APJAI
Asian Pacific Journal of
Allergy and Immunology

ORIGINAL ARTICLE

Gene expression profiles of mRNA, lncRNA, miRNA, and circRNA and their clinical implications in chronic rhinosinusitis with nasal polyps

Hui Li,[1,2] Muhan Shi,[1]* Yuxiao Wu,[1] Shien Huang,[1] Congli Geng,[1] Yan Liu,[1] Xiaopei Yuan,[1] Zhimin Xing,[1] Deyun Wang,[3] Lisheng Yu,[1] Min Wang[1]*

Abstract

Background: Chronic rhinosinusitis with nasal polyps (CRSwNP) is a chronic inflammatory disease with complex pathophysiology and therapeutic strategies. Moreover, the molecular mechanisms underlying the pathogenesis of CRSwNP are incompletely understood.

Objective: This study aimed to investigate the transcriptomic characteristics, ceRNA networks, and whether these molecular markers play a role in the occurrence and development of CRSwNP.

Methods: Following RNA sequencing, a ceRNA network was predicted and constructed based on the sequencing results and multiple databases. Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and disease ontology (DO) were applied to analyze the potential mechanisms in relation to the pathogenesis of CRSwNP. CIBERSORT was used to evaluate the immune cell infiltration levels in CRSwNP. The candidate genes of differentially expressed (DE) mRNA, DE-lncRNA, DE-miRNA, and DE-circRNA were verified by RT-qPCR, and the back-splice junction of circRNA was verified using Sanger sequencing. The clinical significance of differentially expressed genes was analyzed with correlation test and receiver operating characteristic curve.

Results: We identified 716 DE-mRNA, 230 DE-lncRNA, 42 DE-miRNA, and 46 DE-circRNA, and GO and KEGG enrichment analyses indicated that they were involved in multiple biological pathways, predominantly those associated with immunity and inflammation. DO analysis revealed CRSwNP is associated with diseases such as gastroesophageal reflux and allergic reactions. High expression of circ_0021727 was significantly and positively correlated with several important clinical indicators, and the area under the curve was 0.741.

Conclusions: This study provides transcriptomic characteristics, which are potential biomarkers or therapeutic targets for the diagnosis and treatment of CRSwNP.

Key words: Chronic rhinosinusitis with nasal polyps, mRNA, lncRNA, miRNA, circRNA, ceRNA

Citation:
Li, H., Shi, M., Wu, Y., Huang, S., Geng, C., Liu, Y., Yuan, X., Xing, Z., Wang, D., Yu, L., Wang, M. (0000). Gene expression profiles of mRNA, lncRNA, miRNA, and circRNA and their clinical implications in chronic rhinosinusitis with nasal polyps. *Asian Pac J Allergy Immunol*, 00(0), 000-000.
https://doi.org/10.12932/ap-140123-1529

Affiliations:
[1] Department of Otorhinolaryngology, Head and Neck Surgery, Peking University People's Hospital, Beijing, China
[2] Department of Rhinology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China
[3] Department of Otolaryngology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

*These two authors contributed equally to this work

Corresponding author:
1. Min Wang
E-mail: minwang333@sina.com
2. Lisheng Yu
E-mail: yulish68@163.com
3. Deyun Wang
E-mail: entwdy@nus.edu.sg

Introduction

Chronic rhinosinusitis with nasal polyps (CRSwNP) is a chronic inflammatory disease with complex pathogenesis and the molecular mechanisms underlying the pathogenesis of CRSwNP are not clear.

Increasing studies have shown that the human genome can be widely transcribed into noncoding RNAs (ncRNAs), which are closely related to the occurrence and development of diseases.[1] ncRNAs consist of small ncRNAs (< 200 nucleotides), such as miRNA, large ncRNA (> 200 nucleotides), such as lncRNA, and ring-shaped ncRNAs, such as circRNA.[2,3]

miRNA is a highly conservative single-stranded ncRNA with a length of approximately 21–24 nucleotides, which inhibits mRNA translation or promotes mRNA degradation by interacting with the 3′-untranslated region (3′-UTR) of mRNA.[4] It has been estimated that > 60% of the protein-encoding genes are regulated by miRNA, which plays an important role in the occurrence and development of inflammatory diseases and the regulation of the immune system.[5,6] For example, miRNA-126 plays a key role in the pathogenesis of allergic asthma and airway hyperresponsiveness and down-regulation of miRNA-126 can effectively inhibit the function of Th2 cells and the progression of allergic airway disease.[7] lncRNA is a type of ncRNA with a length of > 200 nucleotides, which play an important role in regulating the host immune and inflammatory responses by activating immune-related cells, such as dendritic cells, B lymphocytes, and T lymphocytes.[8] lncRNA RP11-86H7.1, as a pro-inflammatory factor, functions through a competing endogenous RNA (ceRNA) mechanism to promote the expression of NFKB1 and the activation of NF-κB, thereby promoting the inflammatory response of the respiratory tract.[9] circRNA is a closed-loop RNA formed by covalent bonds and plays an important role in inflammatory immune response.[10] Moreover, circ0061052 promotes epithelial–mesenchymal transition of bronchial epithelial cells and participates in airway remodeling of chronic obstructive pulmonary disease (COPD).[11]

With the rapid development of high-throughput sequencing technology, researchers can obtain many potential key genes related to diseases, which provides a comprehensive and in-depth understanding of the molecular mechanism of the disease, as well as clues to discover new diagnostic markers or therapeutic targets.

A number of studies have investigated the gene expression profiles in CRSwNP, unveiled various potential mechanisms underlying the disease[12-19] For example, Peng et al. reported the expression profile of mRNA in CRSwNP and proposed that host immunodeficiency, such as cilia dysfunction and immune dysfunction, as well as inflammatory response and abnormal extracellular matrix metabolism, may be related to the occurrence and development of CRSwNP.[12] In addition, the analysis of lncRNA and mRNA expression profiles has elucidated that inflammation, immune response dysfunction, and the extracellular microenvironment are the key pathogenic mechanisms of CRSwNP.[17] Furthermore, DE-circRNA and DE-miRNA, such as circ_0031593 and miR-145-5p, may play a crucial role in the pathogenesis of CRSwNP, as revealed by analysis of circRNA and miRNA expression profiles.[13] However, comprehensive analysis of mRNA, lncRNA, miRNA, and circRNA expression profiles, research on the regulatory ceRNA network of cirRNA/lncRNA–miRNA–mRNA, and the correlation between circRNA expression and clinical indicators have not been reported so far in CRSwNP.

In this study, the comprehensive expression profiles of mRNA, lncRNA, miRNA, and circRNA of CRSwNP were analyzed on the basis of RNA sequencing. The potential functions of differentially expressed genes (DEGs) and ceRNAs (mRNAs regulated by the ncRNAs-miRNAs) were analyzed using Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Disease Ontology (DO). The immune cell infiltration levels in CRSwNP were evaluated by CIBERSORT and the DEGs were verified using RT-qPCR, and their clinical significance was explored by correlation test and receiver operating characteristic (ROC) curve analysis. Our results clarify transcriptomic characteristics of CRSwNP and provide potential biomarkers or therapeutic targets for its diagnosis and treatment.

Methods

Patients and tissue collection

This study was approved by the ethics committee of Peking University People's Hospital, and informed consent was obtained from all patients. We included 32 subjects with CRSwNP who underwent endoscopic surgery in the hospital and 22 patients with septal deviation (**Supplementary Table 1**). CRSwNP was diagnosed according to the European Position Paper on Rhinosinusitis and Nasal polyps (EPOS) 2020 guidelines.[20] Subjects who had taken corticosteroids or other immunomodulatory drugs in the previous 4 weeks and subjects with antrochoanal polyps, fungal sinusitis, or tumors of the nasal cavity and paranasal sinus were excluded. Nasal polyps and ipsilateral non-polypoid inferior turbinate samples from patients with CRSwNP and inferior turbinate from patients with septal deviation were collected.

Hematoxylin-eosin staining

Hematoxylin-eosin staining was used to determine the eosinophil count. Nasal polyp specimens were fixed in 10% formaldehyde and placed in alcohol at low to high concentrations to remove water. The specimens were then embedded in paraffin wax, sliced using a microtome in sections < 0.5-mm thick, and deparaffinized. After rehydration, the tissue sections were stained with hematoxylin and eosin. The eosinophils in the lamina propria of nasal polyps from five random high-power fields were counted, and the average value was taken as the eosinophil number.

RNA extraction, quality control, library construction, and RNA sequencing

Total RNA from each nasal mucosa was isolated using TRIzol Reagent (Ambion, USA). In brief, RNA degradation and contamination were monitored with 1% agarose gels. RNA purity was measured using a NanoPhotometer spectrophotometer (IMPLEN, USA). The RNA concentration and integrity were determined using the Qubit RNA Assay Kit with the Qubit 2.0 Fluorometer (Life Technologies, USA) and the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 System (Agilent Technologies, USA), respectively. Sequencing libraries of miRNA were generated using NEBNext®Multiplex Small RNA Library Prep Set for Illumina®(NEB, USA.) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. The library preparations were sequenced on an Illumina SE50 platform and 50bp single-end reads were generated. Sequencing libraries of mRNAs, lncRNAs and circRNAs were generated by NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations. The libraries were sequenced on an Illumina PE150 platform and 150 bp pairedend reads were generated. After removing adaptor reads and low-quality tags, clean reads were used for all subsequent analyses. Library construction and sequencing were performed by Novogene Bioinformatic Technology (Beijing, China).

Differential expression analysis

Differential expression analysis was determined by DEseq2.[21] The p-value < 0.05 and $|\log_2$ fold change$|$ > 1 were considered the thresholds for screening differential expressed mRNAs, lncRNAs, miRNAs, and circRNAs (DE-mRNAs, DE-lncRNAs, DE-miRNAs, and DE-circRNAs). Heatmaps were generated using the Complexheatmap R package, and the volcano plots were created using the ggplot2 R package.

Protein–protein interaction (PPI) network

The STRING database (https://cn.string-db.org/) was used to analyze interactions between DEG-encoded proteins and the PPI confidence score was set as 0.4. The PPI network was built using Cytoscape software (version 3.8.2), and the plugin CytoHubba was used to screen the top 20 hub genes in the PPI network, which were ranked by the maximum correlation criterion (MMC).

ceRNA network construction

The miRNA–mRNA regulatory relationships of DE-miRNA were predicted using the Targetscan database (http://www.targetscan.org/vert_72/), MicroT-CDS database (http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=microT_CDS/index), Mirtarbase database (https://maayanlab.cloud/Harmonizome/resource/MiRTarBase), miRDB database (http://mirdb.org/index.html), and multiMiR package.miRNA-lncRNA predictive analysis of the DE-miRNA was performed using the LncSEA database (http://www.licpathway.net/LncSEA/), starBase database (http://starbase.sysu.edu.cn/), and lncBaseV2database (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncbasev2/index-predicted). The miRNA-circRNA regulatory relationship of DE-miRNA was predicted using the Circbank database (http://www.circbank.cn/index.html). The predicted miRNA-–mRNA, miRNA–lncRNA, and miRNA–circRNA regulatory relationships were integrated with DE-mRNA, DE-lncRNA, and DE-circRNA, and the ceRNA networks (circRNA–miRNA–mRNA, and lncRNA–miRNA–mRNA) were constructed by Cytoscape. Based on the genes in the ceRNA networks, CytoHubba was used to screen the top ten hub genes of each molecular type (miRNA, miRNA, lncRNA, and circRNA).

Enrichment analyses

GO functional annotation analysis is a widely used method for large-scale functional enrichment research, including biological process (BP), molecular function (MF), and cellular component (CC). KEGG is a commonly used collection of databases dealing with biological pathways. The R package clusterprofiler was used for the GO and KEGG functional enrichment analyses.[22] Disease Ontology (DO) is an open-source ontology that annotates genes based on human disease. DO enrichment analysis was performed using the R package DOSE.[23] *P*.adjust values < 0.05 were considered to be significantly enriched.

Immune infiltration analysis

CIBERSORT is an algorithm designed to estimate the abundances of immune cells using transcriptomic data.[24] We performed the immune infiltration analysis using the CIBERSORT algorithm to identify differentially enriched immune cells in the nasal polyp tissues and controls.

Real-time quantitative PCR (RT-qPCR) and Sanger sequencing

cDNA (for mRNA, lncRNA, and circRNA) was synthesized using a PrimeScript RT reagent kit (Takara, Japan), and cDNA (for miRNA) was synthesized using an Mir-X miRNA FirstStrand Synthesis (Takara Bio, USA). Primers were synthesized by Sangon Biotech (Shanghai, China) and are listed in **Supplementary Table 2**. GAPDH (for mRNA, lncRNA, and circRNA) and U6 (for miRNA) were used as endogenous controls. The RT-qPCR reactions (for mRNA, lncRNA, and circRNA) were conducted using the Universal SYBR Green Supermix (BIO-RAD, USA), and the RT-qPCR reactions (for miRNA) were conducted using TB Green Advantage Premix (Takara Bio, USA). The reaction conditions were 95°C for 30 s, followed by 40 cycles at 95°C for 5 s and 60°C for 30 s. Sanger sequencing was performed by Sangon Biotech (Shanghai, China) to confirm the back-splicing junctions of circRNAs.

Correlation tests

We also investigated the impact of DEGs (DE-mRNA, DE-lncRNA, DE-miRNA, and DE-circRNA) on laboratory and clinicopathological parameters. The expression of DEGs was correlated with the percentage of peripheral blood eosinophils (PPBE), peripheral blood absolute eosinophil value (PBAEV), tissue eosinophil count (TEC), and cytokines in serum, including interleukin-2 (IL-2), IL-4, IL-17, tumor necrosis factor-α (TNF-α), and interferon-γ (IFN-γ). For these analyses, correlation tests were used, and only patients with CRSwNP were considered.

Statistical analysis

For continuous variables, comparisons between two groups were performed using the independent sample *t*-test or Mann–Whitney U test according to the normality status.

For categorical variables, the chi-square test was used or Fisher's exact test when the chi-squared did not meet the conditions. For the correlation, Pearson's correlation was used when the data met the normality test; otherwise, Spearman's correlation was used. Statistical analyses were performed using SPSS software (version 24.0) and R software (version 3.6.3), and a two-sided p-value < 0.05 was considered statistically significant.

Results

Differential expression analysis

We performed whole-transcriptome RNA sequencing on 15 nasal mucosa samples, including six nasal polyps (as CRSwNP group), with six ipsilateral non-polypoid inferior turbinate (as CRSwIT group) samples from patients with CRSwNP and three inferior turbinate tissues (as control group) from patients with nasal septal deviation. Differentially expressed mRNA (DE-mRNA) and ncRNA (DE-ncRNA), including DE-lncRNA, DE-miRNA, and DE-circRNA, in CRSwNP vs CRSwIT and CRSwNP vs control were analyzed (**Supplementary Table 3**). DE-mRNA and DE-ncRNAs are shown using heatmaps and volcano plots (**Figure 1 and Supplementary Figure 1**). The heatmaps showed that the CRSwNP samples could be significantly separated from the CRSwIT/control samples, indicating that the results of the differential expression analysis were reliable.

To reduce the individual differences caused by genetic or environmental factors, intersection of DEGs in CRSwNP vs CRSwIT and in CRSwNP vs control was taken. A total of 716 DE-mRNA, 230 DE-lncRNA, 42 DE-miRNA, and 46 DE-circRNA were obtained (**Supplementary Figure 2**).

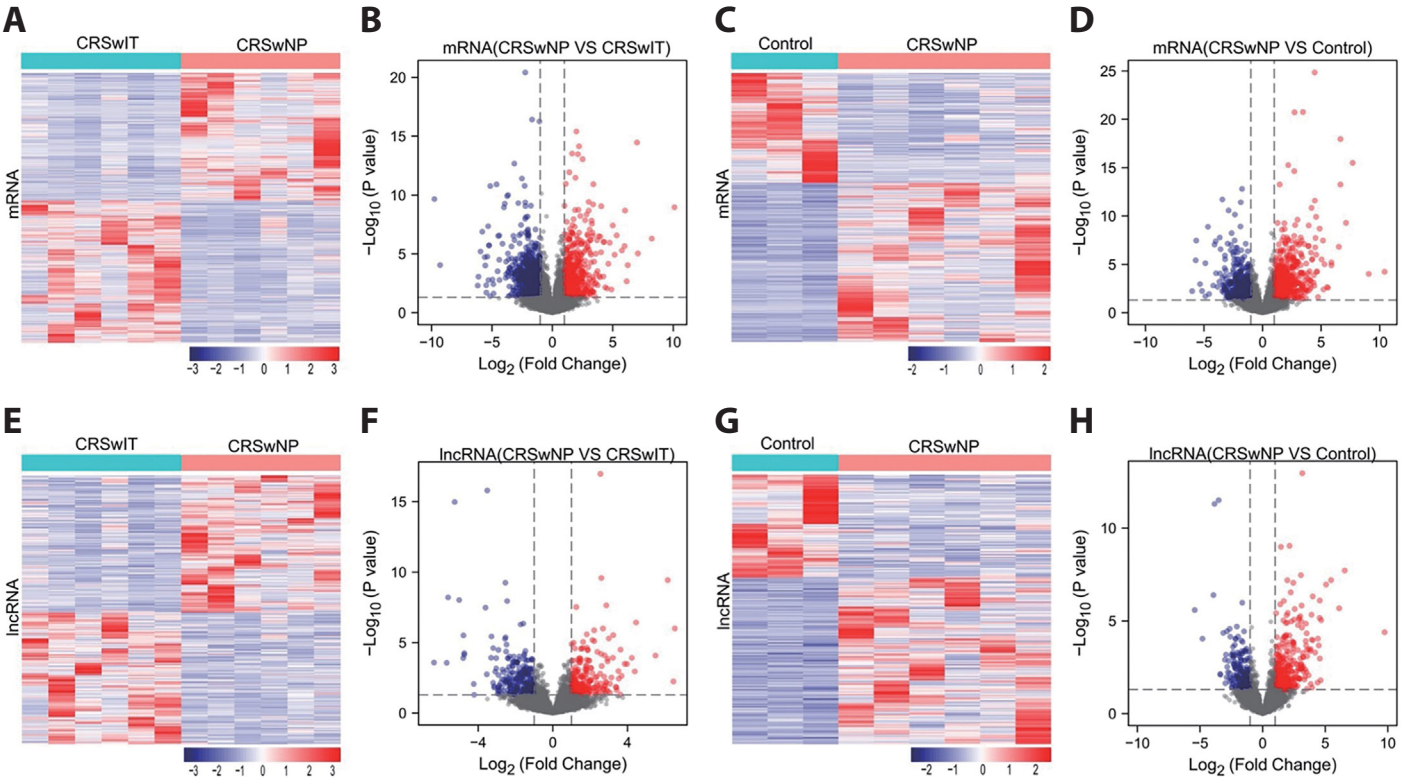


Figure 1. Heatmaps and volcano plots of DEGs. Heatmaps (A, C, E, G, I, K, M, O): red represents highly expressed genes and blue represents lowly expressed genes. Volcano plots (B, D, F, H, J, L, N, P): the x-axis represents the $\log_2$ (fold change), while the y-axis represents the significance degree by –log10 (p value). Red nodes represent upregulated genes and blue nodes indicate downregulated genes; grey nodes represent genes with no significant difference in the expression level.

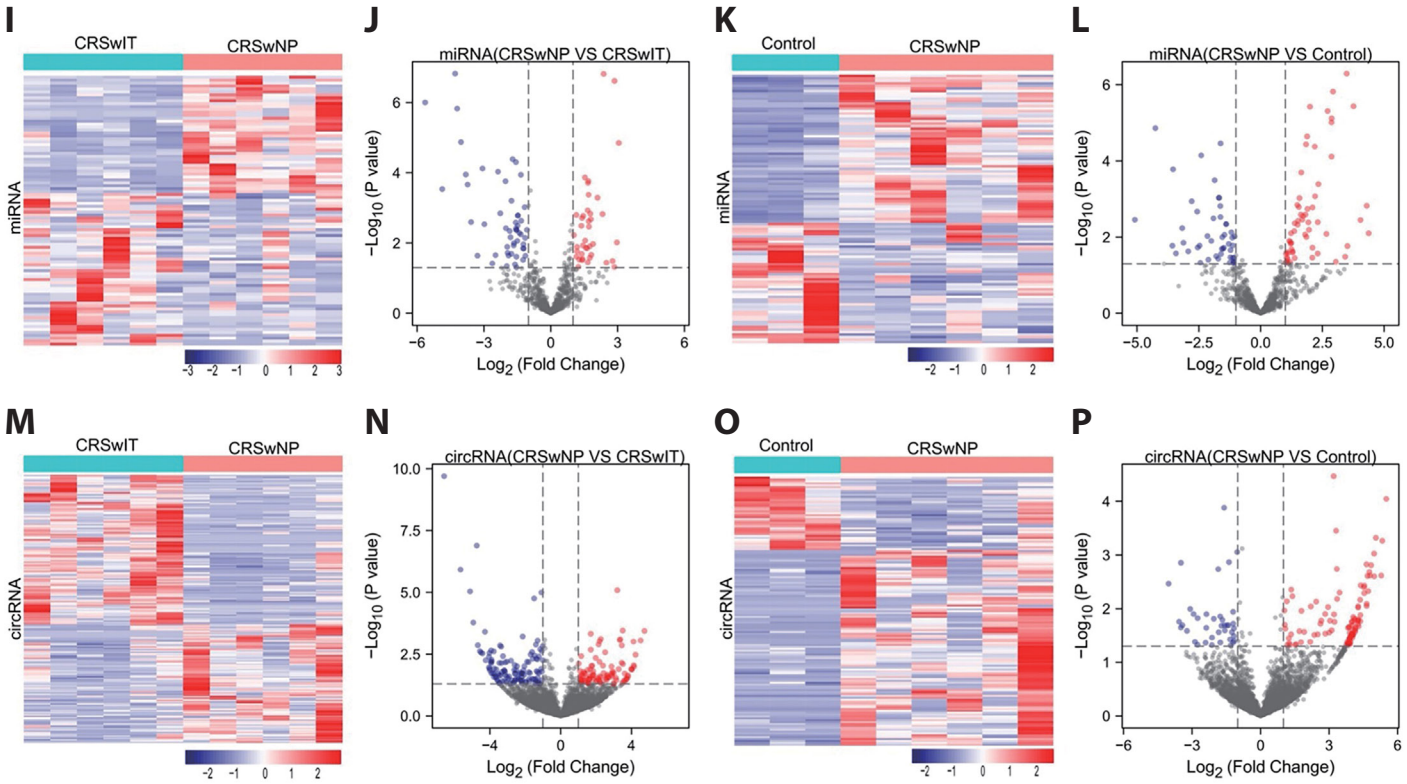


Figure 1. (Continued)

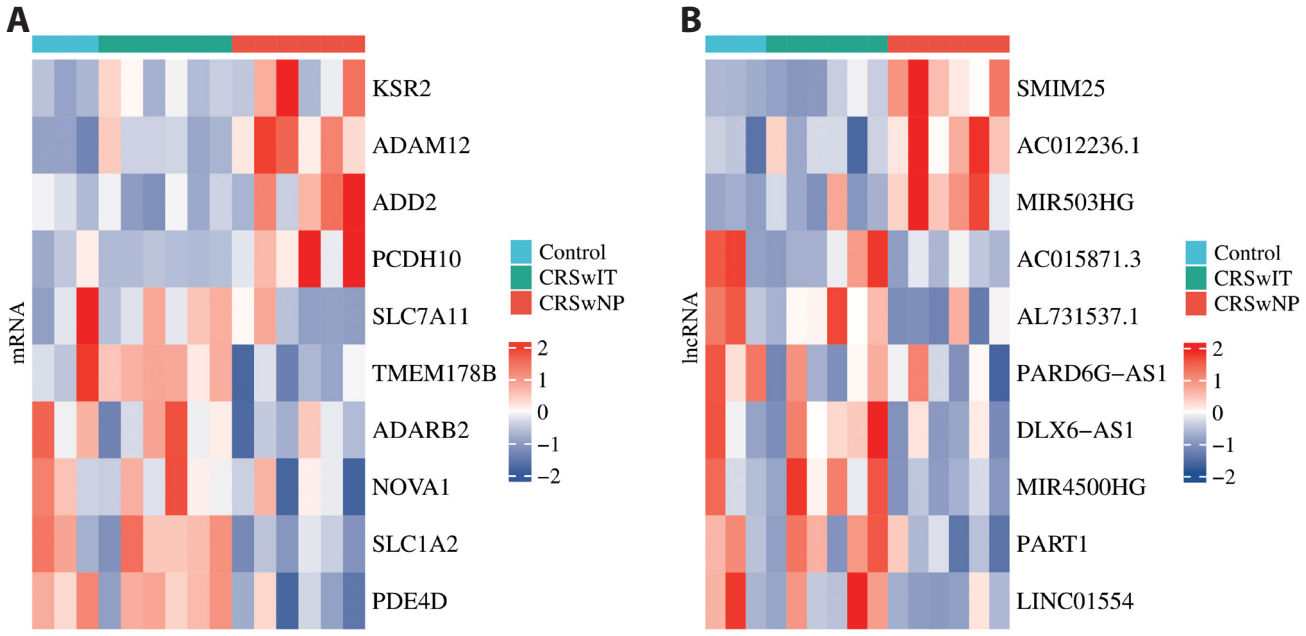


Figure 2. Top ten hub genes in the ceRNA networks. Heatmap of mRNA (A), lncRNA (B), miRNA (C), and circRNA (D). Red indicates upregulated genes and blue represents downregulated genes.

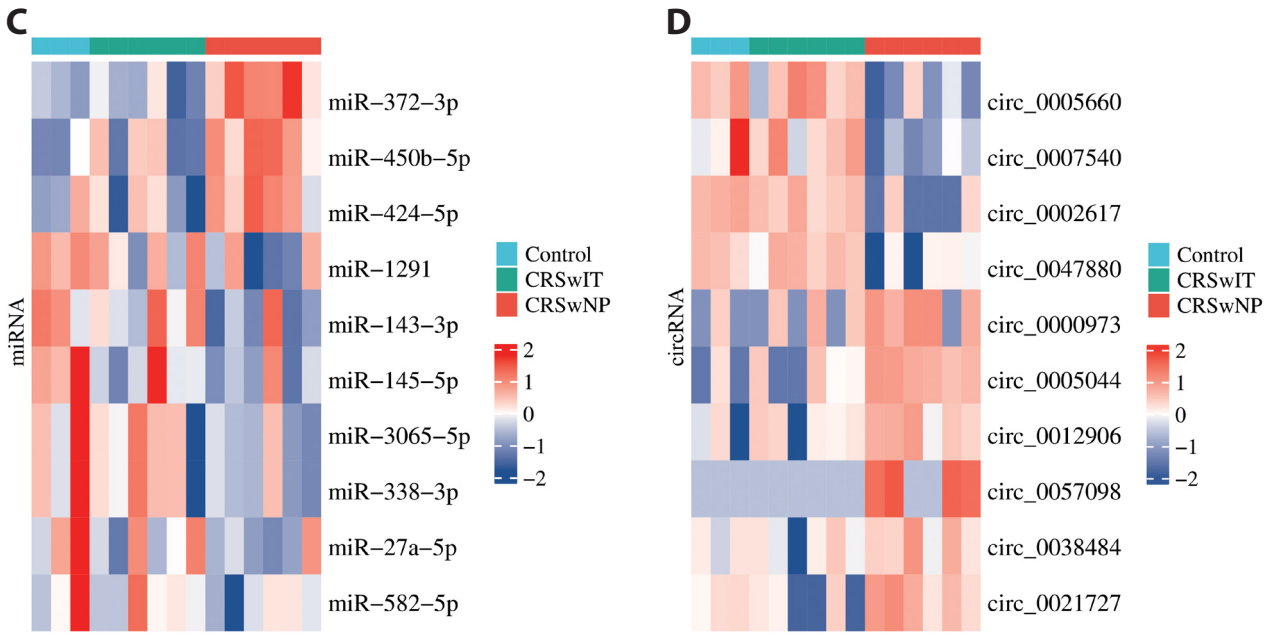


Figure 2. (Continued)

Protein–protein interaction (PPI) network

The PPI network based on DE-mRNA consisted of 2959 interaction pairs (**Supplementary Figure 3A**). Using the Cytoscape plug-in Cytohubba, the top 20 hub genes were screened (**Supplementary Figure 3B**), in which GNG4 interacted with 46 DE-mRNAs, and PMCH interacted with 40 DE-mRNAs.

ceRNA network construction

In the ceRNA hypothesis, the members of ceRNA (circRNAs, lncRNAs, and mRNAs) interact with miRNAs through microRNA response elements (MREs). We constructed lncRNA–miRNA–mRNA (**Supplementary Figure 4A**) and circRNA–miRNA–mRNA (**Supplementary Figure 4B**) networks that integrated the expression profiles and regulatory relationships of DE-circRNA, DE-lncRNA, DE-miRNAs, and DE-mRNAs from our RNA-seq results. The networks consisted of 555 mRNA, 41 miRNA, 26 lncRNA, and 11 circRNA. Using the Cytoscape plug-in Cytohubba, the top 10 hub genes in the networks were identified (**Figure 2**).

Functional enrichment analysis

GO and KEGG pathway analyses were performed on the DE-mRNAs and ceRNAs (mRNAs regulated by the ncRNAs–miRNAs), respectively. The results of GO showed similarities, which were mainly enriched in GO-biological processes (BP) such as neutrophil activation, neutrophil degranulation, eosinophil migration, eosinophil chemotaxis, and IL-5 production (**Figure 3A** and **Supplementary Figure 5A**); in GO-cellular components (CC), such as collagen-containing extracellular matrix, secretory granule membrane, and ion channel complex (**Figure 3B** and **Supplementary Figure 5B**); and in GO-molecular functions (MF), such as receptor ligand activities, cytokine activity, and channel activity (**Figure 3C** and **Supplementary Figure 5C**). The results of KEGG also displayed similarities, which were mainly enriched in cytokine–cytokine receptor interaction, chemokine signaling pathway, and viral protein interaction with cytokine and cytokine receptors (**Figure 3D** and **Supplementary Figure 5D**). In brief, the circRNA- or lncRNA-associated ceRNA network may participate in the pathological and physiological processes of CRSwNP from different aspects.

The results of DO enrichment analysis (**Supplementary Figure 6**) mainly covered diseases such as gastroesophageal reflux disease, allergic reaction, recurrent bacterial, fungus, and viral infections.

Analysis of immune infiltration

The fraction of infiltrating immune cells between the CRSwNP and control groups was evaluated by CIBERSORT. There were differences in 18 out of 22 immune cell types, with significant differences observed in Plasma cells, CD8 T cells, M2 macrophages, and resting mast cells (**Supplementary Figure 7**).

Confirmation of RT-qPCR and Sanger sequencing

RT-qPCR was performed to confirm the RNA-seq data. Twelve differentially expressed transcripts were selected: three mRNA, three lncRNA, three miRNA, and three circRNA (**Figure 4**), including a potential ceRNA (circ_0021727-miR145-5p-PCDH10). Most selected transcripts exhibited significant differential expression and similar expression between RNA-seq and RT-qPCR. Only miR-450-5p was not differentially expressed, and circ_0038484 was expressed in opposing directions in RNA-seq and RT-qPCR (high expression in RNA-seq and low expression in RT-qPCR). We also demonstrated the back-splice junction of circRNAs via Sanger sequencing (**Supplementary Figure 8**).

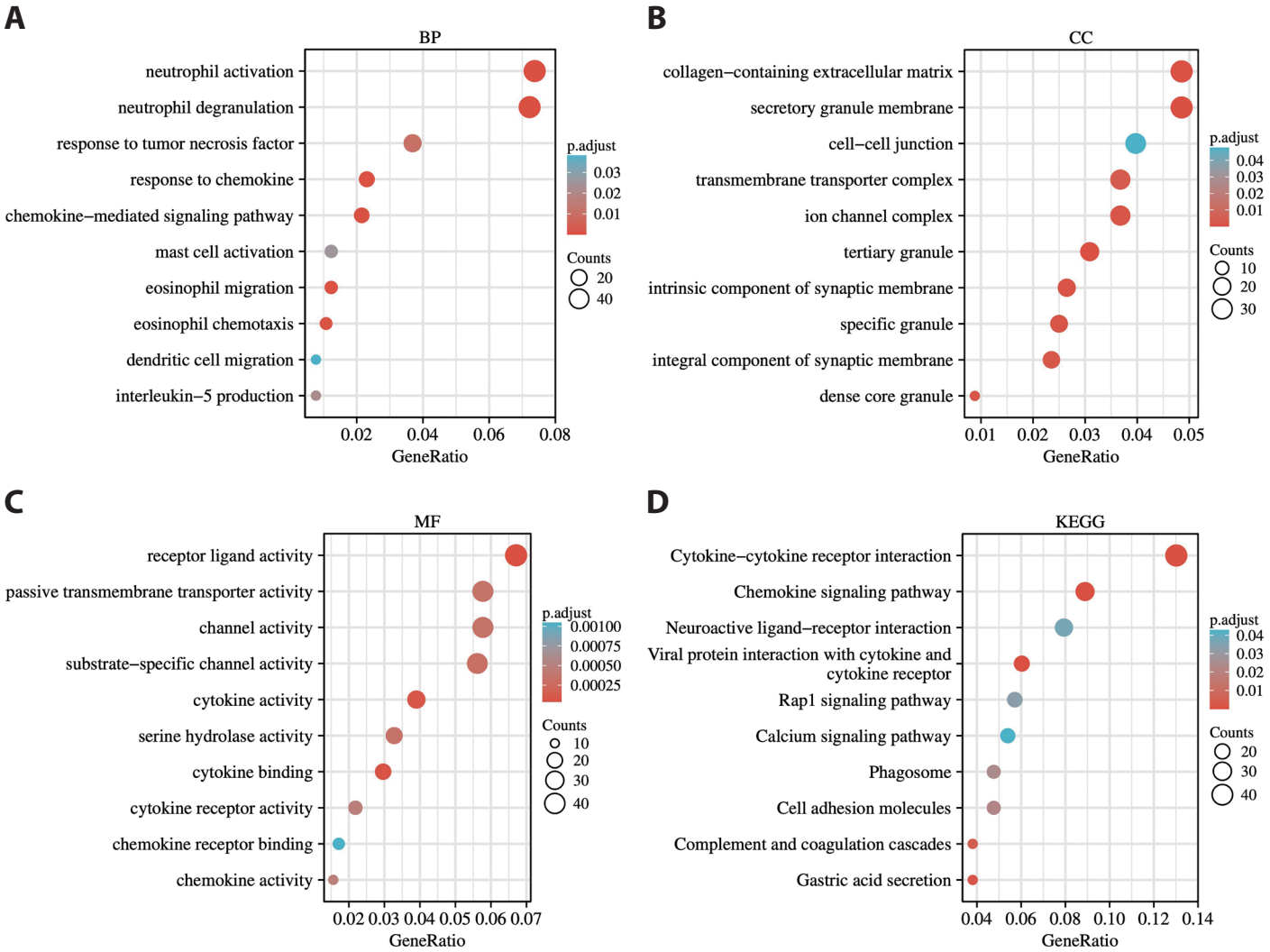


Figure 3. GO and KEGG pathway analyses of DE-mRNA. A. BP enrichment analysis. B. CC enrichment analysis. C. MF enrichment analysis. D. KEGG enrichment analysis. The abscissa indicates the gene ratio, and the ordinate represents the enriched term or pathway. The size and color of the nodes indicate the number of genes in the corresponding term or pathway and the *p*-value, respectively.

BP: Biological process, CC: Cellular component, MF: Molecular function.

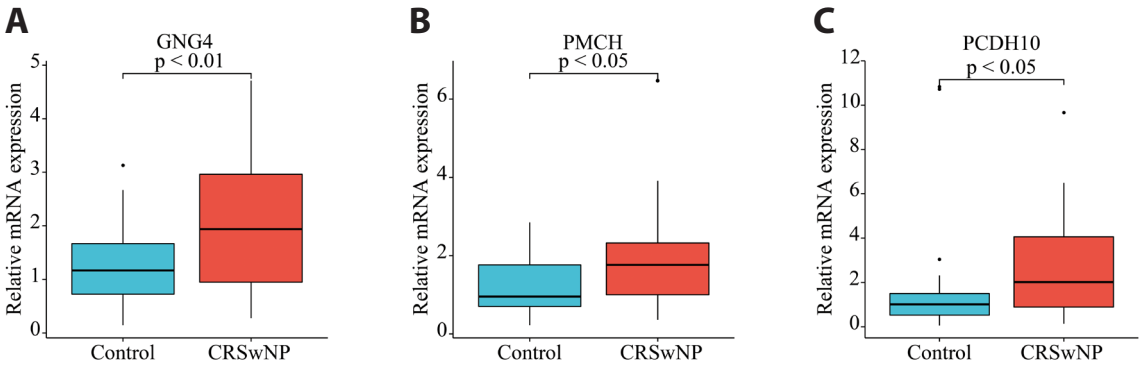


Figure 4. Validation of transcript expression by RT-qPCR. mRNAs (A, B, and C), lncRNAs (D, E, and F), and circRNAs (J, K, L) were validated in 32 CRSwNP subjects and 22 control subjects, and miRNAs (G, H, I) in ten CRSwNP subjects and ten control subjects.

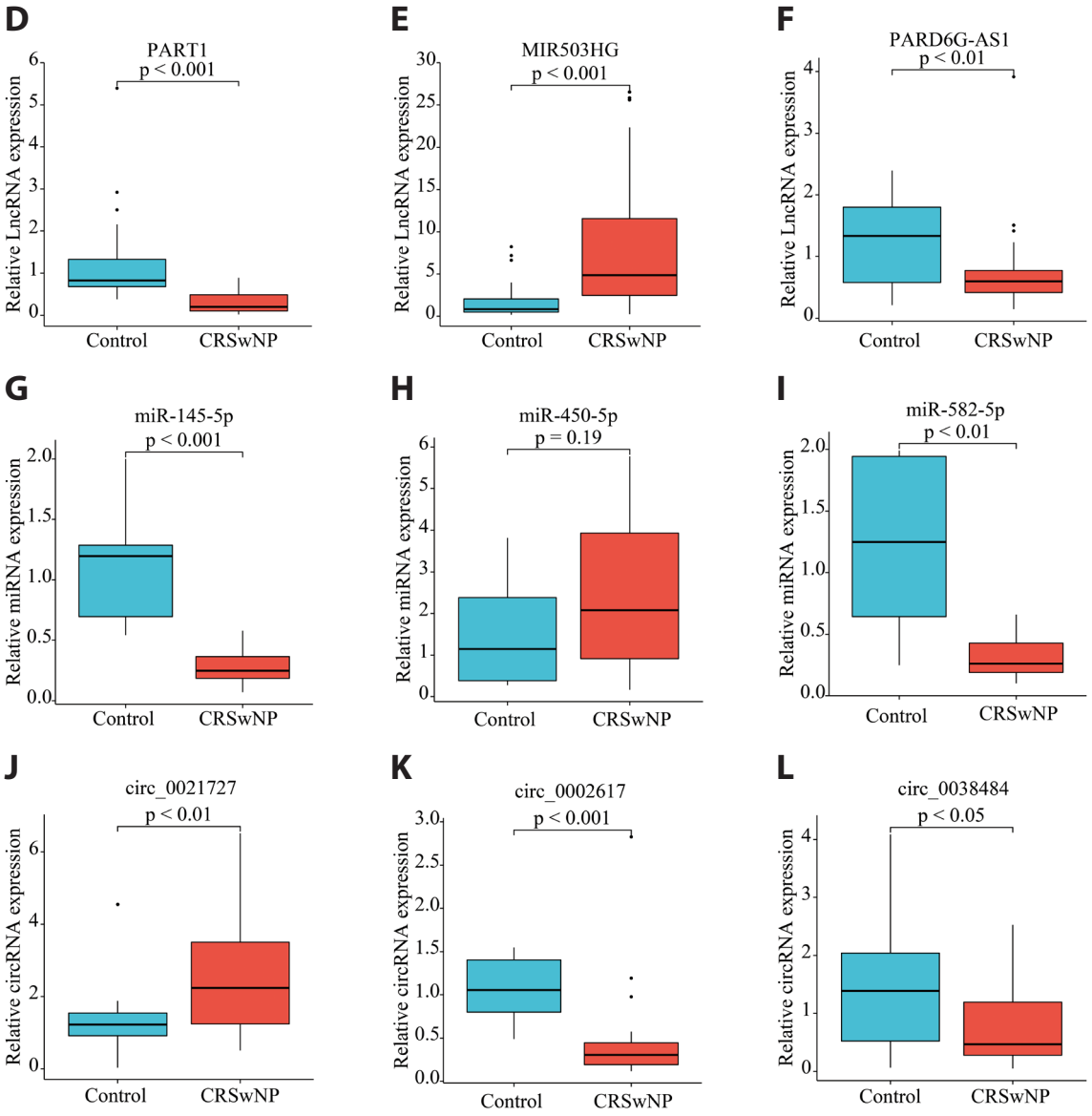


Figure 4. (Continued)

Correlation between selected DEGs and clinical indicators

We further explored the correlation between the above-described verified DEGs and clinical indicators (**Supplementary Table 4**). We found that the expression of circ_0021727 had a significant positive correlation with PPBE, PBAEV, TNF-α, INF-γ, and TEC (**Figure 5**), while circ_0002617 had a significant negative association with PPEB (r = −0.365, $p < 0.05$) and TEC (r = −0.379, $p < 0.05$) (**Supplementary Figure 9**). In addition, there was a significant negative correlation between MIR503HG and TNF-α (r = −0.396, $p < 0.05$) and between GNG4 and IFN-γ (r = −0.442, $p < 0.05$) (**Supplementary Figure 10**). Other correlations were not significant.

ROC curve analysis of selected DEGs

ROC curve analysis was performed to evaluate the functional values of the above-described verified DEGs in the occurrence and development of CRSwNP. As shown in **Figure 6**, the largest area under the curve (AUC) for mRNAs, lncRNAs, miRNAs, and circRNAs was GNG4, PART1, miR-145-5p, and circ_0002617, with values of 0.673, 0.926, 0.980, and 0.926, respectively. More detailed information on the ROC curve can be seen in **Supplementary Table 5**.

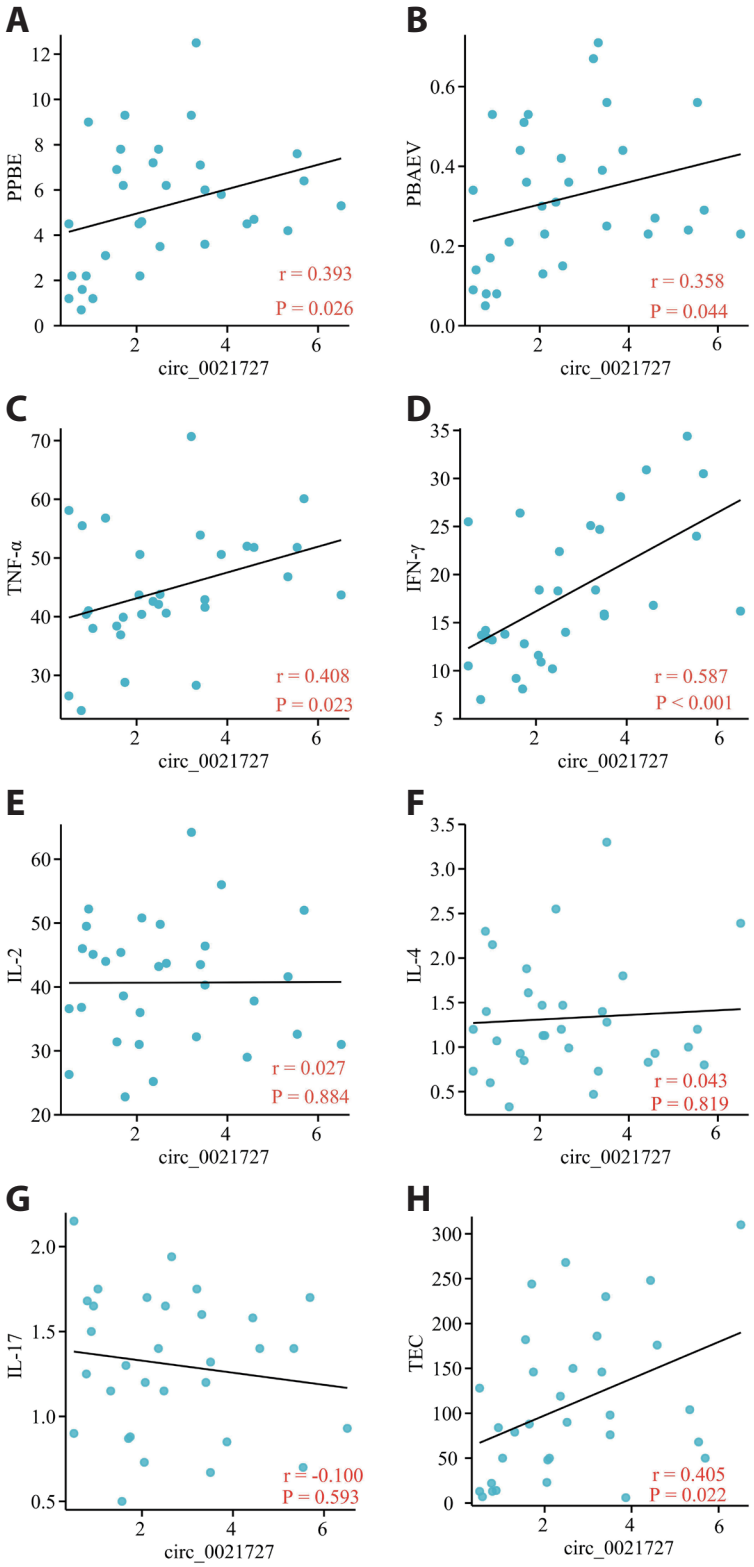


Figure 5. Correlation between circ_0021727 expression and clinical indicators in a sample from CRSwNP (n = 32). The correlation between circ_0021727 and PPBE (A), PBAEV (B), TNF-α (C), IFN-γ (D), IL-2 (E), IL-4 (F), IL-17 (G), and TEC (H).

PPBE: Percentage of peripheral blood eosinophils, PBAEV: Peripheral blood absolute eosinophil value, TEC: Tissue eosinophil count.

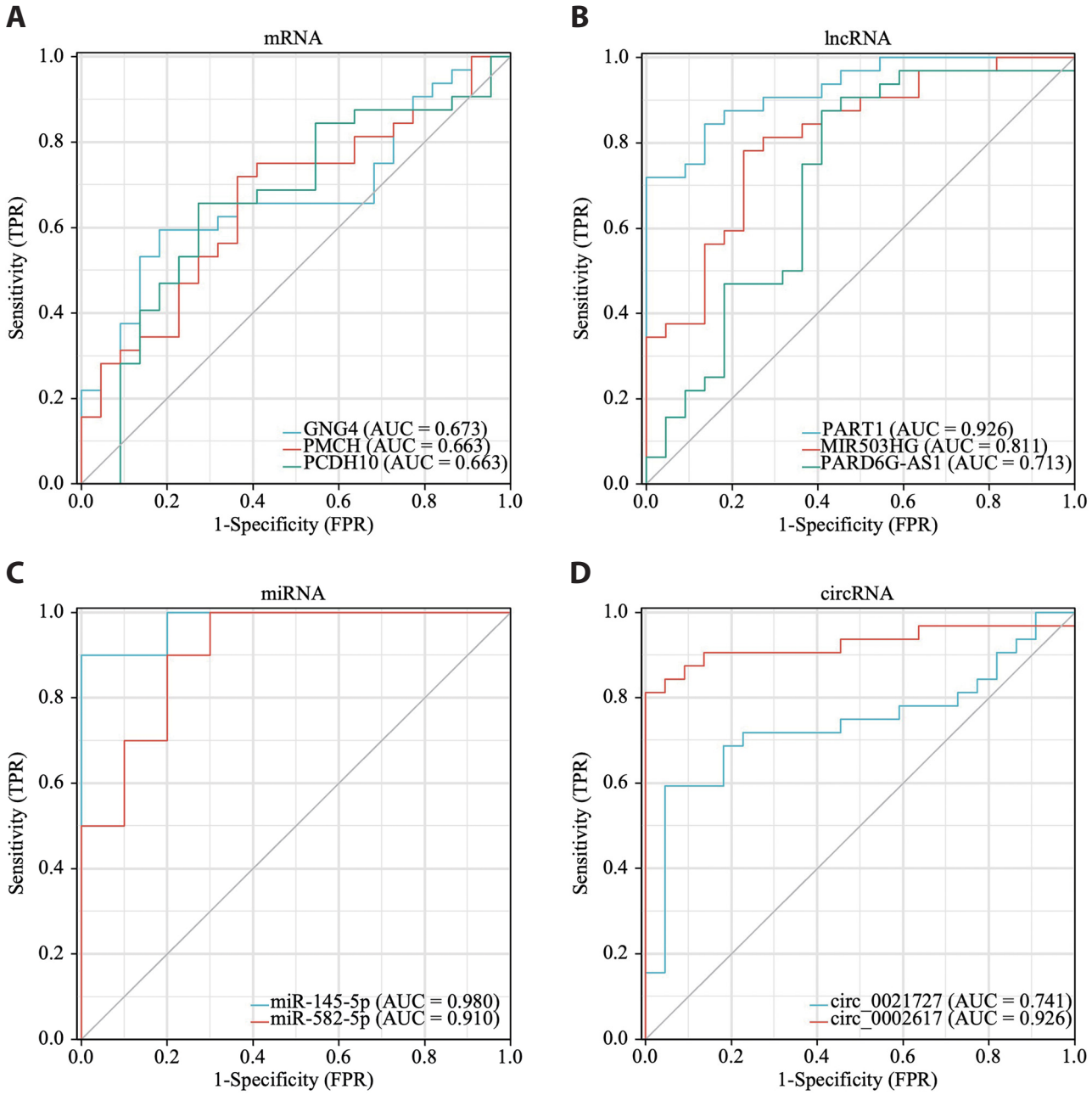


Figure 6. Functional value of the selected DEGs. ROC curve analysis of mRNAs (A), lncRNAs (B), miRNAs (C), and circRNAs (D). The abscissa is 1-specificity, indicating the false positive rate (FPR), and the ordinate is sensitivity, representing the true-positive rate (TPR).

AUC: Area under the curve.

Discussion

Recently, in addition to mRNA, increasing studies have shown that ncRNA, such as lncRNA, miRNA, and circRNA, is widely involved in various diseases. A new regulatory mechanism between mRNA and ncRNA, called ceRNA, was proposed by Salmena et al.[25] The ceRNA regulatory network is complex and intertwined, in which members, including circRNA, lncRNA, mRNA, and pseudogenes, compete for the same miRNA response elements to regulate each other.

In this study, we comprehensively analyzed the expression profiles of mRNA, lncRNA, miRNA, and circRNA in CRSwNP and constructed a ceRNA network. In total, 716 DE-mRNA, 230 DE-lncRNA, 42 DE-miRNA, and 46 DE-circRNA were identified. GO and KEGG were performed to explore the potential functions of DE-mRNA and ceRNA (mRNA regulated by circRNA/lncRNA–miRNA). The results showed that these genes were involved in various biological processes and signaling pathways, demonstrating the complexity of CRSwNP pathogenesis. The similarity of results from enrichment analysis of DEGs and ceRNA also suggested that complex ceRNA was involved in the pathogenesis of CRSwNP, which helps us to understand the interaction between encoding and non-encoding RNA

in the pathogenesis of CRSwNP. We also performed DO analysis and found that these genes could be enriched in allergic reaction, gastroesophageal reflux, ciliary movement disorder, and repeated bacterial, viral, and fungal infections, which were closely related to the pathogenesis of CRSwNP. The analysis of immune cells infiltration revealed that the proportion of eosinophils and neutrophils in CRSwNP was higher than in the control group, although not significantly different. This may be attributed to the relatively small sample size used for sequencing.

In addition, we verified the DEGs using RT-qPCR and explored the clinical significance of the verified DEGs. Sequencing and RT-qPCR results indicated that PMCH was highly expressed in CRSwNP. Sandig et al. considered that Th2 cells, the core of allergic disease pathogenesis, can highly selectively express PMCH.[26] Therefore, the expression of PMCH by activated Th2 cells *in vivo* may directly connect allergic inflammation with energy homeostasis and may lead to asthma. Whether PMCH participates in the pathogenesis of CRSwNP has not yet been reported, which needs to be explored in subsequent experiments. GNG4 encodes genes involved in G-protein coupled receptor signaling and mediates the production of anti-inflammatory cytokines. It has been identified as a critical gene in rhinosinusitis and pediatric asthma through bioinformatic analysis.[27,28] However, functional studies of GNG4 in rhinosinusitis have not been reported.

According to the research conducted by Yu et al,[13] the expression of miR-145-5p is significantly lower in CRSwNP, with an AUC of 0.8690, which is consistent with our study (miR-145-5p is lowly expressed; AUC = 0.980). It has been suggested that the expression of miR-145-5p is low in chronic obstructive pulmonary disease, and that overexpression of miR-145 can significantly reduce the apoptosis and inflammatory response of bronchial epithelial cells.[29] circ_0021727, which was highly expressed in CRSwNP in our study, has a significantly positive correlation with PPBE, PEAEV, TNF-α, IFN-γ, and TEC. In addition, the AUC of circ_0021727 was > 0.7. These results indicate that circ_0021727 can serve as a marker for the diagnosis and prognosis of CRSwNP and may also participate in the pathogenesis of CRSwNP. The host gene of circ_0021727 is CD44, which can encode a cell-surface glycoprotein involved in cell-cell interaction, cell adhesion, and migration, and participates in various cell functions, including T-lymphocyte activation, inflammatory response, and response to bacterial infection.[30] Studies have shown that anti-CD44 antibodies can reduce inflammation and hyper-reactivity of the airway. CD44 may be a target for treating Th2-mediated airway inflammation, including allergic asthma.[31,32] These studies have demonstrated the important role of CD44 in airway inflammation. However, the coverage of circCD44 is rare, and it has been reported in breast cancer and glioblastoma in the past 2 years.[33,34] Whether circ0021727 (circCD44) could play an important role in the pathogenesis of CRSwNP remains to be investigated by subsequent cell and animal experiments.

This study has several limitations. First, the small sample size of sequencing may produce spurious results, but most of the screened hub genes in this study were validated by RT-qPCR. Second, there is no subtype classification of CRSwNP. Studies conducted by Bu et al, Okada et al, and Wang et al. suggested that the gene expression profiles of eosinophilic CRSwNP and non-eosinophilic CRSwNP were inconsistent,[16-18] while data from Yu et al. and Peng et al. indicated that the gene expression profiles of eosinophilic and non-eosinophilic CRSwNP were similar.[12,13] These controversial issues may be related to several factors, such as the type of gene expression profile, site of sampling, sample size, definition of eosinophilic and non-eosinophilic CRSwNP, and the definition of DEGs. Third, the validation of ceRNA is insufficient. Luciferase experiment, RNA immunoprecipitation, and RNA pull-down assay are needed to verify the mutual binding relationship and mutual binding sites among the molecules in the ceRNA networks. Fourth, there is a lack of prognostic analysis regarding circ_0021727 as a potential diagnostic or therapeutic target. Therefore, in our future functional studies on circ_0021727, we plan to conduct a long-term follow-up study and perform prognostic analysis in patients with CRSwNP. Last but not the least, further studies are required to explore the functional properties of DEGs and ceRNA *in vivo* and *in vitro*.

Conclusion

In this study, we clarified the mRNA, lncRNA, miRNA, and circRNA expression profiles of CRSwNP, with the ceRNA regulatory networks of cirRNA–miRNA–mRNA and lncRNA–miRNA–mRNA. We also investigated the functional characteristics of the DEGs and ceRNA networks and explored the molecular mechanisms underlying the immune cell infiltration in CRSwNP. Our results will help us to gain an in-depth understanding of the pathogenesis of CRSwNP, leading the way for subsequent functional studies and providing potential biomarkers and therapeutic targets for the diagnosis and treatment of CRSwNP.

Acknowledgments

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

Author contributions

- Hui Li: conception and design, data collection, analysis and interpretation, manuscript writing, final approval.
- Muhan Shi: conception and design, data collection, analysis and interpretation, review, final approval.
- Yuxiao Wu, Shien Huang, Congli Geng, Yan Liu, Xiaopei Yuan and Zhimin Xing: conception and design, review, final approval.
- Deyun Wang, Lisheng Yu and Min Wang: conception and design, supervision, review, final approval.

Funding

Beijing Health Technologies Promotion Prgogram (BHTPP2022007), Research and Development Foundation of Peking University People's Hospital (RDL2021-05), and Capital Health Research and Development of Special Fund (No: 2020-1-2051) supported this study.

Conflict of interest

No conflicts of interest to be declared.

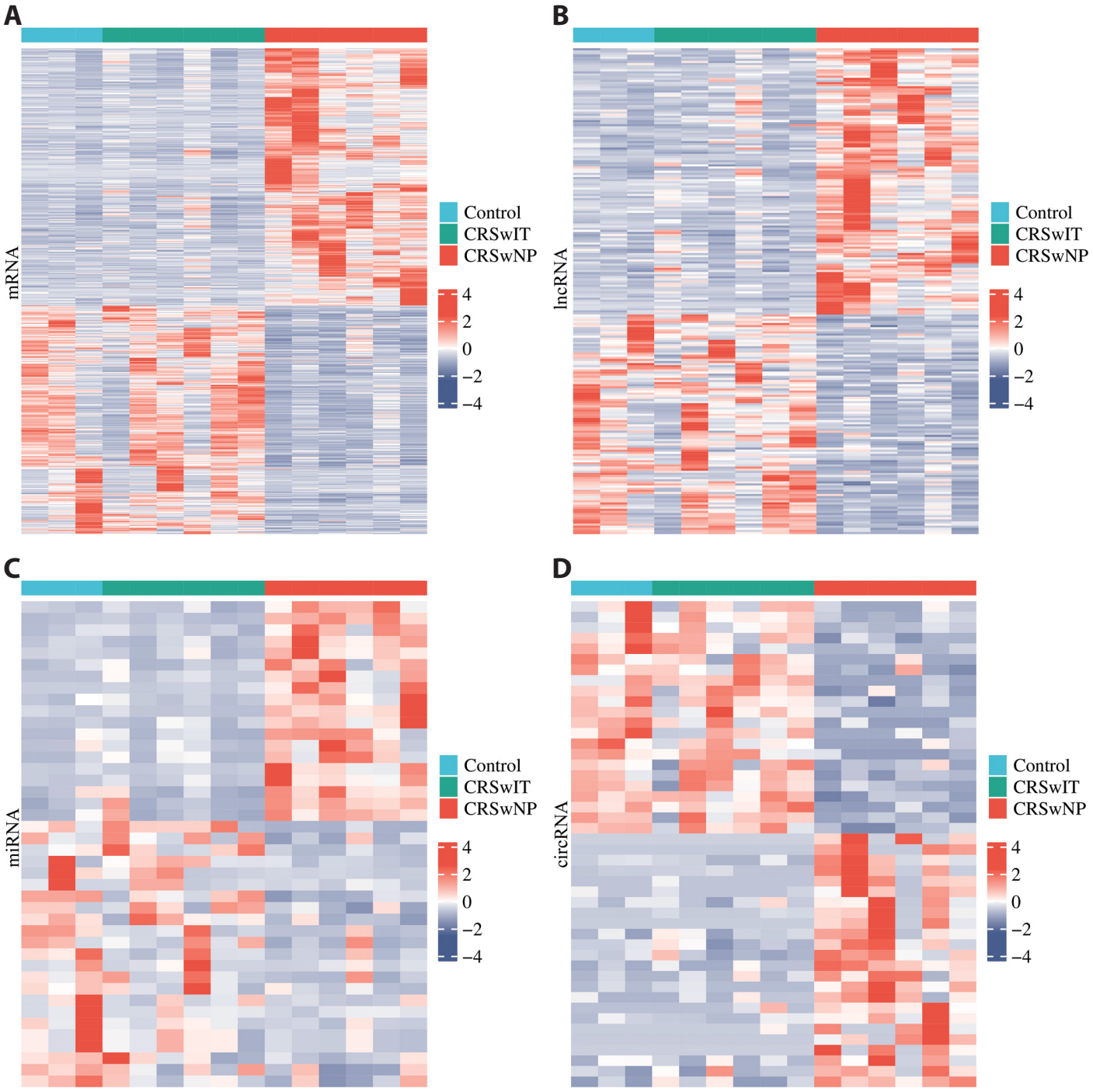
Data availability statement

The data are available on reasonable request.

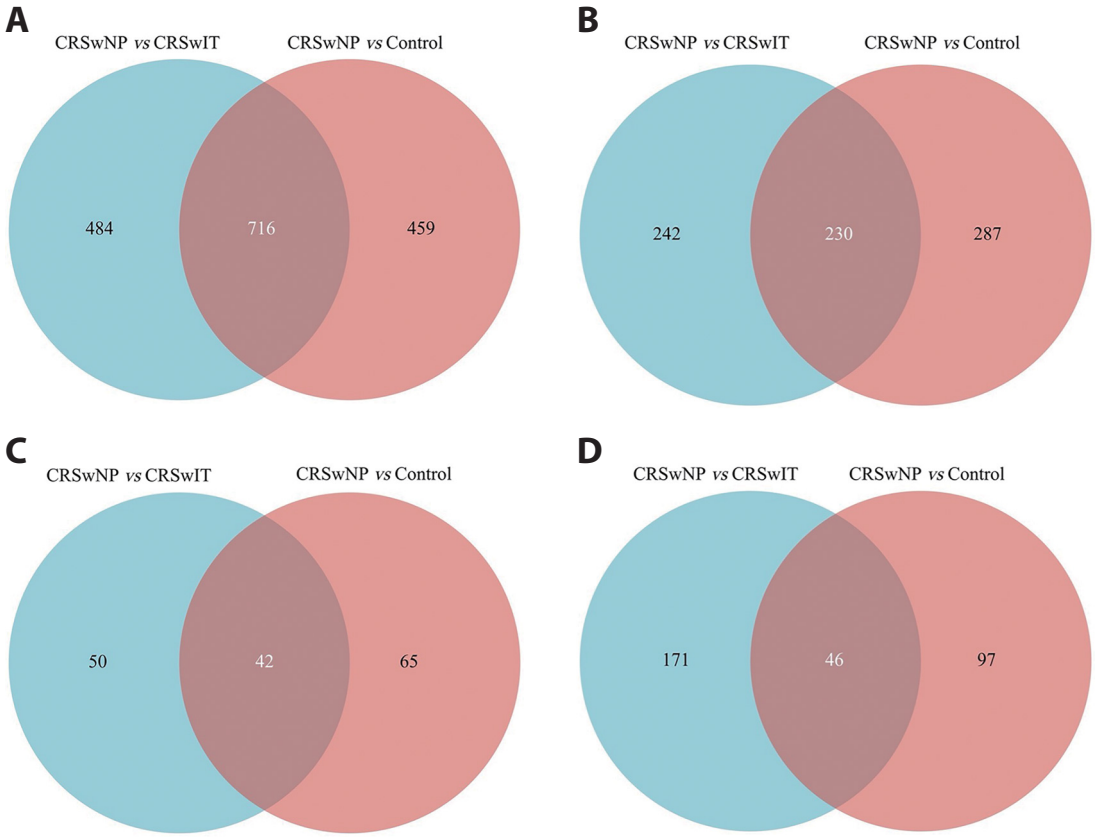
References

1. Beermann J, Piccoli MT, Viereck J, Thum T. Non-coding RNAs in Development and Disease: Background, Mechanisms, and Therapeutic Approaches. Physiol Rev. 2016;96(4):1297-325.
2. Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. Nat Biotechnol. 2014;32(5):453-61.
3. Hombach S, Kretz M. Non-coding RNAs: Classification, Biology and Functioning. Adv Exp Med Biol. 2016;937:3-17.
4. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004;116(2):281-97.
5. Raisch J, Darfeuille-Michaud A, Nguyen HT. Role of microRNAs in the immune system, inflammation and cancer. World J Gastroenterol. 2013;19(20):2985-96.
6. Lu Q, Wu R, Zhao M, Garcia-Gomez A, Ballestar E. miRNAs as Therapeutic Targets in Inflammatory Disease. Trends Pharmacol Sci. 2019;40(11):853-65.
7. Collison A, Herbert C, Siegle JS, Mattes J, Foster PS, Kumar RK. Altered expression of microRNA in the airway wall in chronic asthma: miR-126 as a potential therapeutic target. BMC Pulm Med. 2011;11:29.
8. Aune TM, Spurlock CF 3rd. Long non-coding RNAs in innate and adaptive immunity. Virus Res. 2016;212:146-60.
9. Zhao J, Pu J, Hao B, Huang L, Chen J, Hong W, et al. LncRNA RP11-86H7.1 promotes airway inflammation induced by TRAPM2.5 by acting as a ceRNA of miRNA-9-5p to regulate NFKB1 in HBECS. Sci Rep. 2020;10(1):11587.
10. Li X, Yang L, Chen LL. The Biogenesis, Functions, and Challenges of Circular RNAs. Mol Cell. 2018;71(3):428-42.
11. Ma H, Lu L, Xia H, Xiang Q, Sun J, Xue J, et al. Circ0061052 regulation of FoxC1/Snail pathway via miR-515-5p is involved in the epithelial-mesenchymal transition of epithelial cells during cigarette smoke-induced airway remodeling. Sci Total Environ. 2020;746:141181.
12. Peng Y, Zi XX, Tian TF, Lee B, Lum J, Tang SA, et al. Whole-transcriptome sequencing reveals heightened inflammation and defective host defence responses in chronic rhinosinusitis with nasal polyps. Eur Respir J. 2019;54(5).
13. Yu J, Kang X, Xiong Y, Luo Q, Dai D, Ye J. Gene Expression Profiles of Circular RNAs and MicroRNAs in Chronic Rhinosinusitis With Nasal Polyps. Front Mol Biosci. 2021;8:643504.
14. Silveira M, Tamashiro E, Santos A, Martins RB, Faria FM, Silva L, et al. miRNA-205-5p can be related to T2-polarity in Chronic Rhinosinusitis with Nasal Polyps. Rhinology. 2021;59(6):567-76.
15. Callejas-Díaz B, Fernandez G, Fuentes M, Martínez-Antón A, Alobid I, Roca-Ferrer J, et al. Integrated mRNA and microRNA transcriptome profiling during differentiation of human nasal polyp epithelium reveals an altered ciliogenesis. Allergy. 2020;75(10):2548-61.
16. Bu X, Wang M, Luan G, Wang Y, Wang C, Zhang L. Integrated miRNA and mRNA expression profiling reveals dysregulated miRNA-mRNA regulatory networks in eosinophilic and non-eosinophilic chronic rhinosinusitis with nasal polyps. Int Forum Allergy Rhinol. 2021;11(8):1207-19.
17. Wang W, Gao Z, Wang H, Li T, He W, Lv W, et al. Transcriptome Analysis Reveals Distinct Gene Expression Profiles in Eosinophilic and Noneosinophilic Chronic Rhinosinusitis with Nasal Polyps. Sci Rep. 2016;6:26604.
18. Okada N, Nakayama T, Asaka D, Inoue N, Tsurumoto T, Takaishi S, et al. Distinct gene expression profiles and regulation networks of nasal polyps in eosinophilic and non-eosinophilic chronic rhinosinusitis. Int Forum Allergy Rhinol. 2018;8(5):592-604.
19. Bassiouni A, Ou J, Schreiber A, Geoghegan J, Tsykin A, Psaltis AJ, et al. The global transcriptomic signature in sinonasal tissues reveals roles for tissue type and chronic rhinosinusitis disease phenotype. Rhinology. 2020;58(3):273-83.
20. Fokkens WJ, Lund VJ, Hopkins C, Hellings PW, Kern R, Reitsma S, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. Rhinology. 2020;58(Suppl S29):1-464.
21. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
22. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-7.
23. Yu G, Wang LG, Yan GR, He QY. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. Bioinformatics. 2015;31(4):608-9.
24. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773-82.
25. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language. Cell. 2011;146(3):353-8.
26. Sandig H, McDonald J, Gilmour J, Arno M, Lee TH, Cousins DJ. Human Th2 cells selectively express the orexigenic peptide, pro-melanin-concentrating hormone. Proc Natl Acad Sci U S A. 2007;104(30):12440-4.
27. Hao M, Zan J. The Identification of Childhood Asthma Progression-Related lncRNAs and mRNAs Suitable as Biomarkers Using Weighted Gene Coexpression Network Analysis. Genet Res (Camb). 2021;2021:5511507.
28. Pan Y, Wu L, He S, Wu J, Wang T, Zang H. Identification of hub genes and immune cell infiltration characteristics in chronic rhinosinusitis with nasal polyps: Bioinformatics analysis and experimental validation. Front Mol Biosci. 2022;9:843580.
29. Dang X, Yang L, Guo J, Hu H, Li F, Liu Y, et al. miR-145-5p is associated with smoke-related chronic obstructive pulmonary disease via targeting KLF5. Chem Biol Interact. 2019;300:82-90.
30. Yoshida T, Matsuda Y, Naito Z, Ishiwata T. CD44 in human glioma correlates with histopathological grade and cell migration. Pathol Int. 2012;62(7):463-70.
31. Katoh S, Kaminuma O, Hiroi T, Mori A, Ohtomo T, Maeda S, et al. CD44 is critical for airway accumulation of antigen-specific Th2, but not Th1, cells induced by antigen challenge in mice. Eur J Immunol. 2011;41(11):3198-207.
32. Katoh S. Critical Involvement of CD44 in T Helper Type 2 Cell-Mediated Eosinophilic Airway Inflammation in a Mouse Model of Acute Asthma. Front Immunol. 2021;12:811600.
33. Feng J, Ren X, Fu H, Li D, Chen X, Zu X, et al. LRRC4 mediates the formation of circular RNA CD44 to inhibitGBM cell proliferation. Mol Ther Nucleic Acids. 2021;26:473-87.
34. Li J, Gao X, Zhang Z, Lai Y, Lin X, Lin B, et al. CircCD44 plays oncogenic roles in triple-negative breast cancer by modulating the miR-502-5p/KRAS and IGF2BP2/Myc axes. Mol Cancer. 2021;20(1):138.

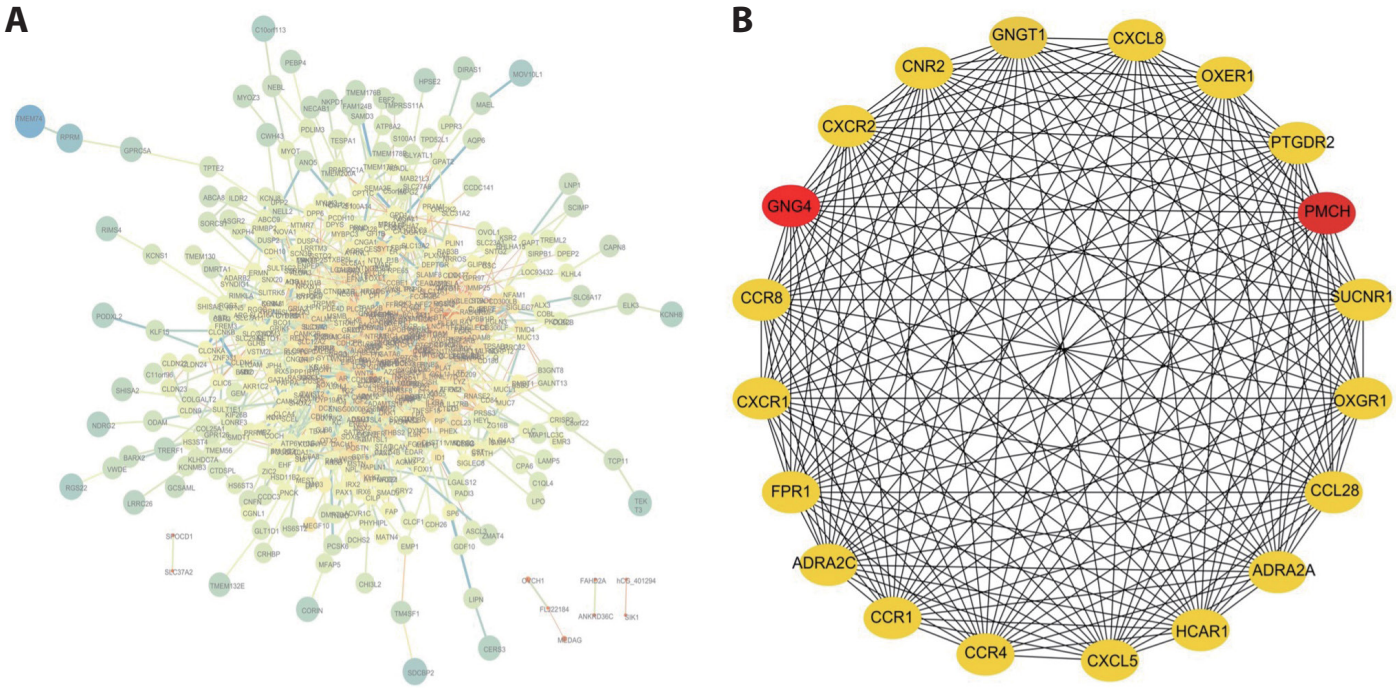
Supplemental material



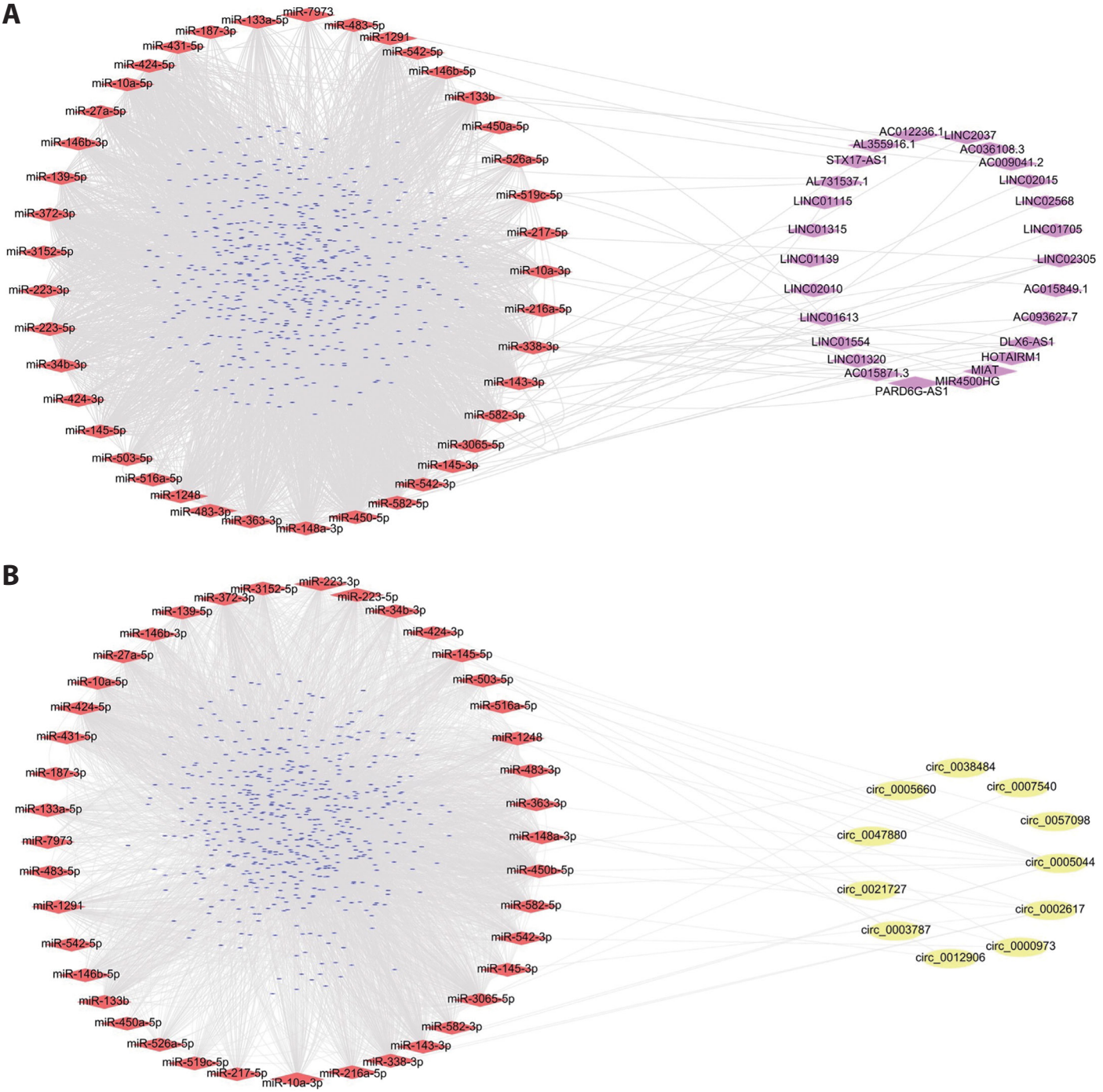
Supplementary Figure 1. Heatmap of mRNA (A), lncRNA (B), miRNA (C), and circRNA (D). Red represents highly expressed genes and blue represents lowly expressed genes.



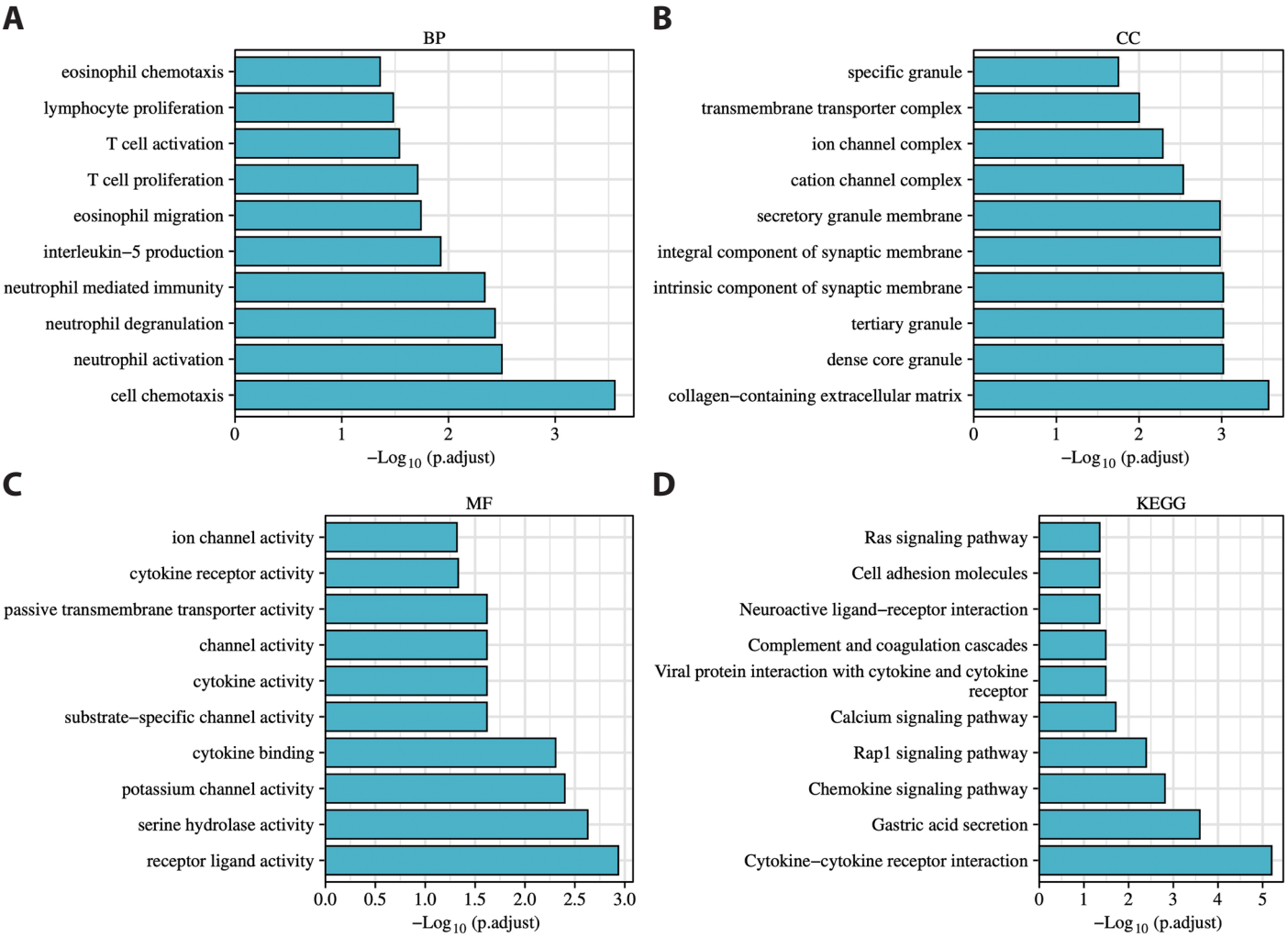
Supplementary Figure 2. Venn diagram showing the intersection of DEGs in CRSwNP vs CRSwIT and in CRSwNP vs control. A. mRNA, B. lncRNA, C. miRNA, and D. circRNA.



Supplementary Figure 3. A. Protein–protein interaction (PPI) network. B. Top 20 hub genes.

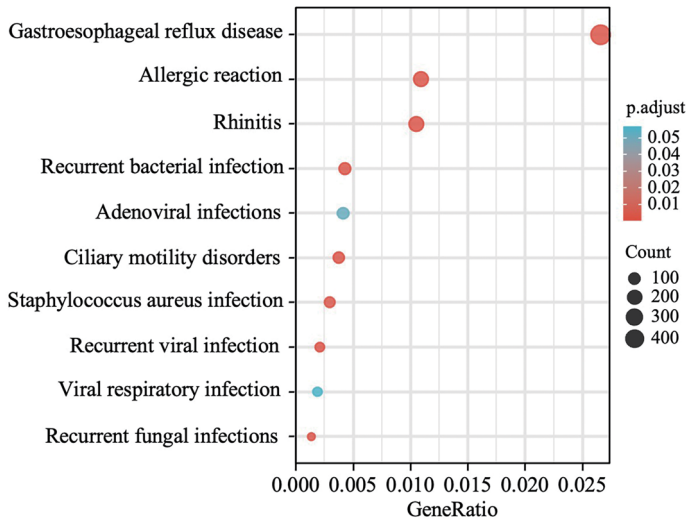


Supplementary Figure 4. ceRNA network. A. lncRNA–miRNA–mRNA network. B. circRNA–miRNA–mRNA network. The blue nodes represent mRNAs, the red frames indicate miRNAs, the purple frames denote lncRNAs, and the yellow nodes represent circRNAs.

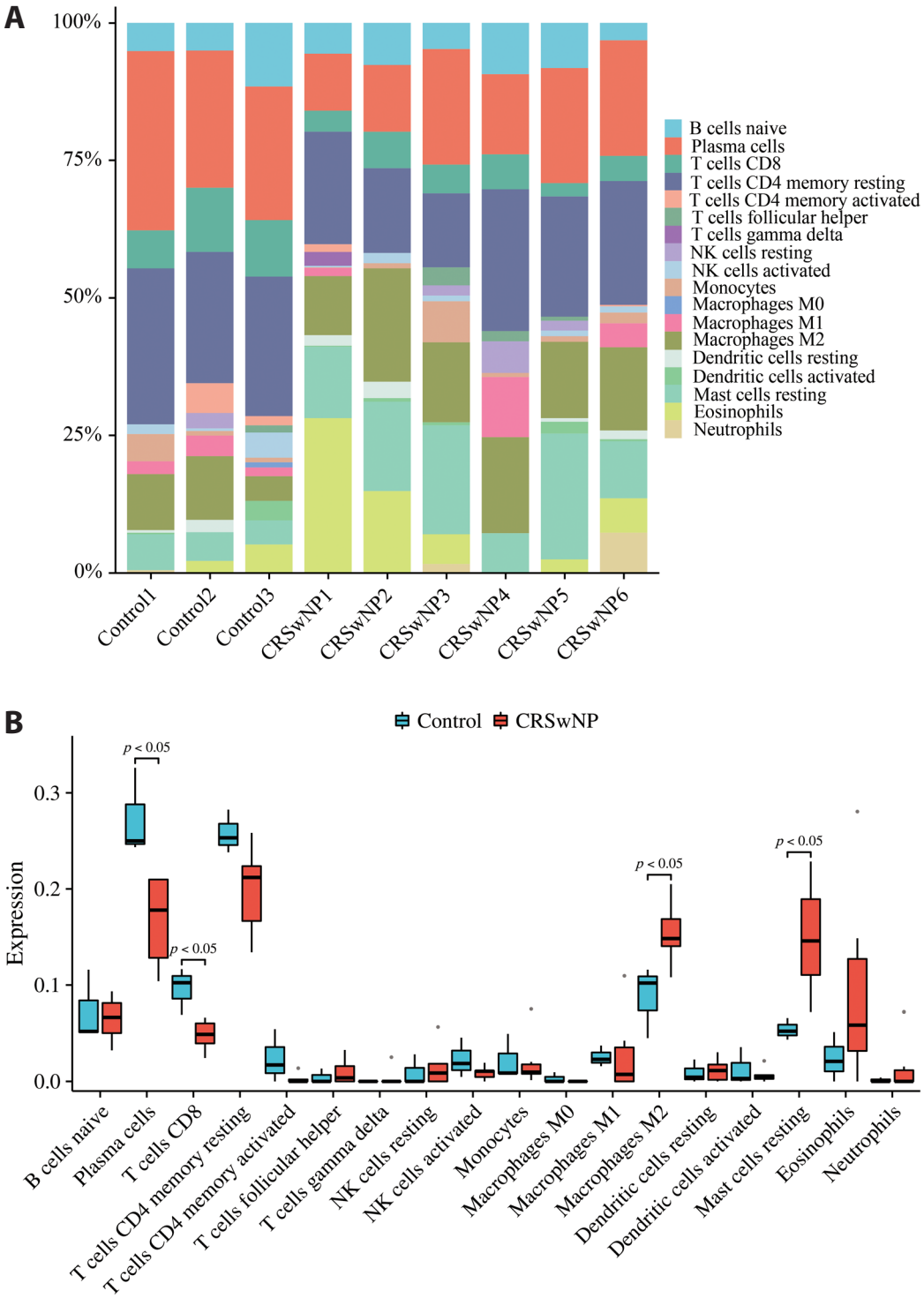


Supplementary Figure 5. GO and KEGG pathway analyses of mRNAs involved in the ceRNA networks. A. BP enrichment analysis. B. CC enrichment analysis C. MF enrichment analysis. D. KEGG enrichment analysis.

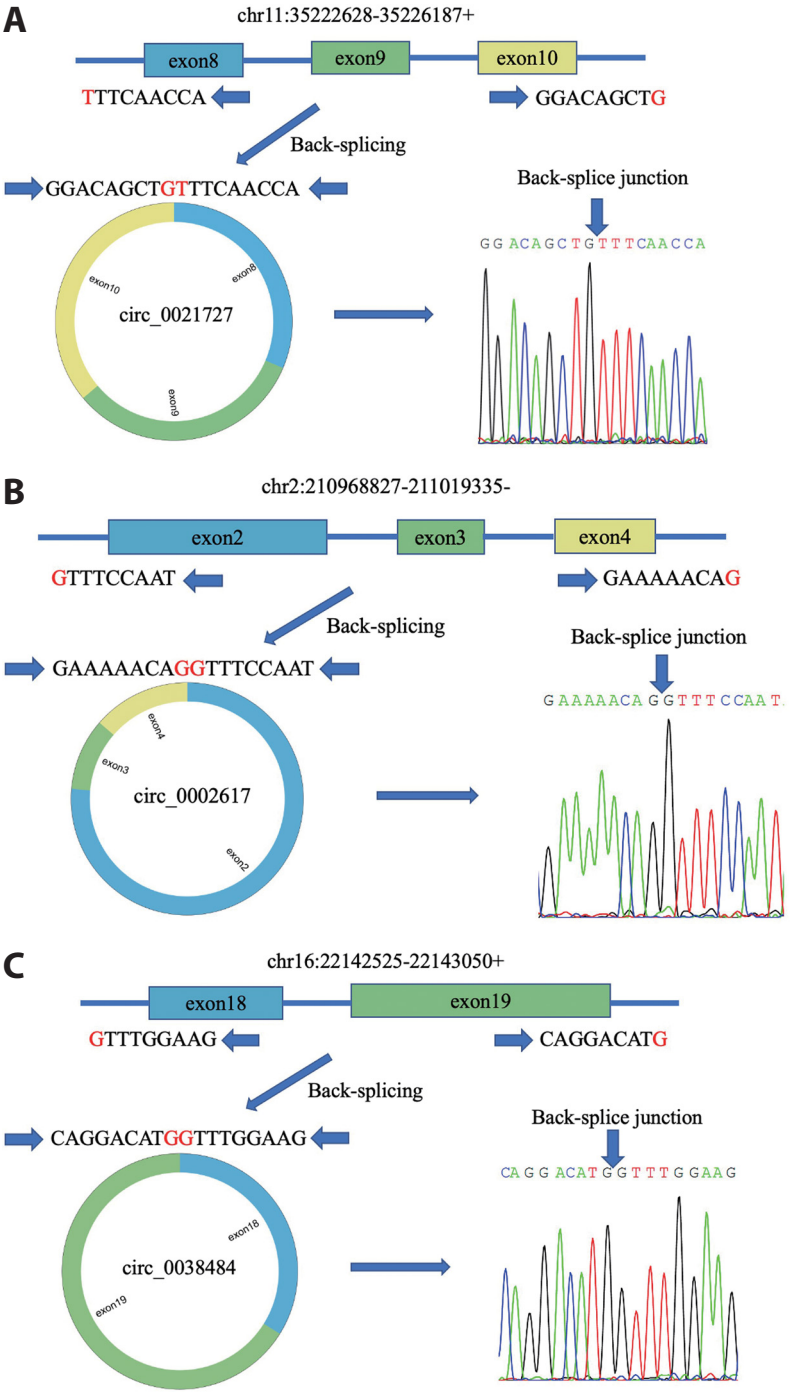
BP: Biological process, CC: Cellular component, MF: Molecular function.



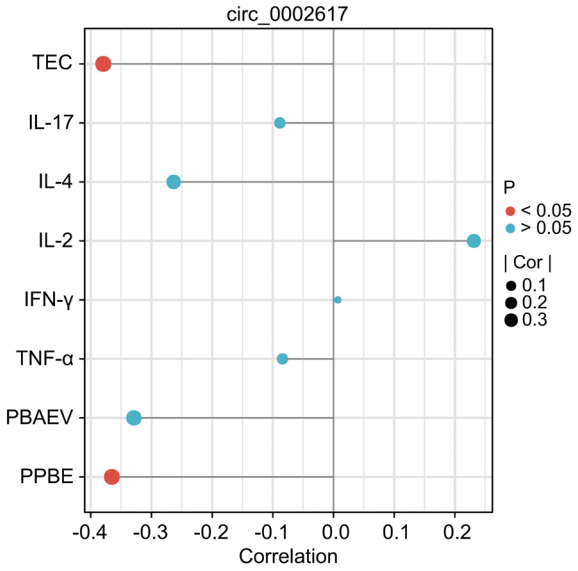
Supplementary Figure 6. DO analysis. The abscissa is the gene ratio, and the ordinate indicates enriched disease. The size and color of the nodes indicate the number of genes in the corresponding term or pathway and the *p*-value, respectively.



Supplementary Figure 7. Analysis of immune infiltration. The distribution of 18 immunocytes in CRSwNP and control groups. Color represents immune cell, the abscissa indicates the samples, and the ordinate represents the proportion of different immune cells.

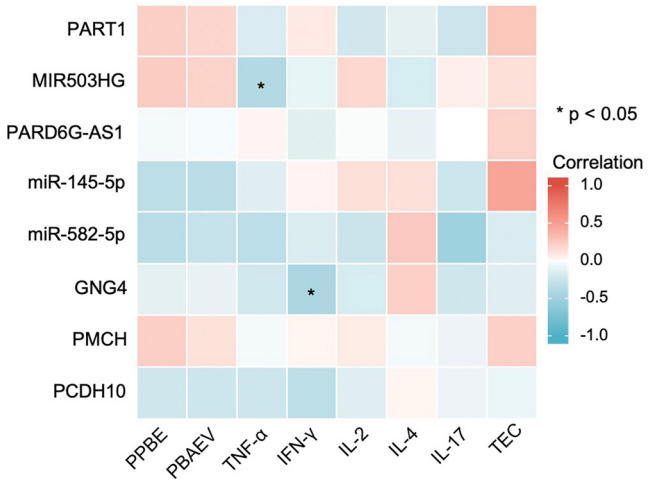


Supplementary Figure 8. Sanger sequencing revealed the back-splice junction sites of circ_0021727, circ_0002617, and circ_0038484.



Supplementary Figure 9. Correlation between circ_0002617 expression and clinical indicators in a sample from CRSwNP (n = 32). The abscissa represents the correlation coefficient, and the ordinate indicates the clinical indices. Red indicates a significant difference, and blue indicates no significant difference. The size of the node represents the absolute value of the correlation coefficient.

PPBE: Percentage of peripheral blood eosinophils, PBAEV: Peripheral blood absolute eosinophil value, TEC: Tissue eosinophil count.



Supplementary Figure 10. Correlation between the expression of mRNAs, miRNAs, lncRNAs, and clinical indicators. Red represents a positive correlation and blue indicates a negative correlation. *$p < 0.05$.

Supplementary Table 1. Clinical features of subjects.

	CRSwNP, n=32	controls, n=22	*p*
Gender (male), n (%)	18 (56.2%)	18 (81.8%)	0.096
Age (year)	46 (35.5, 58.5)	34.5 (32, 39)	0.072
Asthma	10 (31.2%)	0 (0%)	0.003
Allergic rhinitis	4 (12.5%)	0 (0%)	0.137
PBWBCC (10^9/L)	6.19 ± 1.18	6.56 ± 1.77	0.390
PPBL (%)	32.23 ± 7.9	31.8 ± 9.05	0.856
PPBM (%)	5.62 ± 1.12	5.89 ± 1.37	0.435
PPBN (%)	55.55 (49.92, 59.95)	61.05 (53.52, 68.2)	0.078
PPBE (%)	5 (3.4, 7.12)	1.35 (1.2, 2.58)	< 0.001
PPBB (%)	0.7 (0.5, 0.9)	0.6 (0.43, 0.8)	0.138
PBALV (10^9/L)	1.94 ± 0.49	2.04 ± 0.63	0.544
PBAMV (10^9/L)	0.32 (0.27, 0.4)	0.4 (0.28, 0.46)	0.252
PBANV (10^9/L)	3.52 ± 1.07	3.96 ± 1.44	0.226
PBAEV (10^9/L)	0.3 (0.2, 0.44)	0.11 (0.06, 0.16)	< 0.001
PBABV (10^9/L)	0.04 ± 0.02	0.04 ± 0.02	0.383
IgE (IU/ml)	59.91 (24.73, 129.5)	31.14 (13.93, 83.53)	0.085

PBWBCC: Peripheral blood white blood cell count; PPBL: The percentage of peripheral blood lymphocyte; PPBM: The percentage of peripheral blood monocyte; PPBN: The percentage of peripheral blood neutrophils; PPBE: The percentage of peripheral blood eosinophils; PPBB: The percentage of peripheral blood basophils; PBALV: The peripheral blood absolute lymphocyte value; PBAMV: The peripheral blood absolute monocyte value; PBANV: The peripheral blood absolute neutrophils value; PBAEV: The peripheral blood absolute eosinophil value; PBABV: The peripheral blood absolute basophils value;

Supplementary Table 2. Primer names and sequences.

Primer names	Primer sequences
GNG4 Forward	ATCCCGCTCCGGGTGAC
GNG4 Reverse	ACGTGAGCTTCACAGTAGGC
PMCH Forward	TTTGACATGCTCAGATGTATGC
PMCH Reverse	GCAGTATGGGCTTCTCCTCC
PCDH10 Forward	GAAATGAACCTCTTTCGCATGG
PCDH10 Reverse	CTCGATCACCAGCTCATAAGG
PART1 Forward	GAAAACGCAGCTACACCTGG
PART1 Reverse	CCTGCCCTTTGGTTTCTGGG
MIR503HG Forward	GAGAGCCACCTTGTGACGC
MIR503HG Reverse	CAGGGGGAACCGTACAGGAC
PARD6G-AS1 Forward	GGAAGAAATCTGGGCAAAGCC
PARD6G-AS1 Reverse	TACAGGACTCGGAACGATGC
circ_0021727 Forward	GAGGAAGAAGAGACCCCACA
circ_0021727 Reverse	CCGTGGTGTGGTTGAAACAG
circ_0002617 Forward	CCCCAGGAGTGTCAAGATCC
circ_0002617 Reverse	TCTTGTCAGACTCCATGGTACT
circ_0038484 Forward	AGACAAACACCAATCGCAGG
circ_0038484 Reverse	CCGCCAGGCACTTTGTAAAT
GAPDH Forward	CACCCACTCCTCCACCTTTG
GAPDH Reverse	CCACCACCTGTTGCTGTAG
miR-145-5p Forward	AACCTCCGTCCAGTTTTCCCA
miR-450-5p Forward	ATGCGCGCTTTTGCAATATGTT
miR-582-5p Forward	AATCGGCGTTACAGTTGTTCAAC
U6 Forward	CTCGCTTCGGCAGCACA
U6 Reverse	AACGCTTCACGAATTTGCGT

Supplementary Table 3. Analysis of differentially expressed mRNA and ncRNA.

	CRSwNP vs CRSwIT			CRSwNP vs Control		
	DEGs	Upregulated DEGs	Downregulated DEGs	DEGs	Upregulated DEGs	Downregulated DEGs
mRNA	1200	567	633	1175	696	479
lncRNA	472	241	231	517	319	198
miRNA	92	40	52	107	60	47
circRNA	217	96	121	143	104	39

DEGs: differentially expressed genes

Supplementary Table 4. Correlation between differentilally expressed genes and clinical indicators.

		circ_0021727	circ_0002617	PART1	MIR503HG	PARD6G-AS1	miR-145-5p	miR-582-5p	GNG4	PMCH	PCDH10
PPBE	r	0.393	-0.365	0.222	0.237	-0.048	-0.334	-0.350	-0.118	0.222	-0.230
	p	0.026	0.040	0.221	0.191	0.796	0.151	0.130	0.520	0.205	0.222
PBAEV	r	0.358	-0.329	0.186	0.201	-0.032	-0.355	-0.285	-0.095	0.129	-0.240
	p	0.044	0.066	0.307	0.269	0.861	0.124	0.223	0.604	0.185	0.482
TNF-α	r	0.408	-0.084	-0.177	-0.396	0.027	-0.140	-0.342	-0.221	-0.043	-0.237
	p	0.023	0.653	0.340	0.027	0.887	0.647	0.253	0.232	0.199	0.819
IFN-γ	r	0.587	0.007	0.082	-0.106	-0.132	0.027	-0.159	-0.442	0.033	-0.337
	p	0.001	0.970	0.662	0.569	0.477	0.929	0.603	0.013	0.064	0.861
IL-2	r	0.027	0.231	-0.216	0.173	-0.018	0.135	-0.267	-0.195	0.072	-0.143
	p	0.884	0.211	0.242	0.353	0.923	0.661	0.378	0.293	0.444	0.700
IL-4	r	0.043	-0.263	-0.118	-0.194	-0.101	0.138	0.253	0.213	-0.047	0.037
	p	0.819	0.152	0.527	0.295	0.590	0.654	0.404	0.249	0.844	0.801
IL-17	r	-0.100	-0.088	-0.250	0.051	0.003	-0.242	-0.500	-0.235	-0.069	-0.072
	p	0.593	0.636	0.175	0.785	0.989	0.426	0.082	0.204	0.700	0.711
TEC	r	0.405	-0.379	0.276	0.139	0.204	0.467	-0.176	-0.140	0.212	-0.078
	p	0.022	0.032	0.126	0.449	0.263	0.174	0.627	0.446	0.672	0.244

PPBE: The percentage of peripheral blood eosinophils; PBAEV: The peripheral blood absolute eosinophil value; TEC: tissue eosinophil count.

Supplementary Table 5. Details of ROC curve.

	AUC	95%CI	Sensitivity	Specificity	Youden index
GNG4	0.673	0.528-0.819	0.594	0.818	0.412
PMCH	0.663	0.516-0.811	0.719	0.636	0.355
PCDH10	0.663	0.509-0.817	0.656	0.727	0.384
PART1	0.926	0.862-0.991	0.719	1.000	0.719
MIR503HG	0.811	0.694-0.928	0.781	0.773	0.554
PARD6G-AS1	0.713	0.561-0.865	0.875	0.591	0.466
miR-145-5p	0.980	0.933-1.000	0.900	1.000	0.900
miR-582-5p	0.910	0.783-1.000	0.900	0.800	0.700
circ_0021727	0.741	0.605-0.878	0.594	0.950	0.548
circ_0002617	0.926	0.849-1.000	0.812	1.000	0.812

AUC: area under the curve; CI: confidence interval. Youden index: Sensitivity + Specificity – 1.