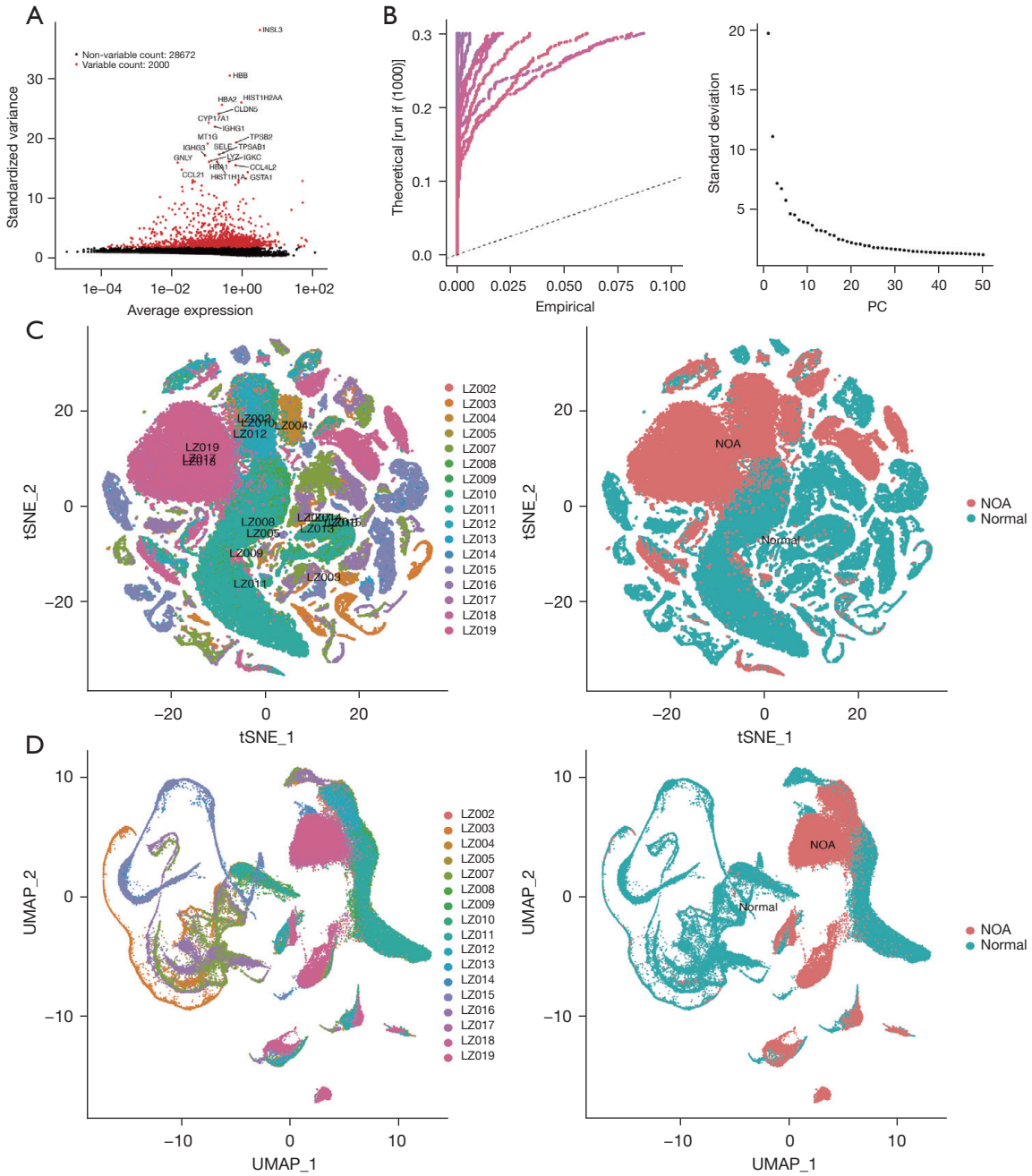

Figure 1 Landscape of the single-cell transcriptome in GSE149512. (A) Top 2,000 genes with large coefficients of variation among cells. (B) PCA clusters of cells. (C,D) Cell clusters through t-SNE (C) and UMAP (D). NOA, non-obstructive azoospermia; t-SNE, t-distributed stochastic neighbor embedding; UMAP, uniform manifold approximation and projection; PC, principal component; PCA, principal component analysis.

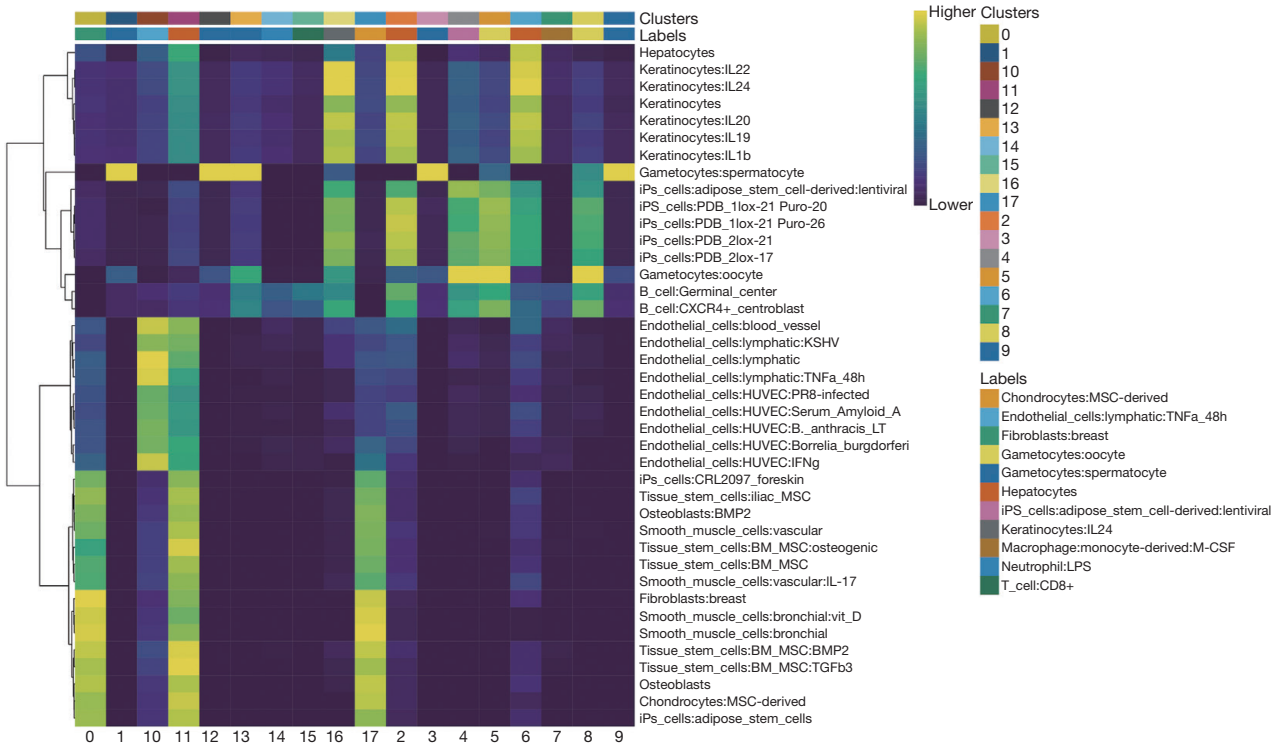

Figure 2 Heatmap showing the identification score of each cell type by SingleR function.

analyze the differences of each cluster (see Figure S1). The singleR algorithm was employed to identify the cell types in each cluster (*Figure 2*), and our results identified cluster7 as macrophages. When the marker gene for cluster7 was compared to the CellMarker database, we discovered that the macrophage marker was also highly expressed in cluster7, and thus designated cluster7 as macrophages (*Figures 3*,*4A*).

Macrophage subpopulations in NOA subtypes

We analyzed the correlations between the macrophage subpopulations and NOA subtypes, and observed that macrophage cluster0 primarily comprised iNOA subtypes (93.43%), cluster1 primarily comprised the KS subtype (48.54%) and the azoospermia factor deficiency (AZFA-Del) subtype (50.38%), and cluster2 comprised 3 subtypes. These findings suggested that the 3 types of macrophages were distinct (*Figure 4B*,*4C*).

Landscape of immune cell composition and their characterization in NOA and normal samples

Using ssGSEA to estimate the abundance of 24 immune cell species, we showed the acquisition of 10 differential immune cell types in the 27 NOA and 4 normal samples based on the GSE45885 data set and found that the macrophages were significantly upregulated in the NOA sample (*Figures 5*,*6*). Our analysis of the correlations among the different immune cell types showed that the macrophages were correlated with other immune cells (*Figure 6B*). In addition, an analysis of variance was conducted to calculate the distribution of the 10 differentially expressed immune cells in the NOA subtypes, and the results showed differences in the distribution of 8 types of immune cells, including macrophages, in the NOA subtypes (*Figure 5B*).

Key macrophage-related module genes in NOA

As no significant outlier samples were found in the GSE45885 data set, we performed sample clustering and constructed a corresponding heatmap (Figure S2). We then used optimal soft-thresholding with a power of 23 (*Figure 7A*) and produced the dynamic tree-cutting segmentation modules. We obtained 17 modules, of which 10 modules were subsequently retrieved after merging the 17 modules by setting the MEDissThres to 0.2 (*Figure 7B*).

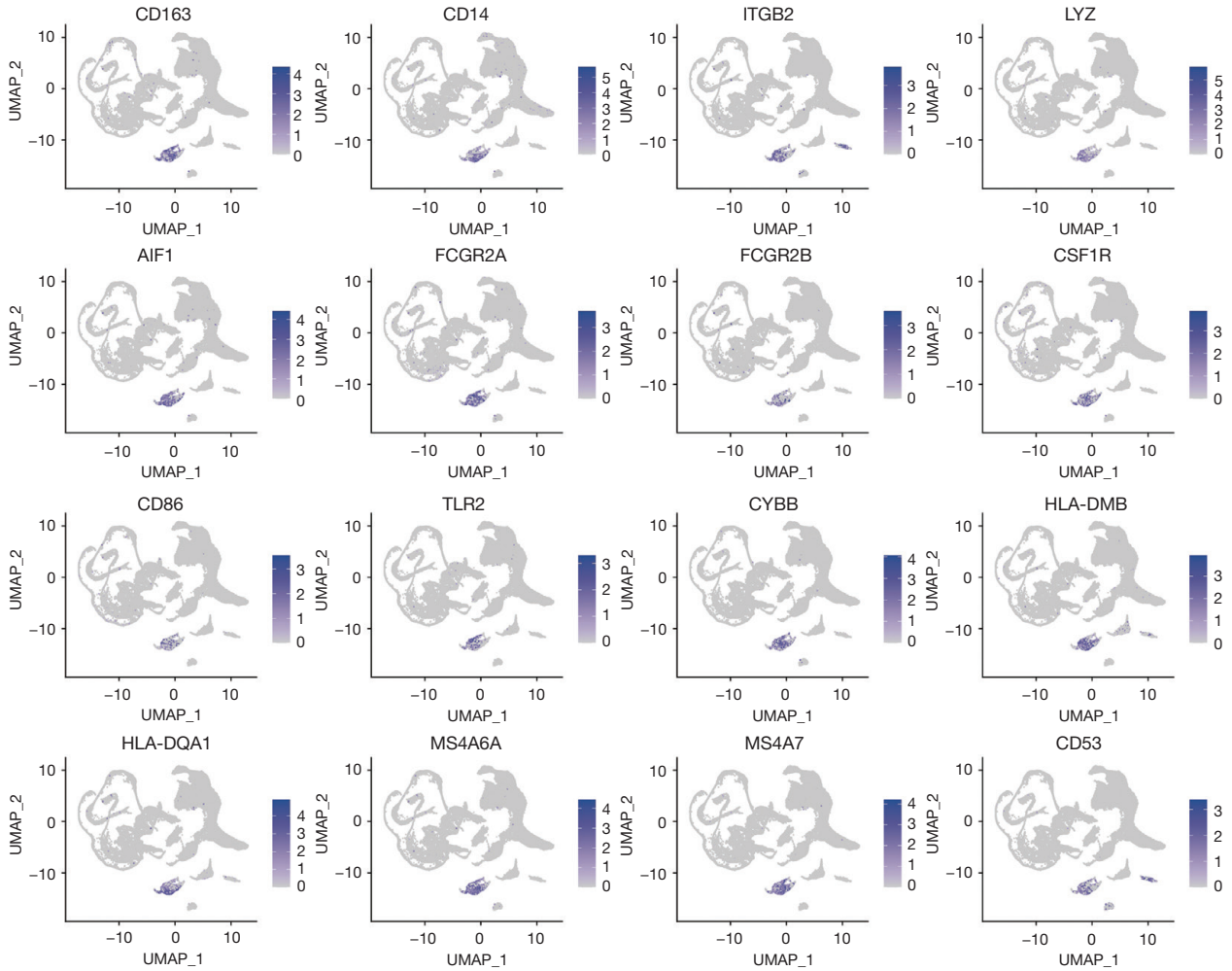

Figure 3 Distribution of the macrophage marker gene in cluster7. CD163, cluster of differentiation 163; CD14, cluster of differentiation 14; ITGB2, integrin beta 2; LYZ, lysozyme; AIF1, allograft inflammatory factor 1; FCGR2A, Fc gamma receptor IIa; FCGR2B, Fc gamma receptor IIb; CSF1R, colony stimulating factor 1 receptor; CD86, cluster of differentiation 86; TLR2, Toll-like receptor 2; CYBB, cytochrome B-245 beta chain; HLA-DMB, HLA class II histocompatibility antigen, DM beta chain; HLA-DQA1, major histocompatibility complex, class II, DQ alpha 1; MS4A6A, membrane spanning 4-domains A6A; MS4A7, membrane spanning 4-domains A7; CD53, cluster of differentiation 53.

The correlations between the 10 modules and traits were then calculated, and an optimal correlation with macrophages in the MEmagenta module was found (Correlation =0.92, P>0.05); thus, MEmagenta, which contained 913 genes, was considered a hub module (*Figure 7C*).

Identification of target genes related to macrophages in NOA

We conducted a macrophage-related gene-differential analysis in the GSE149512 database using the FindAllMarker function and screened a total of 411 upregulated and 245 downregulated mDEGs. An analysis of the DEGs from the NOA and normal samples in the GSE45885 data set revealed a total of 1,780 downregulated and 1,620 upregulated genes. We displayed our results in a volcano plot and heatmap showing the top 50 DEGs (*Figure 8A,8B*). Utilizing intersections between the 1,620 upregulated DEGs in the NOA samples and the 411 significantly upregulated mDEGs, intersections between the 1,780 downregulated DEGs in the NOA samples and

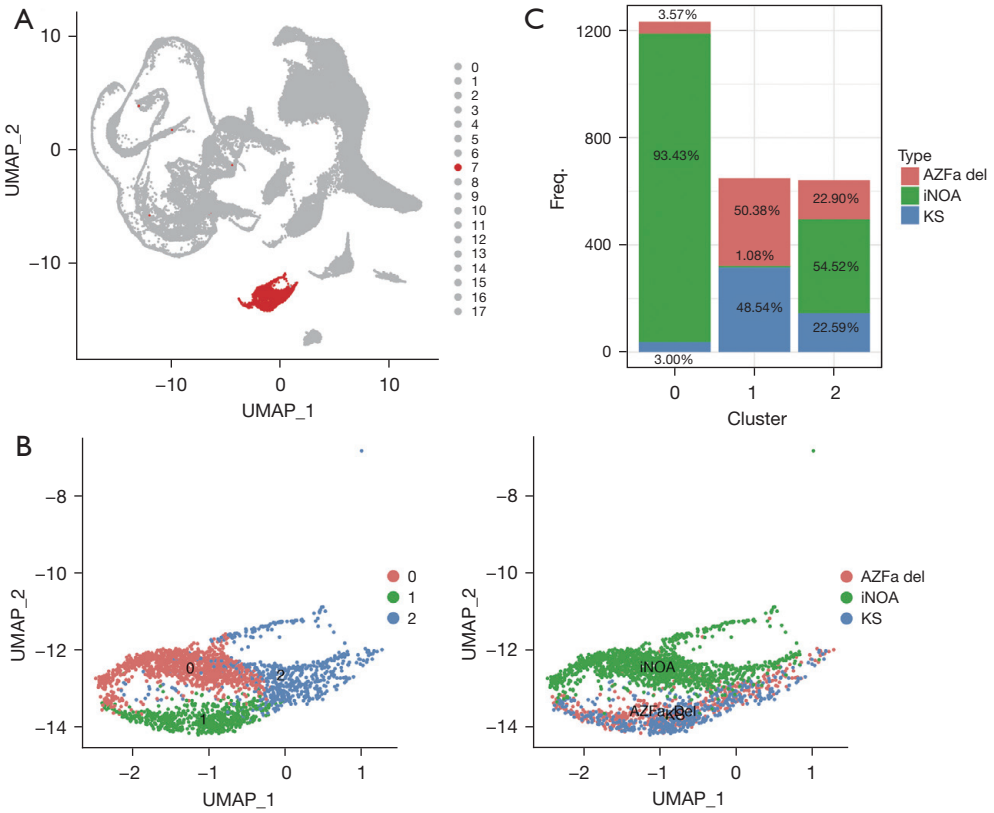

Figure 4 Analysis of the NOA subtypes in the single-cell transcriptome. (A) The position of Cluster7 in UMAP. (B) Macrophage classification in subtypes of NOA. (C) The distribution of the macrophage sub-classification in NOA subtypes. UMAP, uniform manifold approximation and projection; AZFa_del, chromosome Yq azoospermia factor a microdeletion; iNOA, idiopathic non-obstructive azoospermia; KS, Klinefelter syndrome.

the 245 downregulated mDEGs, we then identified 82 macrophage-related genes (*Figure 8C*), and these genes and 918 key macrophage genes derived from WGCNA were then intersected as well. We ultimately identified 7 target genes related to macrophages [i.e., Lipase A (*LIPA*), Major Histocompatibility Complex, Class II, DM Beta (*HLA-DMB*), Cartilage Associated Protein (*CRTAP*), Galectin 3 (*LGALS3*), Decorin (*DCN*), Membrane Spanning 4-Domains A4A (*MS4A4A*), and Chromosome 9 open reading frame 72 (*C9orf72*)] (*Figure 8D*), and selected these as candidate characteristic genes in NOA. An enrichment analysis was conducted with respect to NOA, and we found that the 7 target genes related to macrophages were enriched in 62 GO entries (*Figure 9A*) and 14 Reactome pathways (*Figure 9B*), including 49 BP entries, 11 MF entries, and 2 CC entries.

Markers associated with macrophage infiltration in NOA

To further validate our target genes, we conducted LASSO and SVM-RFE analyses on the 7 target genes related to macrophages. The LASSO analysis identified 3 significant genes (i.e., *C9orf72*, *CRTAP*, and *DCN*) (*Figure 10A*), and our evaluation of the ROC curves showed that these genes could be used to accurately distinguish NOA samples from normal samples (AUC =0.981) (*Figure 10B*). We then conducted an SVM-RFE analysis and identified 5 signature genes (i.e., *LIPA*, *CRTAP*, *MS4A4A*, *HLA-DMB*, and *C9orf72*) (*Figure 10C*).

We subsequently conducted an intersection among the significant genes screened by LASSO and SVM-RFE and identified 2 characteristic target genes (i.e., *C9orf72* and *CRTAP*) (*Figure 10D*). The ROC curves indicated that

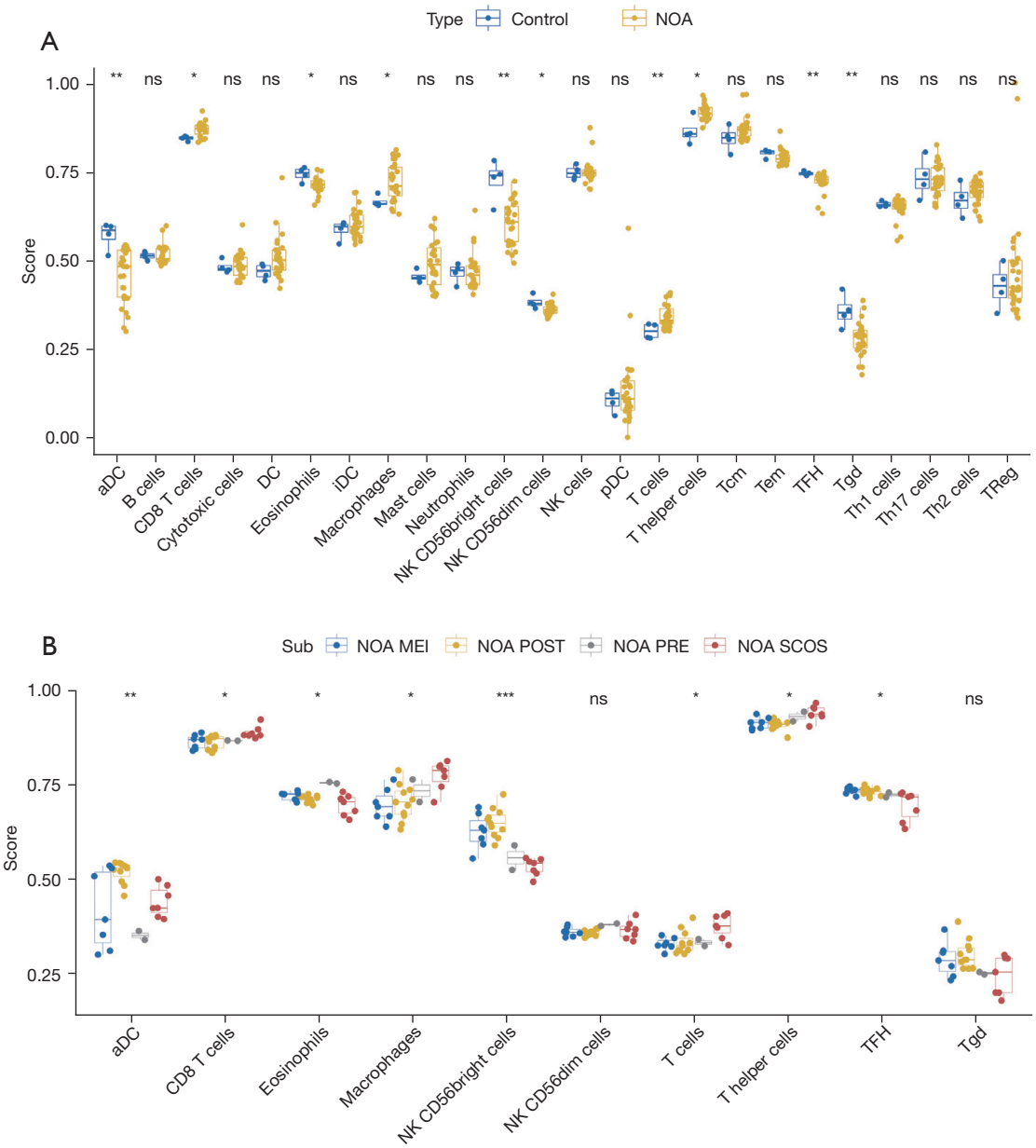

Figure 5 Immune cell infiltration analysis in NOA and the subtypes in GSE45885. (A) Box-plot of the immune cell infiltration between the control samples and NOA samples. (B) Infiltration of the differential immune cells in different subtypes of NOA. *, P<0.05; **, P<0.01; ***, P<0.001. aDC, activated dendritic cell; pDC, plasmacytoid dendritic cell; iDC, immature dendritic cell; NK, natural killer; Tem, effector memory T cell; Tcm, central memory T cell; TFH, T follicular helper cell; Tgd, T gamma delta; TReg, regulatory T cell; NOA, non-obstructive azoospermia; NOA-MEI, NOA meiotic arrest; NOA-POST, post-meiotic arrest in NOA; NOA-PRE, NOA pre-meiotic arrest; NOA-SCOS, NOA sertoli cell-only syndrome; ns, no significant difference.

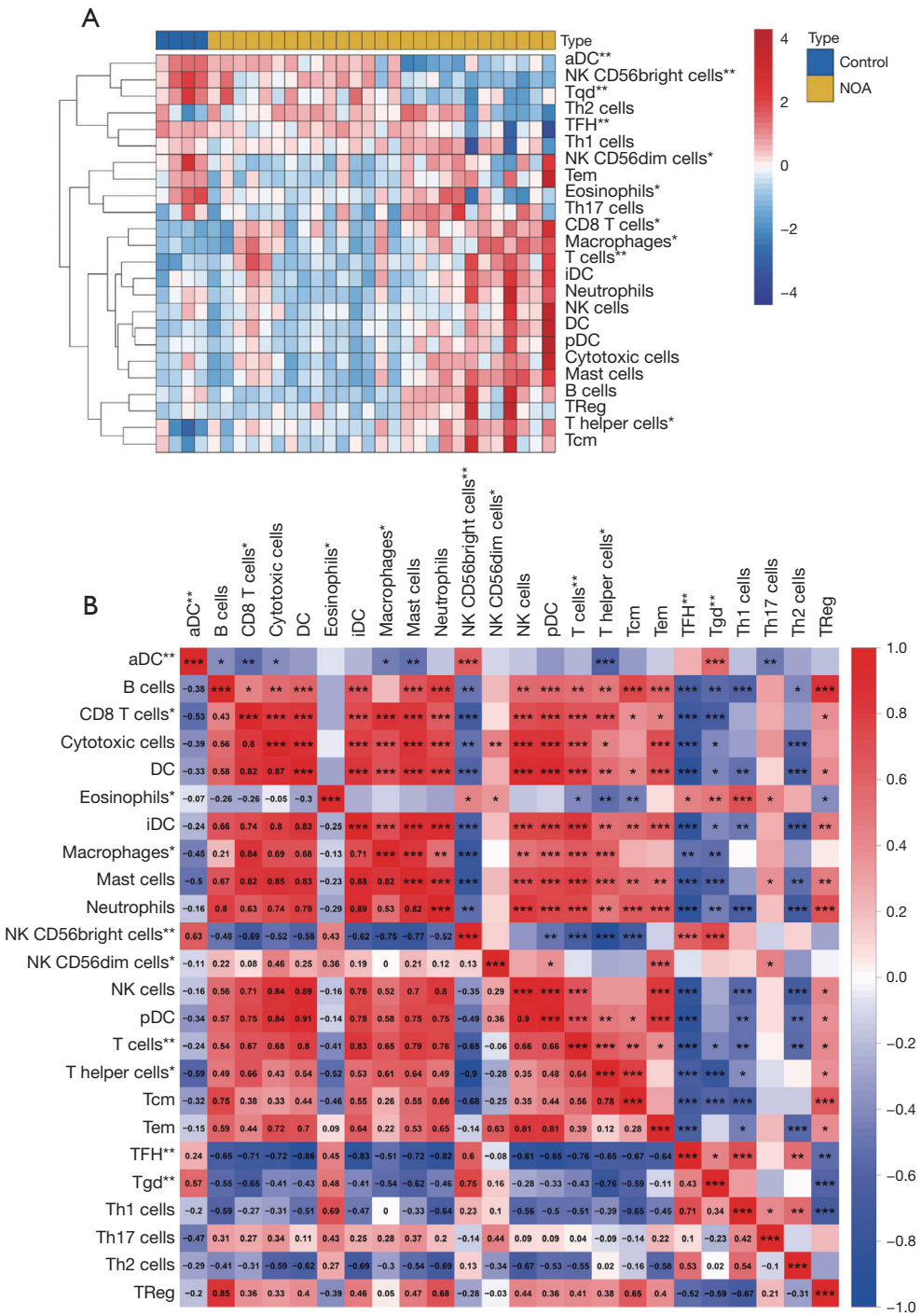

Figure 6 Analysis of immune cell interaction. (A) Heatmap of the immune cell infiltration between the control samples and NOA samples. (B) The correlations among different immune cells. *, P<0.05; **, P<0.01; ***, P<0.001. aDC, activated dendritic cell; pDC, plasmacytoid dendritic cell; iDC, immature dendritic cell; NK, natural killer; Tem, effector memory T cell; Tcm, central memory T cell; TFH, T follicular helper cell; Tgd, T gamma delta; TReg, regulatory T cell; NOA, non-obstructive azoospermia.

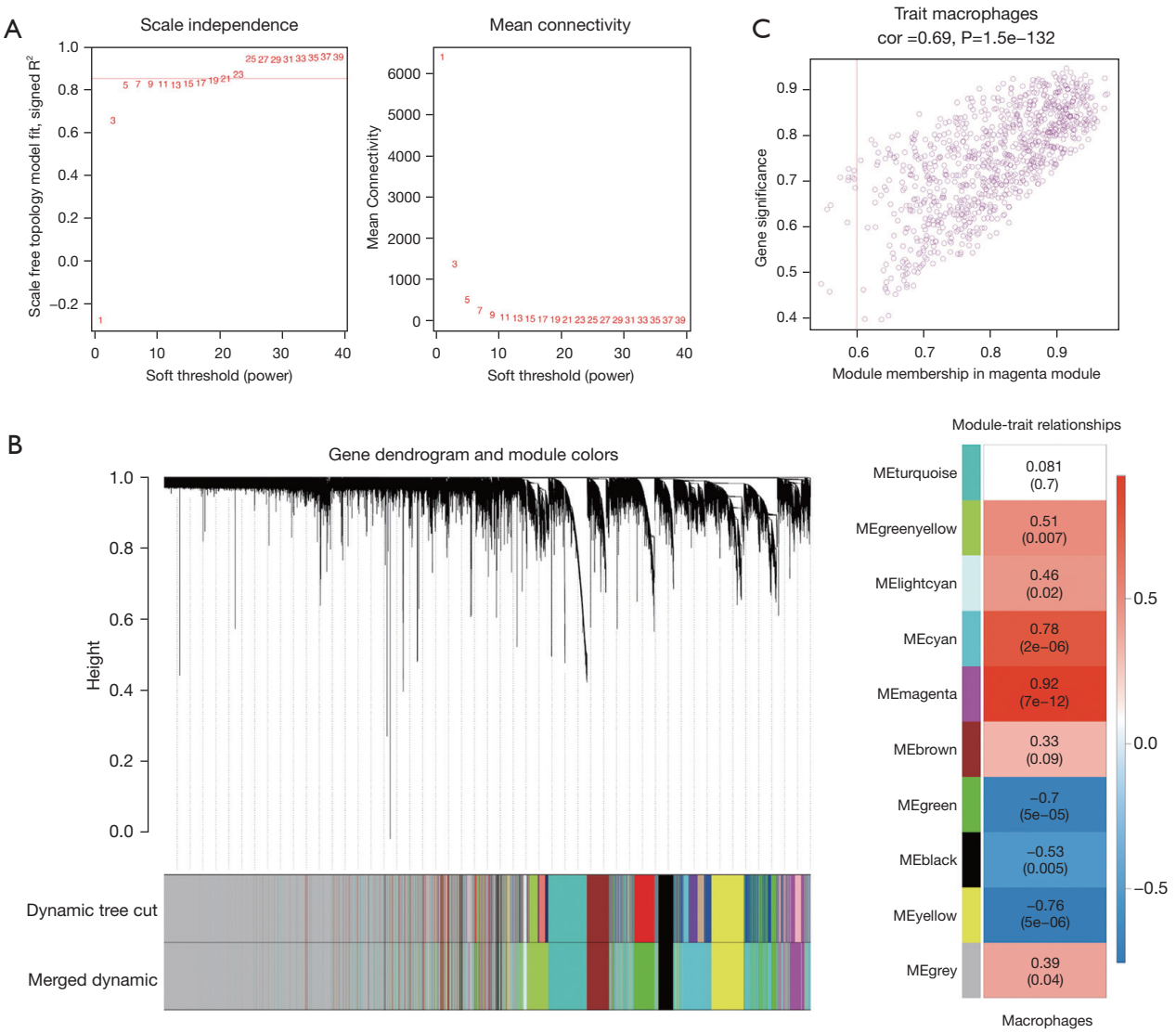

Figure 7 Identification of the macrophage-related hub by WGCNA. (A) Optimal soft threshold screening. (B) Co-expression modules in NOA. The cluster dendrogram represents the 17 modules. The modules represent groups of co-related genes and are colored differently. On the cluster dendrogram, each leaf represents a gene. Module-trait relationship correlation heat map. (C) The correlation between genes in MEmagenta and macrophages. NOA, non-obstructive azoospermia; WGCNA, weighted gene co-expression network analysis.

both the characteristic genes could be used to accurately distinguish between NOA samples and normal samples (AUC for *C9orf72* =0.861; AUC for *CRTAP* =0.917) (*Figure 11A,11B*). To further confirm the expression of *C9orf72* and *CRTAP*, we collected 26 testicular biopsy samples, including 18 from NOA patients and 8 from normal individuals. The RT-qPCR results (Figure S3) showed that both *C9orf72* and *CRTAP* were significantly upregulated in NOA (*Figure 11C*), which suggests that these genes constitute potential markers associated with macrophages in NOA.

Finally, to explore gene function, we conducted a single-gene GSEA and demonstrated that *C9orf72* was highly associated with nucleocytoplasmic transport, cell cycle, and the spliceosome pathway, among others, and that *CRTAP* was highly associated with the Ras-proximate-1 (Rap1) signaling pathway, Ras signaling pathway, and chemokine-signaling pathway, among others (*Figure 11D,11E*) (the

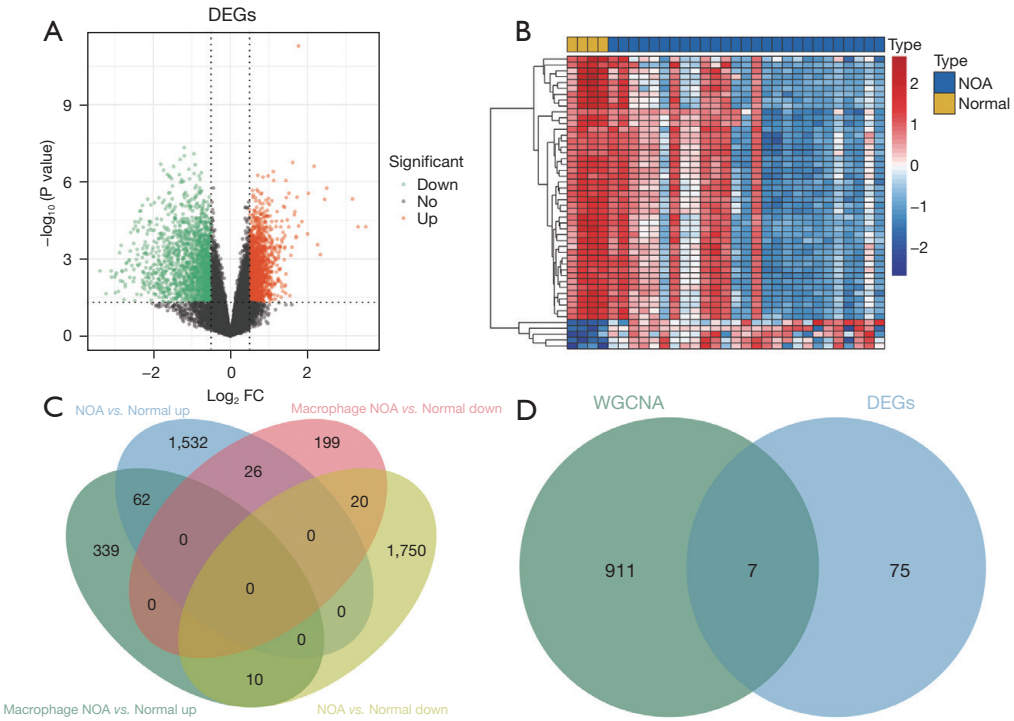

Figure 8 Identification of the biomarkers related to macrophage infiltration in NOA. (A,B) Volcano plot and heatmap of the DEGs between the NOA and normal samples in the GSE45885 data set. The color bar represents the Z-score normalized gene expression. (C) Venn diagram of the DEGs and single-cell transcriptome macrophage differential genes in NOA. (D) Venn diagram of the key macrophage genes and WGCNA macrophage-related genes. NOA, non-obstructive azoospermia; WGCNA, weighted gene co-expression network analysis; DEGs, differential expression genes.

results of the single-gene GSEA of the other candidate genes are depicted in Figure S4).

Discussion

NOA is one of the most severe forms of male infertility (31); however, its etiology remains largely unknown. Apart from well-established risk factors for NOA, such as undescended testis, KS, Kallmann syndrome, and Y-chromosome microdeletion, there exist other causalities, such as gene mutations [e.g., Testis Expressed 11 (*TEX11*) (32)] and gene polymorphisms [in SRY-box transcription factor 5 (*SOX5*) (33)]. Further, numerous acquired factors can lead to NOA, including varicocele, infections, testicular or pituitary tumors, and exposure to various toxicants, and a significant number of NOA patients may also be idiopathic (i.e., have no specific diagnosis). The causes of NOA differ, but its fundamental consequence is impaired spermatogenesis. Thus, the etiology of NOA needs to be elucidated to provide an avenue by which to explore the mechanisms underlying normal spermatogenesis. We undertook the present investigation because the involvement of cytokines in the processes that comprise sperm growth and differentiation remain unclear.

In accordance with previous findings (34), our in-depth analysis of multiple-subtype data sets in NOA indicated that macrophages were involved in its pathogenesis. Macrophages constitute a common type of immune cell that reside in multiple organs and maintain overall bodily homeostasis (35). Their primary function in the immune system is phagocytosis (whereby macrophages defend the host organism against infection and injury), but macrophages are also involved in both innate and adaptive immunity. Macrophages present antigens to T cells and thus activate the cell immune response, and play an anti-inflammatory role by releasing cytokines. Macrophages that promote inflammation are known as M1 macrophages, while those that promote tissue repair and reduce

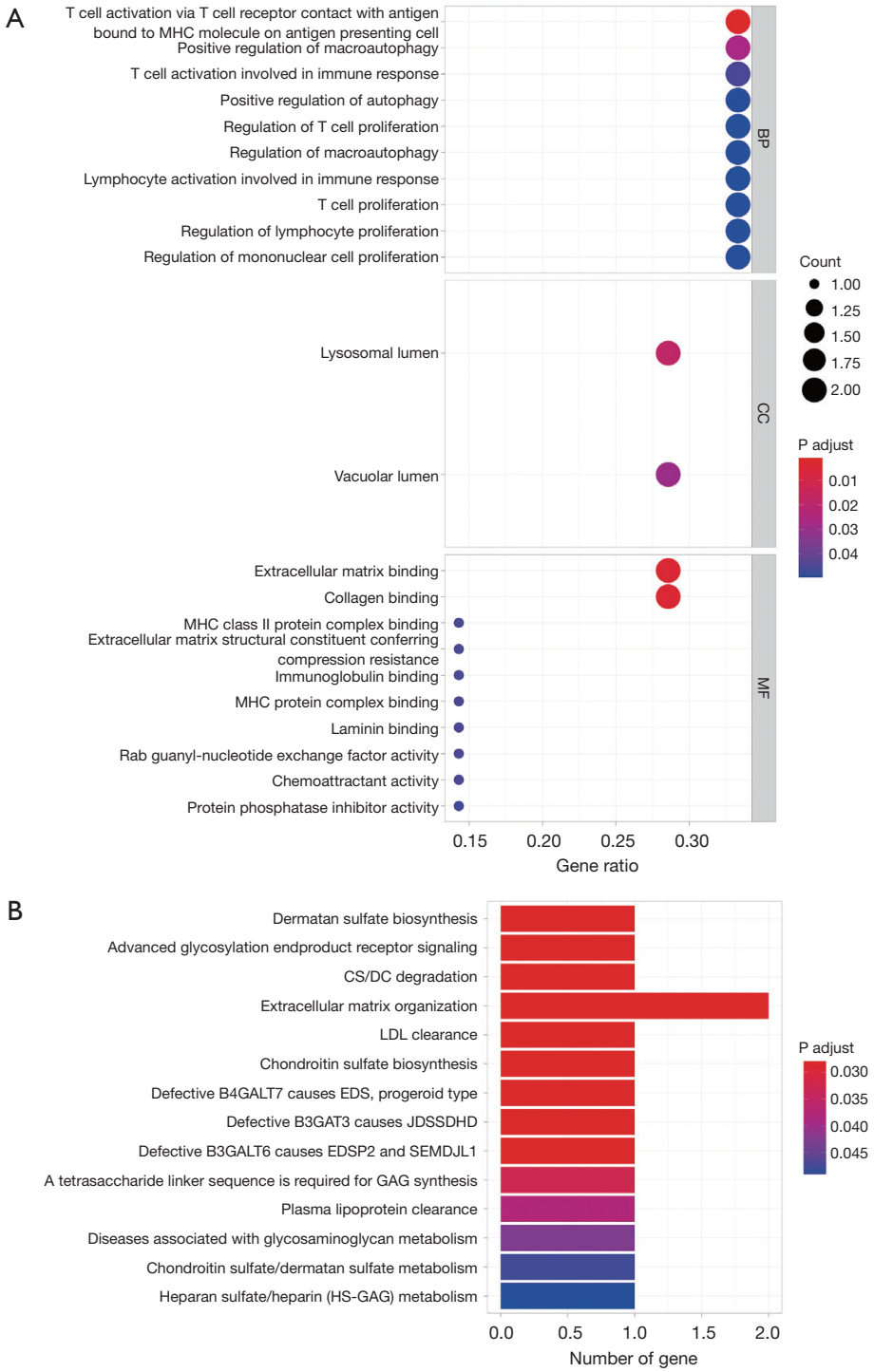

Figure 9 Enrichment analysis of the identified biomarkers. (A) The results of the GO enrichment analysis for the target genes (top 10). (B) The results of the Reactome pathway enrichment analysis for the target genes. BP, biological process; CC, cellular component; MF, molecular function.

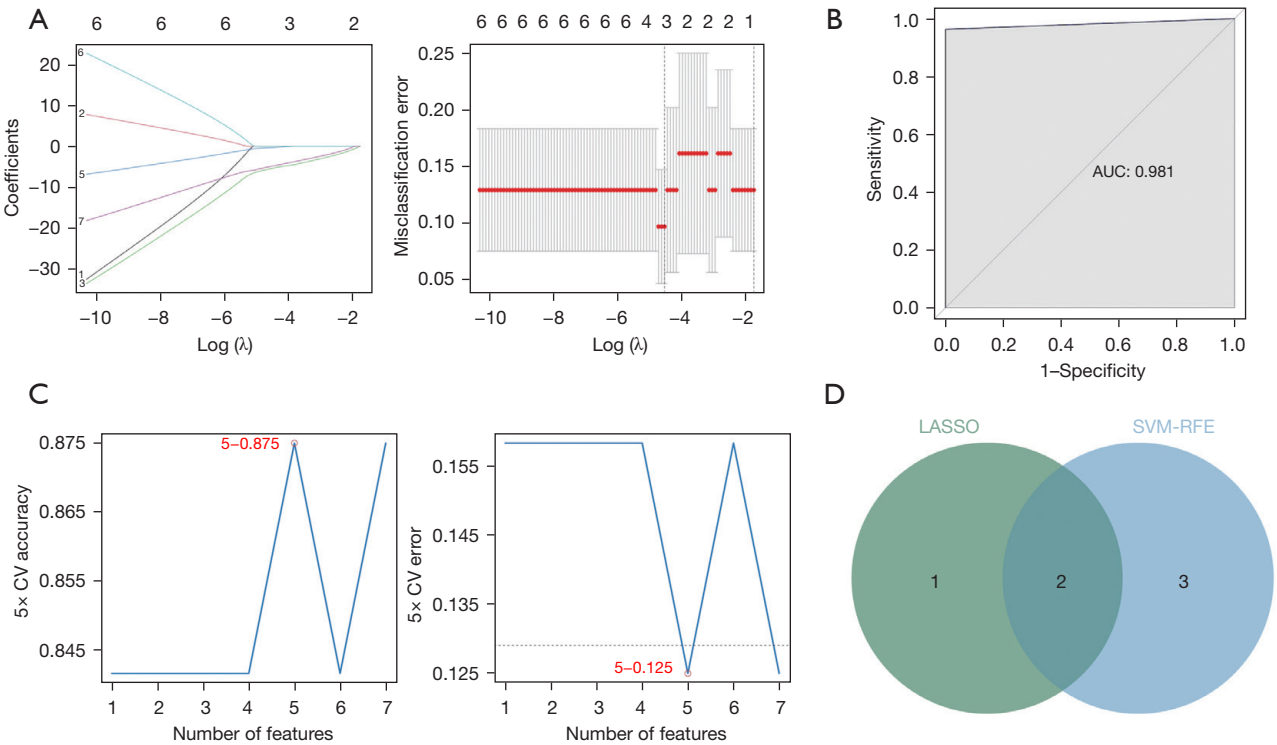

Figure 10 Further identification of the macrophage-related biomarkers in NOA by LASSO and SVM-RFE. (A) Screening of key genes by LASSO regression analysis. Calculation of the LASSO coefficients for the key genes. Selection of the optimal genes (B) ROC curve of the LASSO model. (C) The SVM-RFE algorithm identified 5 genes. Plot of generalization error, prediction accuracy versus number of feature genes (D). Venn diagram of key genes from the LASSO and SVM-RFE model. NOA, non-obstructive azoospermia; LASSO, least absolute shrinkage and selection operator; SVM-RFE, support vector machine recursive feature elimination; AUC, areas under the ROC curves. ROC, receiver-operating characteristic.

inflammation are known as M2 macrophages (36).

In the male primary reproductive organ (i.e., the testis), macrophages protect autoimmune attacks with neo-antigens, and testicular immune privilege is the most significantly protective mechanism from autoimmune (37-39). To maintain this immune-privileged status, macrophages are not typically activated in response to pathologic antigens, and they have been shown to produce anti-inflammatory cytokines constitutively in the rat testis (40). Numerous infiltrates of circulating macrophages in the testis regulate spermatogenesis in orchitis (41). Goluža *et al.* also showed that cluster of differentiation 68-positive macrophages are elevated in NOA (42). Others have reported that macrophage polarization is associated with testicular damage in NOA (34). These findings suggest that macrophage infiltration impairs testicular function by disrupting the immune-privileged state.

Evidence has also revealed that macrophages are critical to spermatogenesis in the testis (17), as they can regulate Leydig cell steroidogenesis by secreting 25-hydroxycholesterol (43,44). In the normal testis, macrophages play a vital role in maintaining immune privilege by suppressing nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) signaling through a nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha (IκBα)-ubiquitination deficiency and a decreased in gene expression in the toll-like receptor (TLR) cascade, and they exhibit a reduced proinflammatory capability (45). Our study also confirmed that macrophages play an important role in NOA.

CRTAP, which was first identified by Castagnola *et al.* in 1997, has been found to primarily function in the differentiation of chondrocytes (46). The GO analysis revealed that the BPs for CRTAP included peptidyl-proline hydroxylation to 3-hydroxy-L-proline, collagen fibril organization, and protein stabilization. Research on

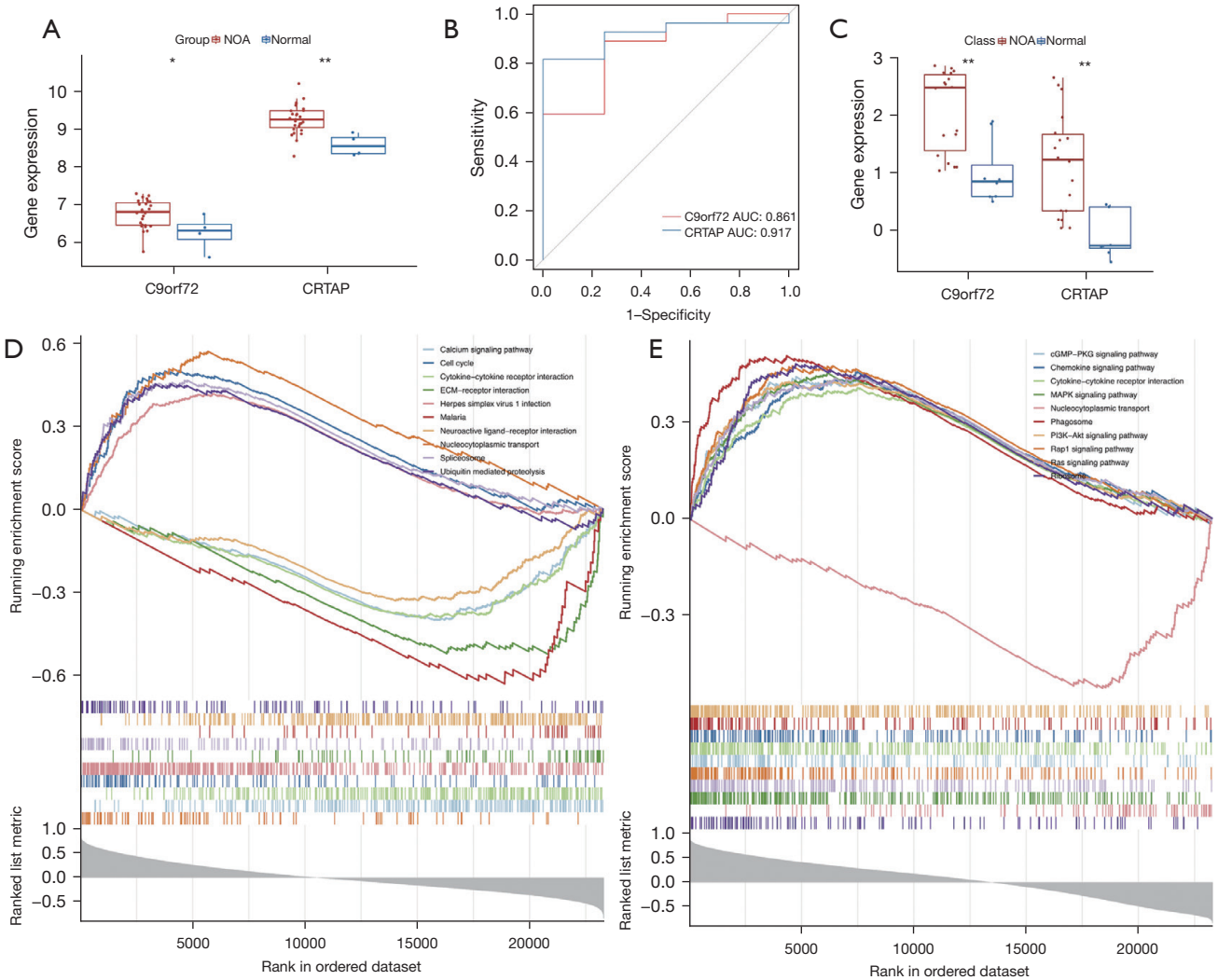

Figure 11 Verification of C9orf72 and CRTAP as the biomarkers of macrophage infiltration in NOA. (A,B) The gene expression of the 2 key genes between the NOA and normal samples, and the ROC curves of the 2 key genes in the GSE45885 data set. (C) The gene expression of the 2 key genes between the NOA and normal samples by RT-qPCR. (D) Single-gene enrichment analysis of C9orf72. (E) Single-gene enrichment analysis of CRTAP. *, P<0.05; **, P<0.01. NOA, non-obstructive azoospermia; C9orf72, open reading frame 72 gene on chromosome 9; RT-qPCR, real-time quantitative polymerase chain reaction; CRTAP, cartilage-associated protein; ROC, receiver-operating characteristic; AUC, areas under the ROC curves.

spermatogenesis in mice has also shown that CRTAP may plays a novel role in spermatogenesis (47). It demonstrated an elevated expression level in spermatogonia, but reduced as differentiation progressed. Researchers have also found that *CRTAP* is highly expressed in the follicles and stroma of the ovary, in testicular interstitial cells at 4 weeks of age, and in germline cells, and mature sperm (but is reduced in grown mice), and increased expression in abnormal mature sperm of Crtap-KO mice has also been observed (48).

In the present study, we found that compared to normal men, the levels of CRTAP in men with NOA were significantly increased, which suggests that CRTAP is positively correlated with the inhibition of spermatogonial differentiation. Our single-gene enrichment analysis revealed that the principal pathways connected with CRTAP were involved in proliferative and immune-response functions. Ras and Rap1 are downstream effectors in their respective signaling pathways, are both GTPases, and primarily function in cell adhesion, proliferation, survival, and differentiation. Further, chemokine-signaling pathways occupy crucial

positions in the induction of the immune response. Given the aforementioned CRTAP functions, it is reasonable to speculate that the CRTAP in testis macrophages may be important in human spermatogenic processes.

C9orf72 has been found to interact with Rab proteins that are involved in autophagy and endocytic transport and to regulate endosomal trafficking. There is evidence that C9orf72 is essential to macrophage function (49), and the C9orf72-enrichment analysis showed that the gene is related to nucleocytoplasmic transport and cell the cycle. Our findings further demonstrated that macrophage differentiation was related to testicular damage in NOA. We hypothesize that C9orf72 affects the differentiation of macrophages; however, this needs to be confirmed by further physiological documentation. The data in this study revealed robust differences in immune cell distribution between the NOA and control samples, and we conjecture that this may be due to the disruption of immune privilege, thus compromising spermatogenesis.

We conducted multiple bioinformatic analyses; however, the present study had a number of limitations. First, we used different data sets to formulate our conclusions. Second, the differences in the sequencing technologies and platforms used might have led to heterogeneity in the data sets. Thus, in the future, we intend to further investigate the relative gene expression in macrophages with NOA, and clarify the fundamental biologic and physiologic underpinnings of NOA and overall spermatogenesis. Our study was the first to ascribe functions to *C9orf72* and *CRTAP* in spermatogenesis in *Homo sapiens* and to provide data that may assist in the elucidation of a novel therapeutic target(s) in NOA, thus extending understandings of the mechanisms underlying spermatogenesis.

Conclusions

Through the combination of tissue transcriptomic and single-cell RNA-sequencing analyses, we concluded that macrophage infiltration is significant in different subtypes of NOA, and we hypothesized that *C9orf72* and *CRTAP* play critical roles in NOA due to their high expression in macrophages.

Acknowledgments

Funding: This work was supported by grants from the National Natural Science Foundation of China (No. 82260292); Yunnan Provincial Reproductive and Obstetrics and Gynecology Clinical Medicine Center (No. zx2019-01-01); Open Project of Yunnan Provincial Key Specialty of Gynecology (No. 2022FKZDZK-13); Open Project of Yunnan Provincial Reproductive and Obstetrics and Gynecology Clinical Medicine Center (Nos. 2022LCZXKF-SZ21, 2022LCZXKF-SZ16, 2020LCZXKF-SZ14, 2020LCZXKF-SZ11, and 2020LCZXKF-SZ06); National Natural Science Foundation of China (No. 82060282); and Program of Application and Fundamental Research of Joint Special project of Yunnan Provincial Science Technology Department and Kunming Medical University (No. 202001AY070001-295).

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at https://atm.amegroups.com/article/view/10.21037/atm-22-5601/rc

Data Sharing Statement: Available at https://atm.amegroups.com/article/view/10.21037/atm-22-5601/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://atm.amegroups.com/article/view/10.21037/atm-22-5601/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013), and approved by the Ethics Committee of The First People's Hospital of Yunnan Province (Approval ID: KHLL2020-KY012). The samples were collected under the condition of fully informed consent before the patient took part in.

References

1. Agarwal A, Mulgund A, Hamada A, et al. A unique view on male infertility around the globe. Reprod Biol Endocrinol 2015;13:37.
2. Thornton A, Binstock G, Yount KM, et al. International fertility change: new data and insights from the developmental idealism framework. Demography 2012;49:677-98.
3. Jungwirth A, Giwercman A, Tournaye H, et al. European Association of Urology guidelines on Male Infertility: the 2012 update. Eur Urol 2012;62:324-32.
4. Vloeberghs V, Verheyen G, Haentjens P, et al. How successful is TESE-ICSI in couples with non-obstructive azoospermia? Hum Reprod 2015;30:1790-6.
5. Westlander G. Utility of micro-TESE in the most severe cases of non-obstructive azoospermia. Ups J Med Sci 2020;125:99-103.
6. McLachlan RI, Rajpert-De Meyts E, Hoei-Hansen CE, et al. Histological evaluation of the human testis--approaches to optimizing the clinical value of the assessment: mini review. Hum Reprod 2007;22:2-16.
7. Kasak L, Laan M. Monogenic causes of non-obstructive azoospermia: challenges, established knowledge, limitations and perspectives. Hum Genet 2021;140:135-54.
8. Cioppi F, Rosta V, Krausz C. Genetics of Azoospermia. Int J Mol Sci 2021;22:3264.
9. Morales A. The long and tortuous history of the discovery of testosterone and its clinical application. J Sex Med 2013;10:1178-83.
10. Heinrich A, DeFalco T. Essential roles of interstitial cells in testicular development and function. Andrology 2020;8:903-14.
11. Li N, Wang T, Han D. Structural, cellular and molecular aspects of immune privilege in the testis. Front Immunol 2012;3:152.
12. Fijak M, Meinhardt A. The testis in immune privilege. Immunol Rev 2006;213:66-81.
13. Bhushan S, Meinhardt A. The macrophages in testis function. J Reprod Immunol 2017;119:107-12.
14. Hedger MP. Macrophages and the immune responsiveness of the testis. J Reprod Immunol 2002;57:19-34.
15. Bhushan S, Theas MS, Guazzone VA, et al. Immune Cell Subtypes and Their Function in the Testis. Front Immunol 2020;11:583304.
16. Mossadegh-Keller N, Sieweke MH. Testicular macrophages: Guardians of fertility. Cell Immunol 2018;330:120-5.
17. DeFalco T, Potter SJ, Williams AV, et al. Macrophages Contribute to the Spermatogonial Niche in the Adult Testis. Cell Rep 2015;12:1107-19.
18. Hussein MR, Abou-Deif ES, Bedaiwy MA, et al. Phenotypic characterization of the immune and mast cell infiltrates in the human testis shows normal and abnormal spermatogenesis. Fertil Steril 2005;83:1447-53.
19. Apa DD, Cayan S, Polat A, et al. Mast cells and fibrosis on testicular biopsies in male infertility. Arch Androl 2002;48:337-44.
20. Dong M, Li H, Zhang X, et al. Weighted Correlation Gene Network Analysis Reveals New Potential Mechanisms and Biomarkers in Non-obstructive Azoospermia. Front Genet 2021;12:617133.
21. Malcher A, Rozwadowska N, Stokowy T, et al. Potential biomarkers of nonobstructive azoospermia identified in microarray gene expression analysis. Fertil Steril 2013;100:1686-94.e1-7.
22. Zhao L, Yao C, Xing X, et al. Single-cell analysis of developing and azoospermia human testicles reveals central role of Sertoli cells. Nat Commun 2020;11:5683.
23. Zhang X, Lan Y, Xu J, et al. CellMarker: a manually curated resource of cell markers in human and mouse. Nucleic Acids Res 2019;47:D721-8.
24. Barbie DA, Tamayo P, Boehm JS, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 2009;462:108-12.
25. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559.
26. Wettenhall JM, Smyth GK. limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics 2004;20:3705-6.
27. Tibshirani R. Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society Series B-Methodological 1996;58:267-88.
28. Guyon I, Weston J, Barnhill S, et al. Gene selection for cancer classification using support vector machines. Machine Learning 2002;46:389-422.
29. Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16:284-7.
30. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.
31. Chiba K, Enatsu N, Fujisawa M. Management of non-obstructive azoospermia. Reprod Med Biol 2016;15:165-73.
32. Yatsenko AN, Georgiadis AP, Röpke A, et al. X-linked

TEX11 mutations, meiotic arrest, and azoospermia in infertile men. N Engl J Med 2015;372:2097-107.
33. Zou S, Li Z, Wang Y, et al. Association study between polymorphisms of PRMT6, PEX10, SOX5, and nonobstructive azoospermia in the Han Chinese population. Biol Reprod 2014;90:96.
34. Zheng W, Zhang S, Jiang S, 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.
35. Wynn TA, Chawla A, Pollard JW. Macrophage biology in development, homeostasis and disease. Nature 2013;496:445-55.
36. Mills CD. M1 and M2 Macrophages: Oracles of Health and Disease. Crit Rev Immunol 2012;32:463-88.
37. Fijak M, Bhushan S, Meinhardt A. Immunoprivileged sites: the testis. Methods Mol Biol 2011;677:459-70.
38. Pérez CV, Theas MS, Jacobo PV, et al. Dual role of immune cells in the testis: Protective or pathogenic for germ cells? Spermatogenesis 2013;3:e23870.
39. Da Silva N, Barton CR. Macrophages and dendritic cells in the post-testicular environment. Cell Tissue Res 2016;363:97-104.
40. Winnall WR, Muir JA, Hedger MP. Rat resident testicular macrophages have an alternatively activated phenotype and constitutively produce interleukin-10 in vitro. J Leukoc Biol 2011;90:133-43.
41. Rival C, Theas MS, Suescun MO, et al. Functional and phenotypic characteristics of testicular macrophages in experimental autoimmune orchitis. J Pathol 2008;215:108-17.
42. Goluža T, Boscanin A, Cvetko J, et al. Macrophages and Leydig cells in testicular biopsies of azoospermic men. Biomed Res Int 2014;2014:828697.
43. Nes WD, Lukyanenko YO, Jia ZH, et al. Identification of the lipophilic factor produced by macrophages that stimulates steroidogenesis. Endocrinology 2000;141:953-8.
44. Lukyanenko YO, Chen JJ, Hutson JC. Production of 25-hydroxycholesterol by testicular macrophages and its effects on Leydig cells. Biol Reprod 2001;64:790-6.
45. Bhushan S, Tchatalbachev S, Lu Y, et al. Differential activation of inflammatory pathways in testicular macrophages provides a rationale for their subdued inflammatory capacity. J Immunol 2015;194:5455-64.
46. Castagnola P, Gennari M, Morello R, et al. Cartilage associated protein (CASP) is a novel developmentally regulated chick embryo protein. J Cell Sci 1997;110 (Pt 12):1351-9.
47. Pang AL, Taylor HC, Johnson W, et al. Identification of differentially expressed genes in mouse spermatogenesis. J Androl 2003;24:899-911.
48. Zimmerman SM, Besio R, Heard-Lipsmeyer ME, et al. Expression characterization and functional implication of the collagen-modifying Leprecan proteins in mouse gonadal tissue and mature sperm. AIMS Genet 2018;5:24-40.
49. O'Rourke JG, Bogdanik L, Yáñez A, et al. C9orf72 is required for proper macrophage and microglial function in mice. Science 2016;351:1324-9.

Cite this article as: Luo X, Zheng H, Nai Z, Li M, Li Y, Lin N, Li Y, Wu Z. Identification of biomarkers associated with macrophage infiltration in non-obstructive azoospermia using single-cell transcriptomic and microarray data. Ann Transl Med 2023;11(2):55. doi: 10.21037/atm-22-5601

Figure S1 The violin plot of the 18 cell clusters (Clusters 0–17).

https://dx.doi.org/10.21037/atm-22-5601

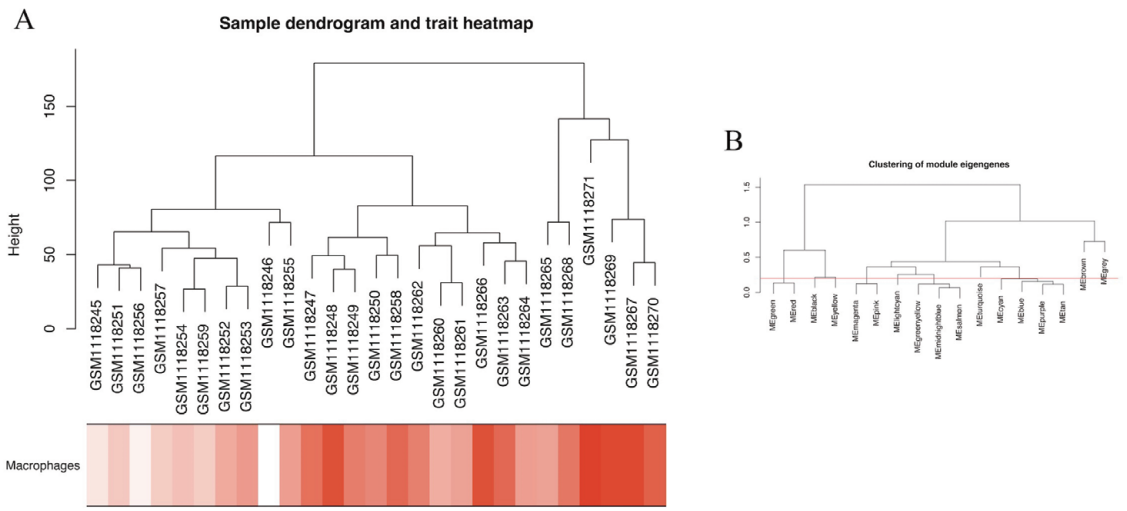

Figure S2 WGCNA data set preprocessing. (A). Sample clustering and outlier checking, and sample clustering heat map based on macrophage scores. (B). The merged dynamic shear-tree algorithm was used to analyze similar modules by setting MEDissThres to 0.2. WGCNA module selection by the soft threshold.

https://dx.doi.org/10.21037/atm-22-5601

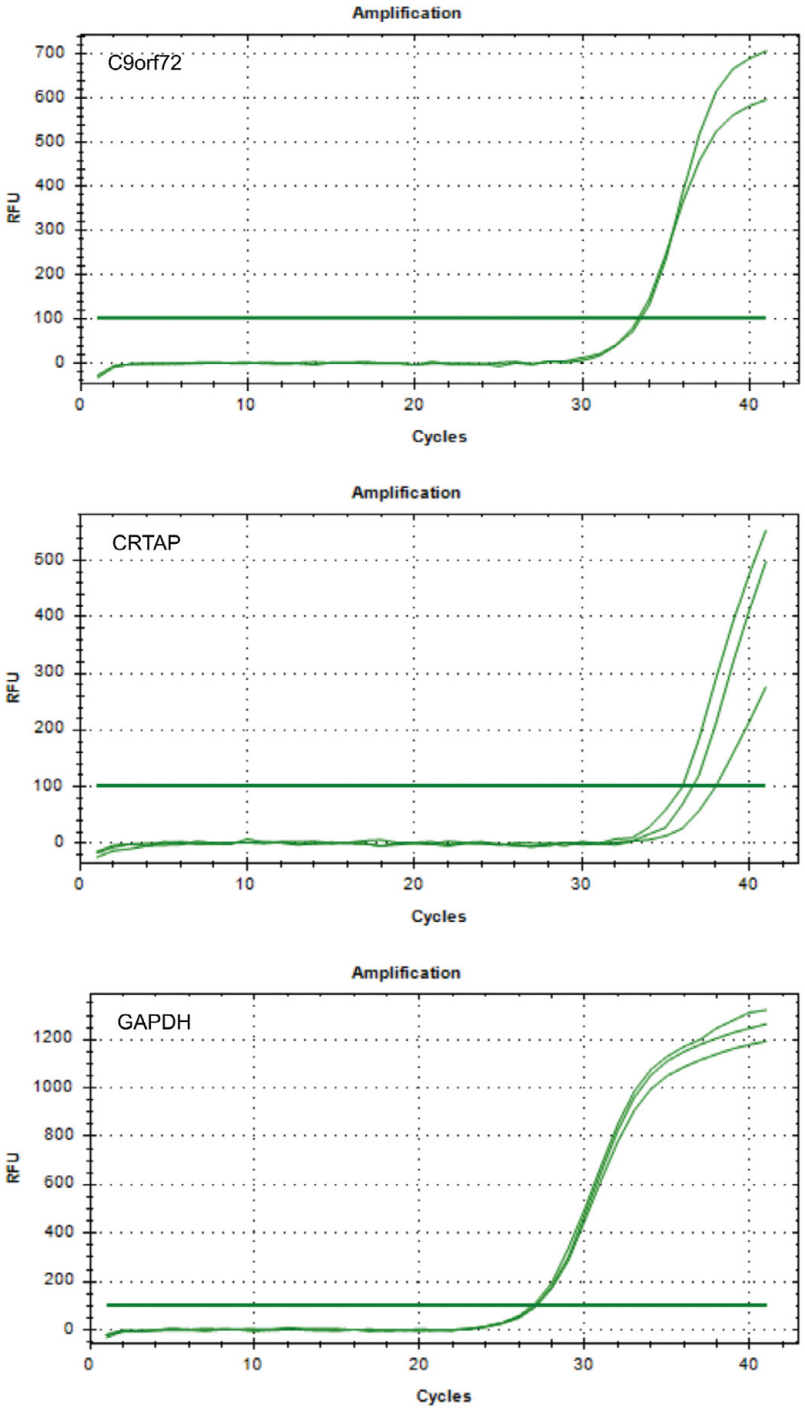

Figure S3 The amplification curves of *C9orf7*, *CRTAP*, and GAPDH in RT-qPCR. RFU, relative fluorescence units.

https://dx.doi.org/10.21037/atm-22-5601

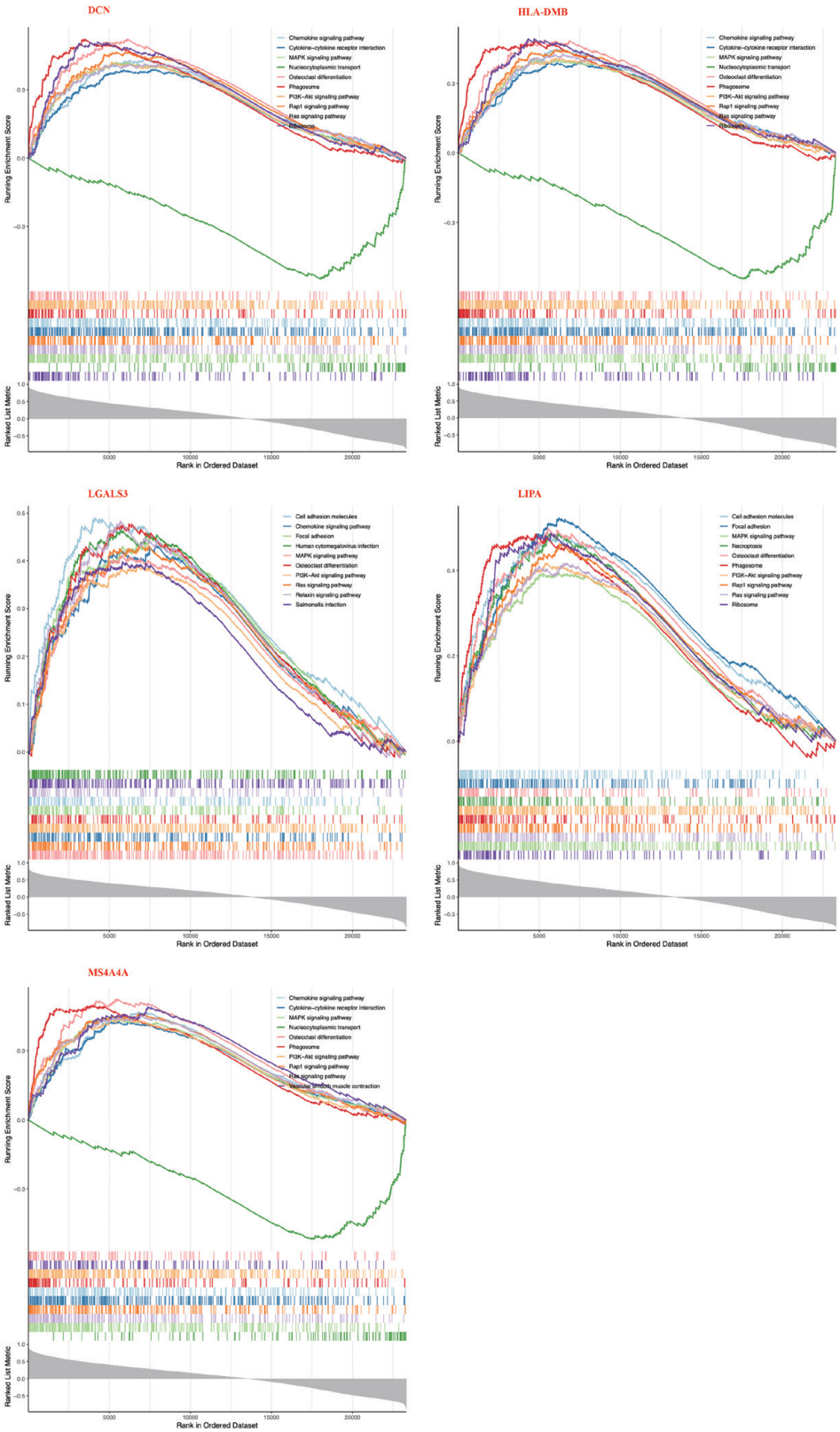

Figure S4 Single-gene enrichment analyses of *DCN*, *HLA-DMB*, *LGALS3*, *LIPA*, and *MS4A4A*.

https://dx.doi.org/10.21037/atm-22-5601