Research Article

Investigation of Several Biomarkers Associated with Asthmatic Chronic Rhinosinusitis with Nasal Polyps

Zhili Zhang,[1,*] Xia Yan,[2] and Jingqin Jiang[1]

[1]Department of Stomatology, Dongying People's Hospital, Dongying, 257091, Shandong Province, P.R. China
[2]Department of Pediatrics, Dongying People's Hospital, Dongying, 257091, Shandong Province, P.R. China

[*]*Corresponding author*: Zhili Zhang, Department of Stomatology, Dongying People's Hospital, No. 317 Dongchengnanyi Road, Dongying 257091, Shandong Province, P.R. China. Tel: +86-5468901503, E-mail: zhangzhili1139@sina.com

Abstract

Background: Asthmatic chronic rhinosinusitis with nasal polyps (aCRSwNP) is a common disruptive eosinophilic disease. However, up to now, there is no effective medical treatment for the disease, which is partly due to that the molecular mechanism of aCRSwNP is still unknown.

Objectives: The aim of this study was to facilitate the systematic discovery of diagnostic biomarkers of aCRSwNP based on integrating pathways, differentially expressed genes (DEGs), and mutual information networks (MINs).

Methods: This was a foundation-application study carried out in Dongying, Shandong Province, P.R. China, in 2016. First, the gene expression profile of aCRSwNP composed of 13 normal samples and 21 aCRSwNP samples was recruited from the gene expression omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) and then, data preprocessing was performed. Second, the attract method was utilized to identify differential pathways. In the following, MINs were constructed and underwent topological analysis. Then, DEGs were examined in aCRSwNP group and normal control group to identify significant genes and key genes. Finally, the support vector machine (SVM) with C-classification was utilized to evaluate the performance of the classification.

Results: A total of 11,100 genes and 273 pathways (gene count > 5) were initially obtained. Then, 5 differential pathways which contained 346 genes were identified. Topological analysis conducted on the MINs revealed 20 hub genes (degree centrality $\geq$ 220). In the following, 795 DEGs were identified ($|\log$ fold change (FC)$| \geq 2.0$, P value ≤ 0.01). Furthermore, 35 significant genes and 14 key genes were detected. Finally, the results of SVM with C-classification indicated that the key genes gave the best result.

Conclusions: Our research identified several key genes (such as IL6R), which might play key roles in the occurrence and development of aCRSwNP. We predicted that these genes might provide additional diagnostic and therapeutic targets for aCRSwNP.

Keywords: Gene Network, Network Meta-Analysis, Metabolic Pathways

1. Background

It is known that both of chronic rhinosinusitis with nasal polyposis (CRSwNP) and asthma are complex inflammatory disorders and are therapeutic challenges for health care system (1). There are approximately 30% with asthma and 15% with aspirin intolerance in patients with CRSwNP (2). Asthmatic chronic rhinosinusitis with nasal polyps (aCRSwNP) is a common disruptive eosinophilic disease with no effective medical treatment (3). This is mainly due to that the molecular mechanism of aCRSwNP is still unknown.

It is also well known that the development and progression of a certain disease are related to accumulated molecular genetics or genomic changes (4). To predict the outcome of certain treatments or classify diseases into subtypes, the gene-expression microarrays can provide a tool to genetically profile diseases (5). The differential expression of several genes has been confirmed to be associated with aspirin-sensitive aCRSwNP and several genes and

gene sets have been indicated to be implicated in the earlier stages of eosinophilic inflammation (6). In spite of expanded efforts to study the genetic bases of aCRSwNP, the exact genes that play key roles in the development and progression are still unknown.

Identification of differentially expressed genes (DEGs) via gene expression analysis across various cell cycle states, biological conditions, subjects, and tissues may help select potential biomarkers for diseases (7). However, as the gene size gets larger, the probability of a false identification may increase (8). The cross validation of datasets can significantly reduce those false findings and increase sensitivity (9). At present, along with the development of high-throughput testing technology, a large-scale of protein interactions has been accumulated. However, there is still a large amount of significant interactions such as key genes in significant pathways that have not been tested (10).

2. Objectives

In the present study, to reveal the molecular mechanism of aCRSwNP, the pathways, DEGs, and mutual information networks (MINs) were integrated together to conducted analysis on the aCRSwNP. Our work demonstrated a practical framework for complex disease candidate biomarker analysis at a comprehensive level; the results might provide an effective avenue to combat this complicated illness and this strategy could be applied to other complex diseases.

3. Methods

This was a foundation-application study carried out in Dongying, Shandong Province, P.R. China, in 2016. In the present study, analysis was conducted on the gene expression of aCRSwNP to identify several biomarkers associated with aCRSwNP. The schematic diagram of analysis procedure is shown in Figure 1.

3.1. Data Recruiting

It is well known that the gene expression omnibus (GEO) at the national center for biotechnology information (NCBI) (http://www.ncbi.nlm.nih.gov/geo/), which is located in Bethesda, MD, USA, is the largest fully public repository for high-throughput molecular abundance data, primarily gene expression data (11). Up to now, there are a total of 4348 datasets in the database. In the present study, under the filter conditions of "Homo sapiens", "transcription profiling by array", and "neither too large nor too small assays samples", the gene expression profile of aCRSwNP with accessing number of GSE23552 (6) was obtained from GEO database. The data of GSE23552, existed on A-AFFY-143 - Affymetrix GeneChip Human Exon 1.0 ST Array version 1, [HuEx-1_0-st-v1] platform, were composed of 13 normal samples and 21 aCRSwNP samples. We downloaded all of the annotation files and microarray data.

3.2. Data Preprocessing

Prior to analysis, data preprocessing was performed to control the quality of gene expression data. First of all, the robust multi-array average (RMA) method was applied to conduct background correction to eliminate the influence of nonspecific hybridization (12). The observed Perfect match (PM) probes were modeled as the sum of a normal noise component N (Normal with mean μ and variance σ^2) and an exponential signal component S (exponential with mean α). The normal was truncated at zero, so as to avoid any possibility of negatives. Given that there was A observed intensity, we adjusted the data according to the following formula.

$$E\left(mA = a\right) = i + j\,\frac{\varnothing\left(\frac{i}{j}\right) - \varnothing\left(\frac{a-i}{j}\right)}{\Phi\left(\frac{i}{j}\right) + \Phi\left(\frac{a-i}{j}\right) - 1}$$

Where i = m-μ-$\sigma^2\alpha$, and j = σ. $\emptyset$ and Φ represented the normal distribution density and distribution functions, respectively.

Then, the quantiles based algorithm was performed to conduct mismatch (MM) correct analysis of normalization (13). The transformation was determined by the following formula:

$$g_n = Y^{-1}\left(T\left(g_n\right)\right)$$

Where the value of T was estimated via each array's empirical distribution and Y was determined via the averaged sample quantiles' empirical distribution.

In the following, the "mas" method was used to carry out PM/MM correction. An ideal mismatch was subtracted from PM for that the ideal MM would always be less than the corresponding PM.

The summarization method was "median polish" (12). A multichip linear model was fit to data from each probe set. In particular, for a probe set r with s = 1, ..., Sr probes and data from q = 1,..., Q arrays, we fitted the following model

$$log_2\left(PM_{sq}^r\right) = \alpha_s^r + \beta_q^r + \epsilon_{sq}^r$$

Where α_s was a probe effect and β_q was the log$_2$ expression value.

Finally, the gene expression profile on probe level was converted into gene symbol level, and the duplicated symbols were wiped off, which resulted in a total of 11,100 gene symbols to be obtained.

3.3. Identifying Differential Pathways

In order to gain further insights into the function enrichment of the genes of the aCRSwNP, differential pathways analysis was conducted on the gene expression profile. In this study, there were two steps for identifying differential pathways. First of all, the database for annotation, visualization and integrated discovery (DAVID) (14) was used to perform kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analysis of the genes, so as to mine the biochemistry pathways which might be involved during the progress of the aCRSwNP. The pathways whose gene count was > 5 were selected for further analysis.

Then, the attract method, a knowledge-driven analytical approach that could best discriminate the conditions between cell phenotypes to identify and annotate the gene-sets (15), was used to determine the differential pathways in the present research. GSEA-ANOVA, an analysis of variance-based implementation of a gene set enrichment
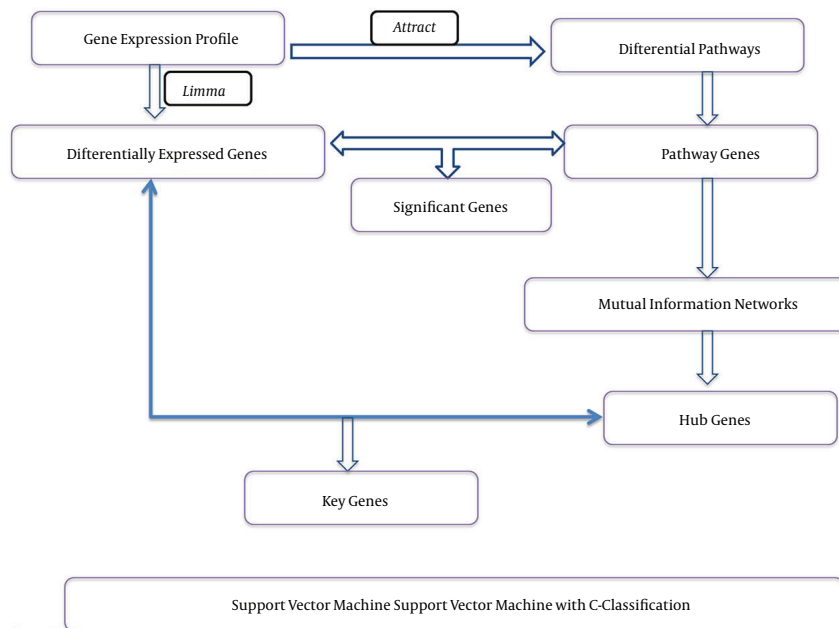
Figure 1. The Schematic Diagram of the Analysis Procedure

algorithm, was utilized to test pathway-level data, so as to identify the values of the F-statistic (15). The larger the F-statistic values, the stronger the associations were, hence a small F-statistic suggested that the gene demonstrated minimal cell type-specific expression changes. To test this relationship more formally, the T-test and a Welch modification were used to adjust the P value, and the adjusted P-value was set in ascending order; the top 5 pathways were regarded as differential pathways. Meanwhile, the genes enriched in the differential pathways were considered as pathway genes.

3.4. Network Construction and Topological Analysis

3.4.1. Network Construction

In the present study, MINs, a subcategory of network inference approaches, were constructed to further disclose the relationships among the pathway genes (16). The fundamental principle of this group of methods is to deduce a link among a series of nodes in case of that there is a high score based on mutual information (MI) (17). In the present study, MIN construction for pathway genes comprised three steps.

First of all, the mutual information matrix (MIM) was calculated. MIM is a square matrix in which, a, b-th elements take the MI between the random genes Xa and Xb, and q is a probability measure.

$$MIM_{ab} = I\left(X_a; X_b\right) = \sum_{a,b} q\left(x_a, y_b\right) log \frac{q\left(x_a, y_b\right)}{q\left(x_a\right) q\left(y_b\right)}$$

Second, an edge score for each pair of nodes was computed by CLR algorithm. CLR algorithm is an extension of the relevance network approach (18). It computed the MI for each pair of genes and derived a score related to the empirical distribution of the MI values (16). Specially, it mainly considered the edge score instead of taking into account the information $I\left(X_a; X_b\right)$ between genes X_a and X_b:

$$z_{ab} = \sqrt{z_a^2 + z_b^2}$$

Of which

$$z_a = max\left(0, \frac{I\left(X_a; X_b - \mu_a\right)}{\sigma_a}\right)$$

Where μ_a and σ_a respectively represented the sample mean and standard deviation of the empirical distribution of the values $I\left(X_a; X_b\right)$.

The last step was inputting the genes and edge scores into the igraph software package (19) to visualize the MIN.

3.4.2. Topological Analysis of MIN

As is well known, the role of a node in a network structure not only is in association with the node itself, but also depends on the adjacent nodes that are in connection together and the topological structure of the network

(20). Further to investigate biological functions and significance of nodes in the network, the indices of topological analysis (degree (21), closeness (22), and betweenness (23)) are often characterized. Degree centrality was always taken into account because it is the simplest topological index for researchers to calculate. The more nodes that a certain gene connected with, the more important the gene was. The genes with high connections were called hub genes. In this research, the genes with degree centrality $\geq$ 220 were considered as hub genes.

3.5. Identifying DEGs

In the last few years, to gain a snapshot of transcriptional activity in different tissues or populations of cells gene, some expression technologies have been frequently used in molecular biology research (24). Limma, an R/Bioconductor software package which can provide an integrated solution for analyzing data from gene expression experiments, was utilized to determine the differential expression of the aCRSwNP group and normal control group (25). In this study, empirical Bayes method (F test) (26) was used for genes with scores greater than an adjustable threshold to differentiate DEGs between aCRSwNP group and normal control group. The false discovery rate (FDR) was employed to proofread the P-values. These genes that met the threshold values of |log fold change (FC)| $\geq$ 2.0, as well as P-value $\leq$ 0.01, were considered as DEGs.

3.6. Identifying Significant Genes and Key Genes

Moreover, to further disclose the differential expression of pathway genes and the hub genes, the intersection of DEGs and either pathway genes or hub genes were separately analyzed. The intersection of DEGs and pathway genes was regarded as significant genes and the intersection of DEGs and hub genes was regarded as key genes.

3.7. Classification and Evaluation

In this paper, there were five kinds of gene data, including pathway genes, hub genes, DEGs, significant genes, and key genes. To determine the stand or fall of classification results, the support vector machine (SVM) (27) with C-classification was utilized to evaluate the performance of the classification. To this end, all of the samples were initially put together and divided randomly in 21:13; the 21 samples were chosen as the experimental group and the 13 samples were selected as the control group. Next, a 5-fold cross-validation was conducted on the train set to evaluate the potential classification strength of the models, and then we estimated its prediction on a separate test set. In order to evaluate the classification results, several measures were employed to provide different insights.

These measures included the area under the receiver operating characteristics curve or in short AUC, true positive rate (TPR), true negative rate (TNR), the Matthews coefficient correlation classification (MCC), and Youden's index (J) calculated using the following formula: J = maximum [sensitivity(c) + specificity(c) - 1] (28). The combination of these measures gave us an adequate overview of the classification performance.

4. Results

4.1. Identifying Differential Pathways

In the present study, the attract method was used to determine the differential pathways. Based on what had been indicated in the method we used two steps of identifying differential pathways analysis in this research. First, the significant dysregulated KEGG pathways for aCRSwNP were analyzed based on DAVID. As we set the threshold of gene count > 5, 273 pathways were obtained. Then, the attract method was applied to analyze these pathways, so as to identify differential pathways. As aforementioned in the methods, the top 5 pathways based on the adjusted P values were regarded as differential pathways. Therefore, five differential pathways were identified as we ranked the pathways in descending order according to their adjusted P values, which were cell adhesion molecules (CAMs), hematopoietic cell lineage, intestinal immune network for IgA production, olfactory transduction, and allograft rejection, in sequence. Meanwhile, there were 346 genes enriched in the differential pathways.

4.2. Network Construction and Topological Analysis

To further disclose the relationships between the pathway genes, MIN was constructed. The MIN construction for pathway genes was visualized by inputting the genes and edge scores obtained via CLR algorithm into the igraph software. All of the 346 pathway genes constituted a MIN relationship. As topological analysis was conducted on the MIN, and threshold values of degree centrality of the genes set at $\geq$ 220, we obtained 20 hub genes, as listed in Table 1.

4.3. Identifying DEGs

The diseases are always associated with the DEGs in the disease conditions compared to normal conditions. To determine the DEGs in the aCRSwNP group and normal control group, limma was utilized. Under the threshold values of |log FC| $\geq$ 2.0, as well as P value $\leq$ 0.01, we obtained a total of 795 DEGs, which 337 genes were down-regulated and 458 genes were up-regulated.

Table 1. The Degree Centrality of the Hub Genes

Gene Symbol	Degree Centrality	Gene Symbol	Degree Centrality
CD4	303.61	CALM2	249.64
ITGB2	294.455	IL2RA	248.265
HLA - DMB	288.155	CD1A	241.36
CSF1R	286.96	CD6	239.535
CD3D	277.275	ADRBK2	237.095
CD1C	270.765	IL6R	229.445
PTPRC	266.635	CSF3R	229.17
CD3E	264.69	SELL	228.59
CD1E	259.19	TIGIT	227.69
ARRB2	257.71	CD28	225.435

4.4. Identifying Significant Genes and Key Genes

As the intersection of the pathway genes and the DEGs was conducted, we gained 35 significant genes, which were as following: CD80, HLA-DMA, HLA-DMB, IL2RA, IL10, CD1A, CD1B, CD1C, CD1E, CCL28, CD6, CD3E, ITGB2, FCER2, PTPRC, IL6R, CSF3R, ITGA1, ITGA2, CLDN11, PECAM1, SELP, ITGA9, SELL, CDH2, NTNG2, NRXN1, NLGN4X, CADM3, NFASC, OR51I1, PRKX, PRKG2, ADRBK2 and ARRB2. Furthermore, 14 key genes including: CD6, IL2RA, CSF3R, PTPRC, ARRB2, SELL, IL6R, ITGB2, CD3E, CD1A, CD1C, CD1E, HLA-DMB, ADRBK2 were identified via taking the intersection of the hub genes and the DEGs.

4.5. Classification and Evaluation

In order to effectively determine whether evaluated test methods were appropriate or not, SVM with C-classification was utilized as a tool for evaluating the performance of the classification. As is observed in the receiver operating characteristic (ROC) curve (Figure 3), the areas of the three components including key genes, significant genes, and DEGs were the same, while the areas of the pathway genes and the hub genes were the same. In addition, we can see from Table 2 that the result of the key genes was the best by giving AUC = 0.94, Accuracy = 95, TNR = 1, TPR = 0.92, MCC = 0.96 and Youden's index = 0.92; hence, we can conclude that the key genes best characterized the disease of aCRSwNP. MIN analysis on the key genes and pathway genes indicated that there were 343 genes and 2872 interactions in the network (Figure 2). The yellow nodes stood for 14 key genes, while it was obvious that all of the key genes were well clustered in the central location of the network, indicating that these 14 genes might play key roles in the occurrence and development of aCRSwNP.

5. Discussion

In this research, by applying the SVM method for analysis of the five types of genes data (DEGs, pathway genes, hub genes, significant genes, and key genes) obtained in the present study, we found that the SVM result was better for the key genes than the others. In other words, the key genes could better distinguish between the aCRSwNP group and normal control group, implying that the key genes could better characterize the disease of aCRSwNP than the others. The key genes were identified by taking the intersection of the DEGs and the hub genes contained in the MIN. There were 14 key genes which were: CD6, IL2RA, CSF3R, PTPRC, ARRB2, SELL, IL6R, ITGB2, CD3E, CD1A, CD1C, CD1E, HLA-DMB, and ADRBK2.

Further to disclose the relationship of these genes with aCRSwNP, we discussed some of these genes in relation to the disease. For example in the case of gene IL6R, known as human interleukin-6 (IL-6), it is a very important proinflammatory cytokine that plays a critical role in the inflammatory response. It has been indicated that IL-6 is implicated in lots of diseases under a series of inflammatory conditions, such as cardiovascular disease, inflammatory arthritis, and inflammatory bowel disease (29-31). Research to elevate IL-6 and sIL-6R has indicated that IL-6 plays a pathogenic role during the progress of chronic rhinosinusitis (CRS). Also, the expression of IL-6 in CRS has been demonstrated to be increased via applying the RT-PCR methods and immunohistochemistry techniques (32, 33). Meanwhile, it has been shown by comparing the middle turbinate with CRSwNP that the level of IL-6 protein in polyp tissue was higher (34, 35). Moreover, IL6R has been identified as a new locus with a genome-wide significant association with asthma risk by Ferreira et al. (36). In the present study, by integrating the DEGs, pathways, and net-

Figure 2. Mutual Information Network (MIN) Analysis on the Key Genes and Pathway Genes
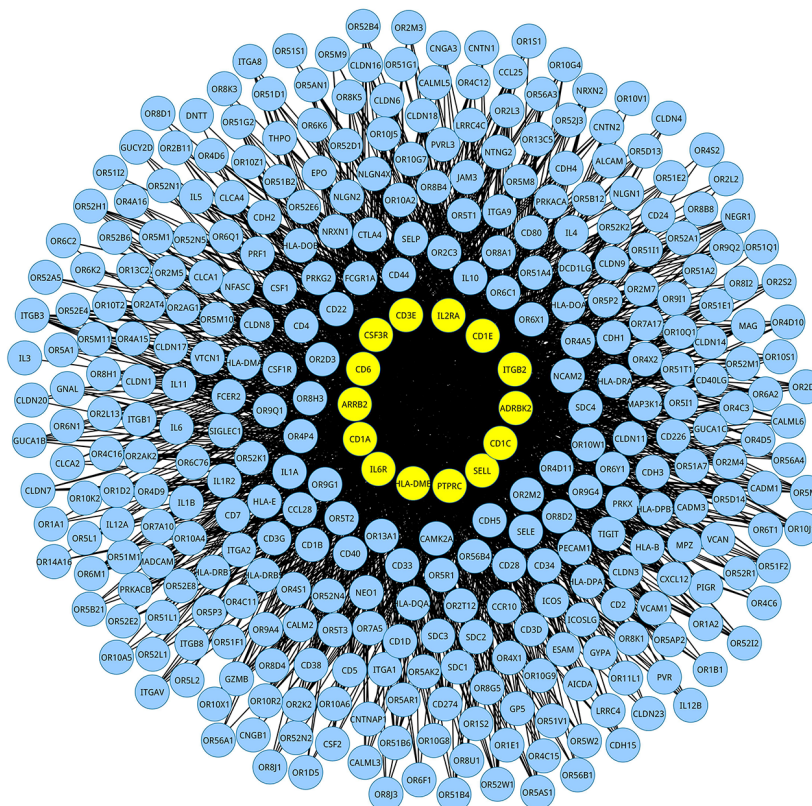
There were 343 genes and 2872 interactions in the network, where nodes represented genes, and edges were the interactions between two genes. The yellow nodes stood for 14 key genes: CD6, IL2RA, CSF3R, PTPRC, ARRB2, SELL, IL6R, ITGB2, CD3E, CD1A, CD1C, CD1E, HLA-DMB, ADRBK2, which were identified via taking the intersection of the hub genes and the DEGs.
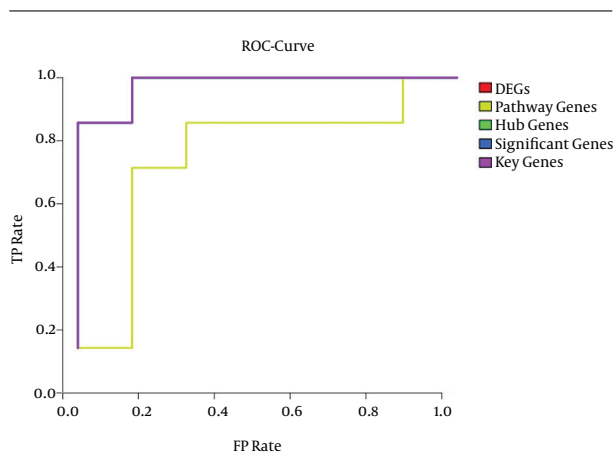

Figure 3. The Receiver Operating Characteristic (ROC) Curve of the Key Genes, Significant Genes and DEGs, Pathway Genes and the Hub Genes

work analysis, we identified that gene IL6R was one of the 14 key genes for aCRSwNP. Based on the abovementioned data reported in literature, we inferred that there was a close relationship between gene IL6R and aCRSwNP.

In addition, some other key genes were also identified to be key genes for aCRSwNP based on the analytical method used in the research. However, there was no direct relationship between these genes and aCRSwNP according to literature. However, there were still some relevance between these genes and aCRSwNP. In the case of gene IL2RA for another example, it has been reported that the T cells produce two distinct surface receptors for interleukin 2 (IL2). One of them is IL2RA that is released into peripheral blood following T cell activation (37). Bachert and colleagues indicated that lymphoid follicle-like structures including B cells, T cells, and plasma cells were associated with CRSwNP (38). In this case, we inferred that there might be some relationships between gene IL6R and aCRSwNP although this hypothesis needs further experimental analy-

Table 2. The Support Vector Machine (SVM) Results of the Five Types of the Genes in the Present Study

Genotype	AUC	Accuracy	MCC	Specificity	Sensitivity	Youden's Index
DEGs	0.94	89	0.85	1	0.83	0.83
Pathway genes	0.89	79	0.58	0.83	0.75	0.58
Hub genes	0.89	93	0.83	1	0.84	0.84
Significant genes	0.94	93	0.87	1	0.88	0.88
Key genes	0.94	95	0.96	1	0.92	0.92

ses to be validated.

5.1. Limitations and Conclusions

This was the first study analyzing aCRSwNP based on integrating pathways, DEGs, and MINs. We successfully identified 14 key genes that might play key roles in the occurrence and development of aCRSwNP. However, there were still several limitations in our work that must be taken into account. For example, all of the data were obtained from databases, and these data might be unstable. Also, the sample size was not enough and the results obtained by bioinformatics method were not verified via animal experiments. In spite of disadvantages, we believe that this method and the predicted key genes provide investigators with valuable resources to not only better understand the mechanisms of aCRSwNP, but also to detect potential biomarkers for early diagnosis and therapy of aCRSwNP. Moreover, this method of analysis might be employed in other related analyses.

In a word, our research identified several key genes (such as IL6R), which might play key roles in the occurrence and development of aCRSwNP. These genes might provide additional diagnostic and therapeutic targets for aCRSwNP.

Acknowledgments

Footnote

Conflict of Interest: We declare that we have no conflict of interest.

References

1. Rix I, Hakansson K, Larsen CG, Frendo M, von Buchwald C. Management of chronic rhinosinusitis with nasal polyps and coexisting asthma: A systematic review. *Am J Rhinol Allergy.* 2015;**29**(3):193–201. doi: 10.2500/ajra.2015.29.4178. [PubMed: 25975250].
2. Gevaert P, Calus L, Van Zele T, Blomme K, De Ruyck N, Bauters W, et al. Omalizumab is effective in allergic and nonallergic patients with nasal polyps and asthma. *J Allergy Clin Immunol.* 2013;**131**(1):110–6 e1. doi: 10.1016/j.jaci.2012.07.047. [PubMed: 23021878].
3. Ho J, Bailey M, Zaunders J, Mrad N, Sacks R, Sewell W, et al. Group 2 innate lymphoid cells (ILC2s) are increased in chronic rhinosinusitis with nasal polyps or eosinophilia. *Clin Exp Allergy.* 2015;**45**(2):394–403. doi: 10.1111/cea.12462. [PubMed: 25429730].
4. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature.* 1998;**396**(6712):643–9. doi: 10.1038/25292. [PubMed: 9872311].
5. Amin SB, Yip WK, Minvielle S, Broyl A, Li Y, Hanlon B, et al. Gene expression profile alone is inadequate in predicting complete response in multiple myeloma. *Leukemia.* 2014;**28**(11):2229–34. doi: 10.1038/leu.2014.140. [PubMed: 24732597].
6. Plager DA, Kahl JC, Asmann YW, Nilson AE, Pallanch JF, Friedman O, et al. Gene transcription changes in asthmatic chronic rhinosinusitis with nasal polyps and comparison to those in atopic dermatitis. *PLoS One.* 2010;**5**(7):11450. doi: 10.1371/journal.pone.0011450. [PubMed: 20625511].
7. Ibanez K, Boullosa C, Tabares-Seisdedos R, Baudot A, Valencia A. Molecular evidence for the inverse comorbidity between central nervous system disorders and cancers detected by transcriptomic meta-analyses. *PLoS Genet.* 2014;**10**(2):1004173. doi: 10.1371/journal.pgen.1004173. [PubMed: 24586201].
8. Christie MR, Tennessen JA, Blouin MS. Bayesian parentage analysis with systematic accountability of genotyping error, missing data and false matching. *Bioinformatics.* 2013;**29**(6):725–32. doi: 10.1093/bioinformatics/btt039. [PubMed: 23365409].
9. McGibbon RT, Pande VS. Variational cross-validation of slow dynamical modes in molecular kinetics. *J Chem Phys.* 2015;**142**(12):124105. doi: 10.1063/1.4916292. [PubMed: 25833563].
10. Nibbe RK, Chowdhury SA, Koyuturk M, Ewing R, Chance MR. Protein-protein interaction networks and subnetworks in the biology of disease. *Wiley Interdiscip Rev Syst Biol Med.* 2011;**3**(3):357–67. doi: 10.1002/wsbm.121. [PubMed: 20865778].
11. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics.* 2007;**23**(14):1846–7. doi: 10.1093/bioinformatics/btm254. [PubMed: 17496320].
12. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* 2003;**31**(4):15. [PubMed: 12582260].
13. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics.* 2003;**19**(2):185–93. [PubMed: 12538238].

14. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.* 2007;**8**(9):R183. doi: 10.1186/gb-2007-8-9-r183. [PubMed: 17784955].

15. Mar JC, Matigian NA, Quackenbush J, Wells CA. attract: A method for identifying core pathways that define cellular phenotypes. *PLoS One.* 2011;**6**(10):25445. doi: 10.1371/journal.pone.0025445. [PubMed: 22022396].

16. Meyer PE, Lafitte F, Bontempi G. minet: A R/Bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinformatics.* 2008;**9**:461. doi: 10.1186/1471-2105-9-461. [PubMed: 18959772].

17. Meyer PE, Kontos K, Lafitte F, Bontempi G. Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J Bioinform Syst Biol.* 2007:79879. doi: 10.1155/2007/79879. [PubMed: 18354736].

18. Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, et al. Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* 2007;**5**(1):8. doi: 10.1371/journal.pbio.0050008. [PubMed: 17214507].

19. Csardi G, Nepusz T. Igraph-package. 2008

20. Koschutzki D, Schreiber F. Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene Regul Syst Bio.* 2008;**2**:193–201. [PubMed: 19787083].

21. Haythornthwaite C. Social network analysis: An approach and technique for the study of information exchange. *Library Information Sci Res.* 1996;**18**(4):323–42.

22. Wolfe AW. Social network analysis: Methods and applications. *American Ethnol.* 1997;**24**(1):219–20. doi: 10.1525/ae.1997.24.1.219.

23. Barthlemy M. Betweenness centrality in large complex networks. *The European Physical Journal B - Condensed Matter.* 2004;**38**(2):163–8. doi: 10.1140/epjb/e2004-00111-4.

24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;**43**(7):47. doi: 10.1093/nar/gkv007. [PubMed: 25605792].

25. Smyth GK. Limma: Linear models for microarray data. 2005 :397–420. doi: 10.1007/0-387-29362-0_23.

26. Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, et al. A comparison of background correction methods for two-colour microarrays. *Bioinformatics.* 2007;**23**(20):2700–7. doi: 10.1093/bioinformatics/btm412. [PubMed: 17720982].

27. Chang CC, Lin CJ. Libsvm. *Acm transactions on intelligent systems and technology.* 2011;**2**(3):1–27. doi: 10.1145/1961189.1961199.

28. Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology.* 2005;**16**(1):73–81. [PubMed: 15613948].

29. Naugler WE, Karin M. The wolf in sheep's clothing: the role of interleukin-6 in immunity, inflammation and cancer. *Trends Mol Med.* 2008;**14**(3):109–19. doi: 10.1016/j.molmed.2007.12.007. [PubMed: 18261959].

30. Perrier d'Hauterive S, Charlet-Renard C, Dubois M, Berndt S, Goffin F, Foidart JM, et al. Human endometrial leukemia inhibitory factor and interleukin-6: control of secretion by transforming growth factor-beta-related members. *Neuroimmunomodulation.* 2005;**12**(3):157–63. doi: 10.1159/000084848. [PubMed: 15905624].

31. Reynaud D, Pietras E, Barry-Holson K, Mir A, Binnewies M, Jeanne M, et al. IL-6 controls leukemic multipotent progenitor cell fate and contributes to chronic myelogenous leukemia development. *Cancer Cell.* 2011;**20**(5):661–73. doi: 10.1016/j.ccr.2011.10.012. [PubMed: 22094259].

32. Bradley DT, Kountakis SE. Role of interleukins and transforming growth factor-beta in chronic rhinosinusitis and nasal polyposis. *Laryngoscope.* 2005;**115**(4):684–6. doi: 10.1097/01.mlg.0000161334.67977.5D. [PubMed: 15805881].

33. Lennard CM, Mann EA, Sun LL, Chang AS, Bolger WE. Interleukin-1 beta, interleukin-5, interleukin-6, interleukin-8, and tumor necrosis factor-alpha in chronic sinusitis: response to systemic corticosteroids. *Am J Rhinol.* 2000;**14**(6):367–73. [PubMed: 11197112].

34. Danielsen A, Tynning T, Brokstad KA, Olofsson J, Davidsson A. Interleukin 5, IL6, IL12, IFN-gamma, RANTES and Fractalkine in human nasal polyps, turbinate mucosa and serum. *Eur Arch Otorhinolaryngol.* 2006;**263**(3):282–9. doi: 10.1007/s00405-005-1031-1. [PubMed: 16456693].

35. Peters AT, Kato A, Zhang N, Conley DB, Suh L, Tancowny B, et al. Evidence for altered activity of the IL-6 pathway in chronic rhinosinusitis with nasal polyps. *J Allergy Clin Immunol.* 2010;**125**(2):397–403 e10. doi: 10.1016/j.jaci.2009.10.072. [PubMed: 20159251].

36. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet.* 2011;**378**(9795):1006–14. doi: 10.1016/S0140-6736(11)60874-X. [PubMed: 21907864].

37. Brown JE, Greenberger PA, Yarnold PR. Soluble serum interleukin 2 receptors in patients with asthma and allergic bronchopulmonary aspergillosis. *Ann Allergy Asthma Immunol.* 1995;**74**(6):484–8. [PubMed: 7788514].

38. Gevaert P, Holtappels G, Johansson SG, Cuvelier C, Cauwenberge P, Bachert C. Organization of secondary lymphoid tissue and local IgE formation to Staphylococcus aureus enterotoxins in nasal polyp tissue. *Allergy.* 2005;**60**(1):71–9. doi: 10.1111/j.1398-9995.2004.00621.x. [PubMed: 15575934].