Original Research Article

Transcriptomic properties Involved in Allergic Rhinitis based on RNA-seq Analysis

Abstract

Allergic rhinitis (AR) is an allergic disease of the upper respiratory tract, becoming a global health problem with its mechanism still not clearly explained. To unveil the pathogenesis and immunology of AR, a transcriptomic assay was executed between the AR patient and health control by RNAseq method and the data were processed by GO, KEGG and functional network analysis. The results showed that 131 common DEGs were obtained between AR patients and control. Based on the 131 DEGs, GO analysis indicated that these DEGs were enriched on 39 GO terms, including two main functions, positive regulation of myeloid cell apoptosis and leukocyte aggregation. Otherwise, KEGG pathway analysis revealed 87 pathways, including two major pathways: *staphylococcus aureus* infection pathway and IL-17 signaling pathway. The biological process of the 131 DEGs can be divided into 3 groups: humoral immune response, positive regulation of endopeptidase activity, regulating protein processing. Functional network of the 131 DEGs sketched by Cytoscape software showed that when the degree of the network was raised to 15, only 11 genes had a linking degree more than 5, with only PLEKHC1 linking the up- and down-regulated genes. Coincidently, when diagramming all the 29 genes related directly with CD68, the gene with the largest linking degree, PLEKHC1 was found to be the only down-regulated gene. These results revealed some characteristic genes and signaling pathways for the pathogenesis of AR, which provides some evidence for further research on the immunology and pathogenesis of AR.

Keywords: Allergic rhinitis; Pathogenesis; Transcriptomic Properties; DEGs

1. Introduction

Allergic rhinitis (AR) is a common inflammatory disease of the upper respiratory tract, raising a global health problem impairing the quality of life of approximately 10%–20% of the population worldwide. AR is also commonly the first step toward the development of conjunctivitis and asthma[1]. AR has multigenetic characteristics involving complex processes that are roughly divided into three stages: sensitization, excitation, and effect. Long-term studies have indicated that the pathological changes observed in AR and bronchial asthma (BA) as well as the pathogenesis and treatment of these conditions have remained unchanged. Settipane et al. proposed that the probability of developing BA in patients with AR is three times higher than that in normal people[2]. Another longitudinal study of children with AR suggested that the risk of developing BA is two times higher in these children than in normal children[3].

Increasing indirect evidence has demonstrated that genetic mutations in the promotors of the genes encoding interleukin (IL)-1, tumor necrosis factor (TNF)-α, IL-4, and IL-6 are involved in the development of AR[4-7]. However, the role of these inflammation-related factors remains unclear with respect to AR pathogenesis, with a small number of factors directly propelling the pathogenesis. AR is difficult to cure, but for long-term disease control, anti-allergic drugs are commonly prescribed. These drugs antagonize inflammatory responses mediated via the AR signaling pathways by blocking B cells from producing IgE inflammatory responses and receptor binding in some cases, or by reducing blood IgE levels. However, as a thorough mechanistic understanding of AR pathogenesis has not been gained, a long-term cure has not been found. Further, AR can occur without allergen stimulation, sometimes even by stimulation from something as simple as cold wind, which makes the pathology of AR more

complicated. In this study, we performed RNA transcriptome analysis on the nasal mucosa of patients with AR to identify characteristic genes associated with AR, followed by a bioinformatics analysis to screen and identify genes and proteins likely involved in the pathogenesis of AR and to understand their relationship with known immunological signal pathways.

2. Materials and methods

2.1 Ethics statement

All experiments in this study were conducted according to a protocol approved by the Experiments Inspectorate under the Yangzhou University Ethics committee (YZU20150110). All study participants provided consent for operation procedure and use of excised tissue specimens. All experiments in this study were performed in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.2 Preparation of samples

Using intraoperative peep endoscopic forceps, tissue samples were collected from the noses of four surgical patients: two with persistent AR (based on symptoms and allergen skin prick test) and the other two with nasal turbinate bone hyperplasia (with endoscopic record). The allergen skin prick test indicated that the two AR patients were both allergic to dust mite, while the control two showed no allergic symptom to any skin prick test allergen. The procedures were performed in the Otolaryngology Head and Neck Department of the Jiangsu Subei People's Hospital, China. Both groups were pre-tested via obtaining preoperative medical history and performing nasal endoscopy and allergen skin prick test to obtain baseline condition of AR in each patient.

2.3 RNA extraction, library preparation, and sequencing

Total RNAs from the collected tissue samples were isolated using Trizol reagent as per manufacturer's instructions (Invitrogen Life Technologies, USA), after which the concentration, quality,

and integrity of RNA were determined to be qualified using a NanoDrop spectrophotometer (Thermo Scientific, USA) all samples had an OD260/280 ratio of >2.0 and RIN value of >8.0. Sample preparation used 3 mg RNA of each sample as input material for RNA. Sequencing libraries were then generated using the TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was performed using divalent cations under elevated temperatures in an Illumina proprietary fragmentation buffer. First-strand cDNA was synthesized using random oligonucleotides and SuperScript II. Second-strand cDNA synthesis was subsequently performed using DNA polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities and the enzymes removed. After adenylation of the 3′ ends of the DNA fragments, Illumina PE adapter oligonucleotides were ligated for hybridization，DNA fragments with ligated adaptor molecules on both ends were selected in size by agarose gel electrophoresis, and then enriched by PCR amplification. The sequencing library was then sequenced on an Illumina HiSeq2500 platform (Illumina) by Biomarker Technologies, Beijing, China.

2.4 Acquisition of the raw data, unigene cluster, and annotation analysis

The original sequence data, or raw data reads, were saved as a FASTQ file, which included detailed read sequences and quality information. FastQC was used for quality control analysis to filter out "dirty" raw reads, such as reads with adapters, reads with > 10% unknown bases, and low-quality reads (reads with > 50% bases with a quality value of ≤5). Sequences that met our filtering requirements were used for subsequent analyses. *De novo* transcriptome assembly was performed using Trinity (Broad Institute, http://trinityrnaseq.sf.net). A K-mer library was constructed with the filtered reads, and the contigs were formed using Inchworm (Broad Institute, http://trinityrnaseq.sf.net). Using Chrysalis (Broad Institute, http://trinityrnaseq.sf.net), a dataset was built from the contigs, and de Bruijn graphs were generated.

Butterfly (Broad Institute, http://trinityrnaseq.sf.net) was then used to optimize the de Bruijn graphs and create the final transcript paths.

Differentially expressed genes (DEGs) were then identified using a false discovery rate (FDR) ≤ 0.001 (FDR no greater than 0.05) and an absolute value of Log2Ratio ≥ 1 (two-fold change) as thresholds for significance. More stringent criteria, with a smaller FDR or greater Log2Ratio change, were used for subsequent analyses. Several DEGs were randomly selected for further qPCR validation.

2.5 GO functional enrichment and KEGG pathway analysis

Blast2go (BioBam Bioinformatics S.L.,USA)was used to annotate unigenes based on GO terms and NR database annotation. Conservation of gene identities in other species was analyzed using BLASTX. To annotate genes with a similar origin or function, unigenes were also aligned with the eggNOG database (http://www.ncbi.nlm.nih.gov/COG/; http://eggnog.embl.de/version_3.0/). To summarize pathway information, the KEGG Automatic Annotation Server was used to perform pathway annotation. Pathways with a Qvalue ≤ 0.05 were considered significantly enriched.

2.6 Intergenic network simulation

To detect the expression level of common DEGs between the experimental group and controls, we calculated the correlation efficiencies between genes by expression levels (the system default intensity between genes in −0.95~0.95). Cytoscape software (www.cytoscape.org/) was used to process the correlation efficiencies among expression levels of DEGS and draw the functional network diagrams.

3. Results

3.1 Information of transcriptome sequencing analysis

Our RNA-seq sequencing provided a total of 26.30 M reads (1.34 gigabases). Quality evaluation of the sequencing demonstrated that the Q30 value of each sample surpassed 80%. After removing the data on rRNA from the original data for accuracy, we obtained 19.72 M high-quality reads (0.98 gigabases; Table 1).

Using standards of FDR < 0.01 and fold change ≥ 2, DEGs were obtained (Table 2), with 131 common DEGs identified. After obtaining a list of DEGs, IL6, PI3KAP1, CCL2, HSPH1, MUC2, and

Table 1 Evaluation for high quality reads data

Samples	Total reads	Total nucleotides(bp)	GC percentage	Q30 percentage*
E1	3,624,510	179,944,230	50.17%	80.58%
E2	3,356,526	167,722,924	49.90%	80.75%
E3	3,888,795	193,523,113	50.94%	80.63%
E4	4,328,857	216,227,165	46.82%	80.63%

*Q30 percentage is the percentage of the bases with the mass value more than 30.

Table 2 Number of DEGs between AR patients and control

group*	number	up	down
E1_vs_E2	820	402	418
E3_vs_E4	818	277	541
E1_E3_vs_E2_E4	131	75	56

* The first sample represents the control, the latter is the AR patient.

Table 3 the 20 most differentiated genes

order	Up-regulated		Down-regulated	
	Gene name	**FDR**	**Gene name**	**FDR**
1	SELE	*	SAA2	*
2	CRISP3	*	SLC34A2	*
3	CA2	*	CAPN13	*
4	PRR4	*	C15orf48	*
5	NR4A3	*	IGLV8-61	*
6	PRB2	*	SLC26A4	1.78E-15
7	MSMB	*	KLHDC7B	2.11E-15
8	ACTC1	*	DNAH5	2.55E-15

9	PIP	*	SAA4	2.78E-15
10	MUC7	*	TMEM190	5.66E-15

*FDR value is smaller than 1.0E-16

FERMT1 were randomly selected for validation using a qPCR method. The qPCR validation results were consistent with the transcriptome sequencing analysis results. The 20 genes with the largest changes in transcription are presented in Table 3; most of these 20 DEGs are closely related with the regulation of inflammation. All reads were statistically investigated (i.e., exon, intron, and intergenic region), with their genomic distribution plotted in Figure 1.

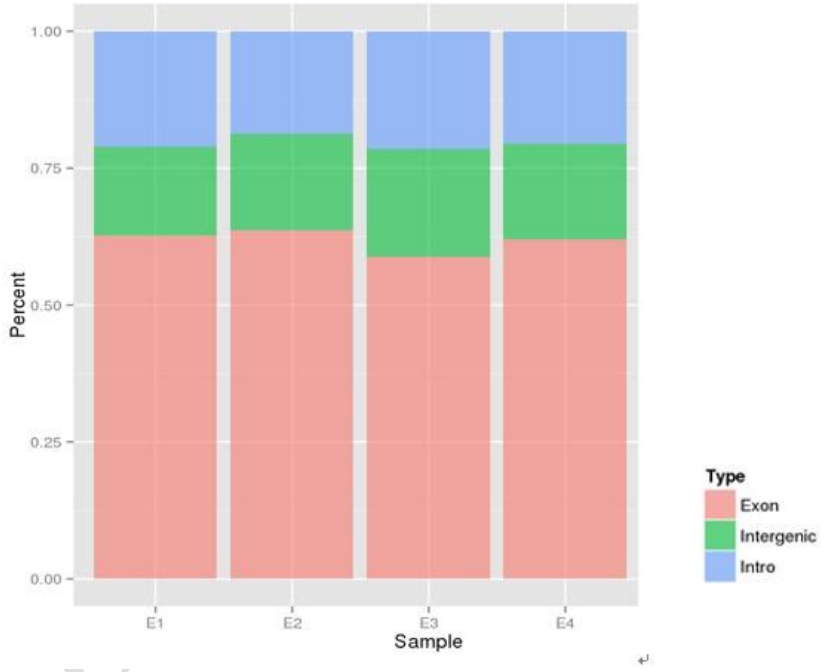

Figure 1. Genomic distribution of reads.

3.2 GO functional enrichment and KEGG pathway analysis

Using the 131 identified DEGs, GO analysis indicated that these DEGs were enriched for 39 GO terms, including two main functions: positive regulation of myeloid cell apoptosis and leukocyte aggregation (Figure 2). KEGG pathway analysis identified 87 potential pathways, including two major pathways: the *Staphylococcus aureus* infection pathway and the Il-17 signaling pathway (Figure 3).

The biological processes of the 131 DEGs can be divided into three groups (Figure 4): humoral immune responses, positive regulation of endopeptidase activity, and the regulation of protein processing.

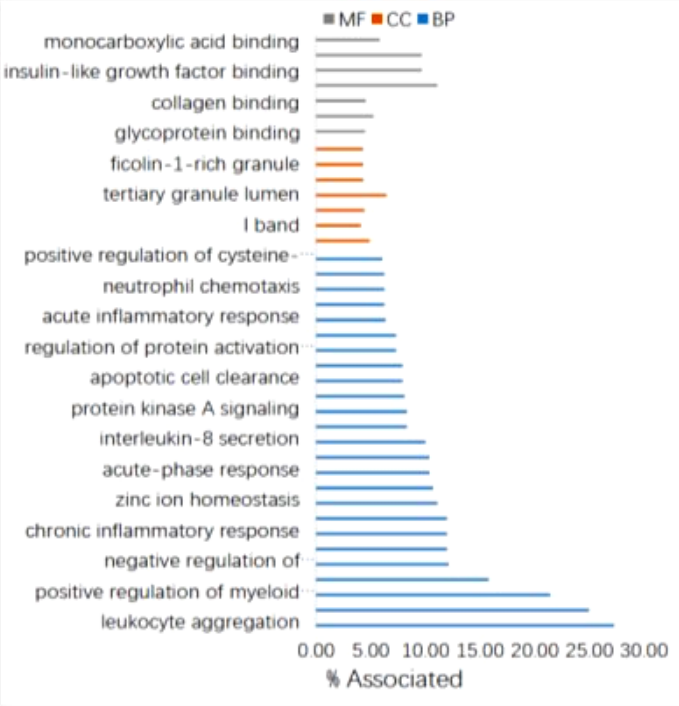

Figure 2 GO terms based on the 131 DEGs

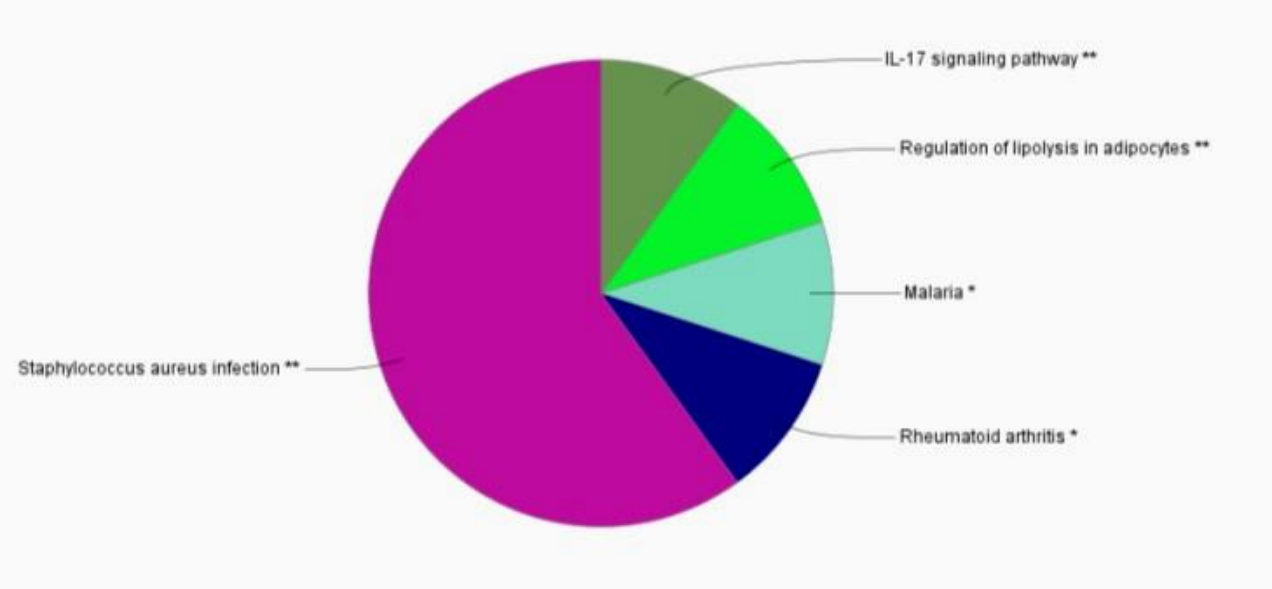

Figure 3 KEGG pathways based on the 131 DEGs

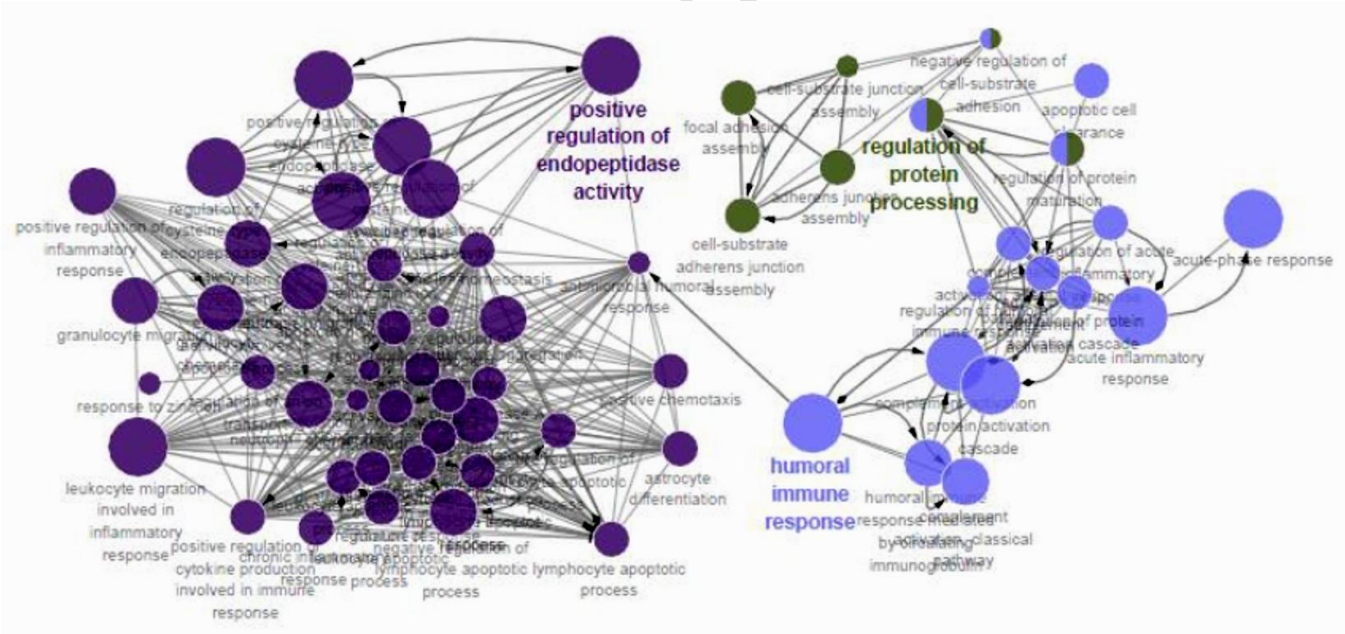

Figure 4 Biological processes among these 131 genes.

3.3 Functional network analysis

Based on the expression levels of the 131 DEGs in each sample, the correlation coefficients were calculated between genes, with the system default intensity between genes being −0.95~0.95. For individually sketching the relationships among upregulated and downregulated genes, the latter linked more closely than the former. When the degree of the network was raised to 10, the downregulated genes remained closely linked, whereas the upregulated genes were parted into two groups (Figure 5). When the degree of the network was raised to 15, only 11 genes had a linking degree of >5 (Figure 6). Among the 11 genes, PLEKHC1 appeared as a hub linking the upregulated and down regulated genes, indicating that it could play a critical role in the network function. Coincidently, when diagramming all 29 genes directly related to CD68 (which had the greatest linkage in our study based on correlation coefficients), PLEKHC1 was found to be the only downregulated gene (Figure 7).

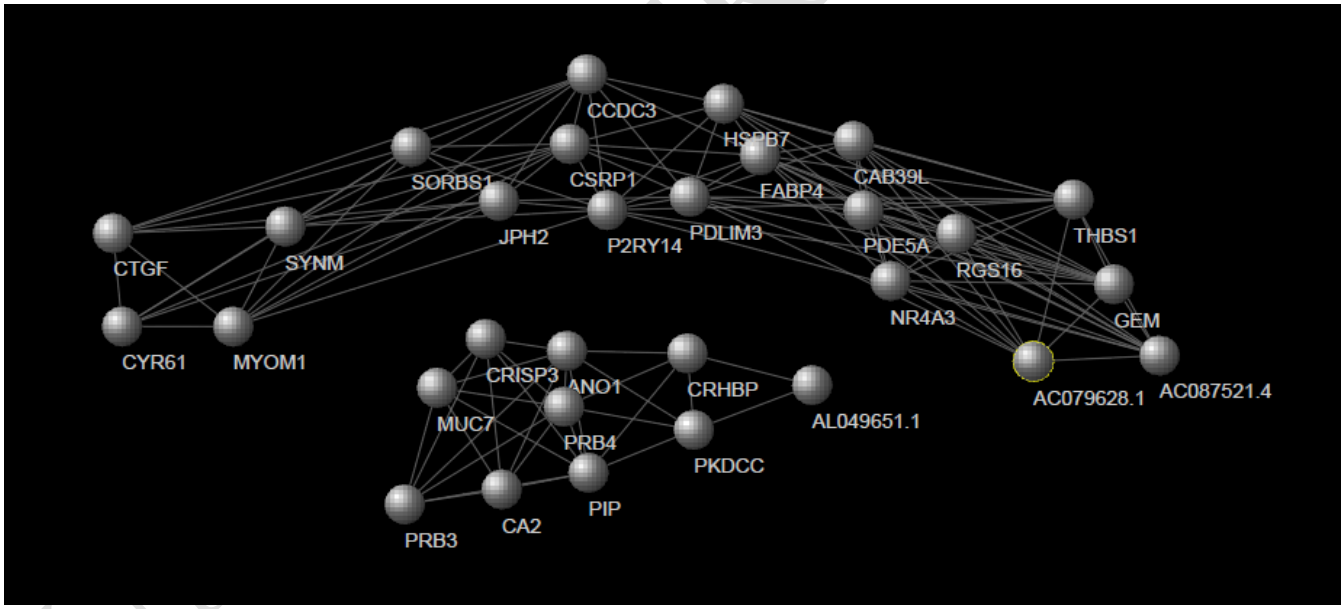

Figure 5 Functional network among the up-regulated genes.

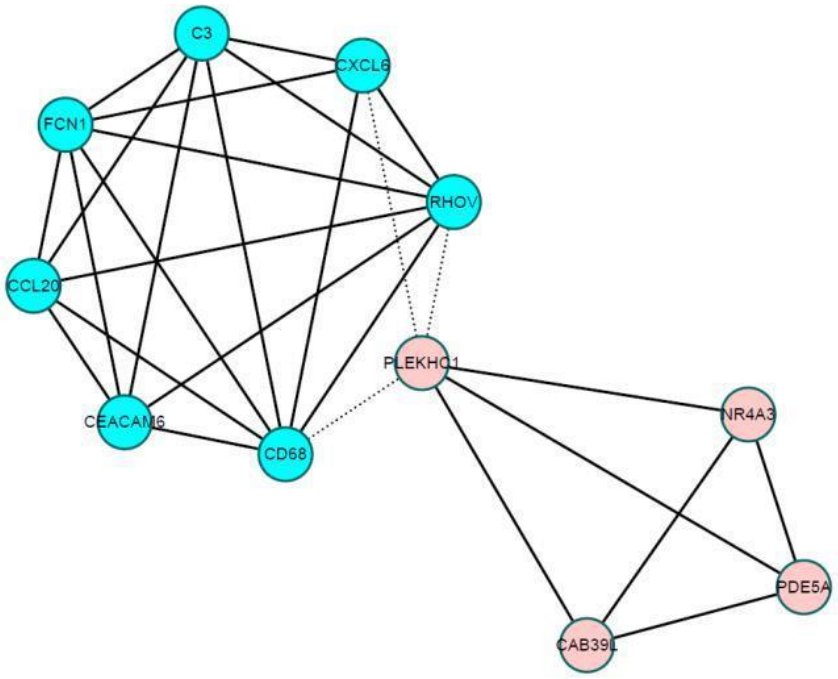

Figure 6 Functional network under the network degree more than 15.

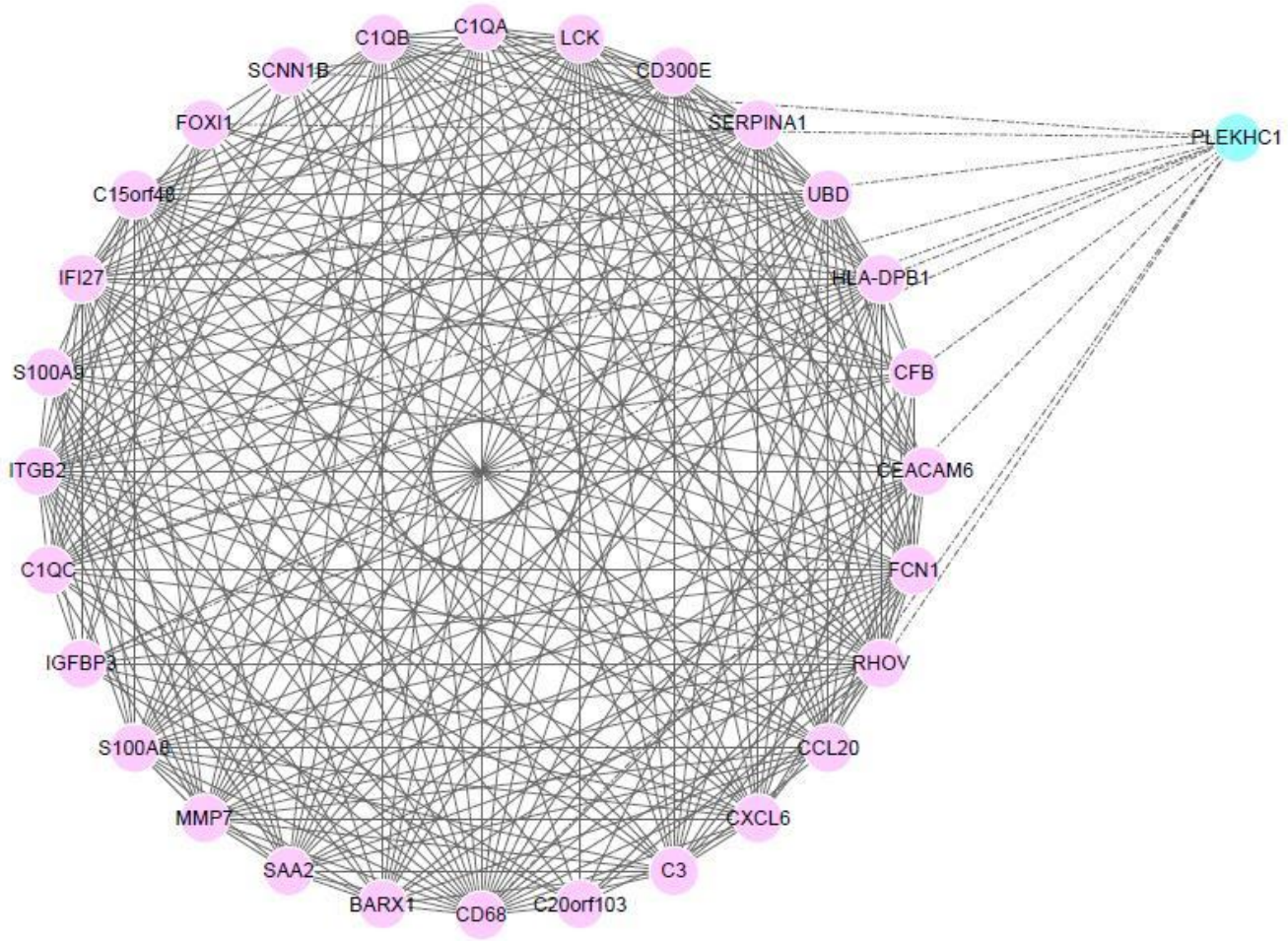

Figure 7 Functional network of the genes related directly with CD68.

4. Discussion

The incidence rates of AR have recently shown an upward trend, the main cause of which is thought to be the interaction between genetic and environmental factors. To prevent the increase in the incidence of AR, drug therapy and immunotherapy have been adopted as the primary treatments for AR; however, a fundamental cure to completely resolve the condition is still lacking.

It has been demonstrated that BCAP assist CD19 in the activation of PI3K and play an important role in immune function regulation[8]. One study has indicated that BCAP plays a peculiar role in the TLR signaling pathway[9], whereby the BCAP gene of the B cell receptor signaling pathway is overexpressed

in patients with AR, which is identical to the results of most studies on AR; this suggests that BCAP is closely related with the pathogenesis of AR within the local populations of Yangtze River basin region of China.

In the present study, CD68 showed the largest number and the greatest strength of linkages, suggesting an important role in AR pathogenesis. CD68 is involved in the phagocytic activity of tissue macrophages through binding to tissue- and organ-specific lectins or selectins, allowing macrophage subsets to identify and migrate to the targeted regions. Rapid recirculation of CD68 from endosomes and lysosomes to the plasma membrane may allow macrophages to crawl over selectin-bearing substrates or other cells. CD68-positive mononuclear cells, rather than or in addition to epithelial cells, are a major source of CCL11, IL-4, and CCL2 expression[10]. Moreover, $CD68^{+}S100A8/A9^{+}$ monocytes are significantly increased in the lower airway mucosa and the alveoli of patients with fatal asthma compared with those in normal individuals. In addition, S100A8 upregulates IL-6 and IL-10 mRNAs, whereas S100A9 upregulates TNF-α and IL-6 mRNAs[11], which coincides with the high expression close relationship of CD68 with S100A8 and S100A9 in this study. However, Ekman et al. reported an increased number of $CD68^{-}CD123^{+}$ dendritic cells along with the positive plasmacytoid dendritic cells in response to inflammatory mediators [12], suggesting that CD68 is not a unique signal for inflammatory stimulation. These controversial views on the function of CD68 remain to be resolved.

Furthermore, our study suggests that the gene PLEKHC1, also known as Fermitin family homolog 2 (FERMT2), links the up- and downregulated DEGs related to AR. FERMT2 is a component of extracellular matrix structures in mammalian cells and is required for the proper control of cell shape and membrane movement[13]. The accumulation of eosinophils and the recruitment of macrophages in the airway mucosa tissue is characteristic of AR, which may explain the higher expression of PLEKHC1 in patients with AR in this study.

Conclusion

In conclusion, the expression levels of BCAP, CD68-S100A8/A9 and PLEKHC1 can play important roles in the pathogenesis of AR in patients from Yangtze River basin, China.

References

[1] Settipane RA, Schwindt C. Chapter 15: allergic rhinitis. Am J Rhinol Allergy. 2013;27(Suppl 1):S52eS55.

[2] Settipane RJ, Hagy G W, Settipane GA. Long-term risk-factors for developing asthma and allergic rhinitis: a 23-year follow-up-study of college students. Allergy Proceedings, 1994, 15 (1): 21-25.

[3] Burgess JA, Haydn Waiters E, Bymes GB, et a1. Childhood allergic rhinitis predicts asthma Incidence and persistence to middle age: a longitudinal study[J]. J Allergy Clin Immunol, 2007, 120 (4):863-869.

[4] Barnes KC, Freidhoff LR, Nickel R, et al. Dense mapping of chromosome 12q13.12 -q23.3 and linkage to asthma and atopy. J Allergy Clin Immunol, 1999, 104 (2): 485-491.

[5] Deichmann KA, Starke B, Schlenther S, et al. Linkage and association studies of atopy and the chromosome l1q13 region. J Med Genet, 1999, 36:379-382.

[6] Barnes KC, Neely JD, Duffy DL, et al. Linkage of asthma and total serum IgE concentration to markers on chromosome 12q: Evidence from Afro-Caribbean and cancasian populations. Genomics, 1996, 37:41-50.

[7] Lydie Holla, Anna Vasku, Vladimir Znojil, et a1. Association of 3 gene polymorphisms with atopy diseases. J Allergy Clin Immunol, 1998, 103(4):702-708.

[8] MacFarlane AW, Yamazaki T, Fang M, Sigal LJ, Kurosaki T, et al. Enhanced NK- cell development and function in BCAP-deficient mice. Blood. 2008, 112:131-140.

[9] Ty Dale Troutman, Wei Hu, Stephanie Fulenchek, et al. Role for B-cell adapter for PI3K (BCAP)as a signaling adapter linking Toll-like receptors (TLRs) to serine/threonine kinases PI3K/Akt. PNAS. 2012, 109:267-272.

[10] J. Kelley Bentley, Uma S. Sajjan, Marta B. Dzaman, et al. Rhinovirus colocalizes with CD68- and CD11b-positive macrophages following experimental infection in humans. J Allergy Clin Immunol. 2013, 132(3): 758–761.

[11] Okada K, Arai S, Itoh H, Adachi S, Hayashida M, Nakase H, Ikemoto M. CD68 on rat macrophages binds tightly to S100A8 and S100A9 and helps to regulate the cells' immune functions. J Leukoc Biol. 2016, 100(5):1093-1104.

[12] Ekman AK, Erjefält JS, Jansson L, Cardell LO. Allergen-induced accumulation of CD68-CD123+ dendritic cells in the nasal mucosa. Int Arch Allergy Immunol. 2011, 155(3):234-242.

[13] Tu Y, Wu S, Shi X, et al. Migfilin and Mig-2 link focal adhesions to filamin and the actin cytoskeleton and function in cell shape modulation". Cell. 2003, 113 (1): 37–47.