



OPEN ACCESS

EDITED BY

Hung-Fat Tse,
The University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY

Weichieh Lee,
Chi Mei Medical Center, Taiwan
Yao-Zong Guan,
Second Affiliated Hospital of Guangxi
Medical University, China

*CORRESPONDENCE

Zheng-Yu Liu
liuzhengyu@hunnu.edu.cn
Hong Wei Pan
panhongwei@hunnu.edu.cn

SPECIALTY SECTION

This article was submitted to
Cardiac Rhythmology,
a section of the journal
Frontiers in Cardiovascular Medicine

RECEIVED 18 April 2022

ACCEPTED 11 July 2022

PUBLISHED 27 July 2022

CITATION

Zheng P-F, Chen L-Z, Liu P, Liu Z-Y and
Pan HW (2022) Integrative
identification of immune-related key
genes in atrial fibrillation using
weighted gene coexpression network
analysis and machine learning.
Front. Cardiovasc. Med. 9:922523.
doi: 10.3389/fcvm.2022.922523

COPYRIGHT

© 2022 Zheng, Chen, Liu, Liu and Pan.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Integrative identification of immune-related key genes in atrial fibrillation using weighted gene coexpression network analysis and machine learning

Peng-Fei Zheng^{1,2,3}, Lu-Zhu Chen⁴, Peng Liu⁴,
Zheng-Yu Liu^{1,2,3*} and Hong Wei Pan^{1,2,3*}

¹Department of Cardiology, Hunan Provincial People's Hospital, Changsha, China, ²Clinical Research Center for Heart Failure in Hunan Province, Changsha, China, ³Hunan Provincial People's Hospital, Institute of Cardiovascular Epidemiology, Changsha, China, ⁴Department of Cardiology, The Central Hospital of ShaoYang, Shaoyang, China

Background: The immune system significantly participates in the pathologic process of atrial fibrillation (AF). However, the molecular mechanisms underlying this participation are not completely explained. The current research aimed to identify critical genes and immune cells that participate in the pathologic process of AF.

Methods: CIBERSORT was utilized to reveal the immune cell infiltration pattern in AF patients. Meanwhile, weighted gene coexpression network analysis (WGCNA) was utilized to identify meaningful modules that were significantly correlated with AF. The characteristic genes correlated with AF were identified by the least absolute shrinkage and selection operator (LASSO) logistic regression and support vector machine recursive feature elimination (SVM-RFE) algorithm.

Results: In comparison to sinus rhythm (SR) individuals, we observed that fewer activated mast cells and regulatory T cells (Tregs), as well as more gamma delta T cells, resting mast cells, and M2 macrophages, were infiltrated in AF patients. Three significant modules (pink, red, and magenta) were identified to be significantly associated with AF. Gene enrichment analysis showed that all 717 genes were associated with immunity- or inflammation-related pathways and biological processes. Four hub genes (*GALNT16*, *HTR2B*, *BEX2*, and *RAB8A*) were revealed to be significantly correlated with AF by the SVM-RFE algorithm and LASSO logistic regression. qRT-PCR results suggested that compared to the SR subjects, AF patients exhibited significantly reduced *BEX2* and *GALNT16* expression, as well as dramatically elevated *HTR2B* expression. The AUC measurement showed that the diagnostic efficiency of *BEX2*, *HTR2B*, and *GALNT16* in the training set was 0.836, 0.883, and 0.893, respectively, and 0.858, 0.861, and 0.915, respectively, in the validation set.

Conclusions: Three novel genes, *BEX2*, *HTR2B*, and *GALNT16*, were identified by WGCNA combined with machine learning, which provides potential new therapeutic targets for the early diagnosis and prevention of AF.

KEYWORDS

weighted gene coexpression network analysis, atrial fibrillation, immune cell subtype distribution pattern, machine learning, hub genes

Introduction

Characterized as rapid disordered atrial electrical activity, atrial fibrillation (AF) was defined as a common persistent arrhythmia in the clinic (1). Based on an existing report, almost 1% of the population is affected by AF worldwide, and its prevalence is positively correlated with age. The incidence of AF in the population older than 80 years reaches 8% (2). AF is significantly correlated with the occurrence of myocardial infarction, stroke, and heart failure, which increases the economic burden not only on the patients' families but also on society (3). Thus, clarifying the pathogenesis and discovering effective therapeutic methods are urgent endeavors. AF is a multifactorial complex disease and is commonly associated with many factors, such as sex, smoking, age, hypertension, diabetes, obesity, valvular heart disease, and ischaemic heart disease (4). However, the exact changes in etiology and pathology in AF are totally unknown. Increasing evidence suggests that immune cells significantly participate in the processes of AF pathogenesis (5). Additionally, a relatively high production of inflammatory markers in the serum of AF patients was observed, such as interleukin 6 (IL-6) and C-reactive protein (CRP) (6, 7). However, to identify potential targets for AF treatment, the connections between immune cells and the molecular pathogenesis mechanisms of AF need to be further explored.

In recent years, the use of CIBERSORT, as a widely used analysis tool, has been applied to RNA-seq data or microarray data to investigate the immune cell infiltration patterns and evaluate the infiltrated immune cell proportions in samples (8). With the continuous promotion of gene chip technology, weighted gene coexpression network analysis (WGCNA), as a powerful biological tool to analyze network relationships and molecular mechanisms, is widely used for massive gene profile analysis (9). WGCNA is often employed to identify coexpressed gene modules and further explore the connections of these identified gene modules with the features of the sample (10). Recently, an increasing number of researchers have applied machine learning to improve the prediction and accuracy of these genes, which were identified through traditional microarrays or next-generation sequencing data (11). The SVM-RFE algorithm and LASSO regression are the most

widely used machine learning methods to identify key genes (12). However, the combined application of WGCNA and machine learning in the identification of AF-associated genes has not been conducted.

In the current research, comprehensive bioinformatics analysis was conducted to investigate the association of immune-associated genes and cells with AF. The composition of the immune cells infiltrated in the tissues was assessed using CIBERSORT, the key gene modules were identified by WGCNA, and the key genes correlated with AF were identified *via* the SVM-RFE algorithm and LASSO regression. Furthermore, the expression of key genes and their diagnostic efficiency were further validated in the training set and validation samples.

Materials and methods

Collection of datasets

Gene expression profiles of left atrial samples of GSE79768 (AF patients = 7, SR individuals = 6) and GSE115574 (AF patients = 14, SR individuals = 15) were extracted from the public Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo>) database. The integrated expression profiles of GSE79768 and GSE115574 were used as training sets. Gene expression profiles were normalized using the *normalize Between Arrays* function in the *limma* package (13). Probes that detected more than one gene were excluded from this study. The expression of genes detected by multiple probes was determined as the average gene expression detected for all probes. Interbatch differences between the GSE79768 and GSE115574 datasets were eliminated using the "sva" package.

Construction of the WGCNA and identification of modules significantly associated with AF

A critical tool in the study of systems biology is WGCNA, which can construct a gene expression data profile-based scale-free network (14). The WGCNA method was used to analyze the

top 25% of genes with high expression variances. The reliability of the constructed scale-free network is ensured by removing outlier samples. First, before the power function was applied, a standard-scale free network was used to approximate the appropriate soft threshold power (soft power = 10) to obtain adjacency values among genes with a variance more significant than all the variance quartiles. Next, we transformed the adjacency values into a topological overlap matrix (TOM) and derived the dissimilarity (1-TOM) values. Finally, the dynamic tree cut method was used to identify modules by hierarchically clustering genes with the 1-TOM as the distance measure with a deep split value of 2 and a minimum size cut-off of 100 for the resulting dendrogram. To evaluate the relationships between clinical shapes and modules for discerning modules of biological significance, we conducted Pearson correlation analysis.

Enrichment analysis of interesting modules

KEGG and GO enrichment analyses of genes in biologically significant modules were carried out by clusterProfiler and the DOSE package in R (15). The threshold was determined to be $FDR < 0.05$.

Identification of key genes by LASSO regression and SVM-RFE algorithm

The LASSO regression and SVM-RFE algorithm were used to identify the key genes with the best prognostic value for AF. LASSO regression (16) was carried out with the package “glmnet”. As a technique for effective feature selection, SVM-RFE can select the best variables by excluding the SVM-generated feature vector (17). Based on the SVM function in the e1071 package of R, the selected biomarkers in the diagnosis of AF were classified and analyzed by the SVM classifier. The common genes identified by these two machine learning methods were defined as key genes for subsequent research.

Evaluation of immune cell subtype distribution

The immune infiltration pattern in AF was explored using a CIBERSORT R script (8). After downloading the immune cell expression matrix, boxplot diagrams, heatmaps, and histograms were generated using the package “ggplot2”. The histogram shows the immune cell proportion infiltrated in AF patients, and the heatmap and boxplot diagrams show the difference in immune cell infiltration in control and AF subjects. The Pearson correlation coefficient between each immune cell was calculated

using the package “corrplot”, and the results are shown in the relevant heatmap.

Correlation between key genes and immune cells and roc curve analysis in training set

The “corrplot” software package was employed to generate the Pearson correlation coefficient between each immune cell and hub gene, and the results are shown in the relevant bar graph. The diagnostic accuracy of the key genes was also tested in the training set.

Study population

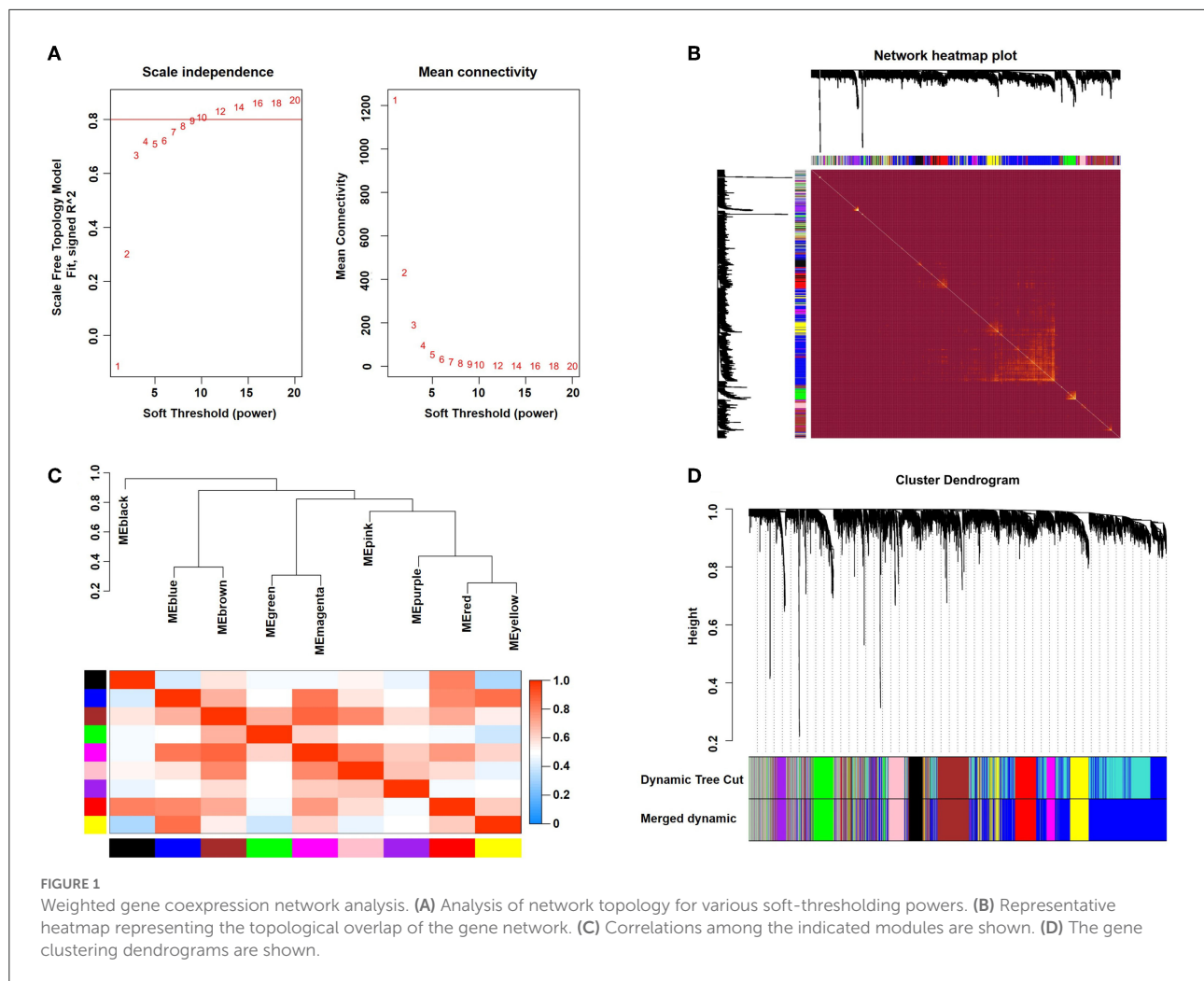
A total of 158 participants, including 82 persistent AF patients and 76 SR subjects, were recruited from Hunan Provincial People’s Hospital from June to December 2021. The disease was continuously sustained for more than 7 days or more than 7 days after cardioversion (automatic, electrical, or drug cardioversion) and was defined as persistent AF (18). Patients with a history of haematologic disease, type 1 diabetes, coronary artery disease, hypertension, autoimmune disease, neoplasia, and renal or liver diseases were excluded. Peripheral blood samples were collected and placed at -80°C for subsequent study. Study protocols were developed based on the Ethics Committee of Hunan Provincial People’s Hospital (No: LL-20210615-144) and the 2008 revision of the Declaration of Helsinki of 1975 (<http://www.wma.net/en/30publications/10policies/b3/>). All subjects provided written and informed consent.

qRT-PCR

Total RNA was extracted from peripheral blood with an RNeasyTM Mini Kit (QIAGEN, Frankfurt, Germany). cDNA was then reverse-transcribed with the PrimeScriptTM RT reagent Kit (Takara, Otsu, Japan). A Taq PCR Master Mix Kit (Takara, Otsu, Japan) was used to perform qRT-PCR using an ABI 7500 instrument (Applied Biosystems, USA). The proprietary qPCR primers used in the experiment were designed and validated by Songon Biotech (Songon Biotech, Shanghai, China).

Statistical analyses

All the data displayed in this study were processed and analyzed with SPSS (Version 22.0). *GALNT16*, *HTR2B*, *RAB8A*, and *BEX2* expression were assessed using independent sample t tests. According to the concentrations of *GALNT16*, *HTR2B*,



RAB8A, and *BEX2* in serum, we constructed receiver operating characteristic (ROC) curves. A nonparametric method was employed to evaluate the respective areas under the curves (AUCs) with 95% CI by MedCalc (MedCalc Software, Mariakerke, Belgium, version 19.7.4) software. R (version 4.1.0) was used to perform the bioinformatics analysis. $P < 0.05$ was considered statistically significant.

Results

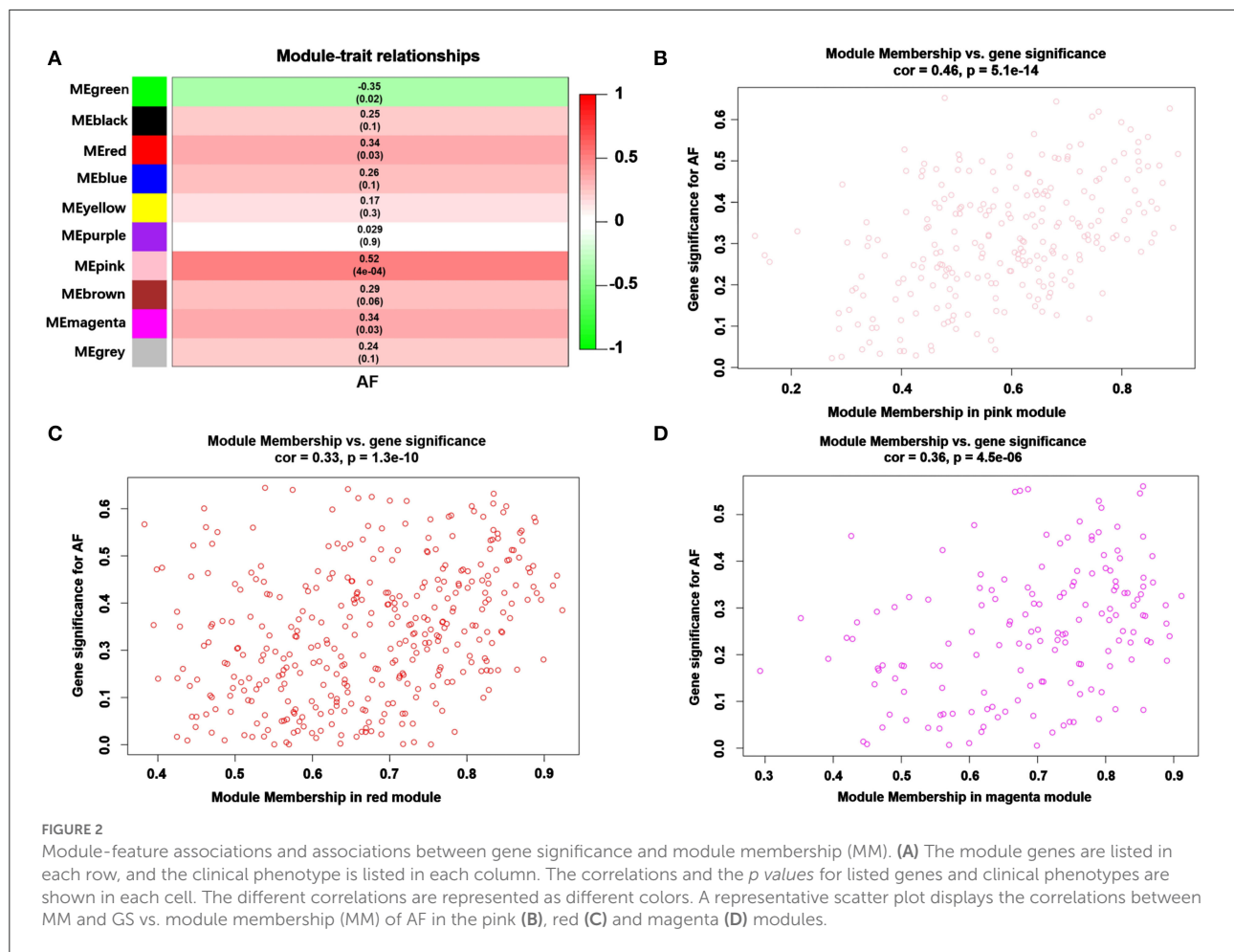
Data preprocessing

First, the normalized gene expression profiles of the GSE79768 and GSE115574 datasets were obtained after standardizing the data format, adding missing values, and removing outliers. Then, after data merging and eliminating the interbatch differences between the GSE79768 and GSE115574 datasets, the combined expression matrix including 21,629 gene symbols was obtained from the 42

left atrial samples in the training set. After removing 1 outlier sample (Supplementary Figure S1), the top 25% of genes with high expression variance in the remaining 41 left atrial samples were selected for subsequent WGCNA and are presented in Supplementary Table S1. In addition, disease grouping information for 41 samples is also presented in Supplementary Table S2.

Weighted gene co-expression networks

After calculation, we revealed that a correlation coefficient of more than 0.8 (the soft threshold β is 10) was highly correlated and suitable for constructing several gene modules (Figure 1A). A TOM was constructed by calculating the correlation and adjacency matrices of the gene expression profiles. The gene cluster tree is depicted in Figure 1B. Next, to select each gene network's gene modules, both TOM and hierarchical average linkage clustering were employed. Figure 1C depicts the



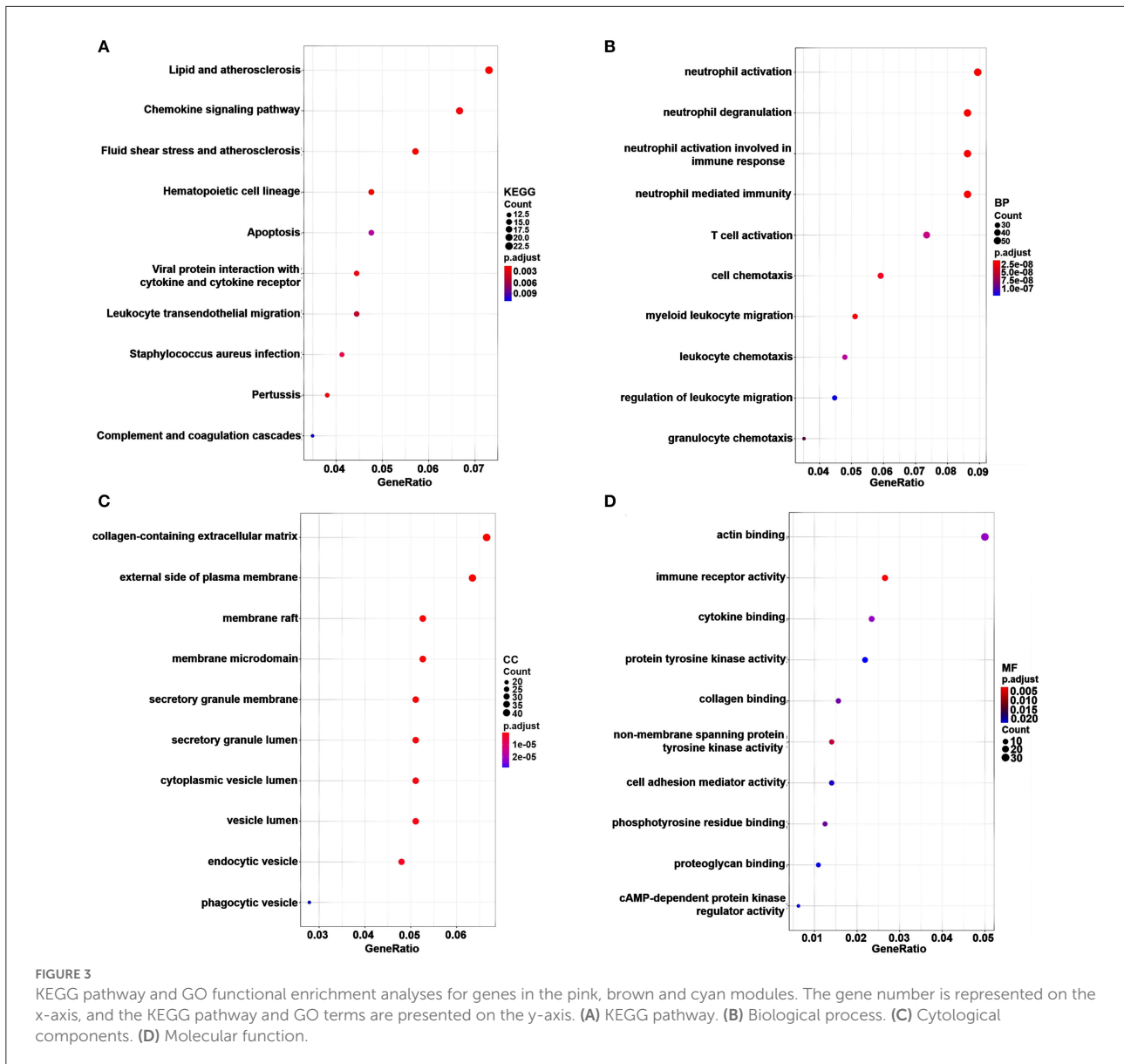
heatmap. The dynamic tree cut algorithm depicted twelve gene modules (Figure 1D).

Identification of modules of interest

Modules closely related to clinical features are often found to carry important and specific biological significance. As shown in Figure 2A, the pink ($r^2 = 0.52$, $P = 4E-04$), red ($r^2 = 0.34$, $P = 0.03$) and magenta ($r^2 = 0.34$, $P = 0.03$) modules appeared to be positively correlated with AF. An in-depth calculation was performed to discern the association between the color module and gene significance (GS). The association between the pink module and gene significance was 0.46 ($P = 5.1E-014$) (Figure 2B), the association between the red module and gene significance was 0.33 ($P = 1.3E-10$) (Figure 2C), and the association between the magenta module and gene significance was 0.36 ($P = 4.5E-06$) (Figure 2D). All gene symbols in the pink, red, and magenta modules and their GS values and corresponding *P* values are described in detail in Supplementary Table S3.

Enrichment analysis of genes in pink, brown, and cyan modules

KEGG pathway and GO enrichment analyses of 717 genes in the pink, red, and magenta modules were carried out to dissect their physiological purposes. Figure 3A shows the top 10 KEGG signaling pathways as follows: hsa04640: haematopoietic cell lineage; hsa05418: fluid shear stress and atherosclerosis; hsa05417: lipid and atherosclerosis; hsa04062: chemokine signaling pathway; hsa04061: viral protein interaction with cytokine and cytokine receptor; hsa05133: pertussis; hsa05150: Staphylococcus aureus infection; hsa04670: leukocyte transendothelial migration; hsa04210: apoptosis; hsa04610: complement and coagulation cascades. Figure 3B shows the top 10 biological processes as follows: GO:0042119: neutrophil activation; GO:0043312: neutrophil degranulation; GO:0002283: neutrophil activation involved in immune response; GO:0002446: neutrophil mediated immunity; GO:0097529: myeloid leukocyte migration; GO:0060326: cell chemotaxis; GO:0042110: T-cell activation; GO:0030595: leukocyte chemotaxis; GO:0071621: granulocyte



chemotaxis; GO:0002685: regulation of leukocyte migration. These signaling pathways and biological processes are mainly related to inflammation and the immune response. In addition, cytoplasmic components and molecular functions are shown in Figures 3C,D. The details of these analyses can also be found in Supplementary Table S4.

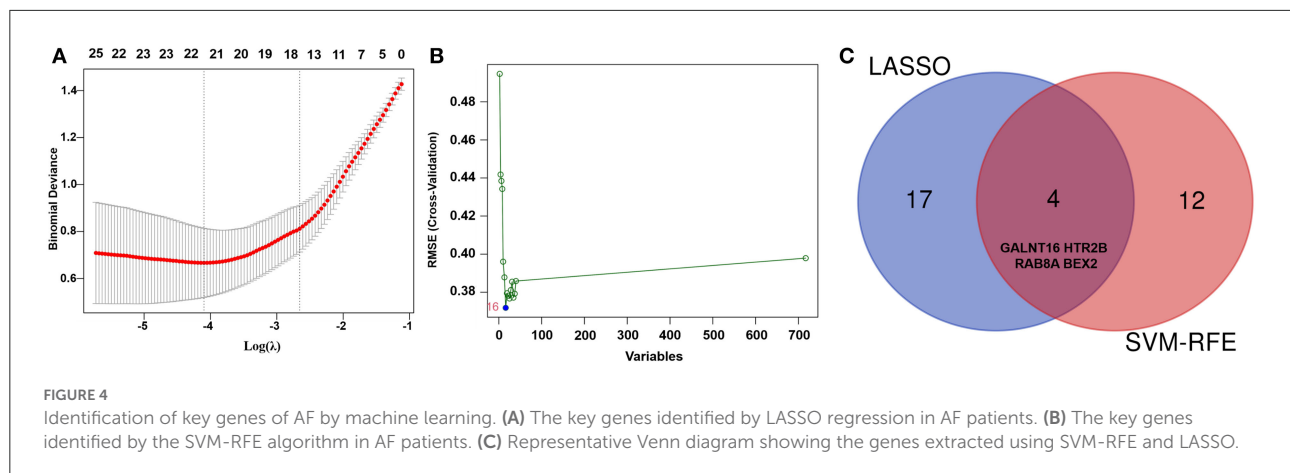
Identification of hub genes

To explore reliable serum biomarkers significantly associated with AF, LASSO regression and the SVM-RFE algorithm were implemented to evaluate the characteristic genes in AF according to the gene expression profile in the key modules. The characteristic genes ($n = 21$) were identified

by LASSO regression (Figure 4A). Another 16 key genes were identified by the SVM-RFE algorithm (Figure 4B). Then, for subsequent investigations, a total of 4 overlapping genes (*BEX2*, *GALNT16*, *RAB8A*, and *HTR2B*) were selected (Figure 4C). In addition, other genes identified with the SVM-RFE algorithm and LASSO regression are listed in Supplementary Table S5.

Profile of immune cell subtype distribution pattern

The CIBERSORT algorithm was utilized to evaluate the differential expression of immune fractions between SR



and AF samples. The cumulative histogram visually shows the relative proportions of various immune cell subtypes (Supplementary Figure S2). As shown in Figure 5A, significant differences in the immune cell proportions were found between the SR and AF groups of samples. A negative association between activated and resting mast cells was observed using a correlation matrix. Meanwhile, we also observed that Tregs were negatively associated with M1 macrophages and positively associated with activated NK cells; M2 macrophages were negatively associated with resting dendritic cells, eosinophils, monocytes, and resting memory CD4 T cells; CD8 T cells were negatively associated with resting memory CD4 T cells; memory B cells were positively associated with plasma cells (Figure 5B). Compared with SR subjects, AF patients generally exhibited decreased infiltration of activated mast cells and regulatory T cells (Tregs) and increased infiltration of resting mast cells, M2 macrophages, and gamma delta T cells (Figure 5C) ($P < 0.05$ – 0.01 , respectively). In addition, the immune cell infiltration pattern in AF is also shown in Supplementary Table S6.

As shown in Figure 6, we observed a positive association between the *BEX2* gene and activated mast cells; additionally, the *HTR2B* gene was positively correlated with gamma delta T cells, M2 macrophages, and M1 macrophages and negatively associated with activated mast cells and resting NK cells. In addition, negative correlations between the *GALNT16* gene and M2 macrophages and gamma delta T cells were also observed. However, there was no significant correlation between the *RAB8A* gene and immune cells ($P < 0.05$ – 0.01).

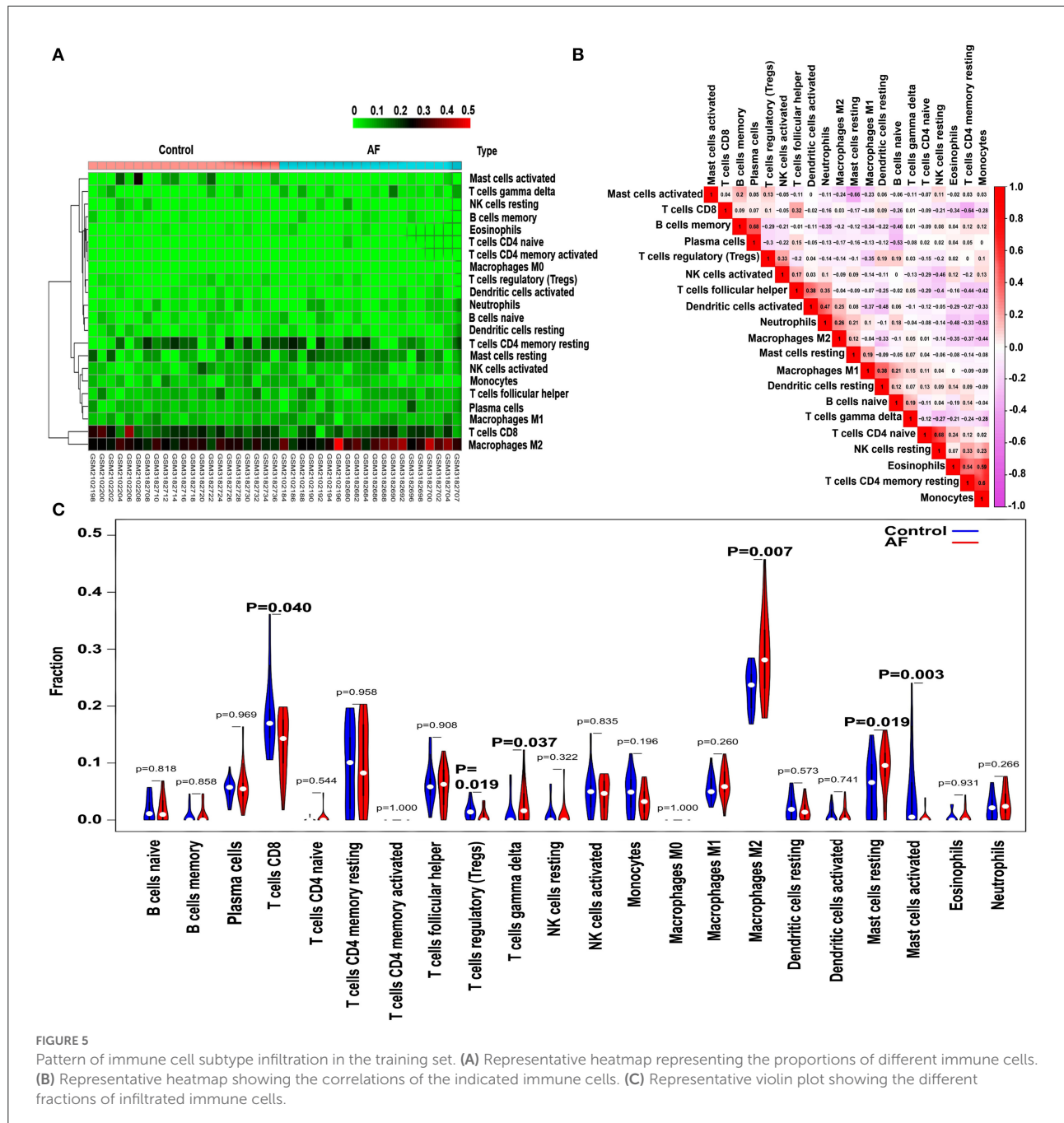
Validation of the key genes in AF patients

The predictive values of *BEX2*, *HTR2B*, *GALNT16* and *RAB8A* for the diagnosis of AF were investigated using ROC curve analysis. The AUC values for *BEX2* (Figure 7A), *HTR2B* (Figure 7B), *GALNT16* (Figure 7C) and *RAB8A* (Figure 7D)

were 0.836 (95% CI 0.709–0.962; $P = 0.0002$), 0.883 (95% CI 0.781–0.986; $P < 0.0001$), 0.893 (95% CI 0.796–0.990; $P < 0.0001$) and 0.841 (95% CI 0.714–0.968, $P < 0.0001$) in the training set, respectively. Further verifying the expression of these key genes in AF patients (Figure 8A), we noticed that *BEX2*, *GALNT16*, and *HTR2B* expression levels were significantly increased in AF patients compared with SR subjects ($P < 0.01$). However, no significant difference in *RAB8A* expression was found between AF patients and SR subjects. Additionally, the AUC values for *BEX2* (Figure 8B), *HTR2B* (Figure 8C), and *GALNT16* (Figure 8D) were 0.858 (95% CI 0.880–0.917; $P < 0.0001$), 0.861 (95% CI 0.803–0.919; $P < 0.0001$) and 0.915 (95% CI 0.875–0.955; $P < 0.0001$), respectively.

Discussion

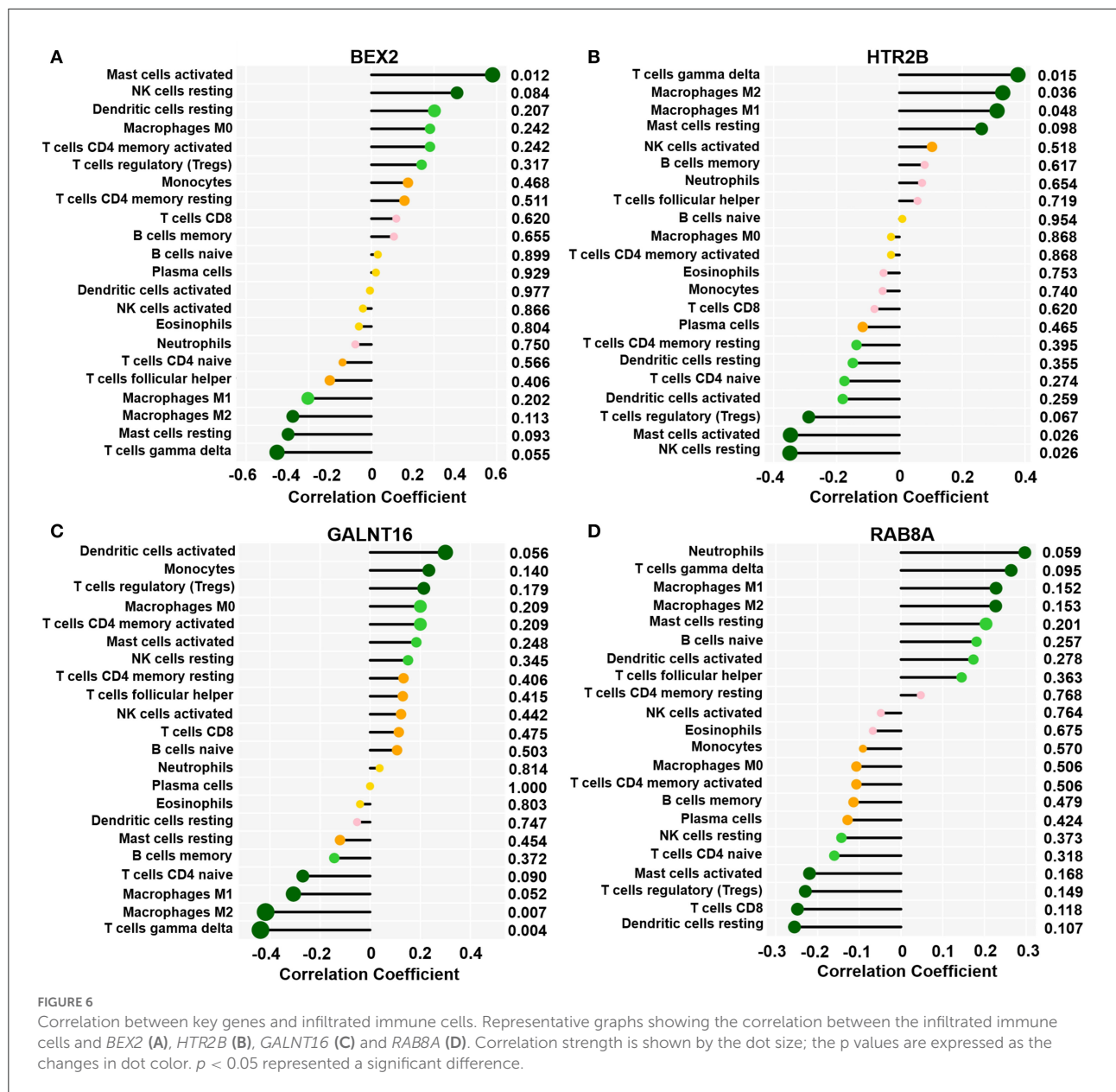
In the current research, GSE79768 combined with GSE115574 as training datasets were downloaded from the GEO database and analyzed using WGCNA. Then, three modules (pink, red, and magenta) were identified to be significantly associated with AF. Four hub genes (*GALNT16*, *HTR2B*, *RAB8A*, and *BEX2*) were revealed to be significantly correlated with AF by LASSO logistic regression and the SVM-RFE algorithm. The CIBERSORT results suggested decreased infiltration of regulatory T cells (Tregs) and activated mast cells and increased infiltration of resting mast cells, M2 macrophages, and gamma delta T cells in AF patients. qRT-PCR results revealed that *BEX2* and *GALNT16* expression levels were significantly decreased and *HTR2B* expression was significantly increased in AF patients compared with SR subjects. ROC analyses based on the training set and our clinical samples revealed that the *HTR2B*, *BEX2*, and *GALNT16* genes remained highly effective in distinguishing AF patients from normal SR subjects. Gene enrichment analysis indicated that these



key genes were mainly involved in several inflammatory- or immune-related signaling pathways and biological processes.

Through a comprehensive search of the NCBI GENE database, we revealed that brain-expressed X-linked 2 (*BEX2*, also known as *BEX1*; DJ79P11.1; HGNC: 30933, gene ID: 84707, OMIM: 300691) acts as a member of the brain expressed X-linked gene family. It has been reported that *BEX* proteins are mainly involved in tumor growth, neurodegeneration, the cell cycle, and transcriptional regulation (19). Kim

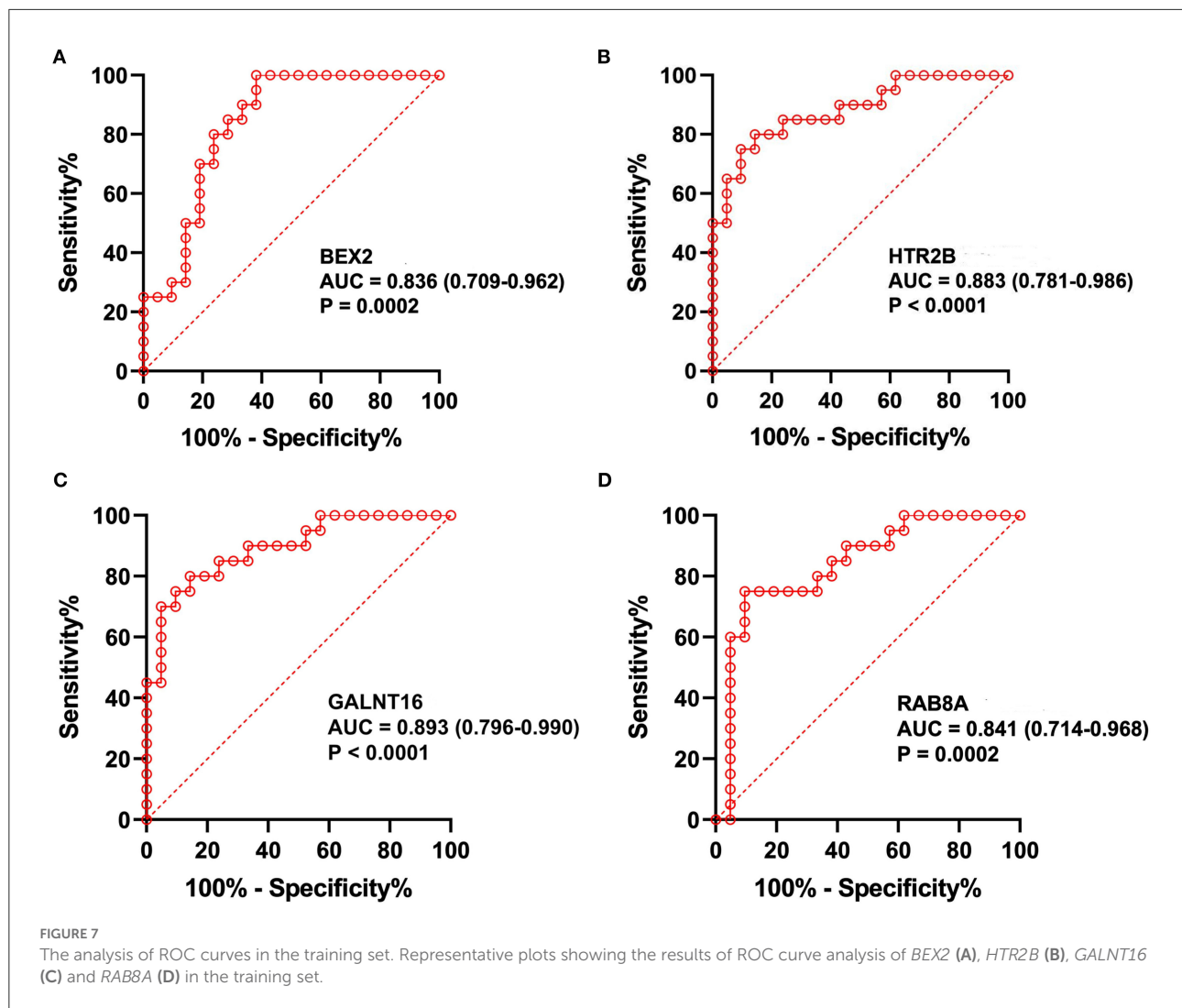
et al. suggested that the *BEX2* gene is closely related to the process and regulation of the immune response, and the expression of *BEX2* was significantly downregulated in asthmatic mouse models induced by ovalbumin, while the expression of *BEX2* was significantly upregulated and could significantly inhibit allergic airway inflammation after treatment with mesenchymal stem cell-derived extracellular vesicles (20). Previous reports documented that the remodeling of atrial structure and electrical properties was significantly associated



with the dysfunction and abnormal structure of mitochondria (21, 22), that the dysfunction of mitochondria was also involved in the altered cardiac electrical properties (23, 24), and that these biological processes ultimately contribute to increased susceptibility to AF. Shao et al. proved that empagliflozin can potentially be useful in the prevention of type 2 diabetes mellitus (T2DM)-related AF by ameliorating mitochondrial dysfunction and improving atrial structure and electrical remodeling in T2DM patients (25). Furthermore, Peng et al. noticed that soybean isoflavones can activate the BNIP3/NIX pathway by upregulating the expression of *BEX2*, thereby alleviating mitochondrial dysfunction by promoting mitochondrial autophagy (26). This evidence suggests that *BEX2*

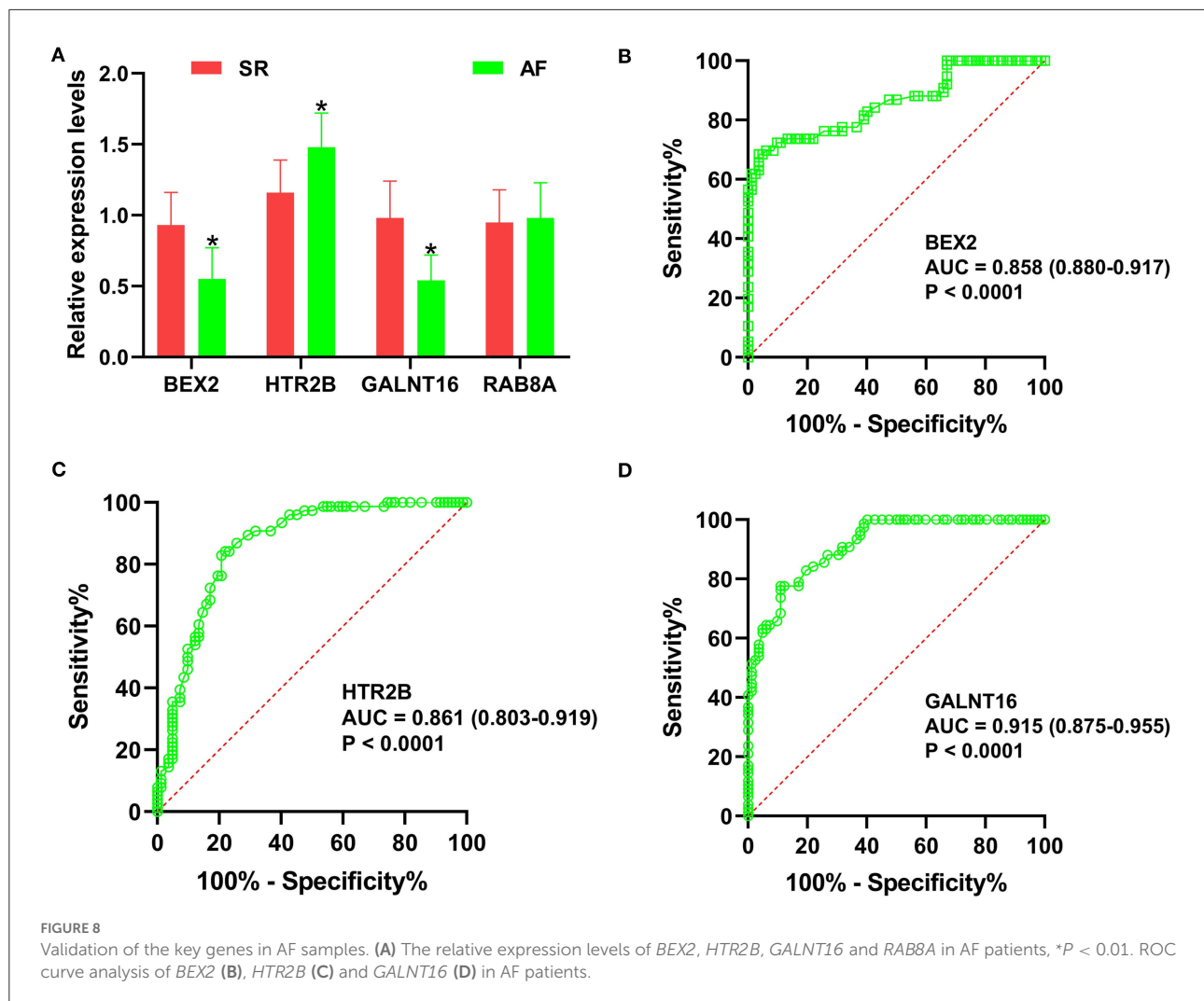
may reduce the incidence of AF by mitigating mitochondrial dysfunction and inducing inflammatory or immune responses. Herein, we found that AF patients exhibited significantly reduced *BEX2* expression in comparison with SR individuals, and a positive correlation between *BEX2* expression and activated mast cells was also observed. However, these findings need to be confirmed by proteomics studies in further research.

Previous studies have found that 5-hydroxytryptamine receptor 2B (*HTR2B*) can be enriched in the calcium signaling pathway, which plays a key role in the processes of myocardial hypertrophy and remodeling (27). Due to excessive release of 5-hydroxytryptamine (5-HT) in platelets or autonomic nervous system activation, Ca^{2+} overloading may trigger AF



(28, 29). As a 5-HT receptor, *HTR2B* activation significantly participates in cardiac remodeling as well as heart failure development and progression (30, 31). By impairing the deposition of collagen, blockage of *HTR2B* can suppress myocardial fibrosis (32). In addition, Nebigil et al. suggested that the overexpression of Htr2b in the mouse heart could lead to abnormal mitochondrial function of cardiac myocytes, resulting in cardiac hypertrophy (33). Furthermore, dramatically elevated expression of *HTR2B* mRNA and protein was observed in 72-h rapid electric-stimulated atrial myocytes (34). This evidence strongly suggests that *HTR2B* is significantly involved in AF development, but its association with immune cells has not been reported. In the present study, obviously increased *HTR2B* expression was observed in AF patients, and we also observed positive correlations between *HTR2B* expression and gamma delta T cells, M1 macrophages, and M2 macrophages and negative correlations with activated mast cells and resting NK cells. However, these observations need to be confirmed by further studies.

Glycosylation is an important posttranslational modification involving N-glycosylation (35) and O-glycosylation (36) and is significantly related to a variety of pathological and physiological processes. Previous studies have found that N-glycosylation is associated with the function of many ion channels, and an association between the dysfunction of glycosylation in potassium channels and long QT syndrome was observed (37). Congenital glycosylation disorder, with effects including defects in the synthesis, processing, or targeting of glycans, is a rare autosomal genetic disease (38). Almost all organs, arrhythmias, and cardiomyopathies are affected by these disorders (39). Moreover, altered glycation patterns in cardiac ion channels, such as hK_{2p}17.1, may contribute to the molecular mechanisms underlying the occurrence of arrhythmogenesis correlated with glycosylation disorders (40). Polypeptide N-acetylgalactosaminyltransferase 16 (*GALNT16*), as a key gene mediating protein glycosylation, can catalyze the initial reaction in the biosynthesis of O-linked oligosaccharides and transfer N-acetyl-D-galactosamine residues to serine or



threonine residues on protein receptors (41). Therefore, we surmise that the downregulation of *GALNT16* expression might lead to the occurrence of AF by causing glycosylation disorder. Moreover, *GALNT16* has been found to be significantly enriched in specific biological functions associated with protein and lipid metabolism, the AMPK signaling pathway, the prolactin signaling pathway, and the insulin/IGF pathway-protein kinase B signaling cascade (42), but the association of *GALNT16* with immune cells and AF susceptibility remains poorly understood. In the current research, we revealed that *GALNT16* expression in AF patients was dramatically decreased compared with that in normal subjects with SR, and the expression of *GALNT16* was also significantly negatively correlated with gamma delta T cells and M2 macrophages.

Recently, many studies have indicated that the inflammatory response also participates in many cardiac pathophysiological processes, such as postinfarction repair and ischaemic injury, which are characterized by immune regulation, cytokine expression, intracellular signaling pathways, and

neuroendocrine system activation. Monocyte subsets are also involved in the inflammatory cascade and atherogenesis in cardiovascular disease. Elevated monocyte counts and activities are significantly related to clinical indices of chronic kidney disease (CKD) and atherosclerosis (43). In addition, the participation of other immune cells, such as neutrophils (44) and mast cells (45), was also identified in the processes of cardiovascular disease occurrence and development. A significantly increased number of CD8+ T cells in AF patients was identified by Wu et al. (46). Liu et al. showed that mast cells and neutrophils were increased in AF atrial tissue in comparison to matched SR atrial tissue (47). In addition, Liu et al. suggested that resting memory CD4 T cells and follicular helper T cells were decreased and that plasma cells, monocytes, resting dendritic cells, and neutrophils were increased in AF samples compared with SR samples (48). Nevertheless, different patterns of immune cell infiltration in AF were observed in the current study compared with previous studies. Our results suggested decreased infiltration of regulatory T cells (Tregs) and activated

mast cells and increased infiltration of gamma delta T cells, M2 macrophages, and resting mast cells in AF patients compared with SR subjects. In addition, we also revealed the interactions among these 22 types of infiltrated immune cells in AF. We noticed that activated mast cells were negatively associated with resting mast cells; Tregs were negatively associated with M1 macrophages and positively associated with activated NK cells; M2 macrophages were negatively associated with resting dendritic cells, eosinophils, resting memory CD4 T cells, and monocytes; CD8 T cells were negatively associated with resting memory CD4 T cells; memory B cells were positively associated with plasma cells. These results suggested that there may be complex immune regulation mechanisms governing the occurrence and development of AF. However, the potential mechanisms of these correlations between infiltrated immune cells need to be verified by experiments *in vivo* and *in vitro*.

This research had several limitations. First, the validation samples included in the current research were recruited from only a single center with small sample sizes. It is unclear whether the expression levels of the key genes differ among individuals in different regions or races. Therefore, the results of this study need to be further tested in multicentre studies and using larger sample sizes. Second, although we found that the expression trends of key genes in blood samples of patients with AF were consistent with the bioinformatics prediction results based on left atrial samples, the expression trends of key genes still need to be further verified in other left atrial samples. Third, more *in vivo* and *in vitro* studies are needed to clarify the underlying mechanisms of these correlations between *GALNT16*, *HTR2B*, *RAB8A*, and *BEX2* as well as infiltrating immune cells in AF.

Conclusions

In summary, we determined that *BEX2*, *HTR2B*, and *GALNT16* may become potential diagnostic markers or novel therapeutic targets in AF. We noticed that resting mast cells, M2 macrophages, and gamma delta T cells might be involved in AF initiation; however, activated mast cells, CD8 T cells, and regulatory T cells (Tregs) may serve protective roles in AF. The interaction mechanisms between *BEX2*, *HTR2B*, and *GALNT16* and immune cells may be of great significance to the pathogenesis and progression of AF.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding authors.

Ethics statement

The study design was approved by the Ethics Committee of Hunan Provincial People's Hospital (No: LL-20210615-144). The patients/participants provided their written informed consent to participate in this study.

Author contributions

P-FZ conceived the study, participated in the design, performed the statistical analyses, and drafted the manuscript. Z-YL and HWP conceived the study, participated in the design, and helped to draft the manuscript. Z-YL carried out the epidemiological survey and collected the samples. L-ZC and PL performed the statistical analyses. All authors read and approved the final manuscript.

Funding

This study was supported by the Key Research and Development Program of Hunan Province (No: 2019SK2021) and the Natural Science Foundation of Hunan Province (NO:2020JJ4406). There was no role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2022.922523/full#supplementary-material>

References

- Zhang YF, Meng LB, Hao ML, Yang JF, Zou T. Identification of co-expressed genes between atrial fibrillation and stroke. *Front Neurol.* (2020) 11:184. doi: 10.3389/fneur.2020.00184
- Lozano-Velasco E, Franco D, Aranega A, Daimi H. Genetics and epigenetics of atrial fibrillation. *Int J Mol Sci.* (2020) 21. doi: 10.3390/ijms21165717
- Panchal G, Mahmood M, Lip GYH. Revisiting the risks of incident atrial fibrillation: a narrative review. *Part 1 Kardiol Pol.* (2019) 77:430–36. doi: 10.33963/KP.14806
- Nguyen TN, Hilmer SN, Cumming RG. Review of epidemiology and management of atrial fibrillation in developing countries. *Int J Cardiol.* (2013) 167:2412–20. doi: 10.1016/j.ijcard.2013.01.184
- Liu Y, Shi Q, Ma Y, Liu Q. The role of immune cells in atrial fibrillation. *J Mol Cell Cardiol.* (2018) 123:198–208. doi: 10.1016/j.yjmcc.2018.09.007
- Hernández Madrid A, Moro C. Atrial fibrillation and C-reactive protein: searching for local inflammation. *J Am Coll Cardiol.* (2007) 49:1649–50. doi: 10.1016/j.jacc.2007.02.009
- Liu L, Zheng Q, Lee J, Ma Z, Zhu Q, Wang Z. PD-1/PD-L1 expression on CD(4+) T cells and myeloid DCs correlates with the immune pathogenesis of atrial fibrillation. *J Cell Mol Med.* (2015) 19:1223–33. doi: 10.1111/jcmm.12467
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* (2015) 12:453–57. doi: 10.1038/nmeth.3337
- Miller JA, Woltjer RL, Goodenbour JM, Horvath S, Geschwind DH. Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med.* (2013) 5:48. doi: 10.1186/gm452
- Radulescu E, Jaffe AE, Straub RE, Chen Q, Shin JH, Hyde TM, et al. Identification and prioritization of gene sets associated with schizophrenia risk by co-expression network analysis in human brain. *Mol Psychiatry.* (2020) 25:791–804. doi: 10.1038/s41380-018-0304-1
- Chen R, Liu X, Jin S, Lin J, Liu J. Machine learning for drug-target interaction prediction. *Molecules.* (2018) 23. doi: 10.3390/molecules23092208
- Wei S, Lu J, Lou J, Shi C, Mo S, Shao Y, et al. Gastric cancer tumor microenvironment characterization reveals stromal-related gene signatures associated with macrophage infiltration. *Front Genet.* (2020) 11:663–63. doi: 10.3389/fgene.2020.00663
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007
- Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol.* (2008) 4:e1000117. doi: 10.1371/journal.pcbi.1000117
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* (2010) 33:1–22. doi: 10.18637/jss.v033.i01
- Wang Q, Liu X. Screening of feature genes in distinguishing different types of breast cancer using support vector machine. *Oncol Targets Ther.* (2015) 8:2311–17. doi: 10.2147/OTT.S85271
- Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. *Eur Heart J.* (2021) 42:373–498. doi: 10.1093/eurheartj/ehab648
- Fernandez EM, Diaz-Ceso MD, Vilar M. Brain expressed and X-linked (Bex) proteins are intrinsically disordered proteins (IDPs) and form new signaling hubs. *PLoS ONE.* (2015) 10:e0117206. doi: 10.1371/journal.pone.0117206
- Kim S-D, Kang SA, Kim Y-W, Yu HS, Cho K-S, Roh H-J. Screening and functional pathway analysis of pulmonary genes associated with suppression of allergic airway inflammation by adipose stem cell-derived extracellular vesicles. *Stem Cells Int.* (2020) 2020:5684250–50. doi: 10.1155/2020/5684250
- Ad N, Schneider A, Khaliulini I, Borman JB, Schwalb H. Impaired mitochondrial response to simulated ischemic injury as a predictor of the development of atrial fibrillation after cardiac surgery: in vitro study in human myocardium. *J Thorac Cardiovasc Surg.* (2005) 129:41–5. doi: 10.1016/j.jtcvs.2004.03.058
- Torrebalba N, Aranguiz P, Alonso C, Rothermel BA, Lavandero S. Mitochondria in structural and functional cardiac remodeling. *Adv Exp Med Biol.* (2017) 982:277–306. doi: 10.1007/978-3-319-55330-6_15
- Zhang X, Zhang Z, Zhao Y, Jiang N, Qiu J, Yang Y, et al. Alogliptin, a dipeptidyl peptidase-4 inhibitor, alleviates atrial remodeling and improves mitochondrial function and biogenesis in diabetic rabbits. *J Am Heart Assoc.* (2017) 6. doi: 10.1161/JAHA.117.005945
- Kurokawa S, Niwano S, Niwano H, Ishikawa S, Kishihara J, Aoyama Y, et al. Progression of ventricular remodeling and arrhythmia in the primary hyperoxidative state of glutathione-depleted rats. *Circ J.* (2011) 75:1386–93. doi: 10.1253/circj.CJ-10-1089
- Shao Q, Meng L, Lee S, Tse G, Gong M, Zhang Z, et al. Empagliflozin, a sodium glucose co-transporter-2 inhibitor, alleviates atrial remodeling and improves mitochondrial function in high-fat diet/streptozotocin-induced diabetic rats. *Cardiovasc Diabetol.* (2019) 18:165. doi: 10.1186/s12933-019-0964-4
- Li P, Yao LY, Jiang YJ, Wang DD, Wang T, Wu YP, et al. Soybean isoflavones protect SH-SY5Y neurons from atrazine-induced toxicity by activating mitophagy through stimulation of the BEX2/BNIP3/NIX pathway. *Ecotoxicol Environ Saf.* (2021) 227:112886. doi: 10.1016/j.ecoenv.2021.112886
- Zhang J, Huang X, Wang X, Gao Y, Liu L, Li Z, et al. Identification of potential crucial genes in atrial fibrillation: a bioinformatic analysis. *BMC Med Genomics.* (2020) 13:104–04. doi: 10.1186/s12920-020-00754-5
- Dobrev D. Electrical Remodeling in Atrial Fibrillation. *Herz.* (2006) 31:108–12. doi: 10.1007/s00059-006-2787-9
- Kaumann AJ, Levy FO. 5-hydroxytryptamine receptors in the human cardiovascular system. *Pharmacol Ther.* (2006) 111:674–706. doi: 10.1016/j.pharmthera.2005.12.004
- Jaffré F, Callebert J, Sarre A, Etienne N, Nebigil CG, Launay JM, et al. Involvement of the serotonin 5-HT2B receptor in cardiac hypertrophy linked to sympathetic stimulation: control of interleukin-6, interleukin-1beta, and tumor necrosis factor-alpha cytokine production by ventricular fibroblasts. *Circulation.* (2004) 110:969–74. doi: 10.1161/01.CIR.0000139856.20505.57
- Monassier L, Laplante MA, Jaffré F, Bousquet P, Maroteaux L, de Champlain J. Serotonin 5-HT(2B) receptor blockade prevents reactive oxygen species-induced cardiac hypertrophy in mice. *Hypertension.* (2008) 52:301–7. doi: 10.1161/HYPERTENSIONAHA.107.105551
- Janssen W, Schymura Y, Novoyatleva T, Kojonazarov B, Boehm M, Wietelmann A, et al. 5-HT2B receptor antagonists inhibit fibrosis and protect from RV heart failure. *Biomed Res Int.* (2015) 2015:438403–03. doi: 10.1155/2015/438403
- Nebigil CG, Jaffré F, Messaddeq N, Hickel P, Monassier L, Launay JM, et al. Overexpression of the serotonin 5-HT2B receptor in heart leads to abnormal mitochondrial function and cardiac hypertrophy. *Circulation.* (2003) 107:3223–9. doi: 10.1161/01.CIR.0000074224.57016.01
- Wang X, Li H, Zhang A, Zhang Y, Li Z, Wang X, et al. Diversity among differentially expressed genes in atrial appendages of atrial fibrillation: The role and mechanism of SPP1 in atrial fibrosis. *Int J Biochem Cell Biol.* (2021) 141:106074. doi: 10.1016/j.biocel.2021.106074
- Esmail S, Manolson MF. Advances in understanding N-glycosylation structure, function, and regulation in health and disease. *Eur J Cell Biol.* (2021) 100:151186. doi: 10.1016/j.ejcb.2021.151186
- Magalhães A, Duarte HO, Reis CA. The role of O-glycosylation in human disease. *Mol Aspects Med.* (2021) 79:100964. doi: 10.1016/j.mam.2021.100964
- Petreccka K, Atanasiu R, Akhavan A, Shrier A. N-linked glycosylation sites determine HERG channel surface membrane expression. *J Physiol.* (1999) 515:41–8. doi: 10.1111/j.1469-7793.1999.041ad.x
- Freeze HH. Genetic defects in the human glycome. *Nat Rev Genet.* (2006) 7:537–51. doi: 10.1038/nrg1894
- Gehrmann J, Sohlbach K, Linnebank M, Böhles HJ, Buderus S, Kehl HG, et al. Cardiomyopathy in congenital disorders of glycosylation. *Cardiol Young.* (2003) 13:345–51. doi: 10.1017/S1047951103000702
- Wiedmann F, Schlund D, Voigt N, Ratte A, Kraft M, Katus HA, et al. N-glycosylation-dependent regulation of hK(2P)171 currents. *Mol Biol Cell.* (2019) 30:1425–36. doi: 10.1091/mbc.E18-10-0687
- Wu H, He G, Song T, Zhang Y, Chen X, Chen H, et al. Evaluation of GALNT16 polymorphisms to breast cancer risk in Chinese population. *Molecular genetics & genomic medicine.* (2019) 7:e848–e48. doi: 10.1002/mgg3.848

42. Gao Y, Jiang J, Yang S, Hou Y, Liu GE, Zhang S, et al. CNV discovery for milk composition traits in dairy cattle using whole genome resequencing. *BMC Genomics*. (2017) 18:265. doi: 10.1186/s12864-017-3636-3
43. Dounousi E, Duni A, Naka KK, Vartholomatos G, Zoccali C. The innate immune system and cardiovascular disease in eskd: monocytes and natural killer cells. *Curr Vasc Pharmacol*. (2021) 19:63–76. doi: 10.2174/1570161118666200628024027
44. Grégory F. Role of mechanical stress and neutrophils in the pathogenesis of plaque erosion. *Atherosclerosis*. (2021) 318:60–9. doi: 10.1016/j.atherosclerosis.2020.11.002
45. Varricchi G, Marone G, Kovanen PT. Cardiac mast cells: underappreciated immune cells in cardiovascular homeostasis and disease. *Trends Immunol*. (2020) 41:734–46. doi: 10.1016/j.it.2020.06.006
46. Wu J, Deng H, Chen Q, Wu Q, Li X, Jiang S, et al. Comprehensive analysis of differential immunocyte infiltration and potential cerna networks involved in the development of atrial fibrillation. *Biomed Res Int*. (2020) 2020:8021208. doi: 10.1155/2020/8021208
47. Liu J, Liu M, Chen X. Identification of potential key biomarkers of atrial fibrillation and their correlation with immune infiltration in atrial tissue. *Comput Math Methods Med*. (2022) 2022:4029840–40. doi: 10.1155/2022/4029840
48. Liu A, Jia K, Liang H, Jin Q. Comprehensive analysis of autophagy-related genes and patterns of immune cell infiltration in valvular atrial fibrillation. *BMC Cardiovasc Disord*. (2021) 21:132–32. doi: 10.1186/s12872-021-01939-1