



## OPEN ACCESS

## EDITED BY

Diego Franco,  
University of Jaén, Spain

## REVIEWED BY

Konstantinos Chatzistergos,  
Aristotle University of Thessaloniki, Greece  
Xumin Guan,  
First Affiliated Hospital of Dalian Medical  
University, China

## \*CORRESPONDENCE

Zhijun Sun,  
✉ 18940251218@163.com

RECEIVED 20 March 2023

ACCEPTED 29 December 2023

PUBLISHED 11 January 2024

## CITATION

Tong F and Sun Z (2024), Identification and validation of potential biomarkers for atrial fibrillation based on integrated bioinformatics analysis.

*Front. Cell Dev. Biol.* 11:1190273.  
doi: 10.3389/fcell.2023.1190273

## COPYRIGHT

© 2024 Tong and Sun. 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.

# Identification and validation of potential biomarkers for atrial fibrillation based on integrated bioinformatics analysis

Fei Tong and Zhijun Sun\*

Department of Cardiology, Shengjing Hospital of China Medical University, Shenyang, China

**Background:** Globally, the most common form of arrhythmias is atrial fibrillation (AF), which causes severe morbidity, mortality, and socioeconomic burden. The application of machine learning algorithms in combination with weighted gene co-expression network analysis (WGCNA) can be used to screen genes, therefore, we aimed to screen for potential biomarkers associated with AF development using this integrated bioinformatics approach.

**Methods:** On the basis of the AF endocardium gene expression profiles GSE79768 and GSE115574 from the Gene Expression Omnibus database, differentially expressed genes (DEGs) between AF and sinus rhythm samples were identified. DEGs enrichment analysis and transcription factor screening were then performed. Hub genes for AF were screened using WGCNA and machine learning algorithms, and the diagnostic accuracy was assessed by the receiver operating characteristic (ROC) curves. GSE41177 was used as the validation set for verification. Subsequently, we identified the specific signaling pathways in which the key biomarkers were involved, using gene set enrichment analysis and reverse prediction of mRNA–miRNA interaction pairs. Finally, we explored the associations between the hub genes and immune microenvironment and immune regulation.

**Results:** Fifty-seven DEGs were identified, and the two hub genes, hypoxia inducible factor 1 subunit alpha inhibitor (*HIF1AN*) and mitochondrial inner membrane protein MPV17 (*MPV17*), were screened using WGCNA combined with machine learning algorithms. The areas under the receiver operating characteristic curves for *MPV17* and *HIF1AN* validated that two genes predicted AF development, and the differential expression of the hub genes was verified in the external validation dataset. Enrichment analysis showed that *MPV17* and *HIF1AN* affect mitochondrial dysfunction, oxidative stress, gap junctions, and other signaling pathway functions. Immune cell infiltration and

**Abbreviations:** AF, Atrial fibrillation; AUC, Area under the receiver operating characteristic curve; DC, Dendritic cell; DEG, Differentially expressed gene; GEO, Gene expression omnibus; GO, Gene ontology; GSEA, Gene set enrichment analysis; GSEA, Gene set variation analysis; GWAS, Genome-wide association study; HIF-1 $\alpha$ , Hypoxia inducible factor 1-alpha; IFN, Interferon; IL, interleukin; LASSO, Least absolute shrinkage and selection operator; MMP, Matrix metalloproteinase; NES, Normalized enrichment score; NK, Natural killer; ROC, Receiver operating characteristic; ROS, Reactive oxygen species; SVM-RFE, Support vector machine based recursive feature elimination; SR, Sinus rhythm; TGF, Transforming growth factor; Th T, helper; TNF, Tumor necrosis factor; TOM, Topological overlap matrix; WGCNA, Weighted gene co-expression network analysis.

immunomodulatory correlation analyses showed that *MPV17* and *HIF1AN* are strongly correlated with the content of immune cells and significantly correlated with *HLA* expression.

**Conclusion:** The identification of hub genes associated with AF using WGCNA combined with machine learning algorithms and their correlation with immune cells and immune gene expression can elucidate the molecular mechanisms underlying AF occurrence. This may further identify more accurate and effective biomarkers and therapeutic targets for the diagnosis and treatment of AF.

#### KEYWORDS

atrial fibrillation, bioinformatics analyses, *MPV17*, *HIF1AN*, weighted gene co-expression network analysis

## 1 Introduction

Atrial fibrillation (AF) is the most prevalent age-associated arrhythmias occurring in approximately 8.5% of patients aged >65 years, and the lifetime AF risk is estimated to be 1 in 3 individuals of European ancestry (Virani et al., 2021). AF substantially affects patients' quality of life and confers the risk of morbidity and mortality through stroke and heart failure. In addition to conventional risk factors such as age, hypertension, diabetes, male gender, obesity, obstructive sleep apnea, and endurance exercise (Myrstad et al., 2014; Schnabel et al., 2015; Virani et al., 2021), a hereditary component for AF risk has also been well recognized (Lubitz et al., 2010). The mechanisms underlying AF are not comprehensively clear, and thus, currently available therapeutic options, such as antiarrhythmic medications and catheter ablation procedures, have limited efficacy and adverse effects (Prystowsky et al., 2015). Emerging gene therapy for AF rhythm control may provide a novel and promising approach to AF (Farraha et al., 2016; Yoo et al., 2020). A preliminary genome-wide association study (GWAS) meta-analysis identified at least 134 genetic loci significantly associated with AF risk (Roselli et al., 2018) and increased our understanding of the pathophysiology of AF; however, much of the heritability of AF has been uncharted or inconclusive thus far (Christophersen and Ellinor, 2016). Therefore, there is a growing interest in developing new therapeutic approaches and exploring effective prognostic models for screening patients with high-risk AF.

With the continuous improvement of high-throughput technologies, several bioinformatics databases have been developed, which provide new perspectives for researchers to explore biomarkers and underlying mechanisms of diseases more precisely and effectively. However, traditional experimental analyses require long-term exploration due to the large amount of data, and the accuracy of bioinformatics analysis is influenced by duplicate data and covariance in database. Weighted gene co-expression network analysis (WGCNA) is a bioinformatics method that aggregates genes with the same or similar expression patterns into a module and analyzes each module in association with phenotypic data to identify potential key genes (Langfelder and Horvath, 2008). WGCNA can identify genes that may have great influences on disease development and is thus widely used to identify disease biomarkers. Machine learning has also garnered increasing interest in the screening of disease biomarkers (Wang et al., 2022). Support vector machine based recursive feature elimination (SVM-RFE) is a machine learning algorithm that

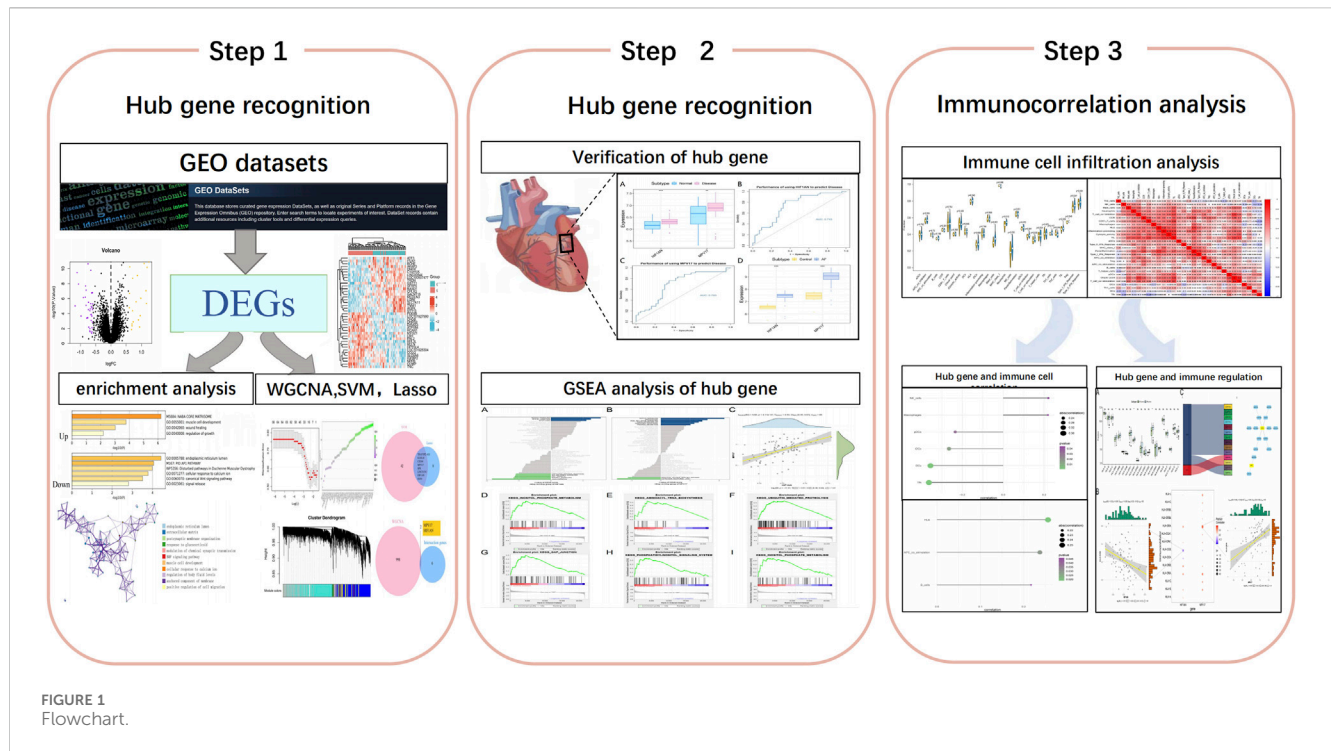
ranks different genes or features based on the sum of squares of feature coefficients, minimizing empirical errors and thus effectively screening genes with significant impact (Duan et al., 2005). Least absolute shrinkage and selection operator (LASSO) is another machine learning algorithm that fits a generalized linear model while performing variable selection and regularization, which can effectively reduce the effect of covariance and thus screen genes with a significant association between variables (Duan et al., 2016). The application of the two machine learning algorithms SVM-RFE and LASSO, in combination with WGCNA, can effectively avoid identical or similar data as well as covariance and thus screen the hub genes from transcriptome data. However, their application in screening for potential AF biomarkers is rarely reported.

In our study, we aimed to screen for key biomarkers potentially related to the development of AF. We also investigated the potential molecular mechanisms of these biomarkers and their association with immune cell infiltration. We screened differentially expressed genes (DEGs) and related pathways between AF and sinus rhythm (SR) samples using two AF microarray datasets retrieved from the Gene Expression Omnibus (GEO) database. WGCNA and the two above-mentioned machine learning algorithms were applied to screen key biomarkers, and the results were validated using an external validation dataset. Subsequently, single-sample gene set enrichment analysis (ssGSEA) was used to identify the specific signaling pathways in which the key biomarkers were involved. We also quantified the percentages of 28 immune cells in the AF samples and SR samples respectively and the subsequent association between identified biomarkers and infiltrating immune cells. These processes may inform studies on the pathogenesis of AF and the development of new immunotherapeutic targets.

## 2 Materials and methods

### 2.1 Data collection

The workflow of this study is shown in Figure 1. In this study, the microarray datasets were downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>), including GSE79768, GSE115574 and GSE41177. GSE79768 and GSE115574 databases as training set contained the gene data of 42 atrial samples from patients with AF and 43 atrial samples from patients with SR, while GSE41177 database as the validation set included 32 left atrial samples of AF and 6 left atrial samples of SR. The “sva” package



in R was used to correct the data between chips for subsequent analysis. All data analyzed in our study were achieved from public databases and thus did not require ethics committee approval.

## 2.2 Enrichment analysis of DEGs and screening of transcription factors

To determine the biological functions and signaling pathways in disease progression, DEGs were annotated and visualized by Metascape ([www.metascape.org](http://www.metascape.org)), and Gene Ontology (GO) analysis was applied for specific genes. A minimum overlap of  $\geq 3$  and  $p \leq 0.01$  were regarded as statistical significance. We explored transcription factors based on the DEGs to elucidate the potential molecular mechanisms by which these genes affect AF development. We extracted upstream motifs based on the DEGs, determined the normalized enrichment score (NES), and performed enrichment analysis for each motif by cumulative recovery curves. The R package RcisTarget was applied to predict relevant transcription factors. We used the rcistarget, hg19, motifdb.cisbpont, 500bp base to access the gene-motif ranking database.

## 2.3 Construction of machine learning-based models

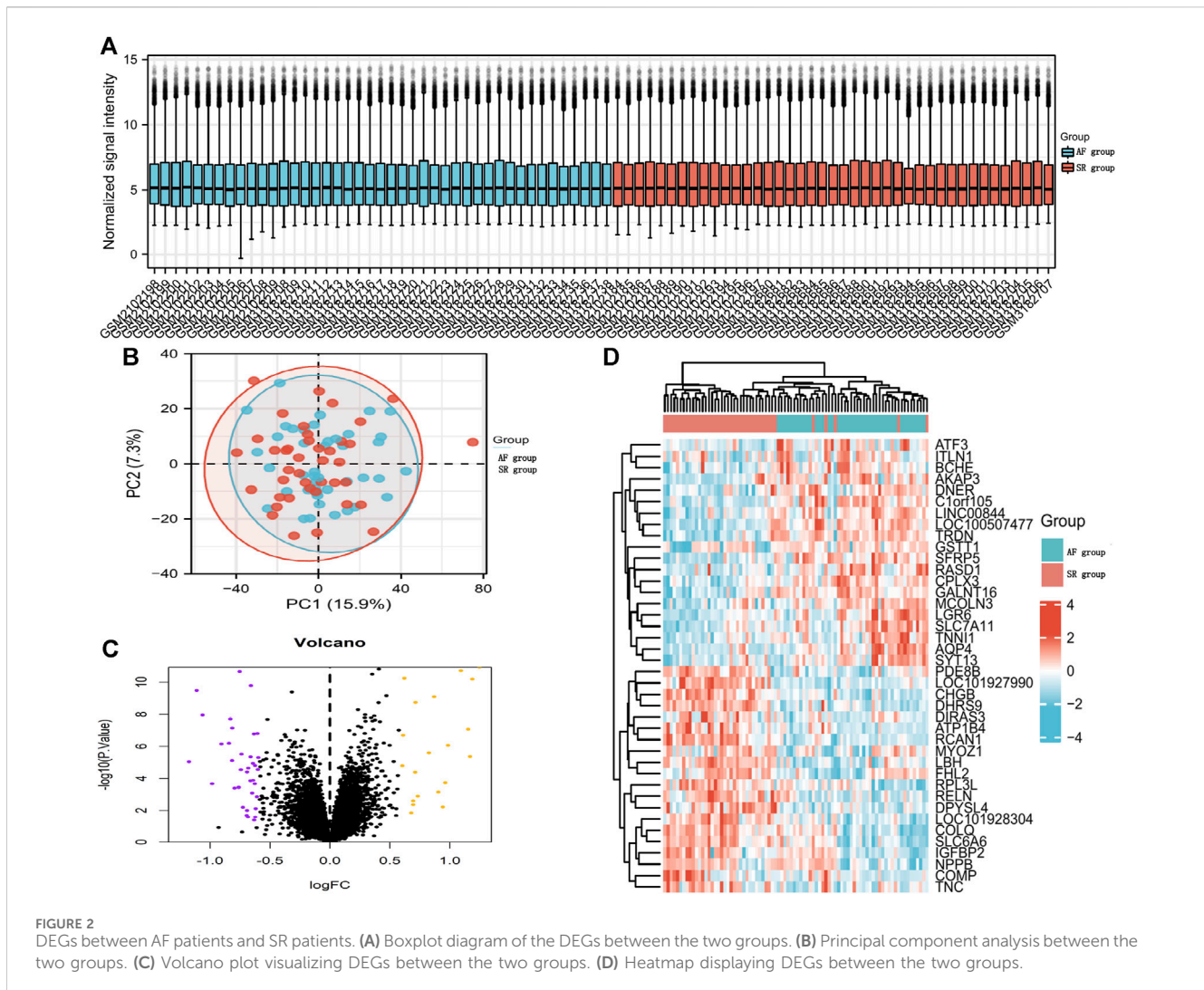
The LASSO regression algorithm and the SVM-RFE algorithm were used to identify the diagnostic biomarkers for diseases. First, we implemented LASSO regression using the “glmnet” R package to filter and visualize the DEGs. Then, we constructed SVM-RFE models using the “e1071” package to further identify the biomarkers for diseases. Finally, we integrated the screened characteristic genes to obtain optimal characteristic genes, using a Venn diagram.

## 2.4 WGCNA

WGCNA is a systematic biology approach to characterize gene association between patients with SR and those with AF, and to identify highly co-varying gene sets and potential biomarker genes or therapeutic targets based on the endogeneity of gene sets. Transcriptome data were read and imported using the R package “WGCNA” (<http://www.r-project.org/>), followed by which the “hclust” function was used to analyze the dataset for any significant outliers. A soft threshold filter was used to make the constructed network compatible with the characteristics of a scale-free network. Next, the weighted adjacency matrix was transformed into a topological overlap matrix (TOM) to assess the connectivity in the network, and a clustering tree structure of TOM was constructed using a hierarchical clustering method. Different branches of the clustering tree represent different gene modules by different colors according to the weighted correlation coefficients of genes. The conservativeness of the modules was evaluated based on the Z-summary score. In our study, the input WGCNA module genes were crossed with optimal characteristic genes from the machine learning to obtain the hub genes.

## 2.5 Identification and validation of hub genes

Box-line plots were used to assess the expression levels of the identified hub genes in AF and SR (control) groups respectively. Subsequently, receiver operating characteristic (ROC) curves were used to assess the diagnostic efficacy of the hub genes by “pROC” package. An external database of GSE41177 was employed to validate the hub genes identified in the training set by the Wilcoxon test.



## 2.6 Gene set enrichment analysis (GSEA)

GSEA was used to assess the distribution of the hub genes in a gene table ranked by disease relevance to determine their contribution to disease development by examining whether the hub genes are enriched at the top or bottom of the table. In this study, GSEA was used to compare the differences in signaling pathways between the high- and low-expression groups and to explore the potential molecular mechanisms underlying the discrepancies of AF and SR groups, where the number of substitutions was set to 1000 and the type of substitution was set to phenotype.

## 2.7 Immune cell infiltration and immune correlation analyses

To further understand the status of infiltrating immune cells in these two groups, we analyzed the RNA-sequencing data of these two groups, using ssGSEA, which inferred the relative proportion of infiltrating immune cells. The interaction between 22 immune cell types was then analyzed by the “corrplot” package to further display the interaction of immune cells. Furthermore, the relative extent of

immune cell infiltration was plotted by the “vioplot” package to evaluate the effect of the hub genes on immune infiltration. Spearman correlation analysis was performed for gene expression as well as the extent of immune cell infiltration.

## 2.8 Statistical analysis

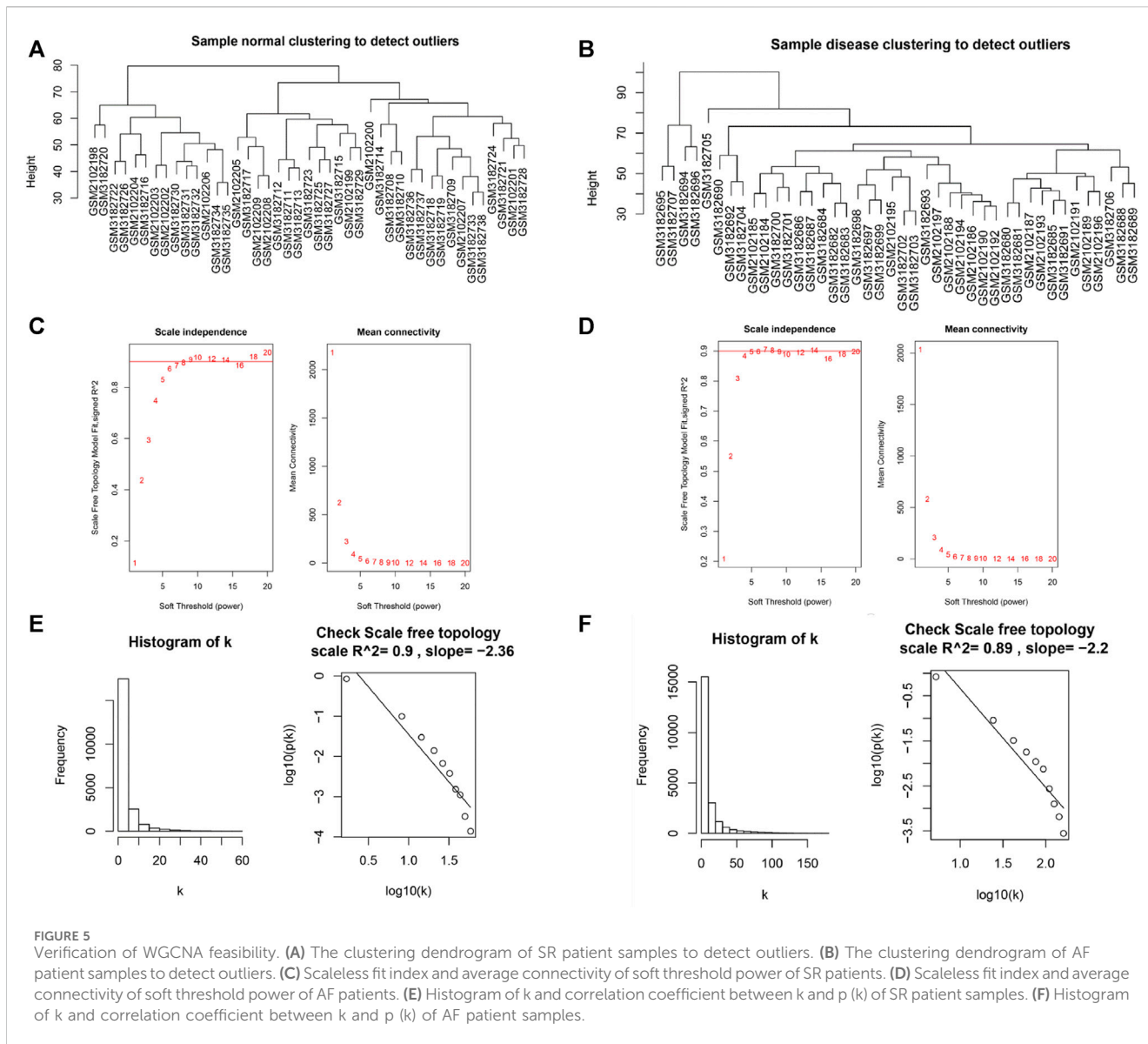
Statistical analyses were performed using R (version 3.6). All tests were two-sided with a significance level of  $P$  of  $<0.05$ .

## 3 Results

### 3.1 Identification of DEGs

We downloaded GSE79768 and GSE115574 AF datasets from GEO and divided 85 patients into SR ( $n = 43$ ) and AF ( $n = 42$ ) groups. After merging microarray data, the batch effects were adjusted by the “sva” package in R. Normalization was displayed by a boxplot (Figure 2A) and clustering between sample subgroups was displayed by a





principal component analysis plot (Figure 2B). DEG analysis was performed by the linear model (limma) package in R, and 57 DEGs (20 upregulated and 37 downregulated) with  $P$  of  $<0.05$  and fold change (FC) of  $\geq 0.58$  were displayed by a volcano plot in Figure 2C. In addition, the top 20 upregulated DEGs versus the top 20 downregulated DEGs were shown in the heat map in Figure 2D.

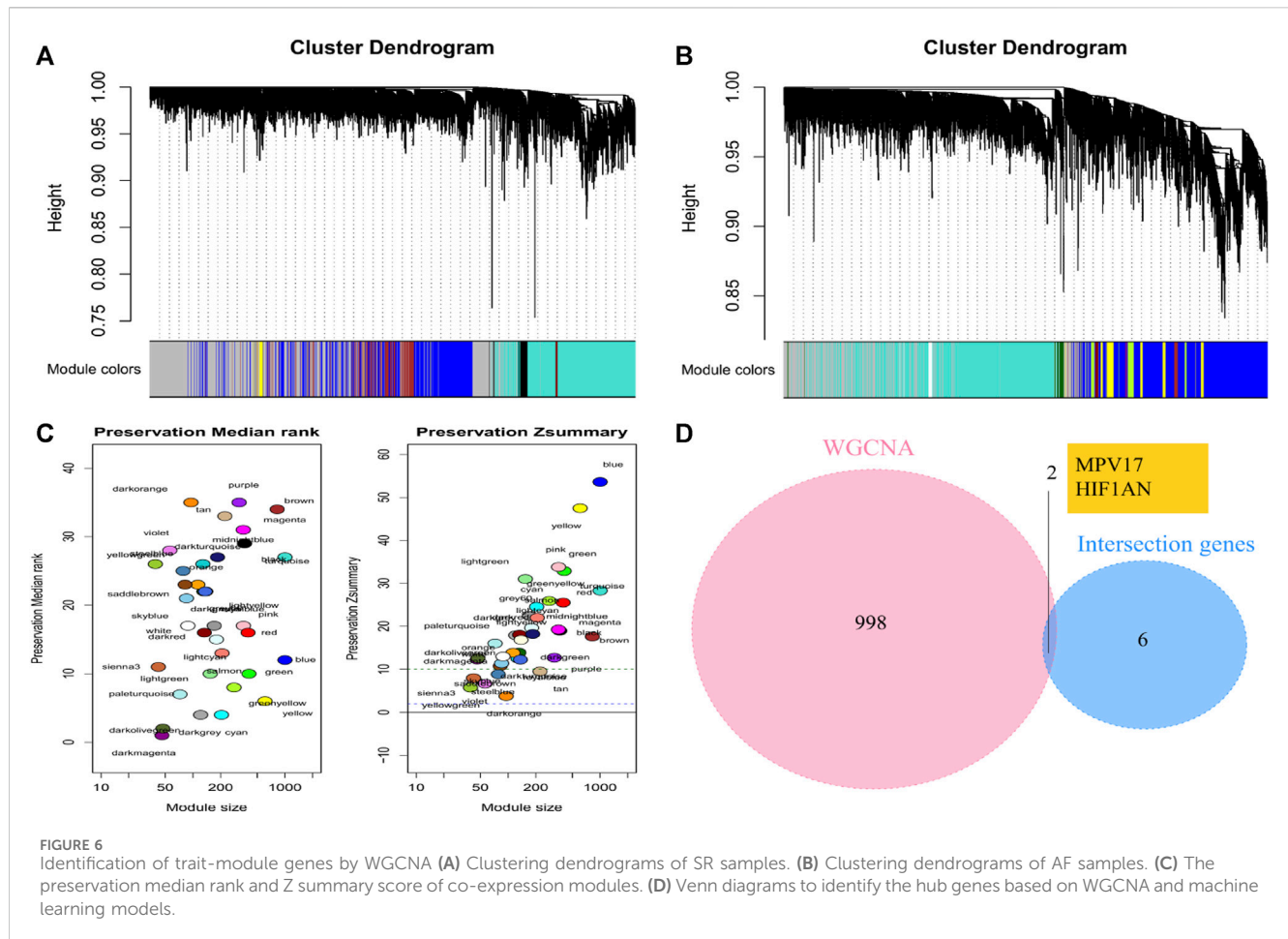
### 3.2 Functional enrichment analysis of DEGs and screening of transcription factors

Metascape analysis indicated that the upregulated DEGs and downregulated DEGs were mostly enriched in the core matrisome pathway and endoplasmic reticulum lumen pathway respectively. These results suggested that different mRNAs fulfilled different physiological functions and were involved in the regulation of AF (Figure 3A). We further explored the downstream regulatory

mechanisms of the DEGs and found them to be regulated by multiple transcription factors via a common mechanism. Therefore, we implemented an enrichment analysis of these transcription factors through cumulative recovery curves and Motif-TF annotation and then screened for important genes. The transcription factor of IKZF1 annotated as a motif of cisbp\_M3450 was the main regulator in the gene set; 10 DEGs were enriched in this motif, and the NES was 4.87. The results showed that all the motifs and corresponding transcription factors were enriched in 10 DEGs (Figure 3B).

### 3.3 Construction of machine learning models

LASSO regression and SVM-RFE were performed to identify the characteristic genes that were most relevant to AF. LASSO identified 16 genes of AF (Figures 4A, B). 50 genes with the highest accuracy screened by SVM-RFE (Figures 4C–E) took the intersection with the



genes identified by LASSO regression, and 8 characteristic genes of AF were ultimately identified (Figure 4F).

### 3.4 Construction of Co-Expression modules and identification of the hub genes

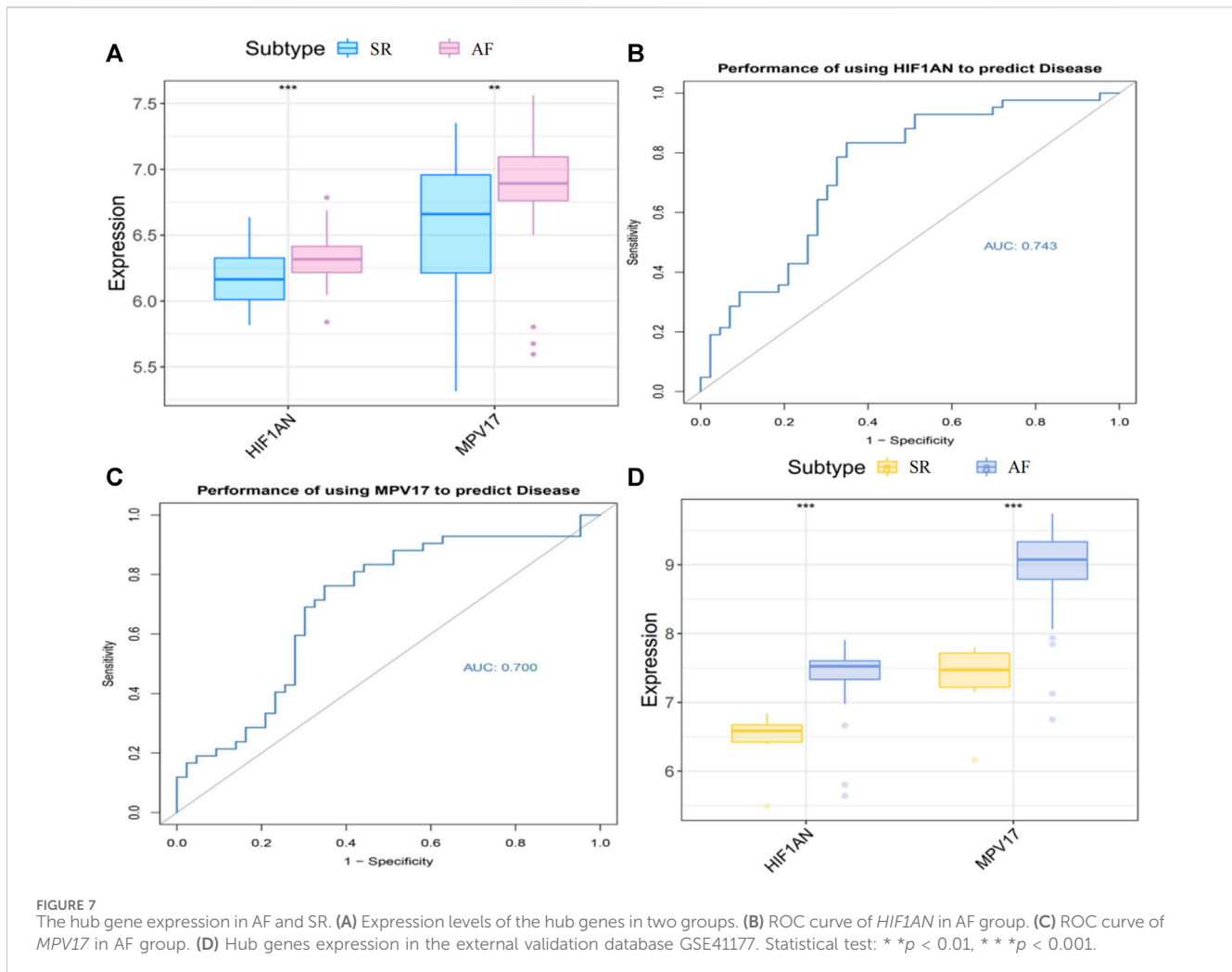
AF is a heterogeneous syndrome, and multiple modulating genes are involved in its pathogenesis. In most diseases, genes with similar expression patterns are prone to have similar biological functions. Thus, gene co-expression networks can facilitate analysis of AF-related biological pathways. We included all genes from 85 patients and found no significant outlier samples in the dataset using the hclust function (Figures 5A, B). A soft threshold power was selected to make the constructed network compatible with the characteristics of scale-free networks. We set the soft threshold to 9 in SR group (Figure 5C) and 7 in AF group (Figure 5D). Through the calculation of the scale-free topology, the  $R^2$  value reached 0.9 (Figures 5E, F). These results further verified the feasibility of WGCNA.

Two co-expression networks of genes from 85 patients were constructed and hierarchical clustering analysis was performed according to weighted correlation coefficients. The clustering results were segmented based on set standards to obtain different gene modules (Figures 6A, B). We identified 35 modules of different sizes represented by the branches of the cluster tree with different colors by WGCNA for the AF group. The network modules in SR group were compared with

those in AF group to identify non-conserved modules that could be interpreted as the change in the network attributed to AF group. These non-conserved modules might be related to disease progression in patients with AF. The median rank and Z-summary scores of conservatism of different color modules were shown in Figure 6C. The Z-summary score of the blue module was the highest, indicating that it retained the network characteristics of SR group. The dark orange module with the lowest Z-summary score was less conservative, indicating that it could be used as a module feature to distinguish patients with AF from those with SR. We extracted the first 1000 genes of the module and took the intersection with 8 characteristic genes screened through LASSO and SVM, and finally obtained two hub genes: hypoxia inducible factor 1 subunit alpha inhibitor (*HIF1AN*) and mitochondrial inner membrane protein 17 (*MPV17*) (Figure 6D). And the enrichment analysis for other genes included in the dark orange module was displayed in Supplementary Figure S1.

### 3.5 Identification and validation of hub gene expression levels and diagnostic value

The expression levels of *HIF1AN* and *MPV17* were assessed by box-line plots and found to be significantly higher in AF group than those in SR group ( $p < 0.001$ ) (Figure 7A). The sensitivity and specificity of two hub genes as diagnostic genes were evaluated by the area under the ROC curve (AUC). The AUCs of *HIF1AN* and



*MPV17* were 0.743 (Figure 7B) and 0.700 (Figure 7C), respectively, which indicated that *HIF1AN* and *MPV17* had high diagnostic value and could better predict disease development. To further validate the results of the bioinformatics analysis, we downloaded GSE41177 dataset from GEO database, which included 32 cases with AF and 6 cases with SR and verified the significant differences in the expression of *HIF1AN* and *MPV17* between AF and SR by Wilcoxon test ( $p < 0.05$ ) (Figure 7D). The enrichment analysis for two hub genes in metabolic processes was displayed in Supplementary Figure S2.

### 3.6 GSEA

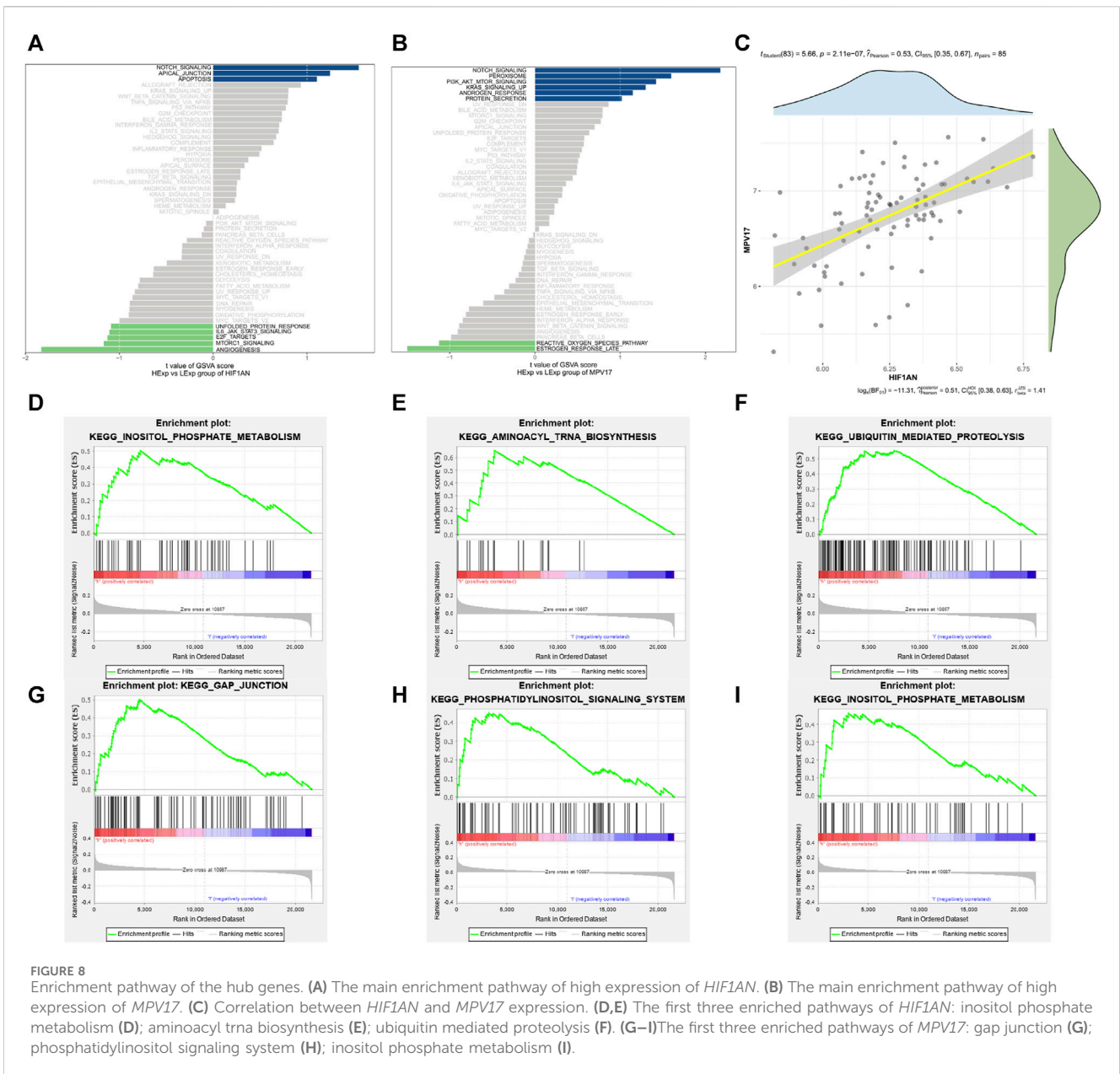
The specific signaling pathways in which the two hub genes got involved were explored to reveal the specific molecular mechanisms associated with AF development. As shown in Figure 8A, the main pathways enriched by the high expression of *HIF1AN* were INOSITOL PHOSPHATE METABOLISM (Figure 8D), AMINOACYL TRNA BIOSYNTHESIS (Figure 8E), and UBIQUITIN MEDIATED PROTEOLYSIS (Figure 8F); as shown in Figure 8B, the pathways enriched by high expression of *MPV17*

were mainly GAP JUNCTION (Figure 8G), PHOSPHATIDYLINOSITOL SIGNALING SYSTEM (Figure 8H), and INOSITOL PHOSPHATE METABOLISM (Figure 8I). And correlation analysis revealed a positive relationship between *HIF1AN* and *MPV17* expression (Figure 8C).

### 3.7 Immune cell infiltration and correlation analysis between hub genes and immune cells

By virtue of analysis of the association between hub genes and immune cells in AF dataset, the potential molecular mechanisms by which *HIF1AN* and *MPV17* influence AF development were revealed. CIBERSORT-based immune cell infiltration analysis indicated that mast cells, neutrophils, natural killer (NK) cells, and type I interferon (IFN) were significantly higher in patients with AF and T helper 2 (Th2) cells were significantly lower in AF (Figure 9A). The immune cell correlation was manifested in Figure 9B, in which red represented positive correlation and blue represented negative correlation. The correlation of *HIF1AN* and *MPV17* with immune cells was analyzed and found that both





**FIGURE 8**  
 Enrichment pathway of the hub genes. (A) The main enrichment pathway of high expression of *HIF1AN*. (B) The main enrichment pathway of high expression of *MPV17*. (C) Correlation between *HIF1AN* and *MPV17* expression. (D,E) The first three enriched pathways of *HIF1AN*: inositol phosphate metabolism (D); aminoacyl trna biosynthesis (E); ubiquitin mediated proteolysis (F). (G–I) The first three enriched pathways of *MPV17*: gap junction (G); phosphatidylinositol signaling system (H); inositol phosphate metabolism (I).

*HIF1AN* and *MPV17* were strongly correlated with quantities of immune cells (Figures 9C,D).

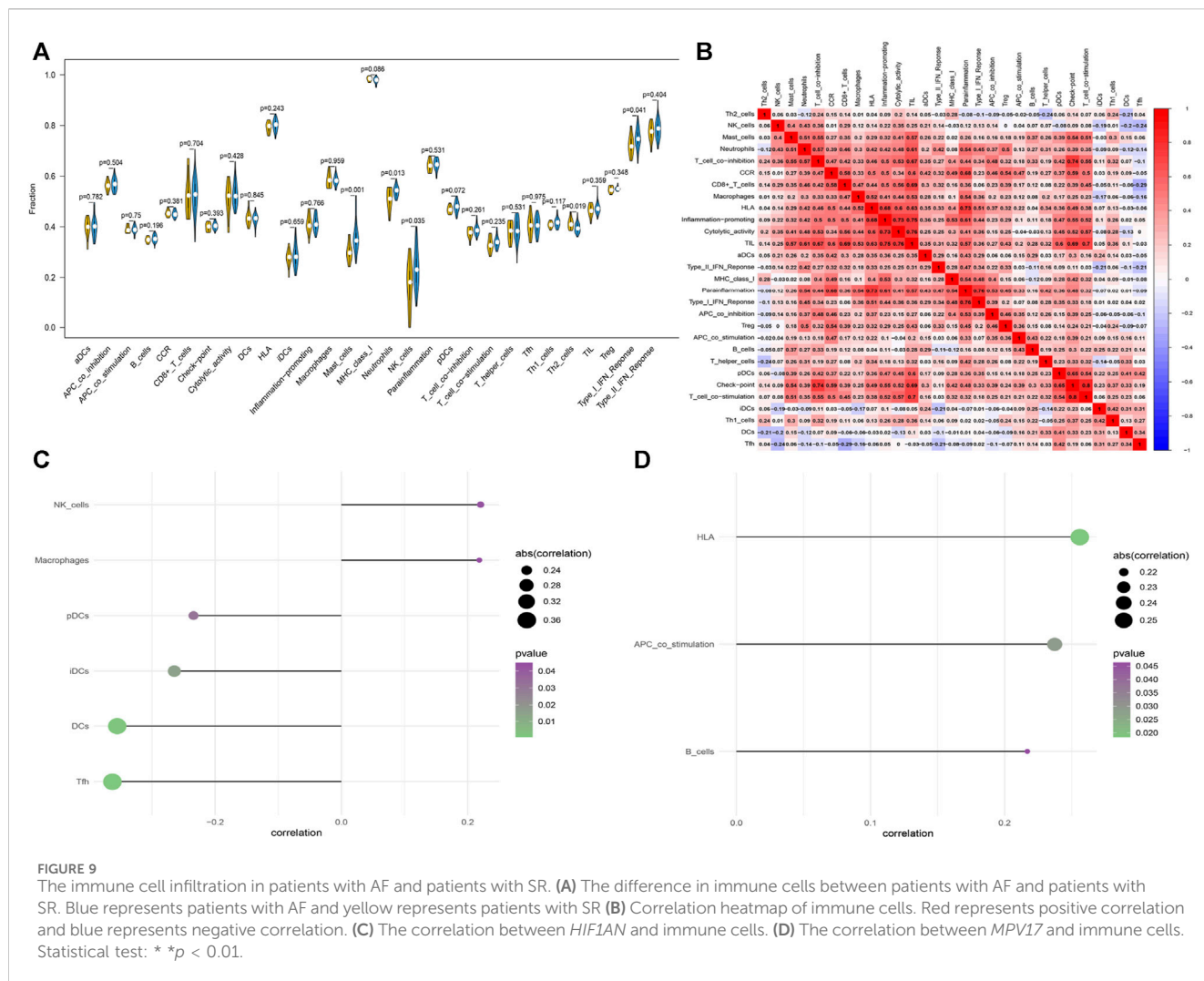
3.1.3 visualized a total of 14 pairs of mRNA–miRNA interaction (Figure 10C).

### 3.8 Correlation analysis between hub genes and immune genes

The differential expression analysis of immune regulatory genes indicated that the significantly different levels of *HLA-DMA*, *HLA-DPB1*, and *HLA-DRA* were observed between the AF group and SR group (Figure 10A). The correlation analysis between the hub genes and the related immune regulation genes indicated that *HIF1AN* was significantly negatively correlated with *HLA-DOB* and *MPV17* was significantly positively correlated with *HLA-DRA* (Figure 10B). In addition, reverse-prediction of the two hub genes by FunRich

## 4 Discussion

AF is a complex disease characterized by environmental and genetic factors that contribute to its pathogenesis. The onset and maintenance of AF require electrical and structural remodeling of the atria (Nattel and Harada, 2014). Atrial electrical remodeling is characterized by shorter effective refractory periods and slowed conduction velocity due to alterations in cardiac action potential, indicating a re-entry substrate in AF (Yeh et al., 2008). Structural remodeling includes atrial dilation, cell hypertrophy, and fibrosis, all of which also contribute to the abnormal electrical signal formation

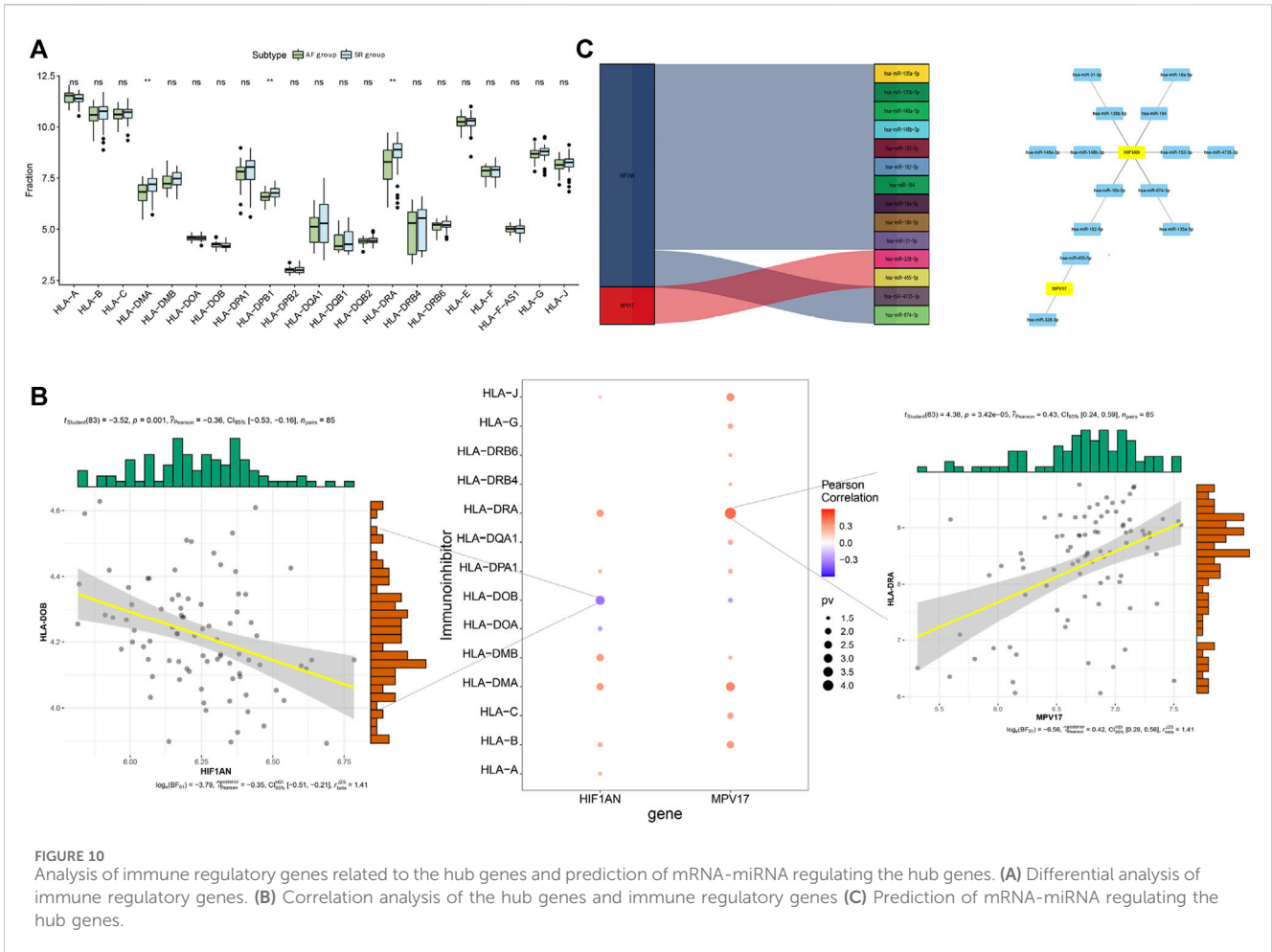


(Alessie et al., 2002). The molecular basis of the mechanisms underlying AF is not completely understood and attributed to fibrosis, calcium dysregulation, inflammation, and oxidative injury (Ferrari et al., 2016; Nattel, 2017). To unravel the mechanism, the genetic variation has increasingly been identified and categorized into three components: rare variation—encoding ion channels and gap junction proteins and influencing cardiac depolarization and repolarization in familial forms of AF; common variation—identified using GWAS and regulating cardiac development and cardiomyocyte contractility and structure; and unidentified variation (Tucker et al., 2016). Genetic variation in AF is heterogeneous. To better understand the pathogenesis and key biomarkers of AF, we compared DEGs between the AF and SR groups based on two GEO datasets and performed enrichment analyses to explore the potential biological functions of DEGs in AF. To more accurately identify central genes associated with AF pathogenesis, we screened two hub genes by combining machine learning algorithms (SVM-RFE and LASSO) based on WGCNA and also confirmed the validity of *HIF1AN* and *MPV17* in both the training and validation sets. Subsequently, using gene enrichment analysis and GSEA, we found that *HIF1AN* and *MPV17* can affect key signaling pathway functions, such as mitochondrial dysfunction,

oxidative stress, and gap junctions. Finally, we performed a correlation analysis between the hub genes and immune cell regulation. We attempted to develop a deep understanding of the pathogenesis and key biomarkers of AF through our study and identify the potential targets for early diagnosis and treatment of AF.

We found that the mRNA expression levels of *HIF1AN* and *MPV17* were significantly upregulated in AF samples compared to those with sinus rhythm, and there was a positive correlation between these two genes. Therefore, *HIF1AN* and *MPV17* could be considered diagnostic biomarker for AF.

*MPV17*, located at 2p23-p21 encodes a mitochondrial inner membrane protein associated with mitochondrial homeostasis and reactive oxygen species (ROS) metabolism (Spinazzola et al., 2006; Casalena et al., 2014). In a swine rapid atrial pacing model of AF, increased activity of ROS in both of the left atria and left atrial appendage (Cai et al., 2002). The ROS can cause atrial fibrosis and increase the susceptibility to AF (Friedrichs et al., 2012). As the major source of intracellular ROS, mitochondrial oxidative stress can lead to arrhythmias such as AF through alterations of ion homeostasis and ion channel behavior (Liu et al., 2022). The precise function of *MPV17* in mitochondria has not been established. An enhanced ROS production was registered in *MPV17* gene-inactivated mice (Binder



**FIGURE 10** Analysis of immune regulatory genes related to the hub genes and prediction of mRNA-miRNA regulating the hub genes. **(A)** Differential analysis of immune regulatory genes. **(B)** Correlation analysis of the hub genes and immune regulatory genes **(C)** Prediction of mRNA-miRNA regulating the hub genes.

et al., 1999; Casalena et al., 2014). The *MPV17* gene product was identified as a membrane protein of peroxisomes playing a major role in the peroxisomal metabolism of ROS (Zwaka et al., 1994). *MPV17* forms a non-selective channel in the inner mitochondrial membrane to decrease mitochondrial membrane potential that may be beneficial under some conditions to preserve mitochondrial homeostasis by preventing excessive production of ROS (Antonovkov et al., 2015). However, elevated *MPV17* in AF in our bioinformatics-based findings appeared to be counterintuitive, and one possible explanation is that it represents a compensatory mechanism in response to excessive ROS in AF. It has been proposed that mitochondria targeted antioxidants may mitigate the progression of arrhythmia (Liu et al., 2022).

*HIF1AN*, also known as *FIH-1*, is located at 10q24 and encodes a protein that binds to hypoxia inducible factor 1-alpha (*HIF-1α*) to negatively modulate *HIF-1α* stability and inhibit *HIF-1α* signaling (Mahon et al., 2001). The *MPV17* lacking cells displayed enhanced *HIF-1α* signaling under normoxia and hypoxia (Shvetsova et al., 2017). And thus, both *HIF1AN* and *MPV17* have a negative effect on *HIF-1α*. It has been widely accepted that *HIF-1α* is a modulator in various organisms for sensing and responding to changes in the oxygen concentration, mediating the cellular adaptation to hypoxia (Menendez-Montes et al., 2016; Zhang et al., 2019). The expression of *HIF1* is essential to vasculogenesis and ventricular development at

early gestational stage under hypoxic conditions. However, sustained *HIF1* activation at midgestation precludes regulation of genes essential for establishment of the cardiac conduction system, such as genes encoding gap-junction protein Connexin 40 essential for the rapid cardiac conducting capacity, and also impedes energy metabolism in mitochondrion, accompanied by increased oxidative stress, resulting in misregulation of genes involved in cardiac conduction system maturation (Menendez-Montes et al., 2016). *HIF-1α* protein expression is elevated in AF atrial tissues *in vivo* than those with SR (Ogi et al., 2010). *HIF-1α* can increase the expression of matrix metalloproteinases (MMPs) and transforming growth factor (TGF)-β in AF through promoting atrial fibrosis (Ogi et al., 2010), and inhibiting the expression of *HIF-1α* can decrease the levels of TGF-β and MMP-9 accompanied by the less myocardial fibrosis in rabbit models (Su et al., 2014). *HIF-1α* was found increased in LAA of patients with AF, suggesting that *HIF* is involved in the inflammatory and fibrotic change of epicardial adipose tissue (Abe et al., 2018). Furthermore, reduced fatty acid oxidation caused by upregulation of *HIF-1α* has been demonstrated in AF models (Bai et al., 2019). *HIF1AN* could be a potential strategy for limiting exacerbated inflammation responses (Palazon et al., 2014). Increased *HIF1AN* in patients with AF in our bioinformatics-based findings manifested the protective

property of *HIF1AN* against atrial fibrosis and inflammation induced by *HIF-1 $\alpha$*  in AF.

GSEA can be used to estimate changes in pathway activity in a sample population (Xie et al., 2022). Using GSEA, we found that high expression of *MPV17* was significantly enriched in GAP JUNCTION pathway. Gap junctions in myocardium are responsible for the intercellular conduction of action potentials; Spach et al. suggested that altered topology of gap junctions took important effects on AF and could be used as a therapeutic strategy for AF (Spach and Starmer, 1995). In recent years, the concept of AF treatment with a normalized gap junction distribution has been widely recognized. It is now believed that various mechanisms affecting gap junctions could lead to an increased susceptibility to AF. Gap junction remodeling during AF can be reversed, which can be accompanied by a decrease in atrial susceptibility to AF (Guo and Yang, 2022). GSEA also revealed that high *HIF1AN* expression was significantly enriched in the inositol phosphate metabolism pathway. This pathway, in which extracellular signaling molecules are activated by binding to cell surface G protein-coupled receptors to release phosphatidylinositol, possibly influences the regulation of cardiac function, especially during the post-ischemic reperfusion of the myocardium (Anderson et al., 1995). The cAMP produced in the activation of phosphatidylinositol metabolism, such as inositol-1,4,5-triphosphate receptors (InsP3Rs), modulates atrial muscle contraction and is thought to contribute to AF. The expression levels of InsP3Rs are upregulated in patients with AF—a process that may be associated with Ca<sup>2+</sup> release and altered homeostasis (Varma et al., 2022). These findings suggest that these signaling pathways are involved in the onset and development of AF and that *MPV17* and *HIF1AN* may be involved in the pathogenesis of AF through these signaling pathways.

During AF, the immune system is greatly altered and plays an important role in the pathophysiology of disease development (Yao et al., 2022). We found neutrophils, mast cells, NK cells, and IFN- $\gamma$  to be significantly elevated in AF and Th2 cells to be significantly decreased among 22 types of immune cell. The number of neutrophilic granulocytes was higher in the atrial fat tissue of AF patients relative to SR individuals (Begieneman et al., 2015). Neutrophils play an important role in AF development by releasing cytokines such as interleukin-6 (IL-6), tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), MMP-2 and ROS that promote atrial remodeling (Liu et al., 2018). Mast cells get involved in cardiac fibrosis by releasing fibrogenic mediators in animal models of AF (Kong et al., 2014; Uemura et al., 2016). However, the number of mast cells in AF patients was similar to those with SR (Smorodinova et al., 2017). The controversial effect of mast cells in AF is due to some anti-fibrotic mediators in mast cells (Yao et al., 2022). The influence of Th2 cells on AF is uncertain. Anti-inflammatory cytokines (IL-4 and IL-10) secreted by Th2 cells downregulate cell-mediated immune responses and cytotoxic inflammatory responses. However, Th2 cells have also been implicated in the pathogenesis of fibrotic conditions (Liu et al., 2018). IFN- $\gamma$  secreted by T cells and NK cells promotes atrial remodeling through macrophage (Liu et al., 2015; Liu et al., 2018).

Hypoxia delays neutrophil apoptosis through HIF-1 $\alpha$ -dependent nuclear factor- $\kappa$ B activity that results in sustained inflammation (Walmsley et al., 2005). The functions of the HIF on macrophage polarization depend on the pathophysiological context. *HIF-1 $\alpha$*  can also drive immunosuppressive functions by macrophages lacking

*HIF-2 $\alpha$*  fail to mount an inflammatory response (Palazon et al., 2014). Hypoxia induces the production of proinflammatory cytokines (TNF- $\alpha$ , IL-1 $\beta$ ) by dendritic cells (DCs) (Mancino et al., 2008). Due to the inhibitory effects of *HIF1AN* and *MPV17* on *HIF-1 $\alpha$*  (Mahon et al., 2001; Shvetsova et al., 2017), the hub genes attenuate the influence of immune response induced by *HIF-1 $\alpha$* . In our bioinformatics-based findings, *HIF1AN* had a positive relationship with NK cells and macrophages, and a negative relationship with DCs in AF; *MPV17* had a positive association with B cells in AF.

Although the hub genes screened based on WGCNA and the machine learning methods showed accurate diagnostic power and were validated using external datasets, there were still some limitations should be announced. First, this study was based on bioinformatics analyses to obtain hub genes, it is necessary to be cautious to conclude gene expression elucidating the molecular mechanisms underlying AF without verification *in vivo* and *in vitro* experiments, and thus, further investigation is warranted to address *in vitro* and *in vivo*. Second, external clinical features of these data were not used in our study. Furthermore, immune cell infiltration analysis was based on limited genetic data, and the specific regulatory mechanisms need further *in vivo* and *in vitro* experimental validation.

## 5 Conclusion

In summary, we used WGCNA combined with machine learning methods to screen and validate the hub genes and identified the correlations between two hub genes and the immune microenvironment and immune regulation. Our results may provide more accurate and effective diagnosis biomarkers and therapeutic targets for 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 author.

## Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements. The manuscript presents research on animals that do not require ethical approval for their study.

## Author contributions

ZS conceived the study and critically revised the content of this manuscript. FT made significant contributions to the study methods, results, and writing of the manuscript. All authors contributed to the article and approved the submitted version.

## 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/fcell.2023.1190273/full#supplementary-material>

## References

- Abe, I., Teshima, Y., Kondo, H., Kaku, H., Kira, S., Ikebe, Y., et al. (2018). Association of fibrotic remodeling and cytokines/chemokines content in epicardial adipose tissue with atrial myocardial fibrosis in patients with atrial fibrillation. *Heart Rhythm*. 15 (11), 1717–1727. doi:10.1016/j.hrthm.2018.06.025
- Allessie, M., Ausma, J., and Schotten, U. (2002). Electrical, contractile and structural remodeling during atrial fibrillation. *Cardiovasc Res*. 54 (2), 230–246. doi:10.1016/s0008-6363(02)00258-4
- Anderson, K. E., Dart, A. M., and Woodcock, E. A. (1995). Inositol phosphate release and metabolism during myocardial ischemia and reperfusion in rat heart. *Circ. Res*. 76 (2), 261–268. doi:10.1161/01.res.76.2.261
- Antonenkov, V. D., Isomursu, A., Mennerich, D., Vapola, M. H., Weiher, H., Kietzmann, T., et al. (2015). The human mitochondrial DNA depletion syndrome gene MPV17 encodes a non-selective channel that modulates membrane potential. *J. Biol. Chem*. 290 (22), 13840–13861. doi:10.1074/jbc.M114.608083
- Bai, F., Liu, Y., Tu, T., Li, B., Xiao, Y., Ma, Y., et al. (2019). Metformin regulates lipid metabolism in a canine model of atrial fibrillation through AMPK/PPAR- $\alpha$ /VLCAD pathway. *Lipids Health Dis*. 18 (1), 109. doi:10.1186/s12944-019-1059-7
- Begieneman, M. P., Rijvers, L., Kubat, B., Paulus, W. J., Vonk, A. B., van Rossum, A. C., et al. (2015). Atrial fibrillation coincides with the advanced glycation end product N<sup>ε</sup>-(carboxymethyl) lysine in the atrium. *Am. J. Pathol*. 185 (8), 2096–2104. doi:10.1016/j.ajpath.2015.04.018
- Binder, C. J., Weiher, H., Exner, M., and Kerjaschki, D. (1999). Glomerular overproduction of oxygen radicals in Mpv17 gene-inactivated mice causes podocyte foot process flattening and proteinuria: a model of steroid-resistant nephrosis sensitive to radical scavenger therapy. *Am. J. Pathol*. 154 (4), 1067–1075. doi:10.1016/S0002-9440(10)65359-X
- Cai, H., Li, Z., Goette, A., Mera, F., Honeycutt, C., Feterik, K., et al. (2002). Downregulation of endocardial nitric oxide synthase expression and nitric oxide production in atrial fibrillation potential mechanisms for atrial thrombosis and stroke. *Circulation* 106 (22), 2854–2858. doi:10.1161/01.cir.0000039327.11661.16
- Casalena, G., Krick, S., Daehn, I., Yu, L., Ju, W., Shi, S., et al. (2014). Mpv17 in mitochondria protects podocytes against mitochondrial dysfunction and apoptosis *in vivo* and *in vitro*. *Am. J. Physiol. Ren. Physiol*. 306 (11), F1372–F1380. doi:10.1152/ajprenal.00608.2013
- Christophersen, I. E., and Ellinor, P. T. (2016). Genetics of atrial fibrillation: from families to genomes. *J. Hum. Genet*. 61 (1), 61–70. doi:10.1038/jhg.2015.44
- Duan, J., Soussen, C., Brie, D., Ider, J., Wan, M., and Wang, Y. P. (2016). Generalized LASSO with under-determined regularization matrices. *Signal Process*. 127, 239–246. doi:10.1016/j.sigpro.2016.03.001
- Duan, K.-B., Rajapakse, J. C., Wang, H., and Azuaje, F. (2005). Multiple SVM-RFE for gene selection in cancer classification with expression data. *IEEE Trans. Nanobioscience* 4, 228–234. doi:10.1109/TNB.2005.853657
- Farraha, M., Chong, J. J., and Kizana, E. (2016). Therapeutic prospects of gene therapy for atrial fibrillation. *Heart Lung Circ*. 25 (8), 808–813. doi:10.1016/j.hlc.2016.04.011
- Ferrari, R., Bertini, M., Blomstrom-Lundqvist, C., Dobrev, D., Kirchhof, P., Pappone, C., et al. (2016). An update on atrial fibrillation in 2014: from pathophysiology to treatment. *Int. J. Cardiol*. 203, 22–29. doi:10.1016/j.ijcard.2015.10.089
- Friedrichs, K., Baldus, S., and Klinke, A. (2012). Fibrosis in atrial fibrillation - role of reactive species and MPO. *Front. Physiol*. 3, 214. doi:10.3389/fphys.2012.00214
- Guo, Y. H., and Yang, Y. Q. (2022). Atrial Fibrillation: focus on myocardial connexins and gap junctions. *Biol. (Basel)* 11 (4), 489. doi:10.3390/biology11040489
- Kong, P., Christia, P., and Frangogiannis, N. G. (2014). The pathogenesis of cardiac fibrosis. *Cell Mol. Life Sci*. 71 (4), 549–574. doi:10.1007/s00018-013-1349-6
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma*. 9, 559. doi:10.1186/1471-2105-9-559
- Liu, C., Ma, N., Guo, Z., Zhang, Y., Zhang, J., Yang, F., et al. (2022). Relevance of mitochondrial oxidative stress to arrhythmias: innovative concepts to target treatments. *Pharmacol. Res*. 175, 106027. doi:10.1016/j.phrs.2021.106027
- Liu, L., Zheng, Q., Lee, J., Ma, Z., Zhu, Q., and Wang, Z. (2015). PD-1/PD-L1 expression on CD 4<sup>+</sup> T cells and myeloid DCs correlates with the immune pathogenesis of atrial fibrillation. *J. Cell Mol. Med*. 19 (6), 1223–1233. doi:10.1111/jcmm.12467
- Liu, Y., Shi, Q., Ma, Y., and Liu, Q. (2018). The role of immune cells in atrial fibrillation. *J. Mol. Cell Cardiol*. 123, 198–208. doi:10.1016/j.yjmcc.2018.09.007
- Lubitz, S. A., Yin, X., Fontes, J. D., Magnani, J. W., Rienstra, M., Pai, M., et al. (2010). Association between familial atrial fibrillation and risk of new-onset atrial fibrillation. *JAMA* 304 (20), 2263–2269. doi:10.1001/jama.2010.1690
- Mahon, P. C., Hirota, K., and Semenza, G. L. (2001). FIH-1: a novel protein that interacts with HIF-1 $\alpha$  and VHL to mediate repression of HIF-1 transcriptional activity. *Genes Dev*. 15 (20), 2675–2686. doi:10.1101/gad.924501
- Mancino, A., Schioppa, T., Larghi, P., Pasqualini, F., Nebuloni, M., Chen, I. H., et al. (2008). Divergent effects of hypoxia on dendritic cell functions. *Blood* 112 (9), 3723–3734. doi:10.1182/blood-2008-02-142091
- Menendez-Montes, I., Escobar, B., Palacios, B., Gómez, M. J., Izquierdo-García, J. L., Flores, L., et al. (2016). Myocardial VHL-HIF signaling controls an embryonic metabolic switch essential for cardiac maturation. *Dev. Cell* 39 (6), 724–739. doi:10.1016/j.devcel.2016.11.012
- Myrstad, M., Løchen, M. L., Graff-Iversen, S., Gulsvik, A. K., Thelle, D. S., Stigum, H., et al. (2014). Increased risk of atrial fibrillation among elderly Norwegian men with a history of long-term endurance sport practice. *Scand. J. Med. Sci. Sports* 24 (4), e238–e244. doi:10.1111/sms.12150
- Nattel, S. (2017). Molecular and cellular mechanisms of atrial fibrosis in atrial fibrillation. *JACC Clin. Electrophysiol*. 3 (5), 425–435. doi:10.1016/j.jacep.2017.03.002
- Nattel, S., and Harada, M. (2014). Atrial remodeling and atrial fibrillation: recent advances and translational perspectives. *J. Am. Coll. Cardiol*. 63 (22), 2335–2345. doi:10.1016/j.jacc.2014.02.555
- Ogi, H., Nakano, Y., Niida, S., Dote, K., Hirai, Y., Suenari, K., et al. (2010). Is structural remodeling of fibrillated atria the consequence of tissue hypoxia? *Circ. J*. 74 (9), 1815–1821. doi:10.1253/circj.09-0969
- Palazon, A., Goldrath, A. W., Nizet, V., and Johnson, R. S. (2014). HIF transcription factors, inflammation, and immunity. *Immunity* 41 (4), 518–528. doi:10.1016/j.immuni.2014.09.008
- Prystowsky, E. N., Padanilam, B. J., and Fogel, R. I. (2015). Treatment of atrial fibrillation. *JAMA* 314 (3), 278–288. doi:10.1001/jama.2015.7505
- Roselli, C., Chaffin, M. D., Weng, L. C., Aeschbacher, S., Ahlberg, G., Albert, C. M., et al. (2018). Multi-ethnic genome-wide association study for atrial fibrillation. *Nat. Genet*. 50 (9), 1225–1233. doi:10.1038/s41588-018-0133-9
- Schnabel, R. B., Yin, X., Gona, P., Larson, M. G., Beiser, A. S., McManus, D. D., et al. (2015). 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet* 386 (9989), 154–162. doi:10.1016/S0140-6736(14)61774-8
- Shvtsova, A. N., Mennerich, D., Kerätär, J. M., Hiltunen, J. K., and Kietzmann, T. (2017). Non-electron transfer chain mitochondrial defects differently regulate HIF-1 $\alpha$  degradation and transcription. *Redox Biol*. 12, 1052–1061. doi:10.1016/j.redox.2017.05.003
- Smorodina, N., Bláha, M., Melenovský, V., Rozsivalová, K., Pridal, J., Ďurišová, M., et al. (2017). Analysis of immune cell populations in atrial myocardium of patients with atrial fibrillation or sinus rhythm. *PLoS One* 12 (2), e0172691. doi:10.1371/journal.pone.0172691
- Spach, M. S., and Starmer, C. F. (1995). Altering the topology of gap junctions a major therapeutic target for atrial fibrillation. *Cardiovasc Res*. 30 (3), 337–344. doi:10.1016/0008-6363(96)88514-2

- Spinazzola, A., Viscomi, C., Fernandez-Vizarra, E., Carrara, F., D'Adamo, P., Calvo, S., et al. (2006). MPV17 encodes an inner mitochondrial membrane protein and is mutated in infantile hepatic mitochondrial DNA depletion. *Nat. Genet.* 38 (5), 570–575. doi:10.1038/ng1765
- Su, F., Zhang, W., Chen, Y., Ma, L., Zhang, H., and Wang, F. (2014). Significance of hypoxia-inducible factor-1  $\alpha$  expression with atrial fibrosis in rats induced with isoproterenol. *Exp. Ther. Med.* 8 (6), 1677–1682. doi:10.3892/etm.2014.1989
- Tucker, N. R., Clauss, S., and Ellinor, P. T. (2016). Common variation in atrial fibrillation: navigating the path from genetic association to mechanism. *Cardiovasc Res.* 109 (4), 493–501. doi:10.1093/cvr/cvv283
- Uemura, K., Kondo, H., Ishii, Y., Kobukata, M., Haraguchi, M., Imamura, T., et al. (2016). Mast cells play an important role in the pathogenesis of hyperglycemia-induced atrial fibrillation. *J. Cardiovasc Electrophysiol.* 27 (8), 981–989. doi:10.1111/jce.12995
- Varma, D., Almeida, J. F. Q., DeSantiago, J., Blatter, L. A., and Banach, K. (2022). Inositol 1,4,5-trisphosphate receptor - reactive oxygen signaling domain regulates excitation-contraction coupling in atrial myocytes. *J. Mol. Cell Cardiol.* 163, 147–155. doi:10.1016/j.yjmcc.2021.10.006
- Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., et al. (2021). Heart disease and stroke statistics-2021 update: a report from the American Heart Association. *Circulation* 143 (8), e254–e743. doi:10.1161/CIR.0000000000000950
- Walmsley, S. R., Print, C., Farahi, N., Peyssonnaud, C., Johnson, R. S., Cramer, T., et al. (2005). Hypoxia-induced neutrophil survival is mediated by HIF-1 $\alpha$ -dependent NF- $\kappa$ B activity. *J. Exp. Med.* 201 (1), 105–115. doi:10.1084/jem.20040624
- Wang, Y., Li, B., and Zhao, Y. (2022). Inflammation in preeclampsia: genetic biomarkers, mechanisms, and therapeutic strategies. *Front. Immunol.* 13, 883404. doi:10.3389/fimmu.2022.883404
- Xie, L., Huang, G., Gao, M., Huang, J., Li, H., Xia, H., et al. (2022). Identification of atrial fibrillation-related lncRNA based on bioinformatic analysis. *Dis. Markers* 2022, 8307975. doi:10.1155/2022/8307975
- Yao, Y., Yang, M., Liu, D., and Zhao, Q. (2022). Immune remodeling and atrial fibrillation. *Front. Physiol.* 13, 927221. doi:10.3389/fphys.2022.927221
- Yeh, Y. H., Wakili, R., Qi, X. Y., Chartier, D., Boknik, P., Kääb, S., et al. (2008). Calcium-Handling abnormalities underlying atrial arrhythmogenesis and contractile dysfunction in dogs with congestive heart failure. *Circ. Arrhythm. Electrophysiol.* 1 (2), 93–102. doi:10.1161/CIRCEP.107.754788
- Yoo, S., Pfenniger, A., Hoffman, J., Zhang, W., Ng, J., Burrell, A., et al. (2020). Attenuation of oxidative injury with targeted expression of nadph oxidase 2 short hairpin rna prevents onset and maintenance of electrical remodeling in the canine atrium: a novel gene therapy approach to atrial fibrillation. *Circulation* 142 (13), 1261–1278. doi:10.1161/CIRCULATIONAHA.119.044127
- Zhang, Y., Liu, D., Hu, H., Zhang, P., Xie, R., and Cui, W. (2019). HIF-1 $\alpha$ /BNIP3 signaling pathway-induced-autophagy plays protective role during myocardial ischemia-reperfusion injury. *Biomed. Pharmacother.* 120, 109464. doi:10.1016/j.biopha.2019.109464
- Zwaka, R. M., Reuter, A., Pfaff, E., Moll, J., Gorgas, K., Karasawa, M., et al. (1994). The glomerulosclerosis gene Mpv17 encodes a peroxisomal protein producing reactive oxygen species. *EMBO J.* 13 (21), 5129–5134. doi:10.1002/j.1460-2075.1994.tb06842.x