Identification of key genes in allergic rhinitis by bioinformatics analysis

Yunfei Zhang*, Yue Huang*, Wen-xia Chen and Zheng-min Xu [ID]

Abstract

Objective: This study aimed to explore the potential molecular mechanism of allergic rhinitis (AR) and identify gene signatures by analyzing microarray data using bioinformatics methods.

Methods: The dataset GSE19187 was used to screen differentially expressed genes (DEGs) between samples from patients with AR and healthy controls. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were applied for the DEGs. Subsequently, a protein–protein interaction (PPI) network was constructed to identify hub genes. GSE44037 and GSE43523 datasets were screened to validate critical genes.

Results: A total of 156 DEGs were identified. GO analysis verified that the DEGs were enriched in antigen processing and presentation, the immune response, and antigen binding. KEGG analysis demonstrated that the DEGs were enriched in *Staphylococcus aureus* infection, rheumatoid arthritis, and allograft rejection. PPI network and module analysis predicted seven hub genes, of which six (*CD44*, *HLA-DPA1*, *HLA-DRB1*, *HLA-DRB5*, *MUC5B*, and *CD274*) were identified in the validation dataset.

Conclusions: Our findings suggest that hub genes play important roles in the development of AR.

Department of Otolaryngology-Head and Neck Surgery, Children's Hospital of Fudan University, Shanghai, China

*These authors contributed equally to this study.

Corresponding author:
Zheng-min Xu, Department of Otolaryngology-Head and Neck Surgery, Children's Hospital of Fudan University, 399 Wan Yuan Road, Shanghai 201102, P. R. China.
Email: 13916320945@163.com

Introduction

Allergic rhinitis (AR), also known as hay fever, is an inflammatory disorder of the nasal mucosa, which is induced by an immunoglobulin E (IgE)-mediated reaction in allergen-sensitized subjects.[1] Four hundred million people suffer from AR worldwide, and the incidence has been increasing over recent decades because of rises in industrialization and air pollution.[2] AR is characterized by sneezing, rhinorrhea, nasal congestion, and nasal pruritus, which are often accompanied by ocular pruritus, redness, and lacrimation in 60% to 70% of patients.[3] Although not life-threatening, the symptoms of AR are frequently bothersome, and adversely affect work and the quality of life, especially in children and adolescents.[4] Moreover, AR is commonly associated with other conditions such as asthma, anosmia, and otitis media, suggesting that these conditions are closely related.[5] Indeed, studies have suggested that 20% to 50% of patients with AR have clinical asthma, whereas more than 80% of patients with allergic asthma have concomitant rhinitis symptoms.[6] Although numerous efforts have been made to understand the mechanisms underlying AR, available treatments are only partially successful in certain subtypes of patients.[7] Thus, there is an urgent need to identify the key molecules involved in AR pathology, to help provide an accurate diagnosis as well as timely treatment.[8]

Evidence such as population heterogeneity, and interactions between multiple genes as well as environmental and genetic factors clearly supports AR being influenced by a complex immune procedure.[9–11] The world's largest study on AR comparing around 60,000 patients with over 150,000 healthy controls of primarily European ancestry identified 41 loci in the human genome which significantly increased the risk of AR when altered.[12] Additionally, polymorphisms of candidate genes have been associated with the development of AR,[13–15] while several genes related to immunity and transcriptional regulation, such as *FOS*, *JUN*, and *CEBPD*, play crucial roles in seasonal allergic rhinitis (SAR) through interacting with each other.[16] Moreover, a comparison of seven SAR samples and five non-allergic samples found that *CDC42EP5*, *SLC39A11*, and *PRDM10* genes associate with the pathogenesis of AR.[17] However, the prevalence and factors responsible for AR etiology are still not fully understood, especially in China.

In the present study, we downloaded the original microarray data GSE19187 from the Gene Expression Omnibus (GEO) database. Subsequently, we compared gene expression profiles of nasal airway epithelium cells from patients with AR and healthy controls to identify differentially expressed genes (DEGs). These underwent gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, followed by the construction of a protein–protein interaction (PPI) network to identify hub genes. Gene expression profiles GSE44037 and GSE43523 were combined to validate the critical genes, which could be used as molecular biomarkers or diagnostic targets for AR therapy. Collectively, our study will help the development of a genetic diagnosis for AR and more effective measures of prevention and intervention.

Materials and methods

Data source

Gene expression profiles were downloaded from the GEO database (http://www.ncbi. nlm.nih.gov/geo). GSE19187 includes 14 cases of isolated AR without asthma and 11 healthy control samples. The expression profiling of nasal epithelial cells collected

by brushing had been performed on the GPL6244 platform (Affymetrix Human Genome 1.0 ST Array);[18] GSE44037 includes five cases of AR and six healthy control samples, based on the GPL13158 platform (Affymetrix HT HG-U133+ PM Array);[19] and GSE43523 includes seven cases of SAR and five healthy control samples, based on the GPL6883 platform (Illumina HumanRef-8 v3.0 expression beadchip). Datasets GSE44037 and GSE43523 were merged as validation data to validate the critical genes. The expression levels of *CD44*, *HLA-DPA1*, *HLA-DRB1*, *HLA-DRB5*, *MUC5B*, and *CD274* were compared between AR and healthy control samples using the Wilcoxon test. The subject characteristics of the dataset are listed in Supplementary Table 1 and the genes detected by microarray data are listed in Supplementary Table 2. All data are freely available online, and this study did not involve any experiments on humans or animals performed by any of the authors. Ethical approval was not required because this study used publicly available datasets.

Identification of DEGs

Linear Models for Microarray Data, an R package from Bioconductor (http://www.bioconductor.org), was used to identify DEGs between samples from patients with AR and healthy controls. If multiple probes corresponded to the same gene, we used the mean value as the expression of this gene. We defined DEGs as a fold change (FC) >1.5 or <-1.5.[18,19] An adjusted P value <0.05 was considered statistically significant. DEGs were clustered using the hierarchical clustering method to show two groups by the heatmap.2 function from the gplots R package (https://www.rdocumentation.org/packages/gplots/). The Euclidean distance was chosen as a measure of distance between the samples. A volcano plot for the distribution of genes from

patients with AR and healthy controls was made by plotting the negative log10-transformed p values against the log ratios (log2 fold change) in the two groups.

GO and KEGG pathway enrichment analyses of DEGs

GO annotation and KEGG pathway enrichment analyses were performed for DEGs using The Database for Annotation, Visualization and Integrated Discovery (DAVID; https:/david.ncifcrf.gov/), which provides systematic and comprehensive annotation tools to analyze high-throughput gene function. A P value <0.05 was considered statistically significant.

PPI network construction and module analysis

The Search Tool for the Retrieval of Interacting Genes database (STRING; version 10.5; http://www.string-db.org/) was used to assess and integrate PPI with a combined score >0.4 as the cutoff criterion. The PPI network of DEGs was visualized by Cytoscape (version 3.5.0) and the plug-in Molecular Complex Detection (MCODE) was used to screen functional modules to identify hub genes. A P value <0.05 was considered statistically significant.

Statistical analysis

Statistical analysis was performed using SPSS version 16.0 (SPSS Inc., Chicago, IL, USA) and R software version 3.2.5 (MathSoft, USA). The DEGs between the groups of AR and healthy controls were determined using a two-tailed Student's *t*-test. The Wilcoxon test was used to compare the expression level of critical genes between the groups of AR and healthy controls in the validation datasets. P values <0.05 were considered statistically significant.

Results

Identification of DEGs

Gene expression profiles from GSE19187 identified 156 DEGs, of which 94 were up-regulated and 62 were down-regulated in patients with AR relative to healthy controls (Figure 1). Detailed DEG information is shown in Supplementary Table 3. Hierarchical clustering analysis revealed a clear distinction of DEGs between patients with AR and healthy controls (Figure 2).

Gene ontology enrichment analysis

The online analysis tool DAVID was used to identify significantly enriched GO terms for DEGs. DEGs that were down-regulated in patients with AR were more enriched than those that were up-regulated (Table 1). The down-regulated genes were mainly enriched in the biological process ontology, including antigen processing and presentation, the T cell receptor signaling pathway, and the immune response. For the cellular component ontology, down-regulated genes were enriched in the major
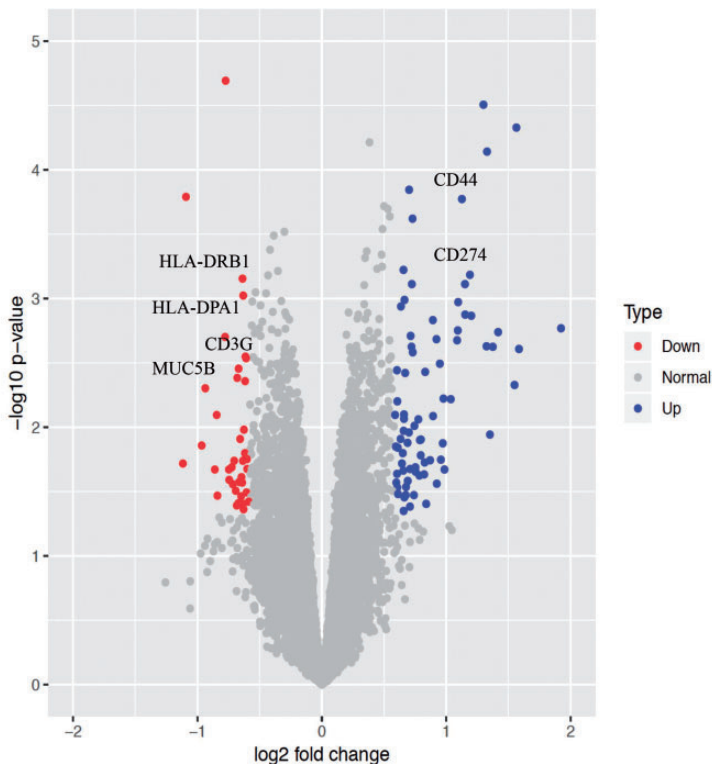

Figure 1. Volcano plot for the distribution of gene expression between patients with AR and healthy controls from the GSE19187 dataset. Negative $\log_{10}$-transformed *P* values are plotted against log ratios (log2 fold change) in the two groups. Red and blue points represent down-regulated and up-regulated genes in patients with AR compared with healthy controls, respectively. Gray points represent non-differentially expressed genes. Critical genes are labeled.
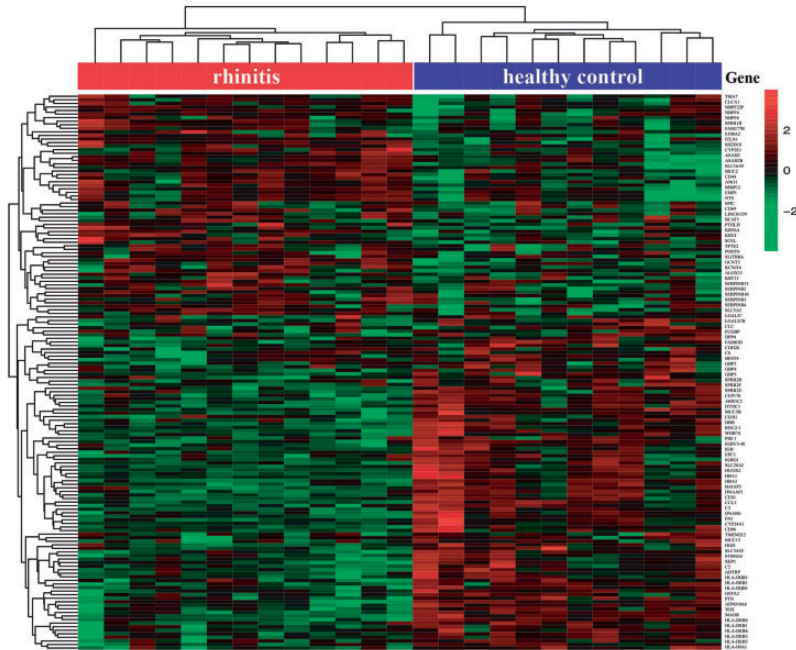AR, allergic rhinitis.

Figure 2. Heatmap of DEG hierarchical clustering between patients with AR and healthy controls. Rows represent genes, while columns represent samples. The red bar represents patients with AR while the blue bar represents healthy controls. Values are gene expression levels.
DEGs, differentially expressed genes; AR, allergic rhinitis.

histocompatibility complex (MHC) class II protein complex, haptoglobin–hemoglobin complex, and various vesicle membranes. For the molecular function ontology, down-regulated genes were significantly enriched in peptide antigen binding, MHC class II receptor activity, and haptoglobin binding.

KEGG pathway analysis

The significantly enriched pathways of DEGs showed that up-regulated genes were only enriched in amoebiasis (Table 2), while down-regulated genes were enriched in various pathways including *Staphylococcus aureus* infection, rheumatoid arthritis, asthma, graft-versus-host disease, leishmaniasis, allograft rejection, and herpes simplex infection (Table 2).

Module analysis and hub genes selection in the PPI network

After assessing the DEG PPI by STRING, a high-quality PPI network was constructed by Cytoscape containing 59 nodes and 92 edges. The MCODE plug-in was used to screen functional modules of the PPI network, which represent a sub-network of highly interconnected proteins. The largest module contained 13 nodes and 33 edges (Figure 3). The seven highest hub nodes evaluated by connectivity degree in the PPI network were identified (Table 3), and their expression level was examined in an independent validation dataset (GSE44037 and GSE43523) and compared with the original dataset (GSE19187). Six of the seven hub genes, *CD44*, *HLA-DPA1*,

Table 1. Gene ontology analyses of the DEGs

Regulation	Category	Term	Gene count	Adjusted P value	Gene symbol
Up-regulated	GOTERM_BP_ FAT	GO:0010951~negative regulation of endopeptidase activity	7	4.0E-03	SERPINB10, SERPINB3, SERPINB4, CSTA, SERPINB13, SERPINB2, FETUB
	GOTERM_MF_ FAT	GO:0004869~cysteine- type endopeptidase inhibitor activity	6	3.3E-05	SERPINB3, CST2, CSTA, CST1, SERPINB13, FETUB
	GOTERM_MF_ FAT	GO:0002020~protease binding	7	2.3E-04	DPP4, SERPINB3, SERPINB4, CST2, CSTA, CST1, SERPINB13
	GOTERM_MF_ FAT	GO:0005198~structural molecule activity	7	1.7E-02	CLDN22, CSTA, UPK1B, KRT4, KRT13, SPRR1B, KRT6A
Down-regulated	GOTERM_BP_ FAT	GO:0002504~antigen processing and presentation of peptide or polysaccharide antigen via MHC class II	6	2.1E-07	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_BP_ FAT	GO:0031295~T cell costimulation	7	9.2E-06	HLA-DRB5, HLA-DRB4, CD3G, HLA-DRB3, HLA- DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_BP_ FAT	GO:0050852~T cell receptor signaling pathway	8	1.3E-05	HLA-DRB5, HLA-DRB4, CD3G, HLA-DRB3, HLA- DRB1, SKP1, HLA-DPA1, HLA-DQB1
	GOTERM_BP_ FAT	GO:0002381~immunoglobulin production involved in immunoglobulin mediated immune response	4	1.4E-05	HLA-DRB5, HLA-DRB4, HLA-DRB1, HLA-DQB1
	GOTERM_BP_ FAT	GO:0019882~antigen processing and presentation	6	2.2E-05	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_BP_ FAT	GO:0002455~humoral immune response mediated by circulating immunoglobulin	4	3.3E-05	HLA-DRB5, HLA-DRB4, HLA-DRB1, HLA-DQB1
	GOTERM_BP_ FAT	GO:0060333~interferon-gamma-mediated signaling pathway	6	5.7E-05	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1

(continued)

Table 1. Continued

Regulation	Category	Term	Gene count	Adjusted P value	Gene symbol
	GOTERM_BP_FAT	GO:0006955~immune response	10	5.7E-05	C3, CD96, HLA-DRB5, HLA-DRB4, CCL5, IGHV3-48, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_BP_FAT	GO:0019886~antigen processing and presentation of exogenous peptide antigen via MHC class II	6	5.2E-06	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_BP_FAT	GO:0042744~hydrogen peroxide catabolic process	4	6.1E-04	HBB, HBA2, HBA1, DUOX2
	GOTERM_BP_FAT	GO:2001179~regulation of interleukin-10 secretion	3	6.1E-04	HLA-DRB5, HLA-DRB4, HLA-DRB1
	GOTERM_CC_FAT	GO:0042613~MHC class II protein complex	6	2.2E-06	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_CC_FAT	GO:0071556~integral component of lumenal side of endoplasmic reticulum membrane	6	4.9E-06	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_CC_FA	GO:0030658~transport vesicle membrane	6	1.4E-06	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_CC_FAT	GO:0030669~clathrin-coated endocytic vesicle membrane	6	1.8E-05	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_CC_FAT	GO:0012507~ER to Golgi transport vesicle membrane	6	4.1E-05	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_CC_FAT	GO:0030666~endocytic vesicle membrane	6	1.1E-04	HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1
	GOTERM_CC_FAT	GO:0032588~trans-Golgi network membrane	6	3.1E-03	HLA-DRB5, HLA-DRB4, HLA-DRB3, HLA-DRB1, HLA-DPA1, HLA-DQB1

DEGs, differentially expressed genes.

Table 2. KEGG pathway analyses of the DEGs

Regulation	Pathway description	Gene count	Adjusted P value	Gene symbol
Up-regulated	hsa05146:Amoebiasis	6	3.8E-03	MUC2, SERPINB10, SERPINB2, SERPINB4, SERPINB13, SERPINB3
Down-regulated	hsa05150:Staphylococc us aureus infection	8	2.4E-08	HLA-DQB1, HLA-DRB1, C3, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1
	hsa05323:Rheumatoid arthritis	8	5.10E-07	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA- DRB4, HLA-DRB5, HLA-DPA1, CCL5, ATP6V0A4
	hsa05310:Asthma	6	1.20E-06	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1
	hsa05332:Graft-versus-host disease	6	1.50E-06	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA- DRB4, HLA-DRB5, HLA-DPA1
	hsa05140:Leishmaniasis	7	1.50E-06	HLA-DQB1, HLA-DRB1, C3, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1
	hsa05330:Allograft rejection	6	1.50E-06	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA- DRB4, HLA-DRB5, HLA-DPA1
	hsa05168:Herpes simplex infection	9	1.70E-06	HLA-DQB1, HLA-DRB1, C3, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1, SKP1, CCL5
	hsa04940:Type 1 diabetes mellitus	6	2.30E-06	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA- DRB4, HLA-DRB5, HLA-DPA1
	hsa05322:Systemic lupus erythematosus	8	2.40E-06	HLA-DQB1, HLA-DRB1, C3, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1, C2
	hsa04672:Intestinal immune network for IgA production	6	3.70E-06	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA- DRB4, HLA-DRB5, HLA-DPA1
	hsa04145:Phagosome	8	4.30E-06	HLA-DQB1, HLA-DRB1, C3, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1, ATP6V0A4
	hsa05320:Autoimmune thyroid disease	6	5.20E-06	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA- DRB5, HLA-DPA1
	hsa05416:Viral myocarditis	6	7.60E-06	HLA-DQB1, HLA-DRB1, C3, HLA-DRB3, HLA- DRB4, HLA-DRB5, HLA-DPA1
	hsa05152:Tuberculosis	8	9.00E-05	HLA-DQB1, HLA-DRB1, C3, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1, ATP6V0A4

Table 2. Continued

Regulation	Pathway description	Gene count	Adjusted P value	Gene symbol
	hsa05321:Inflammatory bowel disease (IBD)	6	1.20E-05	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1
	hsa04612:Antigen processing and presentation	6	2.60E-04	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1
	hsa05164:Influenza A	7	1.00E-04	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1, CCL5
	hsa05145:Toxoplasmosis	6	2.00E-03	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1
	hsa04514:Cell adhesion molecules (CAMs)	6	4.60E-03	HLA-DQB1, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1
	hsa05166:HTLV-I infection	7	7.70E-03	HLA-DQB1, CD3G, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DPA1

KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes.

HLA-DRB1, HLA-DRB5, MUC5B, and CD274, were validated (Figure 4).

Discussion

AR is the most common chronic disease in childhood. It can have a deleterious impact on quality of life, affecting sleep, cognitive and psychomotor function, and participation in social activities, and may impair learning.[1–4] Accumulating evidence suggests that genetic variation significantly increases the risk of disease.[12,16,17] In the present study, we used gene expression profile data to identify DEGs between patients with AR and healthy controls and a series of bioinformatics analyses to screen key genes and pathways associated with the development of AR.

KEGG pathway analysis found that genes down-regulated in patients with AR were enriched in *S. aureus* infection, rheumatoid arthritis, and allograft rejection. Many pathways are associated with the pathogenesis of allergic disease. The nasal carriage of *S. aureus* is an important risk factor for nosocomial and community-acquired infections,[20] and Refaat et al. showed that nasal *S. aureus* actively modulates the immune reaction in patients with persistent AR by promoting local IgE production; this has enabled early detection and treatment.[21] Major associations between common allergic diseases and incident rheumatoid arthritis were previously identified using Taiwan's National Health Insurance Research Database.[22] This is consistent with the hypothesis that allergic diseases and rheumatoid arthritis share a similar underlying etiologic pathway related to chronic inflammatory responses.[23,24] Considering that AR and asthma are both chronic heterogeneous disorders, with an overlapping epidemiology of prevalence, health care costs, and social costs, it is not surprising to identify
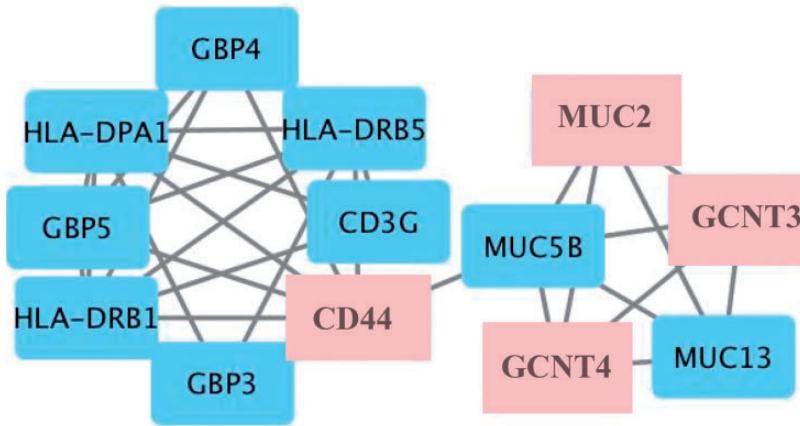
Figure 3. Highest module selected from the DEG PPI between patients with AR and healthy controls. Pink rectangles represent up-regulated genes while blue ones represent down-regulated genes
DEG, differentially expressed gene; PPI, protein–protein interaction; AR, allergic rhinitis.

Table 3. The seven hub nodes with the highest PPI network score

Gene symbol	Gene description	Degree	Genes that interact
CD44	CD44 molecule (Indian blood group)	15	*MUC5B, DPP4, CD36, CD274, HBB, CCL5, HBA1, TIMP1, CD69, GBP3, GBP5, GBP4, HLA-DPA1, HLA-DRB5, HLA-DRB1*
HLA-DPA1	Major histocompatibility complex, class II, DP alpha 1	7	*CD3G, HLA-DQB1, CD274, GBP3, HLA-DRB1, HLA-DRB5, CD44*
HLA-DRB1	Major histocompatibility complex, class II, DR beta 1	7	*HLA-DQB1, GBP3, CD274, CD3G, HLA-DPA1, CD44, HLA-DRB5*
HLA-DRB5	Major histocompatibility complex, class II, DR beta 5	7	*HLA-DQB1, CD274, CD3G, HLA- DPA1, CD44, GBP3, HLA-DRB1*
MUC5B	Mucin 5B, oligomeric mucus/ gel-forming	5	*CD44, GCNT4, GCNT3, MUC2, MUC13*
CD274	CD274 molecular	5	*HLA-DRB5, HLA-DRB1, CD3G, HLA-DPA1, CD44*
MUC2	Mucin 2, oligomeric mucus/ gel-forming	5	*TFF3, MUC5B, GCNT4, GCNT3, MUC13*

PPI, protein–protein interaction.

AR-associated DEGs that are enriched in the asthma pathway.[25–27]

Our PPI network and module analyses predicted seven hub genes, of which six were validated in gene expression profiles GSE44037 and GSE43523 (*CD44, HLA-DPA1, HLA-DRB1, HLA-DRB5,* *MUC5B,* and *CD274*). These key genes might play important roles in the development of AR and act as early diagnosis biomarkers or treatment targets of AR. Interestingly, the cell adhesion molecule CD44, which is involved in lymphocyte infiltration of inflamed tissues, showed the
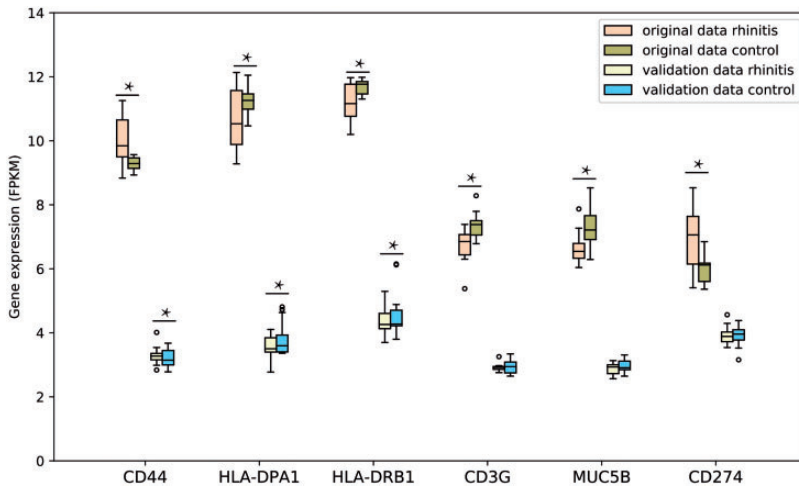
Figure 4. Expression of validated hub genes in original and validation datasets. GSE19187 is the original dataset, while the merger of GSE44037 and GSE43523 was used as the validation dataset. *, $P < 0.05$

highest connectivity degree in the PPI network.[28] CD44 was previously reported to be critical for the development of airway inflammation in a murine model of allergic asthma.[29] To our knowledge, this study is the first to identify an association between CD44 and the development of AR. Recently, a genome-wide association study found that genetic variants in the *HLA-DPA1/HLA-DPB1* locus are associated with the risk of pediatric asthma in Asian populations.[30] Human leukocyte antigen (HLA) genes and alleles are responsible for the display of cell surface proteins and regulation of the immune system,[31] and several studies showed them to be either risk factors or protective factors for allergies involving common allergens.[32–34] Yang et al. used sequence-specific primer PCR to show that two different *HLA-DRB1* alleles associate with AR in northeast China.[35] Moreover, variants of the *HLA-DRA/HLA-DRB5* inter-region were reported to be significant predictors of allergy to penicillins, but not to cephalosporins, suggesting complex gene–environment interactions in which genetic

susceptibility of HLA type 2 antigen presentation plays a central role.[36,37]

MUC5B is an evolutionarily conserved gene that encodes structurally related mucin glycoproteins, which are the principal macromolecules in airway mucus.[38,39] Ali et al. identified MUC5B as a major gel-forming mucin secreted in the airway,[40] while MUC5B was also demonstrated to alter the mucus gel and decrease mucus clearance, or contribute to mechanisms of airway eosinophilia because a deficiency in MUC5B-associated Siglec8 decreased eosinophil apoptosis in the airway lumen.[41] CD274, also known as programmed death ligand 1, emerged as an important immune modulator that can block T cell receptor signaling.[42]

The present study had some limitations. The sample size of our dataset was small, meaning that statistical power was low; we are attempting to collect more data to overcome this. Moreover, our results need to be validated by molecular analyses. Nevertheless, our findings lay the foundation for further study. Micro (mi)RNAs are be a fundamental regulator of gene

expression, with roles identified in numerous immunological and inflammatory disorders, including allergic inflammation.[43] Therefore, our future work will investigate the interactions of genes and miRNA profiles to shed new light on the molecular mechanisms involved in AR pathology.

In conclusion, we identified several key genes that are potentially associated with the development of AR using bioinformatics analyses of DEGs between patients with AR and healthy controls. These genes and their pathways will further our understanding of AR etiology, and help improve diagnosis, prevention, and intervention. Our findings suggest that the up-regulated expression of *CD44* and *CD274*, and the down-regulation of *HLA-DPA1*, *HLA-DRB1*, *MUC5B*, and *HLA-DRB5* can be considered candidate biomarkers or therapeutic targets for AR.

Declaration of conflicting interest

Funding

ORCID iD

Zheng-min Xu https://orcid.org/0000-0002-3505-1866

References

1. Rotiroti G and Scadding GK. Allergic rhinitis- an overview of a common disease. *Paediatrics and Child Health* 2016; 26: 298–303.

2. Kemp AS. Allergic rhinitis. *Paediatr Respir Rev* 2009; 10: 63–68.

3. Szopa W, Burley TA, Kramer-Marek G, et al. Diagnostic and therapeutic biomarkers in glioblastoma: current status and future perspectives. *Biomed Res Int* 2017; 2017: 8013575.

4. Zhang Y and Zhang L. Prevalence of allergic rhinitis in China. *Allergy Asthma Immunol Res* 2014; 6: 105–113.

5. Zou QY, Shen Y, Ke X, et al. Exposure to air pollution and risk of prevalence of childhood allergic rhinitis - A meta-analysis. *Int J Pediatr Otorhinolaryngol* 2018; 112: 82–90.

6. Wilson DR, Torres Lima M and Durham SR. Sublingual immunotherapy for allergic rhinitis: systematic review and meta-analysis. *Allergy* 2005; 60: 4–12.

7. Watson CT, Cohain AT, Griffin RS, et al. Integrative transcriptomic analysis reveals key drivers of acute peanut allergic reactions. *Nat Commun* 2017; 8: 1943.

8. Wang Y, Qi S, Zhu R, et al. The need for clinical indicators in allergic rhinitis. *Revue Française dAllergologie* 2015; 55: 526–529.

9. Suojalehto H, Lindström I, Majuri ML, et al. Altered microRNA expression of nasal mucosa in long-term asthma and allergic Rhinitis. *Int Arch Allergy Immunol* 2014; 163: 168–178.

10. Shiue HS, Lee YS, Tsai CN, et al. DNA microarray analysis of the effect on inflammation in patients treated with acupuncture for allergic rhinitis. *J Altern Complement Med* 2008; 14: 689–698.

11. Seumois G, Zapardiel-Gonzalo J, White B, et al. Transcriptional profiling of Th2 cells identifies pathogenic features associated with asthma. *J Immunol* 2016; 197: 655–664.

12. Waage J, Standl M, Curtin JA, et al. Genome-wide association and HLA fine-mapping studies identify risk loci and genetic pathways underlying allergic rhinitis. *Nat Genet* 2018; 50: 1072–1080.

13. Imoto Y, Tokunaga T, Matsumoto Y, et al. Cystatin SN upregulation in patients with seasonal allergic rhinitis. *PLoS One* 2013; 8: e67057.

14. Ma Y, Shi L and Zheng C. Microarray analysis of lncRNA and mRNA expression profiles in mice with allergic rhinitis. *Int J Pediatr Otorhinolaryngol* 2018; 104: 58–65.

15. Ma Z, Teng Y, Liu X, et al. Identification and functional profiling of differentially expressed long non-coding RNAs in nasal

mucosa with allergic rhinitis. *Tohoku J Exp Med* 2017; 242: 143–150.

16. Li C, Luo L, Wei S, et al. Identification of the potential crucial genes in invasive ductal carcinoma using bioinformatics analysis. *Oncotarget* 2018; 9: 6800–6813.

17. Song Y and Yan Z. Exploring of the molecular mechanism of rhinitis via bioinformatics methods. *Mol Med Rep* 2017: 1–7.

18. Giovannini-Chami L, Marcet B, Moreilhon C, et al. Distinct epithelial gene expression phenotypes in childhood respiratory allergy. *Eur Respir J* 2012; 39: 1197–1205.

19. Wagener AH, Zwinderman AH, Luiten S, et al. The impact of allergic rhinitis and asthma on human nasal and bronchial epithelial gene expression. *PLoS One* 2013; 8: e80257.

20. Baysoy G, Arslan S, Karabay O, et al. Nasal carriage of *Staphylococcus* aureus in children with allergic rhinitis and the effect of intranasal fluticasone propionate treatment on carriage status. *Int J Pediatr Otorhinolaryngol* 2007; 71: 205–209.

21. Damm M, Quante G, Jurk T, et al. Nasal colonization with *Staphylococcus aureus* is not associated with the severity of symptoms or the extent of the disease in chronic rhinosinusitis. *Otolaryngol Head Neck Surg* 2016; 131: 200–206.

22. Jeong HE, Jung SM and Cho SI. Association between rheumatoid arthritis and respiratory allergic diseases in Korean adults: a propensity score matched case-control study. *Int J Rheumatol* 2018; 2018: 3798124.

23. Olsson AR, Wingren G, Skogh T, et al. Allergic manifestations in patients with rheumatoid arthritis. *APMIS* 2003; 111: 940–944.

24. Verhoef CM, Van Roon JAG, Vianen ME, et al. Mutual antagonism of rheumatoid arthritis and hay fever; a role for type 1/type 2 T cell balance. *Ann Rheum Dis* 1998; 57: 275–280.

25. Bergeron C and Hamid Q. Relationship between asthma and rhinitis: epidemiologic, pathophysiologic, and therapeutic aspects. *Allergy Asthma Clin Immunol* 2005; 1: 81–87.

26. Kim H, Bouchard J and Renzi PM. The link between allergic rhinitis and asthma: A role for antileukotrienes? *Can Respir J* 2008; 15: 91–98.

27. Egan M and Bunyavanich S. Allergic rhinitis: the "Ghost Diagnosis" in patients with asthma. *Asthma Res Pract* 2015; 1: 8.

28. Katoh S, Kaminuma O, Hiroi T, 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: 3198–3207.

29. Katoh S, Maeda S, Fukuoka H, et al. A crucial role of sialidase Neu1 in hyaluronan receptor function of CD44 in T helper type 2-mediated airway inflammation of murine acute asthmatic model. *Clin Exp Immunol* 2010; 161: 233–241.

30. Noguchi E, Sakamoto H, Hirota T, et al. Genome-wide association study Identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* 2011; 7: e1002170.

31. Agarwal R. Severe asthma with fungal sensitization. *Curr Allergy Asthma Rep* 2011; 11: 403–413.

32. Ramasamy A, Curjuric I, Coin LJ, et al. A genome-wide meta-analysis of genetic variants associated with allergic rhinitis and grass sensitization and their interaction with birth order. *J Allergy Clin Immunol* 2011; 128: 996–1005.

33. Portelli MA, Hodge E and Sayers I. Genetic risk factors for the development of allergic disease identified by genome-wide association. *Clin Exp Allergy* 2014; 45: 21–31.

34. Muro M, Mondejar-López P, Moya-Quiles MR, et al. HLA-DRB1 and HLA-DQB1 genes on susceptibility to and protection from allergic bronchopulmonary aspergillosis in patients with cystic fibrosis. *Microbiol Immunol* 2013; 57: 193–197.

35. Park H, Ahn K, Park MH, et al. The HLA-DRB1 polymorphism is associated with atopic dermatitis, but not egg allergy in Korean children. *Allergy Asthma Immunol Res* 2012; 4: 143–147.

36. Guéant JL, Romano A, Cornejo-Garcia JA, et al. HLA-DRA variants predict penicillin allergy in genome-wide fine-mapping

genotyping. *J Allergy Clin Immunol* 2015;
135: 253–259.

37. Khor S-S, Morino R, Nakazono K, et al.
Genome-wide association study of self-
reported food reactions in Japanese identi-
fies shrimp and peach specific loci in the
HLA-DR/DQ gene region. *Sci Rep* 2018;
8: 1069.

38. Welsh KG, Rousseau K, Fisher G, et al.
MUC5AC and a glycosylated variant of
MUC5B alter mucin composition in children
with acute asthma. *Chest* 2017; 152: 771–779.

39. Kim DH, Chu HS, Lee JY, et al. Up-
regulation of MUC5AC and MUC5B
mucin genes in chronic rhinosinusitis. *Arch
Otolaryngol Head Neck Surg* 2004; 130:
747–752.

40. Ali MS and Pearson JP. Upper airway
mucin gene expression. *Laryngoscope* 2009;
117: 932–938.

41. Roy MG, Livraghi-Butrico A, Fletcher AA,
et al. Muc5b is required for airway defence.
Nature 2014; 505: 412–416.

42. Kubo S, Yamada T, Osawa Y, et al.
Cytosine-phosphate-guanosine-DNA indu-
ces CD274 expression in human B cells
and suppresses T helper type 2 cytokine
production in pollen antigen-stimulated
CD4-positive cells. *Clin Exp Immunol* 2012;
169: 1–9.

43. Specjalski K and Jassem E. MicroRNAs:
potential biomarkers and targets of therapy
in allergic diseases? *Arch Immunol Ther Exp
(Warsz)* 2019; 67: 213–223.