Am J Transl Res 2023;15(2):710-728
www.ajtr.org /ISSN:1943-8141/AJTR0145700

Original Article

Bioinformatics analysis and verification of key candidate genes influencing the pathogenesis of chronic rhinosinusitis with nasal polyps

Gang Chen[1*], Hong Hao[2*], Lin-E Wang[1]

[1]*Department of Otorhinolaryngology Head and Neck Surgery, Affiliated Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China;* [2]*Department of Otorhinolaryngology Head and Neck Surgery, The 401 Hospital of The China Nuclear Industry, Beijing 102413, China.* [*]*Equal contributors and co-first authors.*

Received August 2, 2022; Accepted December 20, 2022; Epub February 15, 2023; Published February 28, 2023

Abstract: Objectives: Chronic rhinosinusitis (CRS) with nasal polyps (CRSwNP) is a prominent public health issue. Furthermore, the prognosis of eosinophilic CRSwNP is poor, with a high recurrence rate. The underlying molecular mechanisms of eosinophilic CRSwNP remain unclear. Therefore, in this study, we sought to determine the crucial genes underlying eosinophil infiltration in eosinophilic CRSwNP pathogenesis. Methods: We used the Gene Expression Omnibus database (GEO) (GSE36830 and GSE23552 datasets) to mine gene expression profiles of CRSwNP patients and normal subjects. Differentially expressed genes (DEGs) between normal and CRSwNP tissues were identified and subjected to the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment analyses. Co-expression networks were established using a weighted gene co-expression network analysis (WGCNA) and single-sample gene set enrichment analysis (GSEA). Protein-protein interaction networks were developed to detect functional protein modules. Based on the common DEGs, candidate miRNAs and related lncRNAs were predicted using the mirTarBase and StarBase databases. Finally, we generated immune cell subtypes of CRSwNP. Results: A total of 146 DEGs were identified. Of these, 131 genes were upregulated, whereas 15 were downregulated. GO analysis indicated that DEGs primarily participated in leukocyte chemotaxis and migration as well as cell chemotaxis. KEGG pathway analysis suggested that DEGs participated in the interactions between cytokines and viral proteins, osteoclast differentiation, and cytokine-cytokine receptor interactions. Real-time quantitative polymerase chain reaction analysis showed that *Complement C5a Receptor 1 (C5AR1), C-C Motif Chemokine Receptor 3 (CCR3), Complement C3a Receptor 1 (C3AR1)*, and *C-C Motif Chemokine Ligand 13 (CCL13)* expression levels were significantly upregulated in nasal polyps, whereas *C-C Motif Chemokine Ligand 4 (CCL4)* expression levels were significantly downregulated. Conclusions: The candidate genes identified in this study may influence the activation and accumulation of eosinophils, cell chemotaxis, and inflammatory responses, thereby potentially representing molecular targets for future studies of CRSwNP.

Keywords: Eosinophils, biomarkers, chronic rhinosinusitis, bioinformatics, weighted gene co-expression network analysis

Introduction

Chronic rhinosinusitis (CRS) is a prominent public health condition that affects patients' quality of life and imposes a substantial financial burden. CRS is presently divided into two major phenotypes: CRS with nasal polyps (CRSwNP) and without nasal polyps (CRSsNP). Defects in the sinonasal epithelial cell barrier, increased exposure to pathogenic and colonized bacteria, and dysregulation of the host immune system are all thought to play prominent roles in disease pathogenesis [1]. CRSwNP shows evident eosinophilic infiltration, and the mucosal eosinophilia of CRSwNP is related to severe symptoms [2], such as bilateral nasal polyps and sticky mucus secretion in the affected sinuses. These symptoms can cause patients to be susceptible to asthma, dysosmia, and nasal congestion. Patients with eosinophilic CRSwNP reportedly have a less favorable prognosis and higher recurrence rate than patients with

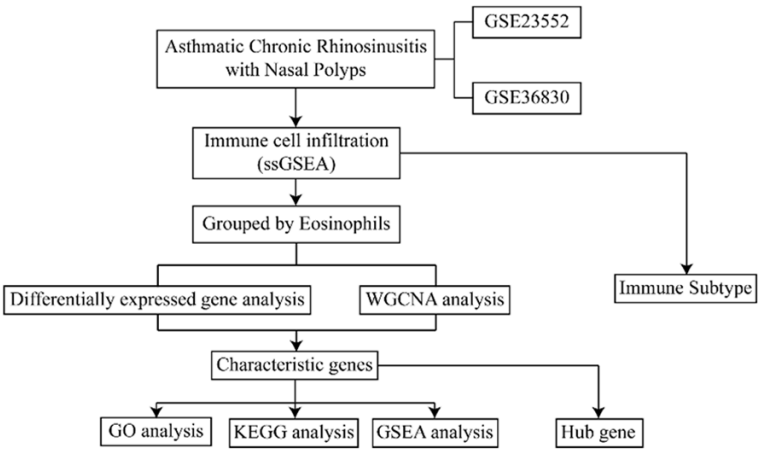


Figure 1. Flow chart of the methodologies used in this study.

CRSsNP. Current research has suggested that eosinophils (EOs) likely influence the pathophysiology of CRS because of their prolonged survival time in CRS; however, the pathogenesis of CRSwNP with infiltration of EOs and other immune cells in the microenvironment remains unclear. EOs can secrete various cytotoxic cationic proteins that seemingly influence the remodeling of the nasal mucosa and the severity of mucosal diseases [3]. Moreover, by secreting chemokines and cytokines, EOs can attract cells, including dendritic cells, neutrophils, Th2 lymphocytes, and macrophages/monocytes, that participate in the mucosal inflammatory response [4, 5]. Additional studies are required to explore EO behavior in CRSwNP pathophysiology and the functions of chemokines and cytokines in eosinophilic CRSwNP are needed.

In the current study, we aimed to identify genes associated with immunity that may be involved in the mucosal inflammatory response and EO accumulation in the pathophysiology of CRSwNP. These genes may also support the initiation, growth, and recurrence of polyps using comprehensive bioinformatics analyses.

Materials and methods

Microarray dataset and differentially expressed gene (DEG) identification

We used the R Bioconductor package 'Gene Expression Omnibus (GEO) query' [6] and downloaded the gene expression datasets Gene Series (GSE) 23552 [7] and GSE36830 [8] from the GEO. These datasets were obtained using the Affymetrix Human Exon 1.0 ST Array (accession number: Gene Platforms (GPL) 5175) and Affymetrix Human Genome U133 Plus 2.0 Array (accession number: GPL570), respectively. The GSE23552 dataset contained 28 samples, including 11 CRSwNP and 17 normal samples. The GSE36830 dataset contained 12 samples, including tissues of 6 nasal polyp (NP) from CRSwNP and 6 normal samples. All samples were used in this study. The data were classified into CRSwNP and normal groups. The raw data of the GSE36830 and GSE23552 datasets were standardized with the RMA algorithm implemented in the 'limma' R package [9]. The 'ComBat' routine from the 'sva' package was applied after combining both datasets. Principal component analysis (PCA), a statistical method used to determine the main variables in a multidimensional dataset, was used to represent differences among observations [10]. A flow diagram for the present analysis is shown in **Figure 1**.

Immune cell infiltration assessment

The single-sample gene set enrichment analysis (ssGSEA) algorithm was used to quantify the relative levels of infiltrating immune cells in CRSwNP [11]. Twenty-eight gene sets of various immune cell types were previously reported through ssGSEA [12]. The gene sets contained various immune cell types, including dendritic cells (DCs), $CD8^+$ T cells (cd8), regulatory T cells, and macrophages. SsGSEA scores were calculated by implementing the 'gsva' function from the 'Gene Set Variation Analysis (GSVA)' R package [13] and were used to represent the infiltration level of each immune cell type.

Differentially expressed genes (DEGs)

To analyze the possible biological impact of the EOs on the occurrence of CRSwNP, patients were classified into low and high infiltration groups in accordance with the median value of eosinophilic infiltration. The DEGs between the

high and low infiltration groups were screened using the R package 'limma' [9]. The results were subsequently filtered using the following parameters: *p*-value less than 0.05, and $|\log_2 FC| > 0.5$. Heatmaps and volcano plots of DEGs were constructed.

Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed based on the R package 'WGCNA' [14]. The data were then normalized to construct co-expression networks. Correlations were performed using the 'bicorrelation' function. Then, genes were clustered into network modules, and networks were constructed using the topological overlap measure. The following parameters of minModuleSize = 50 and mergeCutHeight = 1000 were used for data analysis. The intersection of the genes from the highest module and the DEGs was recorded, and the characteristic genes were identified.

Functional enrichment analysis

Gene Ontology (GO) analysis is a prevalent method used for enrichment analyses of large-scale gene sets, including cellular components, molecular functions, and biological processes. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database resource utilized to detail the advanced functions and applications of biological systems, including diseases, drugs, biological pathways, and genomes. We applied the 'clusterProfiler' R package (v3.0.0) to perform GO and KEGG pathway analyses [15]; a false discovery rate (FDR)-corrected $P < 0.05$ was considered statistically significant. To explore the differences in biological processes between different groups, gene set enrichment analysis (GSEA) was employed according to differential expression patterns between the GSE23552 and GSE36830 datasets. GSEA [11] is a calculation method used to assess whether a group of genes defined *a priori* reveals statistically significant consistent differences. This approach is generally employed to estimate the changes in biological processes and pathways in gene expression datasets. From the Molecular Signatures Database (MSigDB) database, the c2.cp.kegg. v6.2.-symbols [11] were downloaded and utilized for GSEA. An adjusted *p*-value < 0.05 was considered statistically significant.

Construction of the PPI network and identification of hub genes

The online protein interaction database Search Tool for the Retrieval of Interacting Genes (STRING, http://string-db.org) version 11.0 [16] was employed to predict the interactions between proteins and establish a protein-protein interaction (PPI) network for the chosen DEGs. Using the STRING database, genes with scores equal to or greater than 0.7 were selected to construct the network model, which was visualized using Cytoscape (v3.7.2) [17]. In the cooperative expression nets, the maximal clique centrality (MCC) algorithm is the most efficient way to explore the nodes in the nodeset. For each node, the MCC was counted via CytoHubba [18], which is a plug-in in Cytoscape. In the present study, genes containing the first eight values of MCC were regarded as central genes.

Construction of competing endogenous RNA (ceRNA) network

Prior to performing statistical analyses, data on the miRNA-mRNA-target interactions were downloaded from miRTarBase. Based on the common differential gene expression, possible miRNAs and related lncRNAs were predicted using the mirTarBase and StarBase databases. Cytoscape software was utilized to perform ceRNA network analysis (3.7.1) [17].

Construction of immune cell subtypes of chronic sinusitis with nasal polyps

Immune cell subtypes of chronic sinusitis with nasal polyps were calculated, and a nonlinear dimensionality reduction was performed using the R packages 'ConsensusClusterPlus' [19] and 'Rtsne' [20]. Patients were divided into A and B subtypes. Visualization of the immune cell subtypes was performed using the R package 'ggplot2'. Different immune cell types, a statistical analysis of hub genes, and correlations between immune cell subtypes were drawn with R package 'ggpubr'.

Real-time quantitative polymerase chain reaction (qPCR) analysis confirms candidate genes

To validate the expression of candidate genes in nasal polyps and the normal mucosa located in the nasal septum (control groups), total RNA

Table 1. Primers used for qPCR

Gene	Forward primer	Reverse primer
β-actin	CTCGCCTTTGCCGATCC	GAATCCTTCTGACCCATGCC
CXCR4	ACTACACCGAGGAAATGGGCT	CCCACAATGCCAGTTAAGAAGA
C5AR1	TCCTTCAATTATACCACCCCTGA	ACGCAGCGTGTTAGAAGTTTTAT
CCR3	TGGCATGTGTAAGCTCCTCTC	CCTGTCGATTGTCAGCAGGATTA
C3AR1	CCCTACGGCAGGTTCCTATG	GACAGCGATCCAGGCTAATGG
CCL13	CTCAACGTCCCATCTACTTGC	TCTTCAGGGTGTGAGCTTTCC
CXCL2	ACCAGAAGGAAGGAGGAAGC	CTCTGCAGCTGTGTCTCTCT
CCL4	CTGTGCTGATCCCAGTGAATC	TCAGTTCAGTTCCAGGTCATACA
PPBP	GTAACAGTGCGAGACCACTTC	CTTTGCCTTTCGCCAAGTTTC

from the nasal polyps and control groups (three samples for each group) were extracted using the Trizol Universal reagent (TIANGEN, Beijing, China) in accordance with the manufacturer's instructions. The studies involving human participants were reviewed and approved by the Ethics Committee of Affiliated Beijing Friendship Hospital, Capital Medical University. The patients provided their written informed consent to participate in this study (No: 2022-P2-132-01). The cDNA was synthesized using the Evo M-MLV Mix Kit with gDNA Clean for qPCR AG11728 (AG, Hunan, China). qPCR was performed with the ChamQ Universal SYBR qPCR Master Mix (Vazyme Biotech, Nanjing, China) using the ABI Quant Studio 7 Flex fast real-time PCR system. The amplification reaction protocol used was as follows: 95°C for 3 min, followed by 40 cycles at 95°C for 10 s, and 60°C for 30 s. *β-actin* was used as the internal control. Moreover, its relative mRNA expression level was used for quantifying the target gene expression. The relative mRNA expression level was calculated using the relative quantification ($2^{-\Delta\Delta CT}$) method. Primer sequences are listed in **Table 1**.

Statistical analysis

The statistical software R (version 3.5.0, R Foundation for Statistical Computing, http://www.r-project.org) was employed for all analyses. The Shapiro-Wilk normality test was utilized to explore normally distributed data. Two groups of continuous variables were compared using Student's *t*-test to estimate the statistical significance of normally distributed samples or Mann-Whitney U test for non-normal distributed samples. All statistical testing was two-sided, and the statistical significance threshold was set as a p-value < 0.05.

Results

Data preprocessing and gene screening

Gene expression matrix files for the GSE23552 and GSE36830 datasets were downloaded and standardized (**Figure 2A**, **2B**). Subsequently, these two datasets were merged and processed in a single batch. Among these datasets, the expression profiles before and after batch effects are presented as box plots (**Figure 2C**, **2D**). PCA plots of the two different GEO datasets, before and after removing the batch effect, are shown in **Figure 3**.

Analysis of immune cell infiltration-related DEGs

The ssGSEA algorithm was applied to evaluate the changes and impacts of immunologic features in CRSwNP pathogenesis. The relative enrichment scores of 28 different immune cell subtypes in the control (CON) and CRSwNP groups were obtained. The differential expressions between different sets of genes were displayed using heat maps (**Figure 4A**). PCA outcomes indicate differences in immune cell infiltration between the CRSwNP and normal groups (**Figure 4B**). Furthermore, in the merged data sets, the immune infiltration levels of various immune cells differed between the two groups (**Figure 4C**). In addition, the expression of HLA-DMA, HLA-DMA and HLA-DRA significantly differed between the CRSsNP and CRSwNP groups (**Figure 4D**).

Analysis and identification of DEGs

In subsequent analysis, we explored whether EOs exerted a marked effect on CRSwNP onset. Through this process, patients could be categorized into 'high' and 'low' infiltration groups according to the eosinophilic infiltration level. PCA revealed that gene expression levels significantly differed between the 'low' and 'high' infiltration groups (**Figure 5A**). Differential pooled gene expression analysis was performed using the 'limma' package, and 146 DEGs were identified. Of these, 131 genes were markedly elevated, and 15 were remarkably attenuated (**Figure 5B**, **5C**). To identify hub

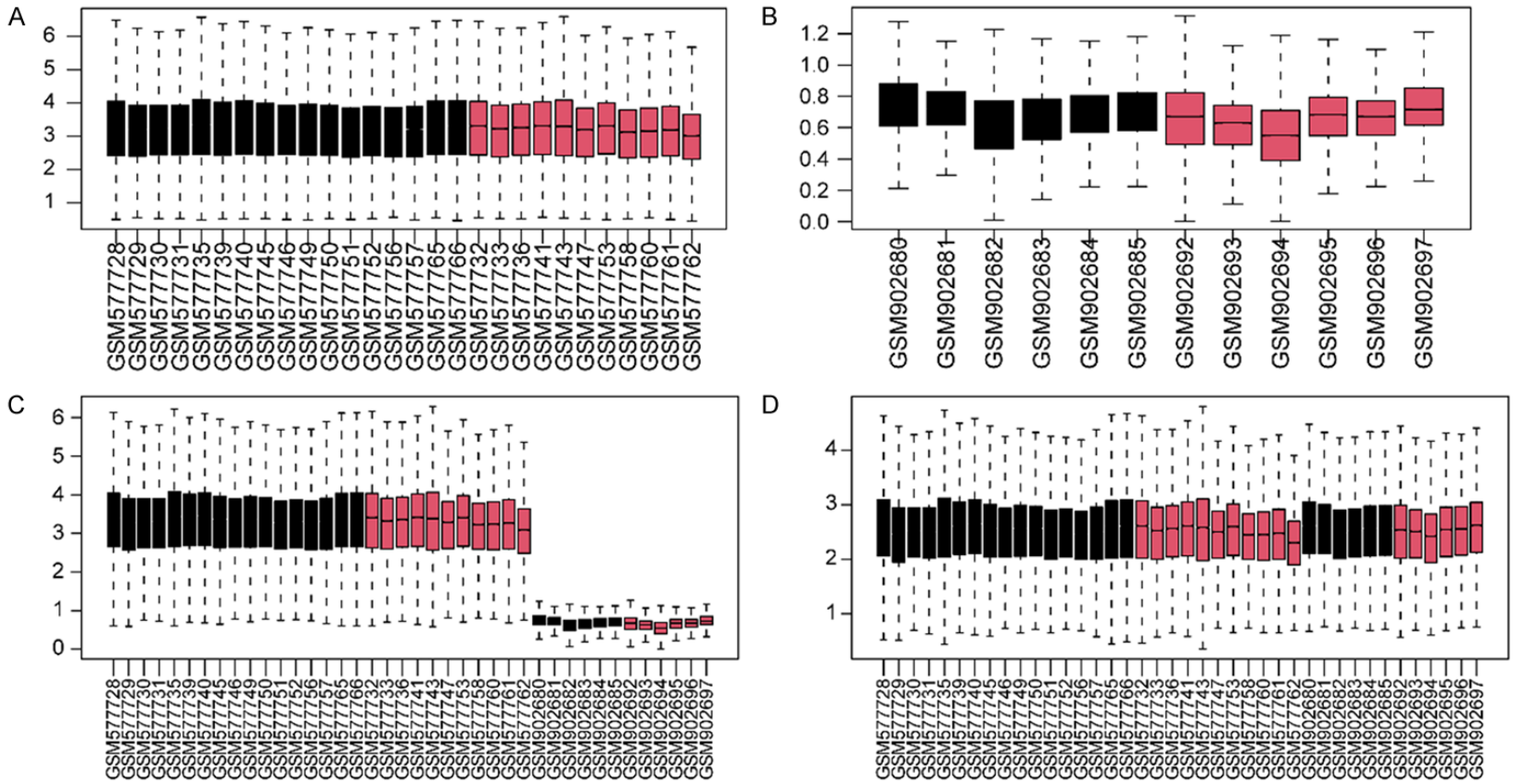


Figure 2. Box plot of unmerged and merged data from the GSE23552 and GSE36830 datasets. A and B. Show box plots of the standardized GSE2411 and GSE36830 datasets, respectively. C and D. Show box plots before and after eliminating the batch effect, respectively.

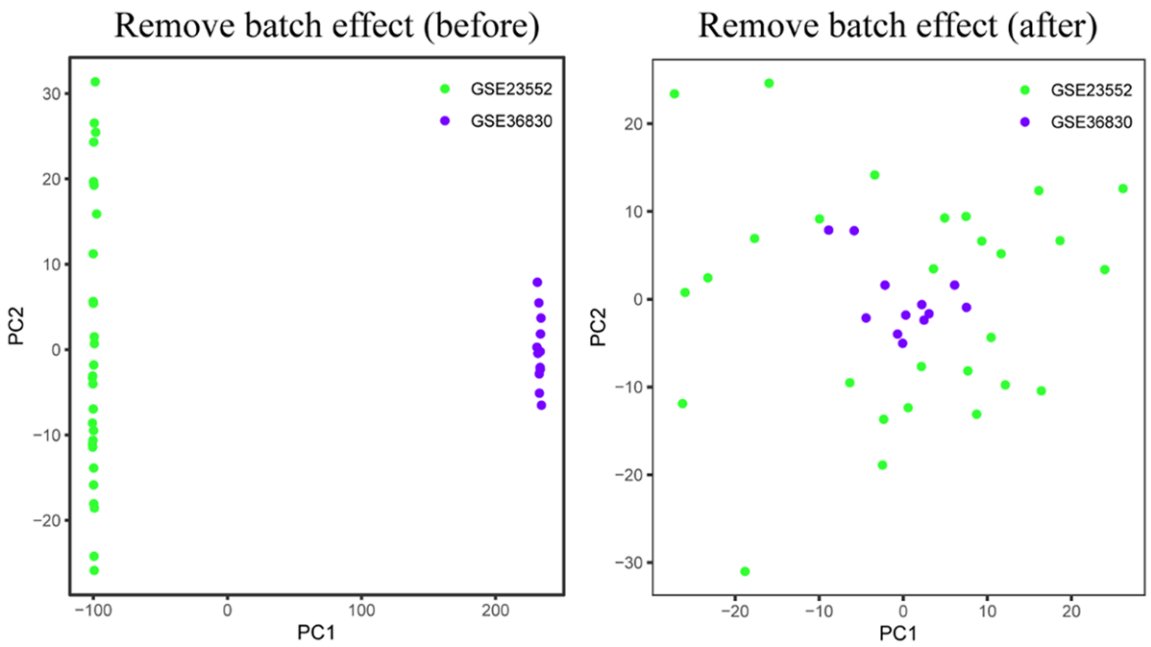


Figure 3. PCA plots of two different GEO datasets before and after removing the batch effect.

genes in the 'low' and 'high' infiltration groups, the pooled samples of the CRSwNP group were investigated using WGCNA. Module characteristic genes were then correlated with grouping information. Finally, 13 characteristic modules were identified. Of these, six modules showed positive correlations, whereas seven revealed negative correlations (**Figure 6A**, **6B**), and the greater the correlation coefficient, the higher the correlation of modules with high eosinophilic infiltration. We, therefore, selected the ME-black module with the maximum number of correlations, which revealed a significant correlation between this module and gene expression signatures in the 'high' infiltration group (correlation coefficient: 0.32, $P < 0.001$; **Figure 6C**). Next, we considered the intersection of the genes in a module and DEGs obtained previously (**Figure 6D**). Ultimately, 131 characteristic genes were used in the subsequent analyses.

Functional enrichment analysis

We implemented a functional enrichment analysis of the formerly acquired characteristic genes. The GO enrichment analysis suggested that the DEGs were principally associated with cell chemotaxis, especially leukocyte chemotaxis and migration (**Figure 7A**, **7B**). The KEGG pathway outcomes indicated that DEGs significantly affected the interactions between cytokines and viral proteins, osteoclast differentiation, and cytokine-cytokine receptor interactions (**Figure 7C**, **7D**). The enrichment outcomes of the interactions of cytokines and viral proteins, as well as the interaction between cytokines and their receptors, are shown in **Figure 7**. Concurrently, based on DEGs grouped by EO infiltration levels, GSEA results further showed that enriched pathways included KEGG cytokine-cytokine receptor interactions, a KEGG hematopoietic cell lineage, and a KEGG JAK-STAT signaling pathway in CRSwNP groups with high eosinophilic infiltration. **Figure 8** shows the KEGG pathway enrichment analysis results.

Construction of PPI network diagram and related regulatory network

STRING was used to develop a PPI network for DEGs (**Figure 9A**). The resulting network between genes was retrieved and imported into Cytoscape to generate a network visualization graph (**Figure 9B**). Downregulated and upregulated genes are shown in blue and red, respectively. Network analysis was performed using the 'CytoHubba' plug-in in Cytoscape. The localizations within high-density regions were selected as significant nodes defined as hub genes, namely *C-X-C Motif Chemokine Ligand 2 (CXCL2)*, *C-X-C Motif Chemokine Receptor 4 (CXCR4)*, *C-C Motif Chemokine*

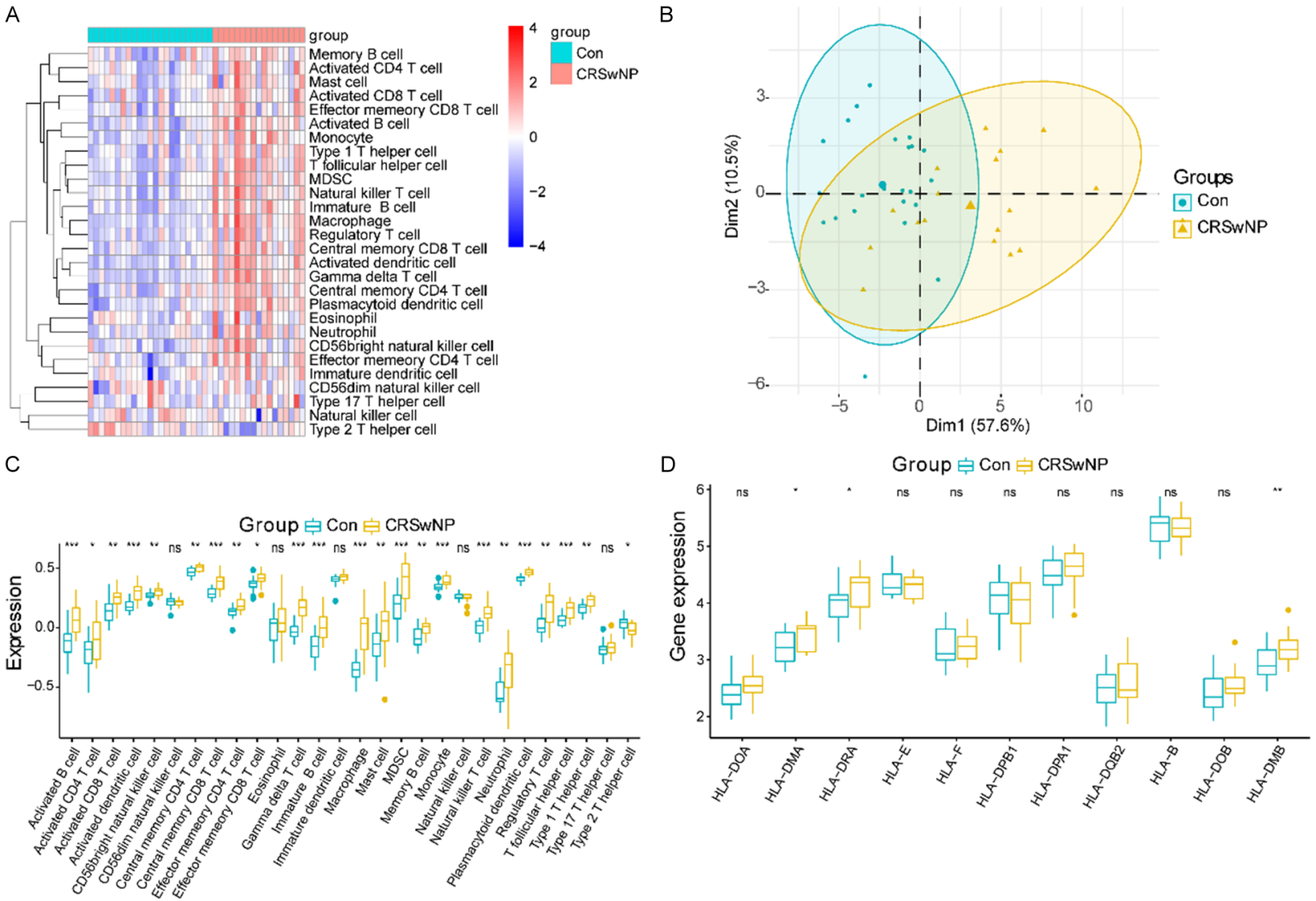


Figure 4. Analysis of the characteristic immunological effect of patients with CRSwNP. A. Infiltration levels of 28 immune cell types in the pooled datasets. B. PCA with different immune cell infiltration levels between the CON and the CRSwNP groups. C. Differential expressions of various types of infiltrating immune cells between the CON and CRSwNP groups. D. Analysis of differential expression of HLA family genes between the CRSwNP and CON groups. Blue-green represents normal tissue samples, and yellow represents tissue samples from the CRSwNP group.

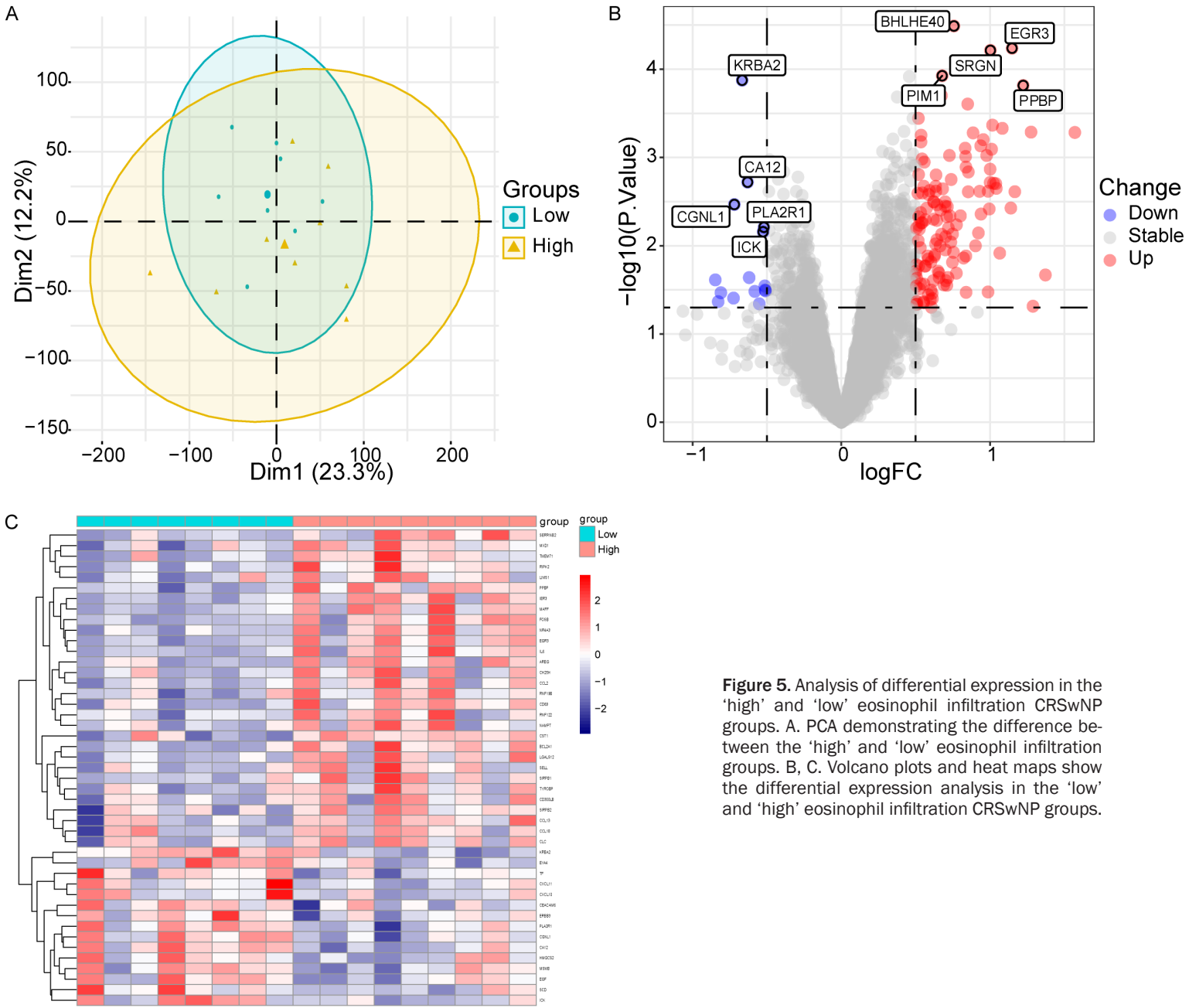


Figure 5. Analysis of differential expression in the 'high' and 'low' eosinophil infiltration CRSwNP groups. A. PCA demonstrating the difference between the 'high' and 'low' eosinophil infiltration groups. B, C. Volcano plots and heat maps show the differential expression analysis in the 'low' and 'high' eosinophil infiltration CRSwNP groups.

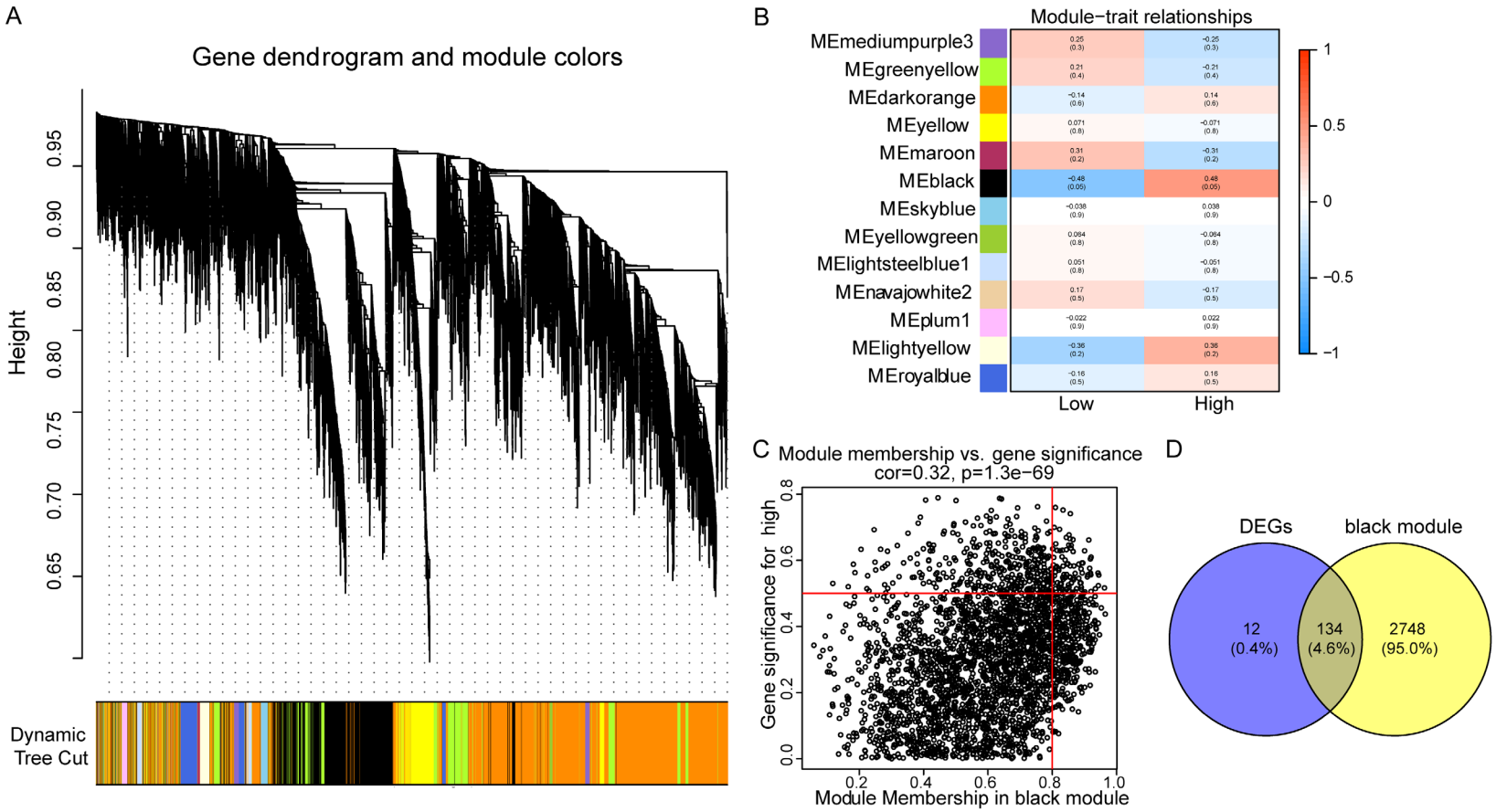
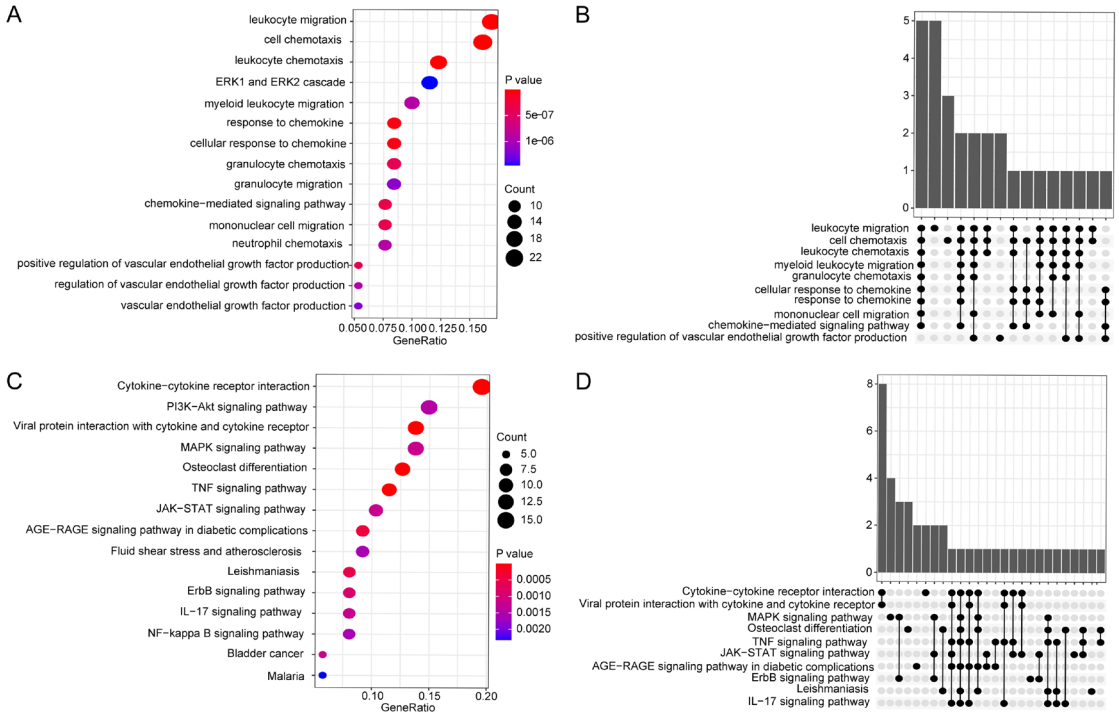
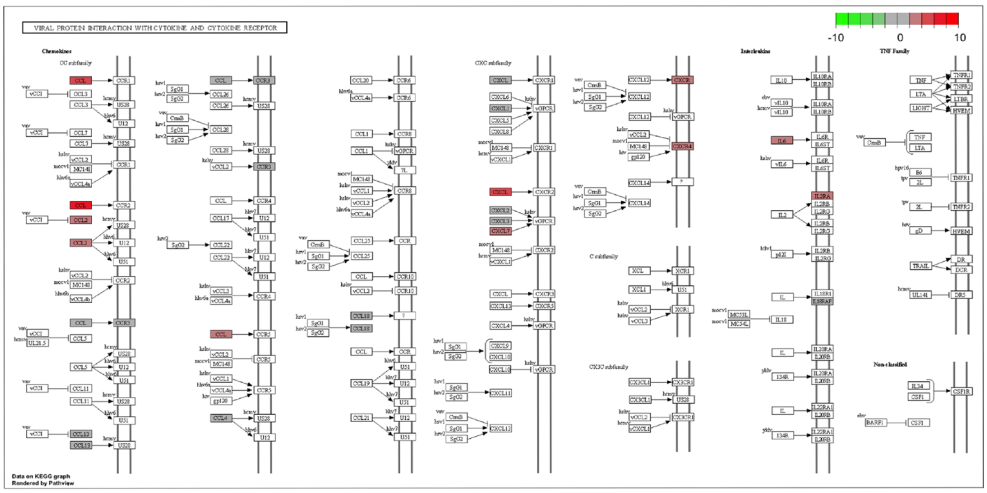


Figure 6. WGCNA results identify characteristic genes. A, B. Identification modules in the 'high' and 'low' eosinophil infiltration CRSwNP groups using WGCNA; 13 characteristic modules were identified. The MEblack module has the highest phenotypic correlation with the 'high' eosinophil infiltration group. C. Correlation analysis of the module MEblack features with the 'high' eosinophil infiltration signatures. D. Venn diagram showing 134 genes in the intersection between differential genes with the module MEblack using WGCNA.

Key candidate genes of CRSwNP



E



F

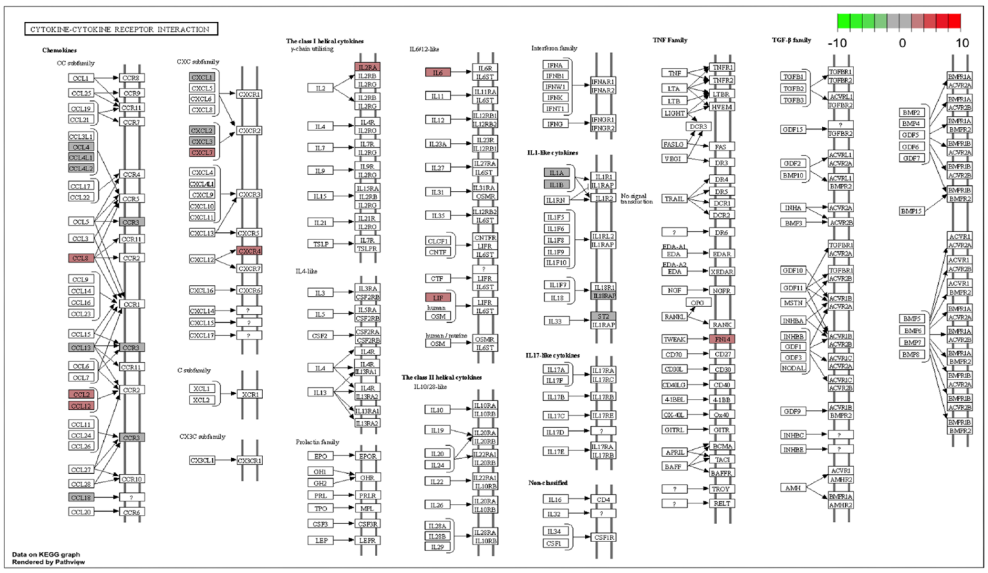


Figure 7. Functional enrichment analysis of feature genes. A, B. GO enrichment analysis suggesting that DEGs are primarily related to leukocyte migration, leukocyte chemotaxis, and cell chemotaxis. C, D. KEGG pathway analysis confirming that the differentially expressed genes are significantly related to the interaction between cytokines and viral proteins, osteoclast differentiation, and cytokine-cytokine receptor interaction. E, F. The two signal pathways with the highest enrichment fraction represent the interactions between viral proteins and cytokine-cytokine receptor interactions.

Ligand 4 (CCL4), *Pro-Platelet Basic Protein (PPBP)*, *Complement C5a Receptor 1 (C5AR1)*, *C-C Motif Chemokine Receptor 3 (CCR3)*, *Complement C3a Receptor 1 (C3AR1)*, and *C-C Motif Chemokine Ligand 13 (CCL13)* (**Figure 9C**). The miRTarBase database was used to download miRNA-mRNA-target interactions. Through constructing a PPI network and identifying hub genes, we obtained eight hub genes. A ceRNA network was constructed based on the miRNA-mRNA-lncRNA interactions (**Figure 9D**).

Identification of CRSwNP-related immune subtypes

Based on the infiltration levels of immune cells in CRSwNP, a consensus clustering method was used to perform a cluster sample analysis. According to the consistent matrix and graph-clustering, when k = 2, interference between two groups was minimal (**Figure 10A**). Therefore, patients were classified into subtypes A and B (**Figure 10B**). To understand the states of the two subtypes, we investigated the differential expression of eight hub genes and different immune cell type infiltration levels in subtypes A and B. The outcomes showed that the levels of different immune cell infiltrates between subtypes A and B varied (**Figure 10C**), including EOs. *C5AR1*, *CCR3*, *C3AR1*, CCL13 and CCL14 expression levels showed significant differences between subtypes A and B (**Figure 10D**).

qPCR validation of candidate genes

To validate these results, we selectively performed a qPCR analysis of eight important DEGs, namely *CXCL2*, *CXCR4*, *CCL4*, *PPBP*, *C5AR1*, *CCR3*, *C3AR1*, and *CCL13*, in three patients with EO infiltration and three healthy controls. The relative mRNA expression levels of *CCL4* were significantly lower in nasal polyps than in the normal mucosa. The expression levels of *C5AR1, CCR3*, *C3AR1*, and *CCL13* were significantly higher in nasal polyps than in the normal mucosa. However, there were no significant differences in the relative mRNA expression levels of *PPBP, CXCR4, and CXCL2* (**Figure 11**).

Discussion

CRS refers to chronic inflammation of the nasal and paranasal sinus mucosa, which affects approximately 12% of the population worldwide [21]. The endotype of inflammation has been determined in recent years. Specifically, the Caucasian populations exhibit the T helper lymphocyte cell type 2 (Th2) inflammatory pathway (known as T2 inflammation) with EO deflection, whereas Asian populations display the Th17 or Th1 inflammatory pathway (T17 and T1 inflammation) [22]. Eosinophilic CRS (ECRS) is a major subset of CRS. CRSwNP (65%-90% of cases) is often associated with the presence of eosinophilic infiltrates in the nasal mucosa and paranasal sinuses [23]. Therefore, there is an urgent need to detail the role of EOs in CRS. Genomics, proteomics, and metabolomics can provide large datasets for extrapolating the function of EOs in CRSwNP. EOs exert several effects on the pathophysiology of CRSwNP: chemotaxis, immune regulation, and cytotoxic function. One of the major functions of EOs is to secrete a variety of known cytotoxic cationic proteins, with more than 30 growth factors, chemokines, and cytokines involved in immune regulation [24]. In our study, eight hub genes (*CXCL2*, *CXCR4*, *CCL4*, *PPBP*, *C5AR1*, *CCR3*, *C3AR1,* and *CCL13*) showed the highest scores in the PPI network analysis. Therefore, it is essential to investigate the hub genes to identify the potential regulatory role of EOs in CRSwNP pathogenesis.

CCR3 is involved in allergic airway inflammation; it is highly expressed in basophils and EOs and is also found in airway epithelial, Th1, and Th2 cells. This receptor may assist in the activation and accumulation of EOs together with other inflammatory cells present in allergic airways. The increased expression of the CCR3 chemokine receptor may be in accordance with a variety of chemokines, thus promoting the

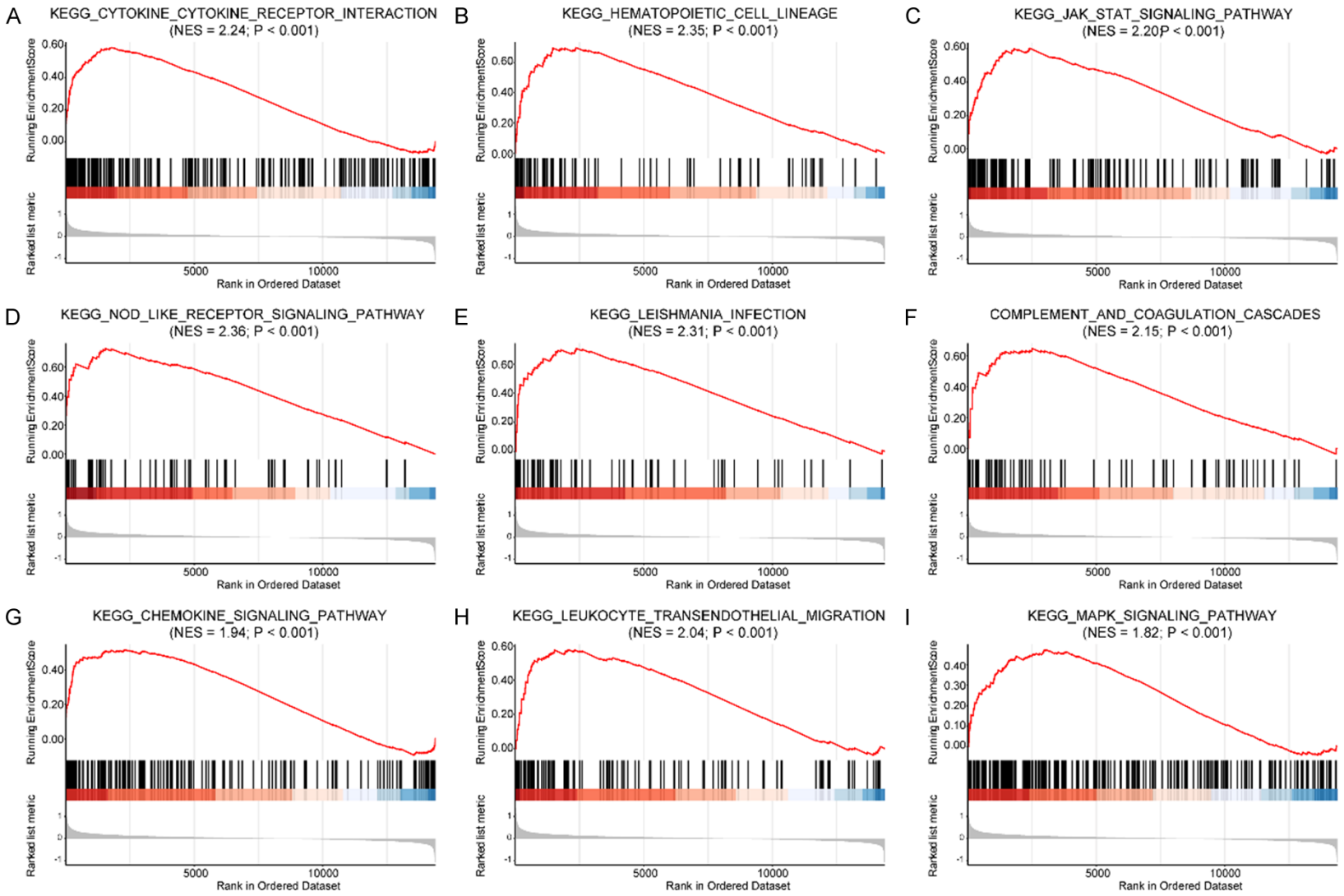


Figure 8. GSEA results based on differential expressions as grouped by EOs. GSEA was implemented to explore differentially expressed genes. GSEA results (A-I) further demonstrated that the pathways enriched included the KEGG cytokine-cytokine receptor interaction, KEGG hematopoietic cell lineage, and KEGG JAK-STAT signaling pathway.

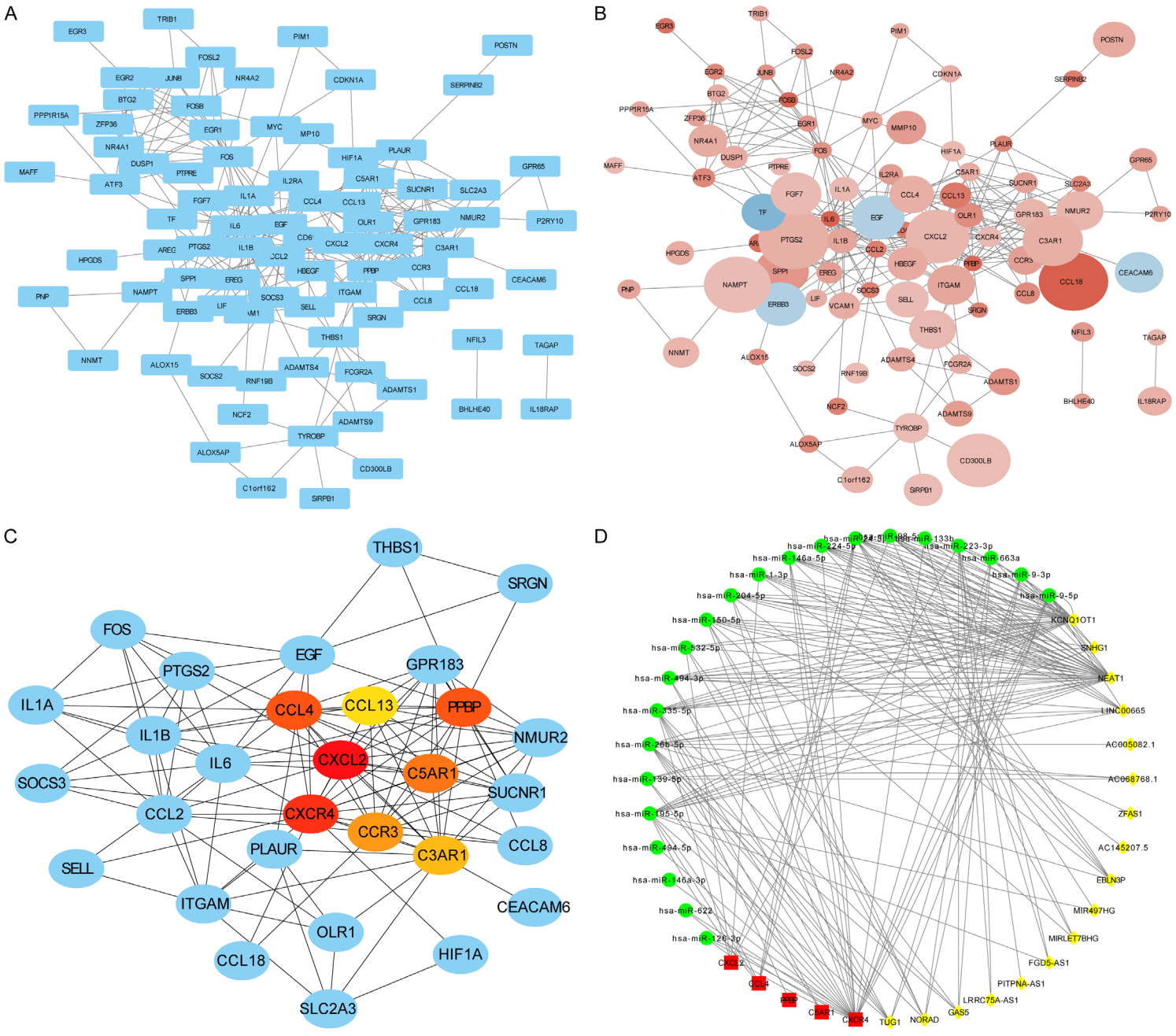

A
TRIB1
EGR3
PIM1
POSTN
FOSL2
EGR2
JUNB
NR4A2
CDKN1A
SERPINB2
BTG2
FOSB
PPP1R15A
ZFP36
EGR1
MYC
MP10
NR4A1
PLAUR
DUSP1
FOS
HIF1A
MAFF
PTPRE
GPR65
ATF3
IL2RA
C5AR1
SUCNR1
SLC2A3
FGF7
IL1A
CCL4
CCL13
TF
IL6
EGF
OLR1
GPR183
NMUR2
P2RY10
AREG
PTGS2
IL1B
CXCL2
CXCR4
C3AR1
HPGDS
SPP1
EREG
HBEGF
PPBP
CCR3
CEACAM6
PNP
NAMPT
SOCS3
ITGAM
CCL18
LIF
CCL8
ERBB3
SELL
SRGN
THBS1
NFIL3
TAGAP
NNMT
ALOX15
SOCS2
RNF19B
ADAMTS4
FCGR2A
ADAMTS1
BHLHE40
IL18RAP
NCF2
ADAMTS9
TYROBP
ALOX5AP
CD300LB
C1orf162
SIRPB1
B
TRIB1
EGR3
PIM1
POSTN
FOSL2
EGR2
JUNB
NR4A2
CDKN1A
SERPINB2
BTG2
FOSB
PPP1R15A
ZFP36
EGR1
MYC
MMP10
NR4A1
PLAUR
DUSP1
FOS
HIF1A
MAFF
PTPRE
GPR65
ATF3
IL2RA
C5AR1
SUCNR1
SLC2A3
FGF7
IL1A
CCL4
CCL13
TF
IL6
EGF
OLR1
GPR183
NMUR2
P2RY10
AREG
PTGS2
IL1B
CXCL2
CXCR4
C3AR1
HPGDS
SPP1
EREG
HBEGF
PPBP
CCR3
CEACAM6
PNP
NAMPT
SOCS3
ITGAM
CCL18
LIF
CCL8
ERBB3
VCAM1
SELL
SRGN
THBS1
NFIL3
TAGAP
NNMT
ALOX15
SOCS2
RNF19B
ADAMTS4
FCGR2A
ADAMTS1
BHLHE40
IL18RAP
NCF2
ADAMTS9
TYROBP
ALOX5AP
CD300LB
C1orf162
SIRPB1
C
THBS1
SRGN
FOS
EGF
PTGS2
GPR183
IL1A
CCL4
CCL13
PPBP
IL1B
NMUR2
IL6
CXCL2
C5AR1
SOCS3
SUCNR1
CCL2
CXCR4
CCR3
PLAUR
C3AR1
CCL8
SELL
ITGAM
CEACAM6
OLR1
CCL18
HIF1A
SLC2A3
D
hsa-miR-224-5p
hsa-miR-133b
hsa-miR-223-3p
hsa-miR-146a-5p
hsa-miR-663a
hsa-miR-1-3p
hsa-miR-9-3p
hsa-miR-204-5p
hsa-miR-9-5p
hsa-miR-150-5p
KCNQ1OT1
hsa-miR-532-5p
SNHG1
hsa-miR-494-3p
NEAT1
hsa-miR-335-5p
LINC00665
hsa-miR-26b-5p
AC005082.1
hsa-miR-139-5p
AC068768.1
hsa-miR-195-5p
ZFAS1
hsa-miR-494-5p
AC145207.5
hsa-miR-146a-3p
EBLN3P
hsa-miR-622
MIR497HG
hsa-miR-126-3p
MIRLET7BHG
CXCL2
FGD5-AS1
CCL4
PITPNA-AS1
PPBP
LRRC75A-AS1
C5AR1
CXCR4
TUG1
NORAD
GAS5

Figure 9. Construction of a PPI network diagram and the ceRNA network. A. PPI networks generated with differentially expressed genes using STRING. B. Data imported into Cytoscape. Red and blue represent upregulated and downregulated genes, respectively. Different logFC values are color-coded (shades of color). The circle size was proportional to the *p*-value. C. Extraction of high-density regions from the PPI network conducted using CytoHubba. The top eight ranking genes are selected as hub genes. D. Using Cytoscape, the ceRNA net was established according to hub genes.

pathogenesis of nasal polyposis via the migration and long-term accumulation of inflammatory cells (for instance, EOs); this ultimately causes inflammatory infiltration of nasal polyps [25]. In mice, Ccr3 expression downregulation inhibits histopathological lesions and EO infiltration of the nasal cavity; this reduces the production of Th2 cytokines in the serum, leading to the remission of allergic symptoms [26]. The results of our study show that the *CCR3* expression was significantly upregulated in nasal polyps compared with that in normal nasal septum mucosa. However, further studies are required to validate the role of *CCR3* in CRSwNP.

The complement system is an important regulator of innate immunity. Once activated, the complement system can trigger an inflammatory response and immunosuppressive functions. Complement C3a receptor 1 (*C3AR1*) is a gene encoding the receptor of the chemotactic and inflammatory peptide anaphylatoxin C3a. Besides modulating the inflammatory response, C3AR1 also influences various crucial cellular processes, including cytokine production and secretion as well as immune cell activation [27], proliferation, migration [26], and differentiation. In some diseased cells, such as osteosarcoma cells, a C3AR1 overexpression has suppressed proliferation, migration, and invasion and induced apoptosis [28]. However, further studies are needed to determine whether there is a relationship between C3AR1 and Eos, additional research is required to analyze the expression profiles of C3AR1 in CRSwNP.

The protein encoded by the C-C motif chemokine ligand 4 (*CCL4*) gene is secreted and possesses inflammatory and chemokinetic properties. There is a positive association between plasmatic CCL4 and the formerly studied inflammatory mediators (vascular endothelial growth factor and tumor necrosis factor-α), which is related to poor prognoses in patients. CCL4 is mainly located in nasal polyp epithelial cells [29]. CCL4 levels were more strongly correlated with eosinophil count and expression of eosinophil granule proteins [30], which highlight the important role of CCL4 in the mechanisms underlying eosinophil recruitment into the airway and may provide a novel insight into this potential therapeutic target [31]. However, in our study, *CCL4* was downregulated in nasal polyps compared with that in normal nasal septum mucosa. Whether CCL4 is involved in eosinophil recruitment and the occurrence and development of nasal polyps is currently further under investigation.

CCL13/Monocyte Chemotactic Protein 4 (MCP-4) is a chemokine of the CC family, and it exerts chemotactic effects on monocytes, basophils, EOs, T cells, immature dendritic cells, and macrophages. This chemokine can cause major immune-regulatory responses via its actions on endothelial, muscle, and epithelial cells. Nasal CCL13 may aggravate nasal mucosal allergies [32]. Numerous studies have suggested that CCL13 participates in various chronic inflammatory diseases. In these diseases, CCL13 acts as a key molecule, participating in the selective recruitment and subsequent activation of inflammatory tissue cell lines [33]. In this study, compared with the results in normal tissues, the *CCL13* expression in nasal polyps was significantly upregulated. However, the biological functions and the effect of CCL13 remain poorly studied in CRSwNP, especially CRSwNP with infiltrations of EOs.

Our study has a few limitations. First, we obtained data from the GEO database, and the microarray data used in the study were generated from a limited number of samples; thus, it is unclear whether the sample size was appropriate for in-depth external validation. Second, our results are based on bioinformatics analyses, and only preliminary experiments were conducted to confirm our results. qPCR analysis showed that the relative expression levels of *C5AR1*, *CCR3*, *C3AR1*, and *CCL13* were significantly upregulated in nasal polyps, whereas those of *CCL4* and *PPBP* were significantly

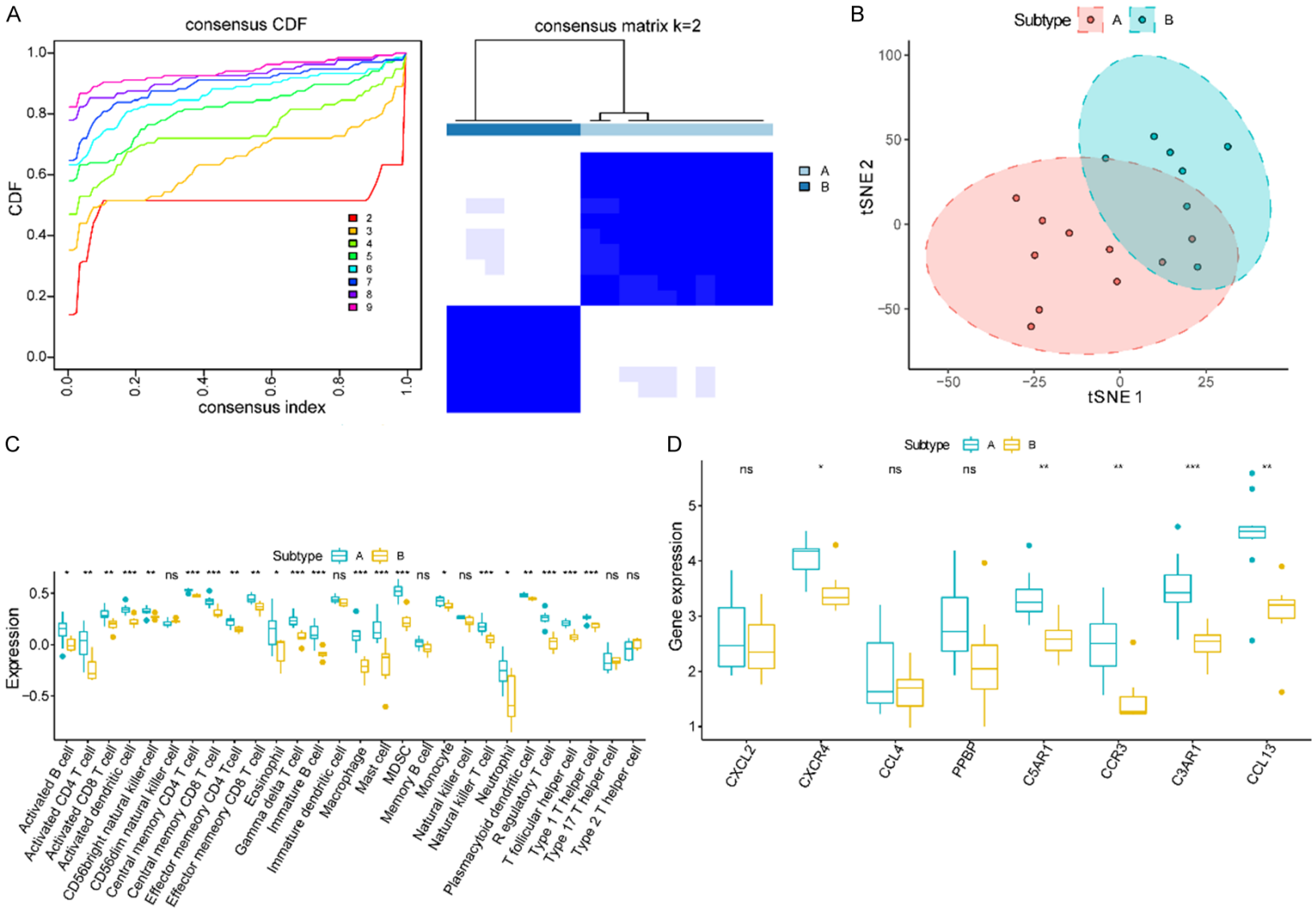


Figure 10. Construction of immune subtype with CRSwNP. A. CDF plot with different K values and a graph for clustering (K = 2). B. tSNE analysis used to show differences between subtypes A and B. C. Differential immune cell infiltration levels between subtypes A and B. D. Differential expression of eight hub genes in subtypes A and B.

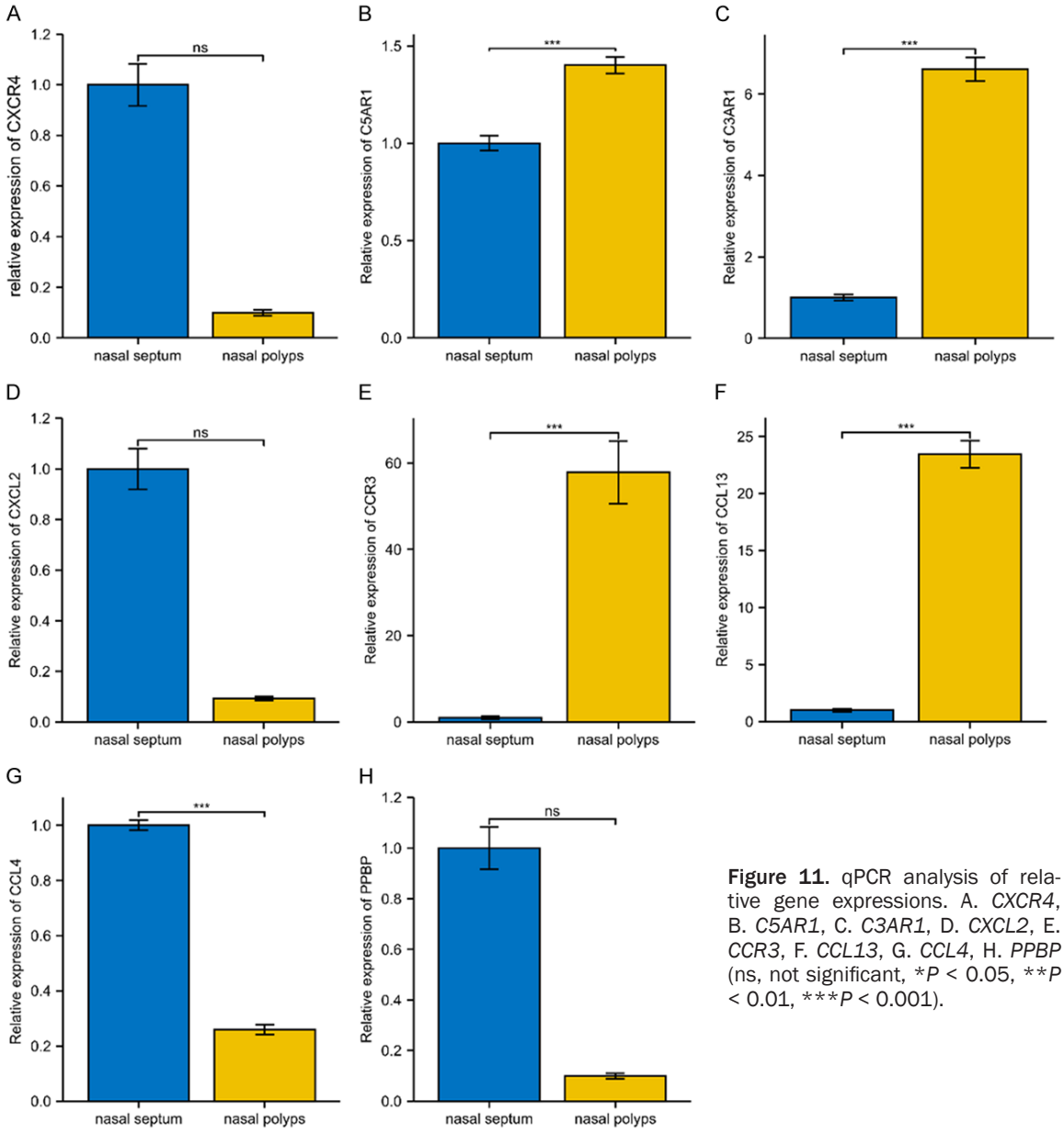


Figure 11. qPCR analysis of relative gene expressions. A. *CXCR4*, B. *C5AR1*, C. *C3AR1*, D. *CXCL2*, E. *CCR3*, F. *CCL13*, G. *CCL4*, H. *PPBP* (ns, not significant, **P* < 0.05, ***P* < 0.01, ****P* < 0.001).

downregulated. However, a small sample size was used for these experiments. Hence, the findings from our study should be validated using larger sample sizes and in-depth experimental validation; for example, immunohistochemical and western blot analyses are needed to systematically clarify the effect of the hub gene and its potential involvement in the pathophysiology of CRSwNP. Third, in the validation experiment, we used nasal polyps with 'high' and 'low' infiltration groups according to the eosinophilic infiltration level. However, considering that the data in our study came from the GEO database of CRSwNP and normal samples, our preliminary validation utilizing the nasal polyps with EO infiltration and the normal mucosa tissues as groups merely comprised a simple grouping of high/low eosinophil infiltration. In follow-up studies, we could remedy this limitation by performing a second-generation sequencing analysis and using all transcriptome data for ssGSEA scores to study the expression of eight hub genes in high/low eosinophil states.

Nonetheless, in this work, we determined that eight hub genes may influence cell chemotaxis, immune cell infiltration, and inflammatory

responses using bioinformatics analyses, and confirmed the role of these hub genes via qPCR analysis. However, more experiments are needed to verify how the key genes regulate eosinophil infiltration and cytokine release, thus explaining the role of eosinophils in the occurrence and development of CRSwNP.

Acknowledgements

The authors would like to thank Zhi-Jie Ma for her supportive contribution to our research development.

Disclosure of conflict of interest

None.

Address correspondence to: Gang Chen, Department of Otorhinolaryngology Head and Neck Surgery, Affiliated Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China. Tel: +86-010-63138417; E-mail: chengang@ccmu.edu.cn

References

[1] Stevens WW, Schleimer RP and Kern RC. Chronic rhinosinusitis with nasal polyps. J Allergy Clin Immunol Pract 2016; 4: 565-572.

[2] Yamada T, Miyabe Y, Ueki S, Fujieda S, Tokunaga T, Sakashita M, Kato Y, Ninomiya T, Kawasaki Y, Suzuki S and Saito H. Eotaxin-3 as a plasma biomarker for mucosal eosinophil infiltration in chronic rhinosinusitis. Front Immunol 2019; 10: 74.

[3] Sun DI, Joo YH, Auo HJ and Kang JM. Clinical significance of eosinophilic cationic protein levels in nasal secretions of patients with nasal polyposis. Eur Arch Otorhinolaryngol 2009; 266: 981-986.

[4] Poposki JA, Uzzaman A, Nagarkar DR, Chustz RT, Peters AT, Suh LA, Carter R, Norton J, Harris KE, Grammer LC, Tan BK, Chandra RK, Conley DB, Kern RC, Schleimer RP and Kato A. Increased expression of the chemokine CCL23 in eosinophilic chronic rhinosinusitis with nasal polyps. J Allergy Clin Immunol 2011; 128: 73-81, e4.

[5] Yang D, Rosenberg HF, Chen Q, Dyer KD, Kurosaka K and Oppenheim JJ. Eosinophil-derived neurotoxin (EDN), an antimicrobial protein with chemotactic activities for dendritic cells. Blood 2003; 102: 3396-3403.

[6] Davis S and Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 2007; 23: 1846-1847.

[7] Plager DA, Kahl JC, Asmann YW, Nilson AE, Pallanch JF, Friedman O and Kita H. Gene transcription changes in asthmatic chronic rhinosinusitis with nasal polyps and comparison to those in atopic dermatitis. PLoS One 2010; 5: e11450.

[8] Stevens WW, Ocampo CJ, Berdnikovs S, Sakashita M, Mahdavinia M, Suh L, Takabayashi T, Norton JE, Hulse KE, Conley DB, Chandra RK, Tan BK, Peters AT, Grammer LC 3rd, Kato A, Harris KE, Carter RG, Fujieda S, Kern RC and Schleimer RP. Cytokines in chronic rhinosinusitis. Role in eosinophilia and aspirin-exacerbated respiratory disease. Am J Respir Crit Care Med 2015; 192: 682-694.

[9] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.

[10] Raychaudhuri S, Stuart JM and Altman RB. Principal components analysis to summarize microarray experiments: application to sporulation time series. Pac Symp Biocomput 2000: 455-466.

[11] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.

[12] Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A, Bruneval P, Fridman WH, Becker C, Pagès F, Speicher MR, Trajanoski Z and Galon J. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013; 39: 782-795.

[13] Hänzelmann S, Castelo R and Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013; 14: 7.

[14] Langfelder P and Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008; 9: 559.

[15] Yu G, Wang LG, Han Y and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.

[16] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ and Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.

[17] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[18] Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014; 8 Suppl 4: S11.
[19] Wilkerson MD and Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010; 26: 1572-1573.
[20] Cieslak MC, Castelfranco AM, Roncalli V, Lenz PH and Hartline DK. t-Distributed Stochastic Neighbor Embedding (t-SNE): a tool for eco-physiological transcriptomic analysis. Mar Genomics 2020; 51: 100723.
[21] Fokkens WJ, Lund VJ, Hopkins C, Hellings PW, Kern R, Reitsma S, Toppila-Salmi S, Bernal-Sprekelsen M, Mullol J, Alobid I, Terezinha Anselmo-Lima W, Bachert C, Baroody F, von Buchwald C, Cervin A, Cohen N, Constantinidis J, De Gabory L, Desrosiers M, Diamant Z, Douglas RG, Gevaert PH, Hafner A, Harvey RJ, Joos GF, Kalogjera L, Knill A, Kocks JH, Landis BN, Limpens J, Lebeer S, Lourenco O, Meco C, Matricardi PM, O'Mahony L, Philpott CM, Ryan D, Schlosser R, Senior B, Smith TL, Teeling T, Tomazic PV, Wang DY, Wang D, Zhang L, Agius AM, Ahlstrom-Emanuelsson C, Alabri R, Albu S, Alhabash S, Aleksic A, Aloulah M, Al-Qudah M, Alsaleh S, Baban MA, Baudoin T, Balvers T, Battaglia P, Bedoya JD, Beule A, Bofares KM, Braverman I, Brozek-Madry E, Richard B, Callejas C, Carrie S, Caulley L, Chussi D, de Corso E, Coste A, El Hadi U, Elfarouk A, Eloy PH, Farrokhi S, Felisati G, Ferrari MD, Fishchuk R, Grayson W, Goncalves PM, Grdinic B, Grgic V, Hamizan AW, Heinichen JV, Husain S, Ping TI, Ivaska J, Jakimovska F, Jovancevic L, Kakande E, Kamel R, Karpischenko S, Kariyawasam HH, Kawauchi H, Kjeldsen A, Klimek L, Krzeski A, Kopacheva Barsova G, Kim SW, Lal D, Letort JJ, Lopatin A, Mahdjoubi A, Mesbahi A, Netkovski J, Nyenbue Tshipukane D, Obando-Valverde A, Okano M, Onerci M, Ong YK, Orlandi R, Otori N, Ouennoughy K, Ozkan M, Peric A, Plzak J, Prokopakis E, Prepageran N, Psaltis A, Pugin B, Raftopulos M, Rombaux P, Riechelmann H, Sahtout S, Sarafoleanu CC, Searyoh K, Rhee CS, Shi J, Shkoukani M, Shukuryan AK, Sicak M, Smyth D, Sindvongs K, Soklic Kosak T, Stjarne P, Sutikno B, Steinsvag S, Tantilipikorn P, Thanaviratananich S, Tran T, Urbancic J, Valiulius A, Vasquez de Aparicio C, Vicheva D, Virkkula PM, Vicente G, Voegels R, Wagenmann MM, Wardani RS, Welge-Lussen A, Witterick I, Wright E, Zabolotniy D, Zsolt B and Zwetsloot CP. European position paper on rhinosinusitis and nasal polyps 2020. Rhinology 2020; 58: 1-464.
[22] Tomassen P, Vandeplas G, Van Zele T, Cardell LO, Arebro J, Olze H, Förster-Ruhrmann U, Kowalski ML, Olszewska-Ziąber A, Holtappels G, De Ruyck N, Wang X, Van Drunen C, Mullol J, Hellings P, Hox V, Toskala E, Scadding G, Lund V, Zhang L, Fokkens W and Bachert C. Inflammatory endotypes of chronic rhinosinusitis based on cluster analysis of biomarkers. J Allergy Clin Immunol 2016; 137: 1449-1456, e4.
[23] Sitarek P, Zielinska-Blizniewska H, Dziki L, Milonski J, Przybylowska K, Mucha B, Olszewski J and Majsterek I. Association of the -14C/G MET and the -765G/C COX-2 gene polymorphisms with the risk of chronic rhinosinusitis with nasal polyps in a Polish population. DNA Cell Biol 2012; 31: 1258-1266.
[24] Davoine F and Lacy P. Eosinophil cytokines, chemokines, and growth factors: emerging roles in immunity. Front Immunol 2014; 5: 570.
[25] Fundová P, Funda DP, Kovář D, Holý R, Navara M and Tlaskalová-Hogenová H. Increased expression of chemokine receptors CCR1 and CCR3 in nasal polyps: molecular basis for recruitment of the granulocyte infiltrate. Folia Microbiol (Praha) 2013; 58: 219-224.
[26] Yuan J, Liu Y, Yu J, Dai M, Zhu Y, Bao Y, Peng H, Liu K and Zhu X. Gene knockdown of CCR3 reduces eosinophilic inflammation and the Th2 immune response by inhibiting the PI3K/AKT pathway in allergic rhinitis mice. Sci Rep 2022; 12: 5411.
[27] Chen H, Sun Q, Zhang C, She J, Cao S, Cao M, Zhang N, Adiila AV, Zhong J, Yao C, Wang Y, Xia H and Lan L. Identification and validation of CYBB, CD86, and C3AR1 as the key genes related to macrophage infiltration of gastric cancer. Front Mol Biosci 2021; 8: 756085.
[28] Zou T, Liu W, Wang Z, Chen J, Lu S, Huang K and Li W. C3AR1 mRNA as a potential therapeutic target associates with clinical outcomes and tumor microenvironment in osteosarcoma. Front Med (Lausanne) 2021; 8: 642615.
[29] Kobayashi Y, Chu HH, Kanda A, Yun Y, Shimono M, Nguyen LM, Mitani A, Suzuki K, Asako M and Iwai H. CCL4 functions as a biomarker of type 2 airway inflammation. Biomedicines 2022; 10: 1779.
[30] Kobayashi Y, Konno Y, Kanda A, Yamada Y, Yasuba H, Sakata Y, Fukuchi M, Tomoda K, Iwai H and Ueki S. Critical role of CCL4 in eosinophil recruitment into the airway. Clin Exp Allergy 2019; 49: 853-860.
[31] Li L, Liu YD, Zhan YT, Zhu YH, Li Y, Xie D and Guan XY. High levels of CCL2 or CCL4 in the

tumor microenvironment predict unfavorable survival in lung adenocarcinoma. Thorac Cancer 2018; 9: 775-784.

[32] Baumann R, Rabaszowski M, Stenin I, Tilgner L, Scheckenbach K, Wiltfang J, Schipper J, Chaker A and Wagenmann M. Comparison of the nasal release of IL-4, IL-10, IL-17, CCL13/MCP-4, and CCL26/eotaxin-3 in allergic rhinitis during season and after allergen challenge. Am J Rhinol Allergy 2013; 27: 266-272.

[33] Mendez-Enriquez E and García-Zepeda EA. The multiple faces of CCL13 in immunity and inflammation. Inflammopharmacology 2013; 21: 397-406.