Mathematical Biosciences and Engineering

MBE, 18(4): 4696–4712.
DOI: 10.3934/mbe.2021238
Received: 21 April 2021
Accepted: 17 May 2021
Published: 27 May 2021

http://www.aimspress.com/journal/MBE

Research article

Uncovering potential novel biomarkers and immune infiltration characteristics in persistent atrial fibrillation using integrated bioinformatics analysis

Shengjue Xiao[1], Yufei Zhou[2], Ailin Liu[1], Qi Wu[1], Yue Hu[3], Jie Liu[1], Hong Zhu[1], Ting Yin[2] and Defeng Pan[1,*]

[1] Department of Cardiology, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, Jiangsu 221004, China
[2] Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu 210029, China
[3] Department of General Practice, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, Jiangsu 221004, China

* **Correspondence**: Email: xzdefengpan@xzhmu.edu.cn.

Abstract: Atrial fibrillation (AF) is the most common cardiac arrhythmia. This study aimed to identify potential novel biomarkers for persistent AF (pAF) using integrated analyses and explore the immune cell infiltration in this pathological process. Three pAF datasets (GSE31821, GSE41177, and GSE79768) from the Gene Expression Omnibus (GEO) database were integrated with the elimination of batch effects. 264 differentially expressed genes (DEGs) were identified using Linear models for microarray data (LIMMA), 12 modules were screened out by weighted gene co-expression network analysis (WGCNA) in pAF compared with normal controls. Subsequently, common genes (CGs) were identified as the intersection of DEGs and genes in the most significant module. Functional enrichment analysis showed that CGs were mainly enriched in the "Calcineurin-NFAT (nuclear factor of activated T-cells)" signaling pathway, particularly regulator of calcineurin 1 (*RCAN1*), and protein phosphatase 3 regulatory subunit B, alpha (*PPP3R1*). Ulteriorly, the microRNA-transcription factor-mRNA network revealed that microRNA-34a-5p could target both *RCAN1* and *PPP3R1* in the pAF pathogenesis. Finally, immune infiltration analysis by CIBERSORT, a versatile computational method, displayed a higher level of monocytes, dendritic cells and neutrophils, as well as a lower level of CD8+ T cells and T cells regulatory (Tregs) in pAF compared with the control group. In conclusion, our present study revealed several novel pAF-associated genes, miRNAs, and pathways, including microRNA-34a-5p, which might target *RCAN1* and *PPP3R1* to regulate pAF through the calcineurin-

NFAT signaling pathway. In addition, there was a difference in immune infiltration between patients with pAF and normal groups and immune cells might interact with specific genes in pAF.

Keywords: persistent atrial fibrillation; differentially expressed genes; pathways; microRNA-transcription factor-mRNA network; immune infiltration

1. Introduction

Atrial fibrillation (AF) is the commonest cardiac arrhythmia in the general population, associated with high mortality and morbidity. Based on AF duration, AF can be divided into paroxysmal AF, persistent AF (pAF), long-standing pAF, and permanent AF. AF is estimated to affect 34 million people worldwide and is increasing as the population ages [1]. AF can induce major adverse cardiovascular events, including higher risk of stroke, heart failure [2]. However, the mechanisms of AF are complex and incompletely understood. Inflammation and atrial fibrosis were reported to be the fundamental mechanisms of the pathophysiological processes in the development of AF [3]. Soluble biomarkers of inflammation, including CRP and IL-6, and fibrosis, containing TGF-1 and matrix metalloproteinases, were used for predicting AF [4].

MicroRNA (miRNA), a small non-coding RNA molecule, plays a crucial role in RNA silencing and post-transcriptional gene expression regulation. Recent studies have elucidated the involvement of miRNAs in the pathological process of cardiovascular disease [5], including regulating arrhythmogenesis and cardiac excitability [6]. Moreover, miRNAs can be involved in AF-induced ion channel remodelling and fibrosis [7]. Therefore, miRNAs can provide useful insights into the pathophysiology of AF. Additionally, regulation of immune cells is critical in the pathogenesis and recurrence of AF [8]. Previous studies proved that increased infiltration of immune cells could be related to the enlargement of the left atrium of AF patients, a risk factor for AF recurrence [9]. Although some researchers also focused on the function of various immune cells, such as macrophages, monocytes, neutrophils [10], how these cells change in specific subtype of AF still remains blank.

At present, microarray technology, together with integrated bioinformatics analysis, have been extensively applied to identify potential novel biomarkers and their roles in various diseases to further explore the pathogenesis and develop potential treatments [11,12]. Additionally, Weighted gene co-expression network analysis (WGCNA) is gradually becoming a mature strategy in bioinformatics applications and can be used to identify underlying mechanisms, potential biomarkers or therapeutic targets in different types of diseases [13].

To our knowledge, few studies have investigated miRNAs, their target genes and immune cells in the heart tissue of patients with pAF. In the present study, three datasets (GSE31821, GSE41177 and GSE79768) were merged into an integrated dataset by the SVA method to remove batch differences. To identify the pivotal genes, miRNAs, pathways associated with pAF, differential expression analysis, WGCNA, enrichment analysis, as well as protein-protein interaction (PPI) network, miRNA-TF-mRNA network were constructed successively. Finally, immune infiltration was investigated in pAF by CIBERSORT. The study focused on novel findings in pAF mechanisms and may help to the identification of potential biomarkers for diagnosis and treatment of pAF.

2. Materials and methods

2.1. Microarray data

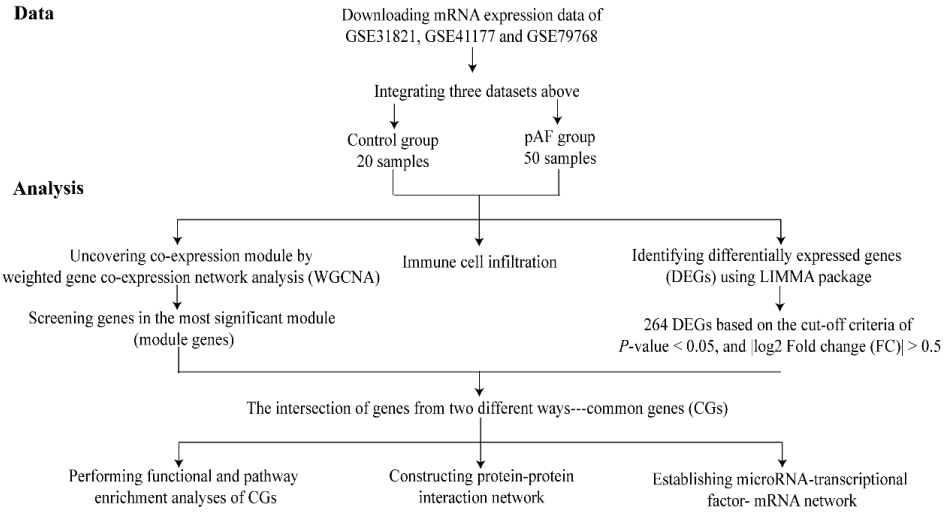

Figure 1. Study flowchart. Abbreviations: pAF, persistent atrial fibrillation; LIMMA, Linear models for microarray data.

The study flowchart is shown in Figure 1. Three pAF datasets (GSE31821, GSE41177, GSE79768) were downloaded from the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) [14]. The above three datasets were all gene expression array and generated using GPL570 (HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA) [15]. The dataset of GSE31821 included the auricle tissues of 4 patients with pAF and 2 normal samples; GSE41177 contained 38 atrial tissue samples, including 32 patients with pAF and 6 normal atrial tissue samples; GSE79768 consisted of 26 heart tissue samples including 14 pAF atrium heart tissue samples and 12 normal heart tissue samples.

2.2 Data processing and identification of differentially expressed genes

The three raw datasets were pre-processed by affy in R, including background calibration, normalization, and log2 transformation [16]. When multiple probes corresponded to one common gene, the average value was taken as its expression value. Besides, the surrogate variable analysis (SVA) package from Bioconductor was used to remove batch effects and other unwanted variation between three datasets [17]. Finally, the LIMMA package was then applied to screen DEGs and a p-value < 0.05, and $|\log_2$ Fold change (FC)| > 0.5 was set as a cut-off point for selecting DEGs [18].

2.3. Weighted gene co-expression network analysis (WGCNA) construction and identification of modules

To explore the interactions between genes, a system biology approach, WGCNA, was applied for gene co-expression network construction. First, genes having upper 25% variation across samples in

the integrated dataset were imported to the WGCNA. Second, to ensure that the results of network construction were reliable, outlier samples were removed. Third, Adjacency was calculated from the soft thresholding power β, which was derived by co-expression similarity, using pick-Soft-Threshold function. Then, the adjacency was transformed into a topological overlap matrix (TOM), and the corresponding dissimilarity (1-TOM) was calculated. Fourth, modules were detected through hierarchical clustering and dynamic tree cut function. To classify genes with similar expression profiles into gene modules, average linkage hierarchical clustering was conducted according to the TOM-based dissimilarity measure with a minimum size (gene group) of 50 for the genes dendrogram [19]. Fifth, for the modules correlated to the clinical attributes, module membership (MM) and gene significance (GS) were calculated. Finally, the eigengene network was visualized. Information on genes in the modules were used for further analysis. The intersection of the DEGs screeded from the integrated dataset and the genes in significant module is common genes (CGs).

2.4. Gene ontology (GO) annotation and kyoto encyclopaedia of genes and genomes (KEGG) pathway enrichment analyses of the CGs

DAVID (Database for Annotation, Visualization and Integrated Discovery) 6.8 (https: //david.ncifcrf.gov/) is a gene functional classification tool that was applied for enrichment analysis to perform the Gene Ontology (GO) for the CGs [20]. The enriched GO terms were divided into Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) ontologies. To understand the biofunction of the CGs, we uploaded the CGs to the DAVID database (Gene Functional Classification Tool, http: //david.abcc.ncifcrf.gov/) [21] and KOBAS database (http: //kobas.cbi.pku.edu.cn/kobas3) [22]. The significant enrichment for GO and KEGG analyses threshold was p-value < 0.05 and count ≥ 2.

2.5. Protein-protein interaction (PPI) network construction

The PPI interaction network of CGs was constructed with the Search Tool for the Retrieval of Interacting Genes (String) database (version 11.0; www.string-db.org) [23] and minimum interaction score ≥ 0.4 was defined as the cut-off value. Finally, the PPI network was visualized by Cytoscape [24].

2.6. MiRNA (microRNA)-TF (transcription factor)-mRNA regulatory network construction

Target miRNAs of the CGs were predicted with the miRTarBase [25], Starbase [26] and Targetscan [27] database. To improve the accuracy of the prediction, only miRNAs predicted by all three databases were retained. Combined with the Enrichr database (http: //amp.pharm.mssm.edu/Enrichr/), the (transcription factors) (TFs) targeted the CGs were predicted. The results with the p-value ≤ 0.05 were chosen as the cut-off values. After miRNA-TF-mRNA regulatory relationships were obtained, the miRNA-TF-mRNA regulatory network was visualised using Cytoscape.

2.7. Immune cell infiltration analysis

To quantify the relative proportions of infiltrating immune cells in pAF, CIBERSORT, a method of analysing the different immune cell types of tissues, was adopted to analyse the merged expression

data and calculate immune cell infiltrations [28]. $p < 0.05$ was used to filter the samples. And the percentage of each immune cell type in the samples was calculated and displayed in a bar plot. The heat map of the 22 immune cells was made using the "pheatmap" package [29]. The "vioplot" package was used to compare and visualize the levels of 22 immune cells between the pAF and control samples [30]. A correlation heatmap, which revealed the correlation of 22 types of infiltrating immune cells, was performed using the "corrplot" package.

2.8. Statistical analysis

All data were expressed as the mean ± standard deviation (SD). Statistical analyses were performed using GraphPad Prism (version 8.3.4). A *p*-value < 0.05 was considered significant for screening DEGs, module genes, enrichment analysis, PPI network, miRNA-TF-mRNA network and immune infiltration analysis.

3. Results

3.1. Identification of DEGs

A total of 264 genes were differentially expressed, of which 179 genes were up-regulated and 85 were down-regulated. The top 5 up-regulated genes involved S100 Calcium Binding Protein A12 (*S100A12*), C-X-C Motif Chemokine Receptor 2 *(CXCR2*), Myeloid Cell Nuclear Differentiation Antigen (*MNDA*), Nebulin Related Anchoring Protein (*NRAP*), and Four And A Half LIM Domains 2 (*FHL2*) while the top 5 down-regulated genes were Proteoglycan 4 (*PRG4*), Delta/Notch Like EGF Repeat Containing (*DNER*), Transcription Elongation Factor A Like 2 (*TCEAL2*), Family With Sequence Similarity 110 Member C (*FAM110C*), Mesothelin (*MSLN*) (Table S1). The heat map of DEGs and volcano plot of DEGs are shown in Figure 2A,B, respectively.

3.2. Identification of gene co-expression networks and modules

Firstly, the genes were ranked by variance from large to small, and the top 25% variance (5044) genes were selected for analysis. Secondly, the flashClust tools package was applied to perform the cluster analysis, by setting the threshold value to 80 and 1 outlier samples were subsequently detected and removed. Cluster 1 contains the 69 samples we want to keep. Thirdly, the power parameter ranging from 1–20 was screened out using the 'pickSoftThreshold' function in WGCNA package. In our research, we selected the power of $\beta = 8$ (scale free $R^2 = 0.85$) as the soft threshold to ensure a scale-free network. Set the threshold to 0.25 to merge similar modules in the cluster tree. As showed in Figure 3A, a total of 12 modules were obtained, in which genes had similar co-expression traits. The colours of all the modules were selected at random to distinguish between modules. The module eigengene (ME) in the magenta module ($r = 0.48$; $p = 3E{-}5$) exhibited the highest positive correlation and had significant correlation with pAF compared with other modules (Figure 3B). Therefore, the magenta module, including 616 genes, was identified as key modules for pAF and retained for further analysis. Additionally, we calculated correlation between gene's module memberships (MMs, correlation between a specific gene and module's eigengene) and gene significances (GSs, correlation between a specific gene and clinical variable) in the magenta module. As expected, significant positive correlations

between gene's MM and GS were obtained in the magenta module (cor = 0.44, P = 1.5E-30) as shown in Figure 3C. The intersection of the DEGs screened from the integrated dataset and the genes in magenta module is common genes (CGs) (Figure 3D). The genes in the magenta module were clearly most significantly associated with pAF.

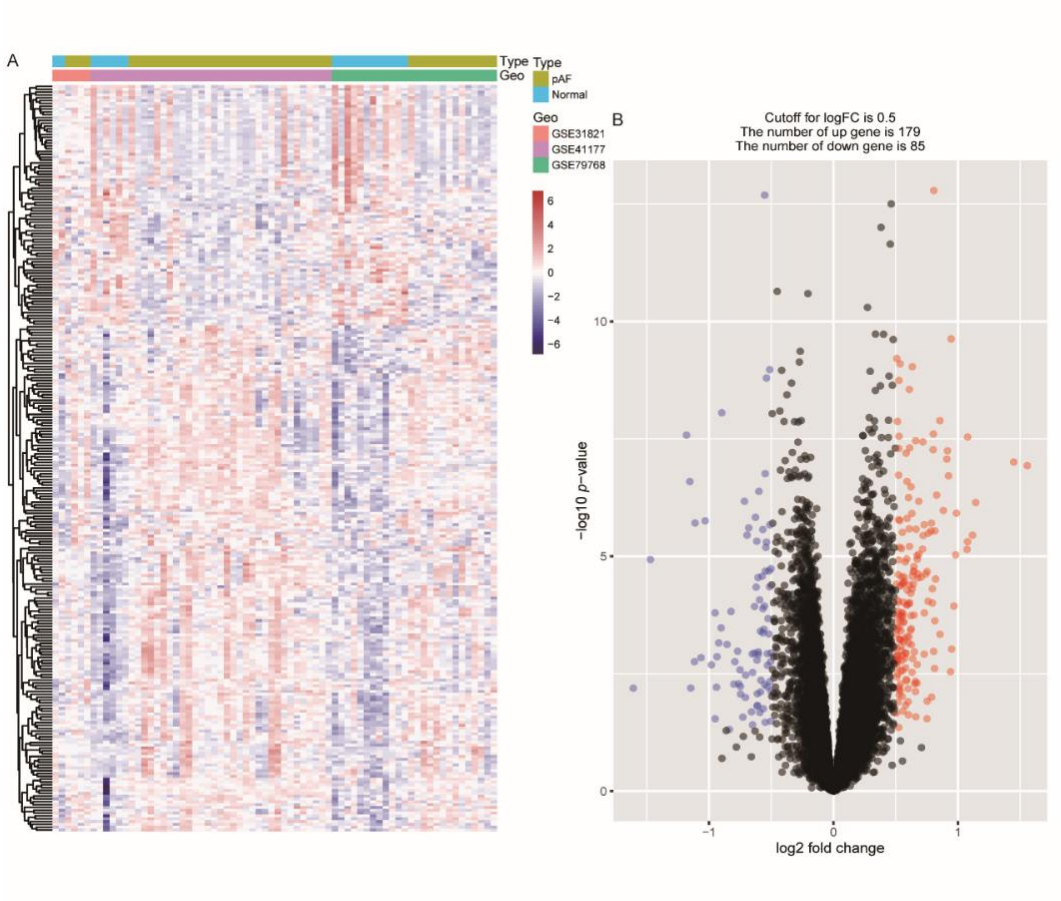

Figure 2. Heat map and Volcano plot for the DEGs identified from the integrated dataset. (A) Each row of the heat map represents one DEG, and each column represents one sample, either normal or pAF. The red and blue colors represent upregulated and downregulated DEGs, respectively. (B) Red plot points represent upregulated DEGs, and blue plot points show downregulated DEGs. Abbreviations: DEG, differentially expressed gene; pAF, persistent atrial fibrillation.

3.3. Enrichment analyses of the common genes

Functional enrichment analysis indicated that the CGs were mainly involved in biological process (BP) terms, including "angiogenesis", "osteoblast differentiation", "cellular protein modification process", "calcineurin-NFAT signalling cascade" and "endothelial cell morphogenesis". In the cell component (CC) ontology, the CGs were significantly enriched in "extracellular exosome", "Golgi membrane" and "Golgi apparatus". Molecular function (MF) analysis showed that the CGs were mainly enriched in "TPR domain binding", "protein binding", "integrin binding" (Table S2). Additionally, the KEGG pathway of the CGs was found to be enriched in "metabolic pathways",

"protein processing in endoplasmic reticulum", "Oxytocin signalling pathway", "Hippo signalling pathway", "NOD-like receptor signalling pathway", "Rap1 signalling pathway" and "MAPK signalling pathway". (Figure 4) (Table S3). The results of BP revealed that two upregulated DEGs (*RCAN1* and *PPP3R1*) were enriched in calcineurin-NFAT signalling cascade (*p* value = 0.023112).

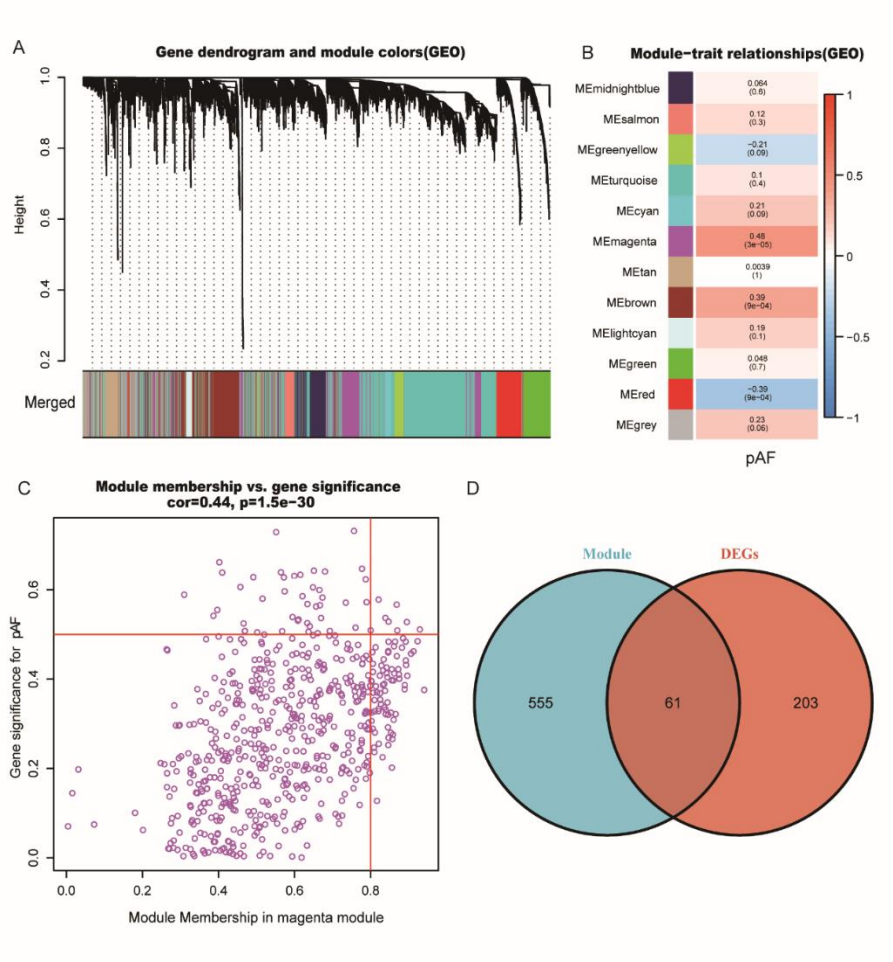

Figure 3. Weighted gene co-expression network analysis and Venn diagram. (A) GCMs, represented by different colors under the gene tree. (B) Heat map of the association between modules and pAF. The magenta module was significantly correlated with pAF. The numbers inside and outside of the brackets represent *p*-values and correlation coefficients, respectively. (C) Correlation plot between MM (X-axis) and GS (Y-axis) of genes contained in the magenta module. (D) Venn diagram showing overlapping genes between the DEGs and the genes in the magenta module. Abbreviations: GCMs, gene co-expression modules; GS, gene significance; MM, module membership. Color images are available online; DEGs, differentially expressed genes; pAF, persistent atrial fibrillation.

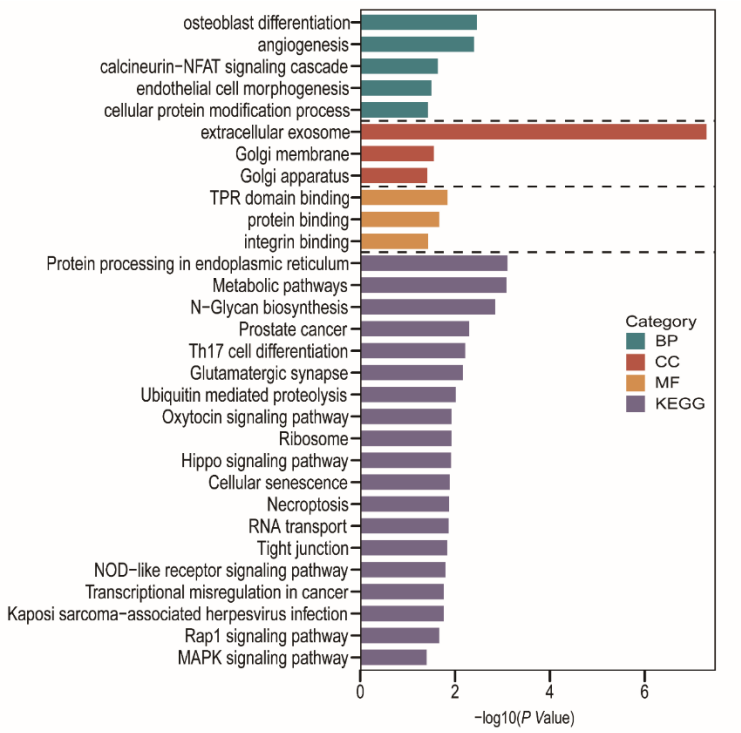

Figure 4. Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of common genes. The green bars represent biological processes, the red bars represent cellular components, the brown lines represent molecular functions, and the purple lines represent KEGG pathways. Abbreviations: BP, biological process; CC, cellular component; MF, molecular function; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

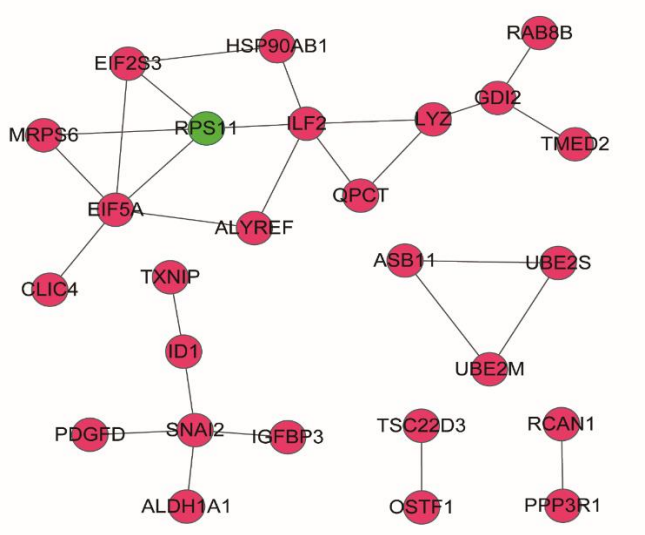

Figure 5. Protein-protein interaction network. Red nodes represent upregulated genes, and the green node represents a downregulated gene.

3.4. PPI network

After removing the isolated nodes, the PPI network was constructed with 26 nodes and 27 edges, including 25 up-regulated genes and 1 down-regulated genes. The CGs in the PPI network were defined as hub genes, including interleukin enhancer binding factor 2 (*ILF2*), eukaryotic translation initiation factor 5A (*EIF5A*), ribosomal protein S11 (*RPS11*), snail family transcriptional repressor 2 (*SNAI2*), protein phosphatase 3 regulatory subunit B, alpha (*PPP3R1*), regulator of calcineurin 1 (*RCAN1*) and so on (Figure 5). In the network, the *ILF2* and *EIF5A* had the highest degrees of connectivity; both connecting 5 nodes.

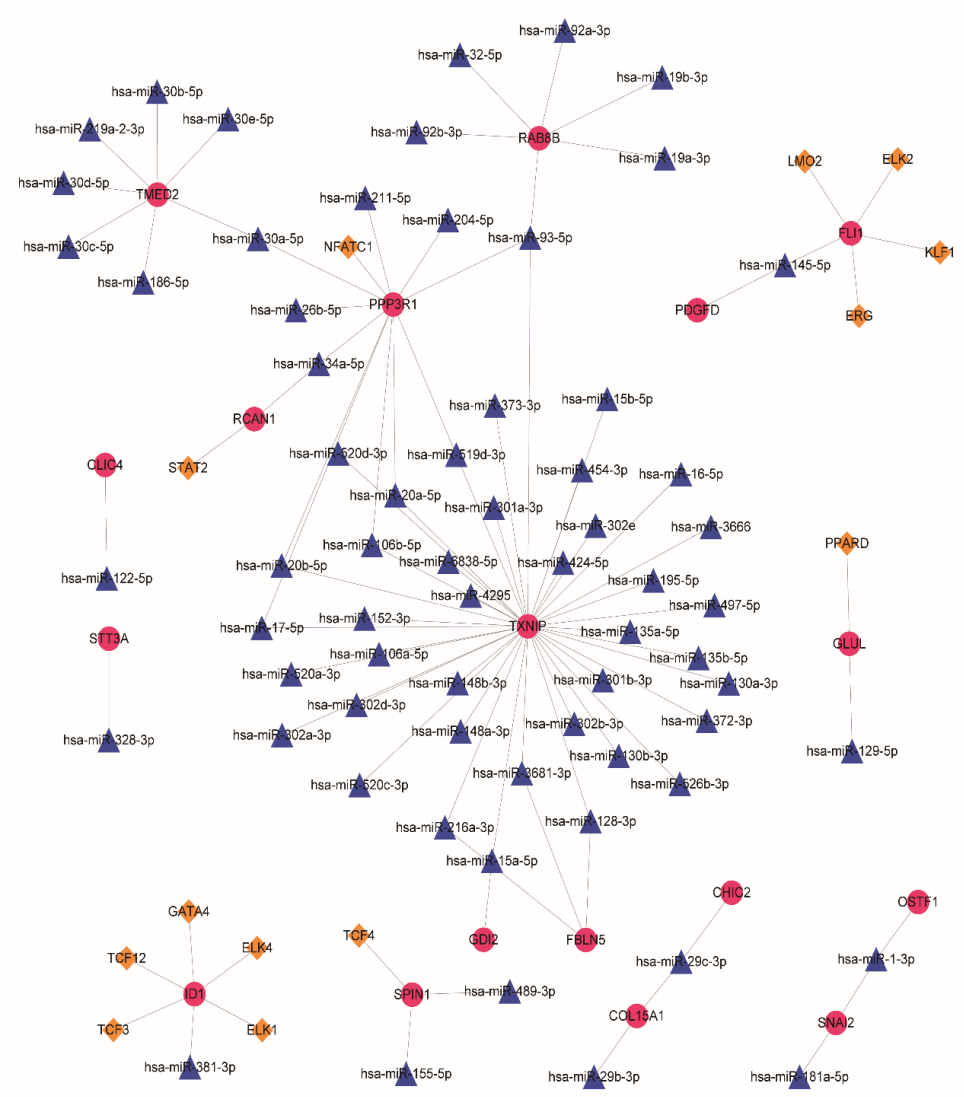

Figure 6. The miRNA-TF-mRNA regulatory network. The red balls represent upregulated genes, the blue triangles are predicted miRNAs, and the brown rhombuses represent TFs. Abbreviations: mRNAs, messenger RNAs; miRNAs, microRNAs; TFs, transcription factors.

3.5. MiRNA-TF-mRNA regulatory network analysis

After miRNA-gene and TF-gene pairs were predicted, 97 miRNA-TF-mRNA regulatory relationships were obtained. The regulatory network (including 65 miRNAs, 14 transcription factors, 18 co-upregulated genes, and 39 co-downregulated genes) was established by Cytoscape (Figure 6). Within the network, we discovered that *PPP3R1* was targeted by Nuclear Factor of Activated T Cells 1 *(NFATC1)*, whilst *RCAN1* was targeted by Signal Transducer And Activator Of Transcription 2 (*STAT2*). We also found that miR-34a-5p could regulate *RCAN1* and *PPP3R1* at the same time.

3.6. Immune cell infiltration analysis

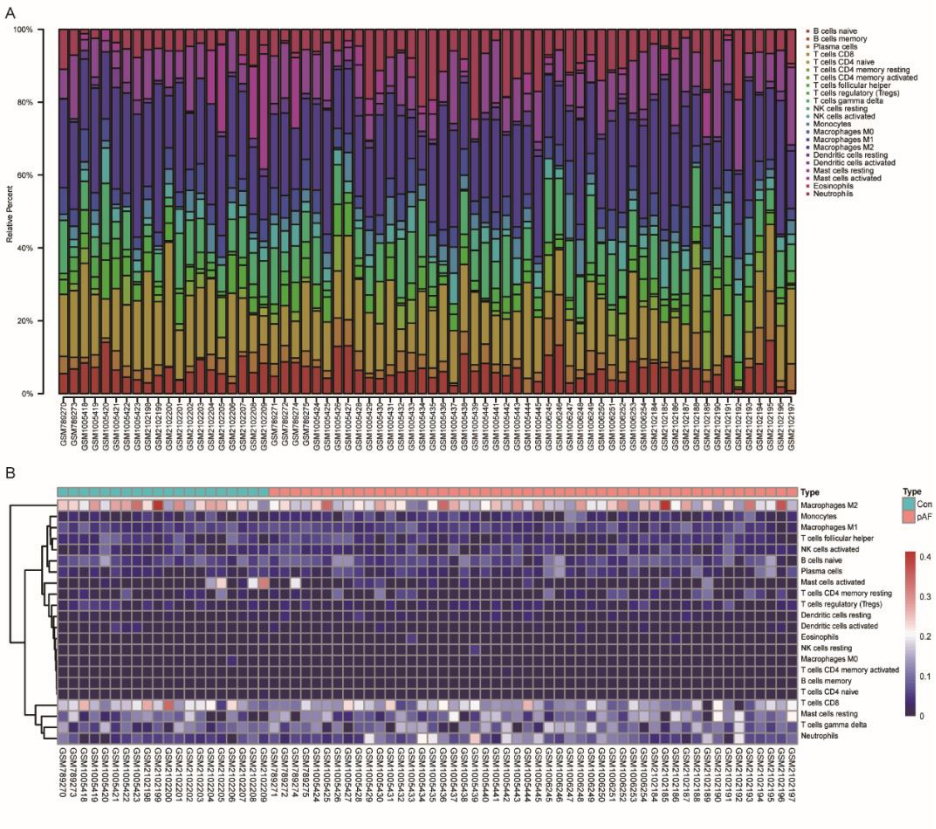

Figure 7. The landscape of immune infiltration between pAF and normal controls. (A) The relative percentage of 22 types of immune cells. (B) The heat map of 22 types of immune cells. Abbreviations: pAF, persistent atrial fibrillation.

We further performed the CIBERSORT algorithm to predict the immune cells infiltration between patients with pAF and the control group. The percentage of each of the 22 types of immune cells in each sample was shown in the bar plot and heat map (Figure 7A,B). The vioplot of the immune cell infiltration difference demonstrated that pAF patients had a higher level of monocytes, dendritic cells and neutrophils, and a lower level of CD8+ T cells and T cells regulatory (Tregs) compared with the control group (Figure 8A). The correlation of 22 types of immune cells revealed that activated Mast cells were positively related with T cells CD4 naive (r = 0.59) and

Plasma cells were positively related with B cells naive (r = 0.41), whereas activated Mast cells were negatively related to resting Mast cells (r = −0.61), T cells CD8 were negatively related to resting T cells CD4 memory and Macrophages M2 negatively correlated with T cells gamma delta (r = −0.45) (Figure 8B). There is a significant difference in immune cell infiltration between heart tissue in patients with pAF and the normal rhythm group. Therefore, monocytes, dendritic cells, neutrophils, CD8+ T cells and T cells regulatory (Tregs) may be potential core immune cells, involved in stimulating the progression of pAF.

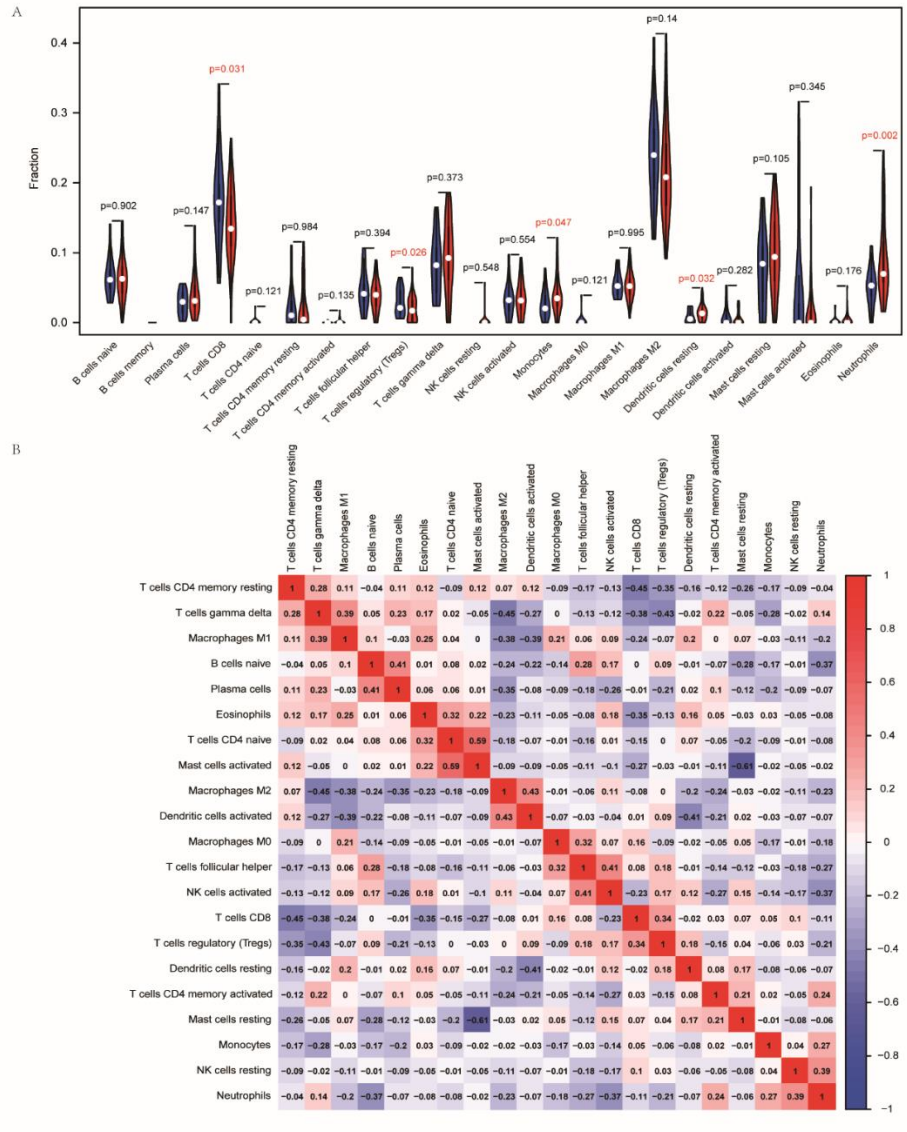

Figure 8. Distribution and visualization of immune cell infiltration. (A) Comparison of 22 immune cell subtypes between patients in pAF and controls. Blue and red colors represent normal and pAF samples, respectively. (B) Correlation matrix of all 22 immune cell subtype compositions. Both horizontal and vertical axes demonstrate immune cell subtypes. Immune cell subtype compositions (higher, lower, and same correlation levels are displayed in red, blue, and white, respectively).

4. Discussion

AF is the most common cardiac arrhythmia worldwide and the mechanisms of AF has not been fully excavated. Previous studies of gene expression files were limited by single method or small samples. In our study, we integrated three persistent AF datasets, used two different methods to identify common genes (WGCNA and LIMMA method). With the help of the LIMMA method, we identified a total of 264 DEGs between pAF samples and normal controls, amongst which 179 genes were up-regulated and 85 were down-regulated. Genes such as *S100A12* and *CXCR2* were significantly upregulated whereas *PRG4* and *DNER* were significantly downregulated. The mechanisms for the occurrence and maintenance of AF are associated with atrial fibrosis, inflammation and oxidative stress [31]. *S100A12* belongs to the S100/calgranulin subfamily of Ca2+-binding proteins, which induces an inflammatory response leading to the accumulation of immune cells in damaged tissue. The chemokine receptor *CXCR2* is a critical regulator in hypertension and cardiac remodeling. Zhang, Y.L. identified that blocking *CXCR2* prevents and reverses the development of AF in spontaneously hypertensive rats. Consequently, *CXCR2* may be a potential therapeutic target for hypertensive AF. [32]. Proteoglycan is a structural constituent of the aortic valve and accumulates in areas of fibrosis and moderate calcification. *PRG4* has been found in fibrotic parts of the aortic valves. Herder, C. et al. identified that *DNER* might operate as a biomarker of inflammation associated with the pathogenesis of the distal sensorimotor polyneuropathy [33]. In the WGCNA, we identified 12 gene modules, amongst which the magenta module, with 616 genes, has the highest correlation with pAF.

To make the result more accurate and reliable, we made the intersection of the genes in the most significant module and the differentially expressed genes in LIMMA method. 61 common genes were identified, enrichment and pathway analyses were further performed. Moreover, PPI network and ceRNA network were constructed. The biological process analysis revealed pAF was mainly related to calcineurin-NFAT signaling cascade.

Calcineurin-NFAT (nuclear factor of activated T-cells) signaling is critical for numerous biological processes, He RL and his colleagues [34] proved Calcineurin-NFAT could regulate pulmonary arterial hypertension rats through modulating the proliferation, migration and apoptosis of pulmonary artery smooth muscle cell. Zhou X et al [35] demonstrated that the activation of calcineurin-NFAT signaling pathway could lead to cardiac hypertrophy. In a rabbit AF model, the expression level of calcineurin and NFAT was up-regulated [36], speculating its role in regulating AF. However, the relationship of this signaling pathway and AF has not been fully reported. With the consideration of PPI network and gene enrichment, *RCAN1* and *PPP3R1* were enriched in the signaling pathway. *RCAN1* (regulator of calcineurin 1) belongs to RCAN family, of which, *RCAN2* and *RCAN3* are homologous [37]. *RCAN* was reported to be associated with various diseases including Down Syndrome [38], Alzheimer's disease [39] and diabetes [40]. *RCAN* could bind to calcineurin, a Ca2+/calmodulin-dependent phosphatase, inhibits calcineurin activity. *PPP3R1* is an important subunit of calcineurin [41]. Therefore, we conjectured that *RCAN1* and *PPP3R1* could regulate pAF through calcineurin-NFAT signaling pathway.

Atrial electrical and structural remodeling is the fundamental pathogenesis of AF [42]. MicroRNAs (miRNAs) are essential post-transcriptional modulators of gene expression. Lv, X., et al. found overexpression of miR-29b-3p reduced atrial structural and electrical remodeling in a rat fibrosis model, through the targeted regulation of the PDGF-B signaling pathway [43]. Taken miRNA-TF-mRNA network into account, miR-34a-5p functions as the bridge to connect *RCAN1* and *PPP3R1*.

Broad evidence proved the vital role of miR34a in cardiac disorders. Researchers found silencing miR-34 could attenuate cardiac dysfunction in moderate hypertrophic cardiomyopathy [44]. Zhu Y et al [45] demonstrated that miR-34a may serve an important role in early electrophysiological remodeling and the development of AF via the regulation of Ank-B expression. Furthermore, Diener C et al proved that *PPP3R1* could target miR-34a-5p in regulating intracellular calcium signaling [46], one of the basic mechanisms of AF. While *RCAN1* was identified as a novel direct target of miR-34a in regulating endothelial inflammation after cardiac bypass in goat placenta [47]. We speculated that miR-34a-5p might target *RCAN1* and *PPP3R1* in regulating AF through Calcineurin-NFAT signaling pathway. Therefore, miR-34a-5p/*RCAN1* and miR-34a-5p/*PPP3R1* may be viewed as potential therapeutic targets for pAF.

Additionally, the critical role of immune response in the pathogenesis of AF has attracted more attention, various immune cells, including innate and adaptive immune cells, infiltrated the atrium and interacted with atrial cardiomyocytes or secreting several chemokines and cytokines to regulate the microenvironment of the heart [48]. However, regarding the specific subtype pAF, immune infiltration analyses are not abundant, thus with the help of CIBERSORT, we evaluated the expression level of 22 kinds of immune cells in pAF. The results showed that monocytes, dendritic cells (DCs) and neutrophils were upregulated and CD8+ T cells and T cells regulatory were downregulated.

Recent studies have confirmed the function of monocyte in atrial structural remodelling after catheter ablation of AF. The activation of monocyte especially the enhanced migration ability could lead to the enlargement of atrial structure [49]. DCs are essential for regulating innate and adaptive immunity owing to its antigen-presenting role, the activation of DCs were related to rheumatic heart disease [50] and systematic inflammation response after coronary artery bypass grafting [51]. A prospective clinical study found the elevated ratio of neutrophils to lymphocyte (NTL) had an intimate positive corelation with the incidence of AF. The role of CD8+ T cell remained controversial recently, Kazem N et al. [52] constructed a prospective clinical study and found the fraction of CD8+ lymphocytes was significantly higher in individuals developing postoperative atrial fibrillation (POAF) comparing with individuals free of POAF. While with the bioinformatic analysis, Li and his colleagues [53] found the expression of CD8+ T cell decreased in AF patients, which is consistent with our findings. We speculated that the contrary results might be determined by different timepoints of the AF onset. Whether the patient catch AF initially or after surgery might lead to different findings, which needs to be further explored. Moreover, the expression of T reg cells were reported to be prominently decreased in peripheral whole blood and right atrial tissue of AF patients [54].

In this study, we found that pAF patients had an elevated level of *CXCR2* and monocytes. Zhang, Y.L. also found that *CXCR2* plays a critical role in driving monocyte infiltration of the atria, causing atrial remodelling and AF after hypertension. This is consistent with our study [55]. In addition, *S100A12* and neutrophils are both highly expressed in pAF patients. Therefore, we surmise that *S100A12*, serving as a biomarker of inflammation, may induce neutrophil infiltration, leading to atrial remodelling.

However, there were some limitations to our study. Firstly, we only concentrated on several important enrichment results and genes associated with the enrichment, which might neglect potential genes related to pAF. Secondly, for immune infiltration, further interaction between immune cells and miRNAs should be focused. Besides, the conclusions obtained from this paper should be investigated in *vivo* and *vitro*.

5. Conclusions

We conjectured that miR-34a-5p might target *RCAN1* and *PPP3R1* in regulating pAF through Calcineurin-NFAT signaling pathway. *S100A12* may drive neutrophil infiltration through induction of inflammatory response, leading to atrial remodeling. In addition, Monocytes, dendritic cells, neutrophils, CD8+ T cells and T cells regulatory (Tregs) may be involved in the pathogenesis of pAF.

6. Highlight

1) We combined the LIMMA differential gene expression analysis with WGCNA in order to identify the hub genes and pathways related to the pathogenesis of pAF.

2) MiR-34a-5p may regulate *RCAN1* and *PPP3R1* in controlling pAF through the Calcineurin-NFAT signaling pathway, which provides a molecular mechanism and potential therapeutic targets for pAF.

3) *S100A12* may lead to atrial remodeling through driving neutrophil infiltration.

4) Monocytes, dendritic cells, neutrophils, CD8+ T cells and T cells regulatory (Tregs) could potentially be developed as targets of immunotherapy in patients with pAF.

Conflicts of interest

The authors declare that they have no competing interests.

Acknowledgments

We would like to express our gratitude to the providers of the datasets GSE31821, GSE41177 and GSE79768 for sharing the data online. This work was supported by the grants from Xuzhou Science and Technology Bureau to Defeng Pan [grant number: KC20097].

References

1. S. S. Chugh, R. Havmoeller, K. Narayanan, D. Singh, M. Rienstra, E. J. Benjamin, et al., Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 study, *Circulation*, **129** (2014), 837–847.
2. R. S. Wijesurendra, B. Casadei, Mechanisms of atrial fibrillation, *Heart*, **105** (2019), 1860–1867.
3. E. Zacharia, N. Papageorgiou, A. Ioannou, G. Siasos, S. Papaioannou, M. Vavuranakis, et al., Inflammatory biomarkers in atrial fibrillation, *Curr. Med. Chem*., **26** (2019), 837–854.
4. C. Tsioufis, D. Konstantinidis, I. Nikolakopoulos, E. Vemmou, T. Kalos, G. Georgiopoulos, et al., Biomarkers of atrial fibrillation in hypertension, *Curr. Med. Chem*., **26** (2019), 888–897.
5. E. M. Small, E. N. Olson, Pervasive roles of microRNAs in cardiovascular biology, *Nature*, **469** (2011), 336–342.
6. Y. Lu, Y. Zhang, N. Wang, Z. Pan, X. Gao, F. Zhang, et al., MicroRNA-328 contributes to adverse electrical remodeling in atrial fibrillation, *Circulation*, **122** (2010), 2378–2387.

7. J. Andrade, P. Khairy, D. Dobrev, S. Nattel, The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms, *Circ. Res*., **114** (2014), 1453–1468.
8. Y. F. Hu, Y. J. Chen, Y. J. Lin, S. A. Chen, Inflammation and the pathogenesis of atrial fibrillation, *Nat. Rev. Cardiol*., **12** (2015), 230–243.
9. T. Yamashita, A. Sekiguchi, S. Suzuki, T. Ohtsuka, K. Sagara, H. Tanabe, et al., Enlargement of the left atrium is associated with increased infiltration of immune cells in patients with atrial fibrillation who had undergone surgery, *J. Arrhythm*., **31** (2015), 78–82.
10. T. Yan, S. Zhu, M. Zhu, C. Wang, C. Guo, Integrative identification of hub genes associated with immune cells in atrial fibrillation using weighted gene correlation network analysis, *Front. Cardiovasc. Med*., **7** (2020), 631775.
11. E. Zhao, C. Zhou, S. Chen, A signature of 14 immune-related gene pairs predicts overall survival in gastric cancer, *Clin. Transl. Oncol*., **23** (2021), 265–274.
12. E. Zhao, H. Xie, Y. Zhang, Identification of differentially expressed genes associated with idiopathic pulmonary arterial hypertension by integrated bioinformatics approaches, *J. Comput. Biol*., **28** (2021), 79–88.
13. P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, *BMC Bioinf.*, **9** (2008), 559.
14. E. Clough, T. Barrett, The gene expression omnibus database, *Methods Mol. Biol*., **1418** (2016), 93–110.
15. T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, et al., NCBI GEO: archive for functional genomics data sets-update, *Nucleic Acids Res*., **41** (2013), D991–995.
16. R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, et al., Exploration, normalization, and summaries of high density oligonucleotide array probe level data, *Biostatistics (Oxford, England)*, **4** (2003), 249–264.
17. J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, J. D. Storey. The sva package for removing batch effects and other unwanted variation in high-throughput experiments, *Bioinformatics*, **28** (2012), 882–883.
18. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, *Nucleic Acids Res.*, **43** (2015), e47–e47.
19. E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, A. L. Barabási, Hierarchical organization of modularity in metabolic networks, *Science*, **297** (2002), 1551–1555.
20. W. H. Da, B. T. Sherman, R. A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, *Nucleic Acids Res*., **37** (2009), 1–13.
21. W. H. Da, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, *Nat. Protoc*., **4** (2009), 44–57.
22. J. Wu, X. Mao, T. Cai, J. Luo, L. Wei, KOBAS server: a web-based platform for automated annotation and pathway identification, *Nucleic Acids Res*., **34** (2006), W720–724.
23. D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, et al., STRING v10: protein-protein interaction networks, integrated over the tree of life, *Nucleic Acids Res*., **43** (2015), D447–452.
24. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res*., **13** (2003), 2498–2504.

25. C. H. Chou, S. Shrestha, C. D. Yang, N. W. Chang, Y. L. Lin, K. W. Liao, et al., miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions, *Nucleic Acids Res*., **46** (2018), D296–d302.
26. J. H. Yang, J. H. Li, P. Shao, H. Zhou, Y. Q. Chen, L. H. Qu. starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data, *Nucleic Acids Res.*, **39** (2011), D202–D209.
27. V. Agarwal, G. W. Bell, J. W. Nam, D. P. Bartel, Predicting effective microRNA target sites in mammalian mRNAs, *Elife*, **4** (2015).
28. A. M. Newman, C. L. Liu, M. R. Green, A. J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, *Nat. Methods*, **12** (2015), 453–457.
29. A. L. Dailey, Metabolomic bioinformatic analysis, *Methods Mol. Biol*., **1606** (2017), 341–352.
30. K. Hu. Become competent within one day in generating boxplots and violin plots for a novice without prior R experience, *Methods Protoc*., **3** (2020).
31. S. Xiao, Y. Zhou, Q. Liu, T. Zhang, D. Pan, Identification of pivotal microRNAs and target genes associated with persistent atrial fibrillation based on bioinformatics analysis, *Comput. Math. Methods Med*., **2021** (2021), 6680211.
32. Y. L. Zhang, H. J. Cao, X. Han, F. Teng, C. Chen, J. Yang, et al., Chemokine receptor CXCR-2 initiates atrial fibrillation by triggering monocyte mobilization in mice, *Hypertension*, **76** (2020), 381–392.
33. C. Herder, J. M. Kannenberg, M. Carstensen-Kirberg, A. Strom, G. J. Bönhof, W. Rathmann, et al., A systemic inflammatory signature reflecting cross talk between innate and adaptive immunity is associated with incident polyneuropathy: KORA F4/FF4 Study, *Diabetes*, **67** (2018), 2434–2442.
34. R. L. He, Z. J. Wu, X. R. Liu, L. X. Gui, R. X. Wang, M. J. Lin, Calcineurin/NFAT signaling modulates pulmonary artery smooth muscle cell proliferation, migration and apoptosis in monocrotaline-induced pulmonary arterial hypertension rats, *Cell Physiol. Biochem*., **49** (2018), 172–189.
35. X. Zhou, Q. Zhang, T. Zhao, X. Bai, W. Yuan, Y. Wu, et al., Cisapride protects against cardiac hypertrophy via inhibiting the up-regulation of calcineurin and NFATc-3, *Eur. J. Pharmacol*., **735** (2014), 202–210.
36. L. Y. Li, Q. Lou, G. Z. Liu, J. C. Lv, F. X. Yun, T. K. Li, et al., Sacubitril/valsartan attenuates atrial electrical and structural remodelling in a rabbit model of atrial fibrillation, *Eur. J. Pharmacol*., **881** (2020), 173120.
37. S. K. Lee, J. Ahnn, Regulator of calcineurin (RCAN): beyond down syndrome critical region, *Mol. Cells*, **43** (2020), 671–685.
38. V. Parra, F. Altamirano, C. P. Hernández-Fuentes, D. Tong, V. Kyrychenko, D. Rotter, et al., Down syndrome critical region 1 gene, Rcan1, helps maintain a more fused mitochondrial network, *Circ. Res*., **122** (2018), e20–e33.
39. H. Peiris, D. J. Keating, The neuronal and endocrine roles of RCAN1 in health and disease, *Clin. Exp. Pharmacol. Physiol*., **45** (2018), 377–383.
40. J. Chen, H. Zhang, D. Niu, H. Li, K. Wei, L. Zhang, et al., The risk factors related to the severity of pain in patients with chronic prostatitis/chronic pelvic pain syndrome, *BMC Urol*., **20** (2020), 154.
41. L. Bär, C. Großmann, M. Gekle, M. Föller, Calcineurin inhibitors regulate fibroblast growth factor 23 (FGF23) synthesis, *Naunyn Schmiedebergs Arch. Pharmacol*., **390** (2017), 1117–1123.

42. M. S. Dzeshka, G. Y. Lip, V. Snezhitskiy, E. Shantsila, Cardiac fibrosis in patients with atrial fibrillation: mechanisms and clinical implications, *J. Am. Coll. Cardiol*., **66** (2015), 943–959.
43. X. Lv, P. Lu, Y. Hu, T. Xu. Overexpression of miR-29b-3p inhibits atrial remodeling in rats by targeting PDGF-B signaling pathway, *Oxid. Med. Cell Longev*., **2021** (2021), 3763529.
44. B. C. Bernardo, X. M. Gao, Y. K. Tham, H. Kiriazis, C. E. Winbanks, J. Y. Ooi, et al., Silencing of miR-34a attenuates cardiac dysfunction in a setting of moderate, but not severe, hypertrophic cardiomyopathy, *PLoS One*, **9** (2014), e90337.
45. Y. Zhu, Z. Feng, W. Cheng, Y. Xiao. MicroRNA-34a mediates atrial fibrillation through regulation of Ankyrin-B expression, *Mol. Med. Rep*., **17** (2018), 8457–8465.
46. C. Diener, M. Hart, D. Alansary, V. Poth, B. Walch-Rückheim, J. Menegatti, et al., Modulation of intracellular calcium signaling by microRNA-34a-5p, *Cell Death Dis*., **9** (2018), 1008.
47. H. Y. Yuan, C. B. Zhou, J. M. Chen, X. B. Liu, S. S. Wen, G. Xu, et al., MicroRNA-34a targets regulator of calcineurin 1 to modulate endothelial inflammation after fetal cardiac bypass in goat placenta, *Placenta*, **51** (2017), 49–56.
48. Y. Guo, G. Y. Lip, S. Apostolakis. Inflammation in atrial fibrillation, *J. Am. Coll. Cardiol*., **60** (2012), 2263–2270.
49. H. Suehiro, K. Kiuchi, K. Fukuzawa, N. Yoshida, M. Takami, Y. Watanabe, et al., Circulating intermediate monocytes and atrial structural remodeling associated with atrial fibrillation recurrence after catheter ablation, *J. Cardiovasc. Electrophysiol*., 2021.
50. M. Shiba, Y. Sugano, Y. Ikeda, H. Okada, T. Nagai, H. Ishibashi-Ueda, et al., Presence of increased inflammatory infiltrates accompanied by activated dendritic cells in the left atrium in rheumatic heart disease, *PLoS One*, **13** (2018), e0203756.
51. A. J. Perros, A. Esguerra-Lallen, K. Rooks, F. Chong, S. Engkilde-Pedersen, H. M. Faddy, et al., Coronary artery bypass grafting is associated with immunoparalysis of monocytes and dendritic cells, *J. Cell Mol. Med*., **24** (2020), 4791–4803.
52. N. Kazem, P. Sulzgruber, B. Thaler, J. Baumgartner, L. Koller, G. Laufer, et al., CD8+CD28null T lymphocytes are associated with the development of atrial fibrillation after elective cardiac surgery, *Thromb Haemost*., **120** (2020), 1182–1187.
53. S. Li, Z. Jiang, X. Chao, C. Jiang, G. Zhong, Identification of key immune-related genes and immune infiltration in atrial fibrillation with valvular heart disease based on bioinformatics analysis, *J. Thorac. Dis*., **13** (2021), 1785–1798.
54. Y. Chen, G. Chang, X. Chen, Y. Li, H. Li, D. Cheng, et al., IL-6-miR-210 suppresses regulatory T cell function and promotes atrial fibrosis by targeting Foxp3, *Mol. Cells*, **43** (2020), 438–447.
55. Y. L. Zhang, F. Teng, X. Han, P. B. Li, X. Yan, S. B. Guo, et al., Selective blocking of CXCR2 prevents and reverses atrial fibrillation in spontaneously hypertensive rats, *J. Cell Mol. Med*., **24** (2020), 11272–11282.

AIMS Press