

RESEARCH

Open Access



# Hub microRNAs and genes in the development of atrial fibrillation identified by weighted gene co-expression network analysis

Qiang Qu<sup>1,2†</sup>, Jin-Yu Sun<sup>1,2†</sup>, Zhen-Ye Zhang<sup>1†</sup>, Yue Su<sup>1,2</sup>, Shan-Shan Li<sup>1</sup>, Feng Li<sup>1</sup> and Ru-Xing Wang<sup>1\*</sup>

## Abstract

Co-expression network may contribute to better understanding molecular interaction patterns underlying cellular processes. To explore microRNAs (miRNAs) expression patterns correlated with AF, we performed weighted gene co-expression network analysis (WGCNA) based on the dataset GSE28954. Thereafter, we predicted target genes using experimentally verified databases (ENOCRI, miRTarBase, and Tarbase), and overlapped genes with differentially expressed genes (DEGs) from GSE79768 were identified as key genes. Integrated analysis of association between hub miRNAs and key genes was conducted to screen hub genes. In general, we identified 3 differentially expressed miRNAs (DEMs) and 320 DEGs, predominantly enriched in inflammation-related functional items. Two significant modules (red and blue) and hub miRNAs (hsa-miR-146b-5p and hsa-miR-378a-5p), which highly correlated with AF-related phenotype, were detected by WGCNA. By overlapping the DEGs and predicted target genes, 38 genes were screened out. Finally, 9 genes (i.e. *ATP13A3*, *BMP2*, *CXCL1*, *GABPA*, *LIF*, *MAP3K8*, *NPY1R*, *S100A12*, *SLC16A2*) located at the core region in the miRNA-gene interaction network were identified as hub genes. In conclusion, our study identified 2 hub miRNAs and 9 hub genes, which may improve the understanding of molecular mechanisms and help to reveal potential therapeutic targets against AF.

**Keywords:** Atrial fibrillation, Weighted gene co-expression network analysis, Hub microRNAs, Hub genes, Inflammation

## Introduction

Atrial fibrillation (AF), the most common sustained cardiac arrhythmia, is expected to affect one quarter of middle-aged adults worldwide, primarily at the age of > 70 years [1]. According to the epidemiological investigation, the overall prevalence was estimated to be 1–4% in Australia, Europe and the USA [1, 2]. The number continued to increase as the population is ageing, reaching 4

per 1000 person-years in 2016 [3, 4]. AF, which accounts for the majority ( $\approx 70\%$ ) of tachyarrhythmia, places a substantial economic burden on healthcare systems [5].

AF has been associated with electrophysiological and structural remodeling, characterized by rapid and uncoordinated atrial activity [6]. The mechanisms underlying AF are complex and multifactorial, and are broadly classified as trigger and substrate, which lead to an increased risk of stroke, heart failure, and premature death [7]. Among them, several molecular factors such as abnormal  $\text{Ca}^{2+}$  handling, fibrosis, and inflammation, are known to be relevant to the development of AF at systemic level [8–11]; however, AF treatment remains challenging because its exact mechanisms at cellular/molecular level are not fully defined. Therefore, further basic investigation to

\*Correspondence: ruxingw@aliyun.com

<sup>†</sup>Qiang Qu, Jin-Yu Sun and Zhen-Ye Zhang have contributed equally to this work

<sup>1</sup> Department of Cardiology, The Affiliated Wuxi People's Hospital of Nanjing Medical University, No. 299, Qingyang Road, Wuxi 214023, China

Full list of author information is available at the end of the article



examine predictive and prognostic biomarkers in AF management will be of great interest.

MicroRNAs (miRNAs) are short (19–25 nucleotides) non-coding RNAs that regulate gene expression by binding to complementary sequences of mRNAs [12, 13]. One miRNA may target hundreds of different mRNAs, while one mRNA may be regulated by multiple miRNAs [14]. The miRNA-gene interaction network can be used to uncover potential biomarkers for disease detection and treatment. A large number of studies have demonstrated the important role of miRNAs in regulating cardiac excitability and arrhythmogenesis in various cardiovascular diseases [15–18]. However, data regarding regulating effects of miRNAs on the transition from sinus rhythm to AF, including potential mechanisms, are limited.

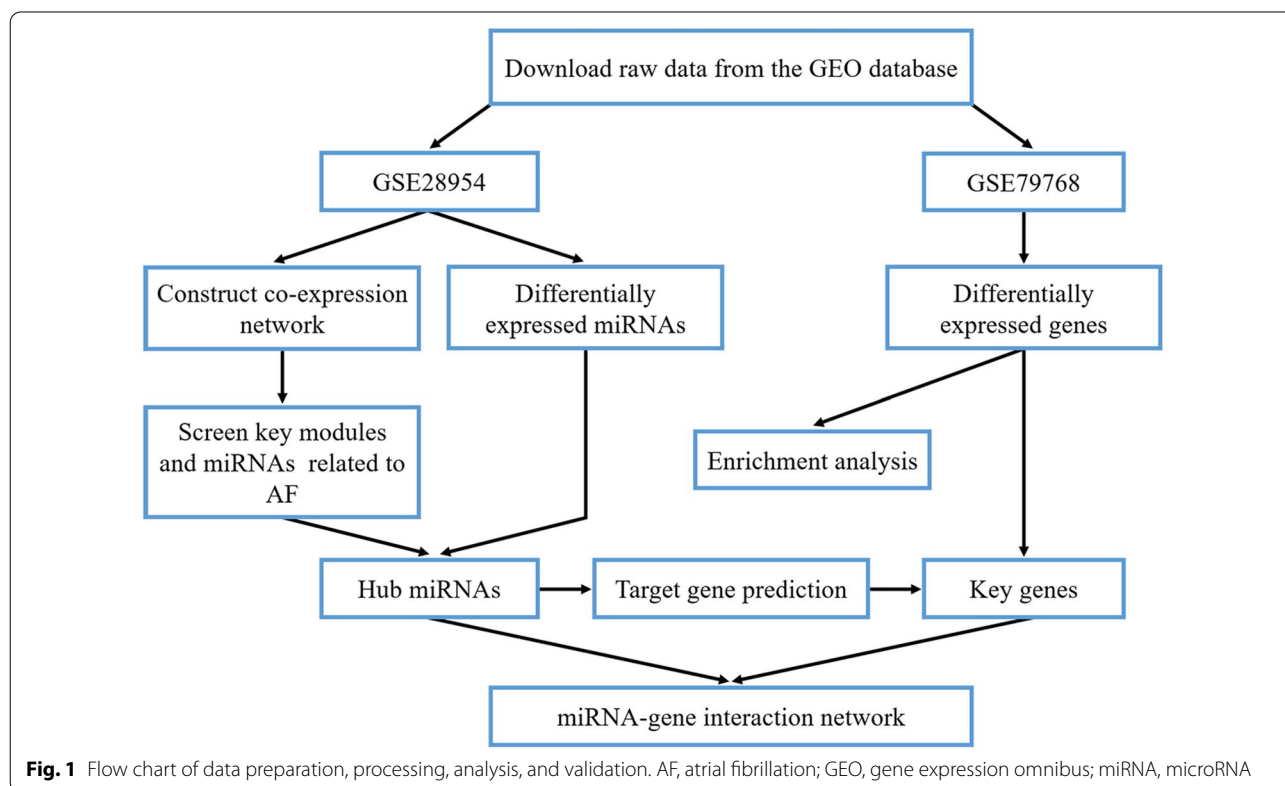
Weighted gene co-expression network analysis (WGCNA) is a bioinformatics algorithm to detect causal relationship between genome and clinical features based on alterations in transcriptome expression patterns under pathological conditions [19]. Co-expression network, composing various biological processes, is grouped into several modules to explore the association between genes and clinical features [20]; meanwhile, genes in the same module are usually functionally-linked and have similar regulating effects on signal pathways [21]. The most central genes tend to play a vital role within a

specific module and are generally regarded as hub genes. WGCNA, unlike differential expression analysis, has been confirmed as a powerful systematic analysis method to recognize higher-order association across genes rather than detect disease-related individual genes. This study aimed to investigate hub miRNAs and genes in patients with AF by combining differential expression analysis and WGCNA.

### Materials and methods

#### Data sources

The workflow of this study is shown in Fig. 1. Microarray-based miRNA and mRNA expression data were obtained from the Gene Expression Omnibus (GEO) database if they met the following selection criteria [22]: (1) expression data were from atrial tissue samples in non-AF and AF patients; (2) study type was restricted to expression profiling by array; (3) organism was restricted to Homo sapiens; and (4) raw/processed expression data were public and available. The expression profiles of GSE28954 and GSE79768, separately containing miRNA and gene expression data, were included to screen hub miRNAs and genes related to the development of AF [23, 24].



### Probe reannotation and data preprocessing

The expression profiles were acquired by reannotating probes in the datasets, as previously described [25]. The probe set was mapped to RefSeq transcript ID based on the latest version of the annotation file. The average expression value was used if multiple probes were applied to detect one miRNA or mRNA. Then, we performed background correction,  $\log_2$ -transformation, and quantile normalization using the 'normalize Between Arrays' function of the 'linear models for microarray data (limma)' package in R software v3.6.3 [26, 27]. No missing expression value was detected during data processing.

### Identification of differentially expressed miRNAs (DEMs) and genes (DEGs)

We used the 'limma' package to screen the DEMs and DEGs between normal rhythm and AF samples with a threshold of false discovery rate (FDR) < 0.05 and  $|\log_2$  fold-change (FC)| > 0.5. Thereafter, we visualized the DEMs and DEGs as a volcano plot and a heatmap using the 'ggplot2' and 'pheatmap' package, respectively.

### Functional enrichment analysis of the DEGs

To identify characteristic biological and biochemistry pathways in the development of AF, we used the 'clusterProfiler' package [28], together with Metascape [29], to perform gene ontology (GO) [30, 31] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis [32–34] for the DEGs, with cut-offs of  $q$  value < 0.05 and  $q$  value < 0.20 established for significant biological processes and biological processes, respectively. The  $q$ -value of a test measures the proportion of false positives incurred (FDR) control when that particular test is called significant [35].

### Weighted gene co-expression network construction and module detection

To explore the relationship between genome and clinical features, we used the 'WGCNA' package to construct a co-expression network based on the DEMs. After assessing the presence of obvious outliers by sample clustering, we used the 'one-step network construction' function to construct the miRNA co-expression network. We used the 'pick-Soft-Threshold' function to calculate the soft-thresholding power  $\beta$  and converted the adjacency into a topological overlap matrix. Then, we conducted average linkage hierarchical clustering and dynamic tree cut at a merging threshold of 0.25. The co-expression network was visualized as a heatmap based on topological overlap dissimilarity and the cluster dendrogram.

### Screening key modules

We used module eigengene (ME), the first principal component of module expression, to represent the expression profile of module miRNAs [36]. To validate module-trait relationships (MTRs, defined as the correlation between MEs and clinical features) of miRNA modules, we categorized miRNAs into corresponding modules according to the constructed modules [37]. We calculated the ME of each module and included related clinical features. Then, we calculated gene significance (GS, defined as the  $\log_{10}$ -transformation of  $P$ -value in the linear regression slope between gene expression and clinical features) and module significance (MS, described as the average GS of all genes in the module) to further assess correlation intensity between a miRNA module and a clinical feature [25, 38]. In general, a module with the highest MS value was regarded as the key module for further analysis.

### Hub miRNA identification and key gene validation

After overlapping miRNAs from the key module with the DEMs from the dataset GSE28954, all of which were regarded as key miRNAs, we further calculate module membership (MM, representing the association between expression and ME) and GS (representing the association between expression and clinical features) to identify hub miRNAs. In general, key miRNAs with high values of MM and GS were regarded as hub miRNAs. The hub miRNAs were uploaded to 3 experimentally verified miRNA target prediction databases (ENOCRI [39, 40], miRTarBase [41], Tarbase [42]) to map corresponding target genes, predict the interaction between miRNAs and genes, and construct a miRNA-target gene regulatory network. Further, we overlapped predicted target genes and the DEGs to determine key genes.

### Construction of the miRNA-gene interaction network

To reveal the principle of cellular organization and predict protein functions, we constructed a protein-protein interaction (PPI) network of the DEGs by Metascape, an online tool for analyzing system-level datasets. Then, we constructed a miRNA-gene interaction network based on upstream and downstream relationship between hub miRNAs and key genes. After analyzing the interaction between key genes using the STRING database v11.0 with a threshold of confidence score > 0.150, we illustrated the network graph and merged the key gene interaction network and miRNA-target gene regulatory network into a miRNA-gene interaction network by Cytoscape v3.6.2 [43].

However, the STRING database, which contains both positive and negative regulatory relationship between miRNAs and genes, does not provide their causal association. The core region with a dense connection in the network was detected by the 'cytoHubba' plugin via mixed character calculation.

## Results

### Microarray data normalization and identification of the DEMs and DEGs

In the dataset GSE28954, 10 AF and 18 non-AF samples were included, with no further information given; in the dataset GSE79768, 14 AF and 12 non-AF samples were included with a mean age of 55.4 years, comprising 10 (38.5%) men and 16 (61.5%) women. After data preprocessing of AF chip expression datasets GSE28954 and GSE79768 (Fig. 2A–D), 3 DEMs and 320 DEGs were identified at cut-off levels of  $FDR < 0.05$  and  $|\log_2 FC| > 0.5$ . Among the 3 DEMs, hsa-miR-146b-5p was up-regulated in atrial tissue samples from AF patients, whereas hsa-miR-378a-5p and hsa-miR-490-3p were down-regulated; among the 320 DEGs, 203 DEGs were up-regulated in atrial tissue samples from AF patients, whereas 117 DEGs were down-regulated. The volcano plot and heatmap are shown in Fig. 2E–H.

### Functional enrichment analysis of the DEGs

We performed functional enrichment analysis of the DEGs based on the GO and KEGG databases. As shown in Fig. 3A, the enriched biological processes were mainly involved in nephron development, T cell mediated immunity, regulation of T cell mediated immunity, regulation of T cell cytokine production, T cell cytokine production, and positive regulation of neuroinflammatory response. Moreover, the KEGG enrichment analysis showed that cytokine-cytokine receptor interaction was the most enriched pathway, followed by neuroactive ligand-receptor interaction, viral protein interaction with cytokine and cytokine receptor, and chemokine signaling pathway (Fig. 3B).

Figure 3C summarizes the top 20 clusters of pathway and process enrichment analysis in Metascape, of which 5 representative enrichment terms were closely associated with inflammation, including regulation of T cell cytokine production, positive regulation of cell migration, regulation of cell adhesion, regulation of leukocyte migration, and heterotypic cell–cell adhesion. In general, functional enrichment

analysis of the DEGs supported the involvement of inflammation in the development of AF.

### Construction of the co-expression network

We used average linkage method, together with Pearson's correlation method, to cluster the samples of GSE28954; meanwhile, we used distance across samples in Pearson's correlation matrices to assess quality of microarray and locate obvious outliers. The samples were included if they met the inclusion criteria: data cut-off height of  $< 20$  (Fig. 4A, B). The soft-thresholding power was set at 5 since scale-free topology index reached 0.9 and mean connectivity was relatively high (Fig. 4C). In general, 7 modules were identified by average linkage hierarchical clustering, with none merged due to dissimilarity (data cut-off height of  $< 0.25$ ). Turquoise module was the largest module with 123 miRNAs, followed by blue module with 106 miRNAs; conversely, red module was the smallest module with 39 miRNAs. Grey module, which included miRNAs belonging to no other module, were dismissed in the following analysis. The cluster dendrogram of the DEMs is shown in Fig. 4D. Moreover, we analyzed the interaction association between modules. The network heatmap, in combination with eigengene adjacency heatmap of miRNAs, demonstrated a high level of independence across co-expression clusters (Fig. 4E, F).

### Constitution of module-trait relationship and detection of key modules

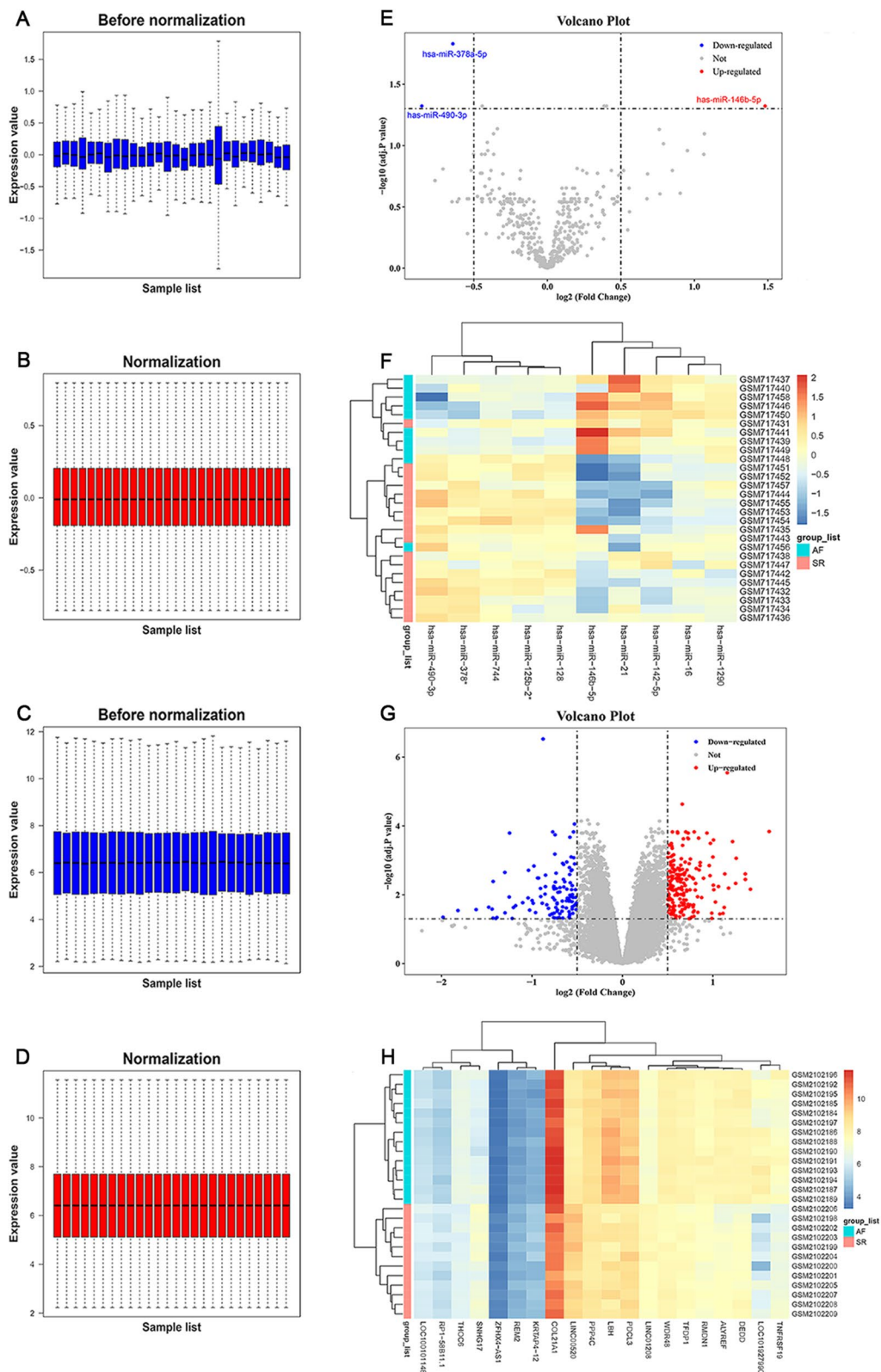
To identify miRNAs related to a specific clinical feature, we used correlation analysis to explore the association of modules with overall health status (healthy, valvular heart disease, coronary artery bypass graft, and AF). Red module had the highest negative correlation with the development of AF ( $r = -0.46$ ,  $P = 0.02$ ), whereas blue module had the highest positive correlation ( $r = 0.36$ ,  $P = 0.06$ ; Fig. 4G). Moreover, Fig. 4H shows that they had the highest MS values across all modules. Therefore, red module and blue module were regarded as key modules for AF. Significant correlation between MM and GS were also seen for the development of AF (Fig. 4I).

### Identification of hub miRNAs

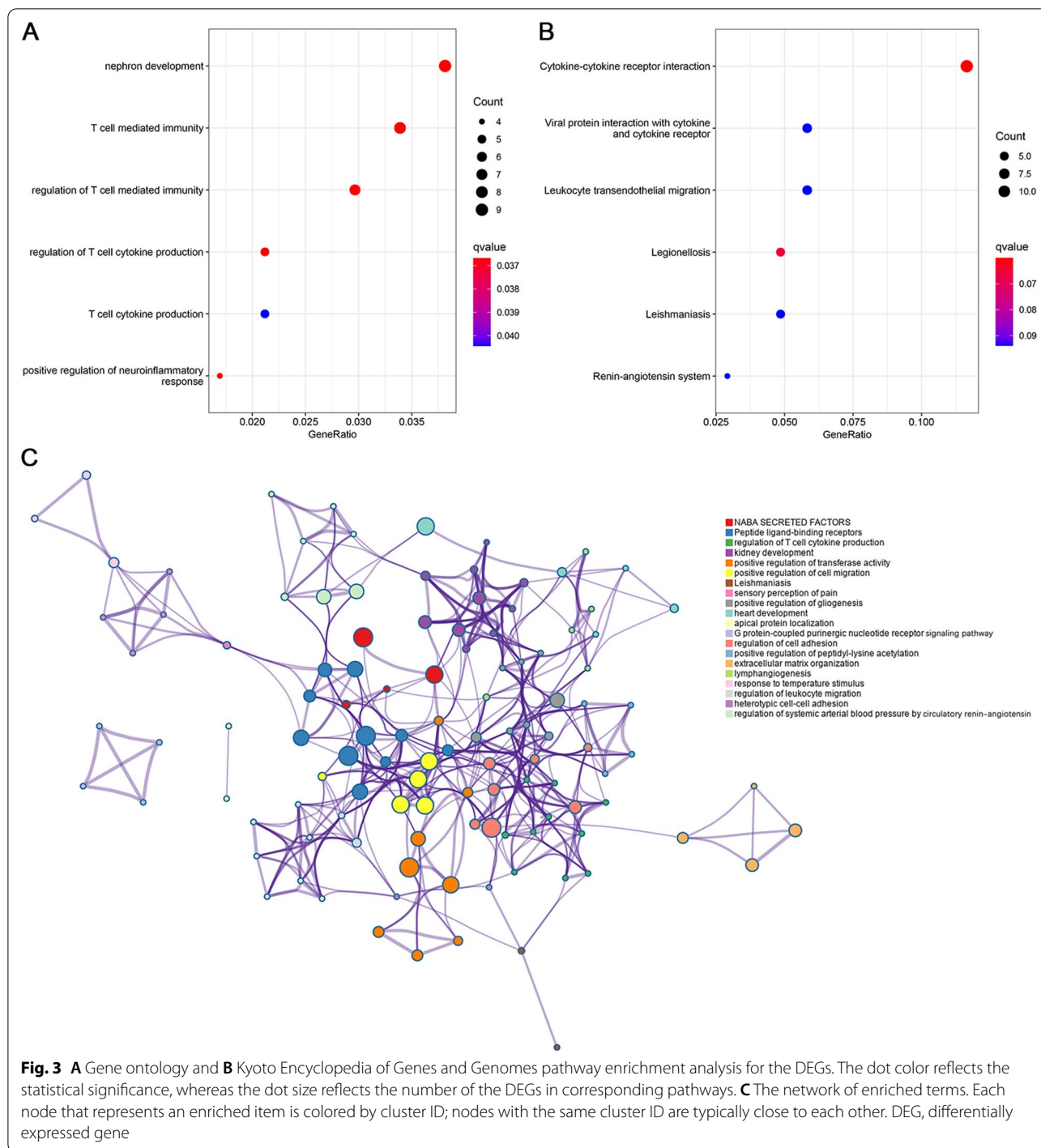
After overlapping miRNAs in key modules with the DEMs from the dataset GSE28954, hsa-miR-146b-5p and hsa-miR-378a-5p were regarded as the 2 key miRNAs.

(See figure on next page.)

**Fig. 2** The normalization of microarray expression profiles and visualization of differentially expressed miRNAs and genes between normal rhythm and AF samples. **A, B** The normalization of the dataset GSE28954. **C, D** The normalization of the dataset GSE79768. **E, F** The volcano plot and heatmap to show differential miRNA expression from GSE28954 with a threshold of false discovery rate ( $FDR < 0.05$ ) and  $|\log_2$  fold-change ( $FC| > 0.5$ ). **G, H** The volcano plot and heatmap to show differential gene expression from GSE79768 with a threshold of  $FDR < 0.05$  and  $|\log_2 FC| > 0.5$ . AF, atrial fibrillation; miRNA, microRNA



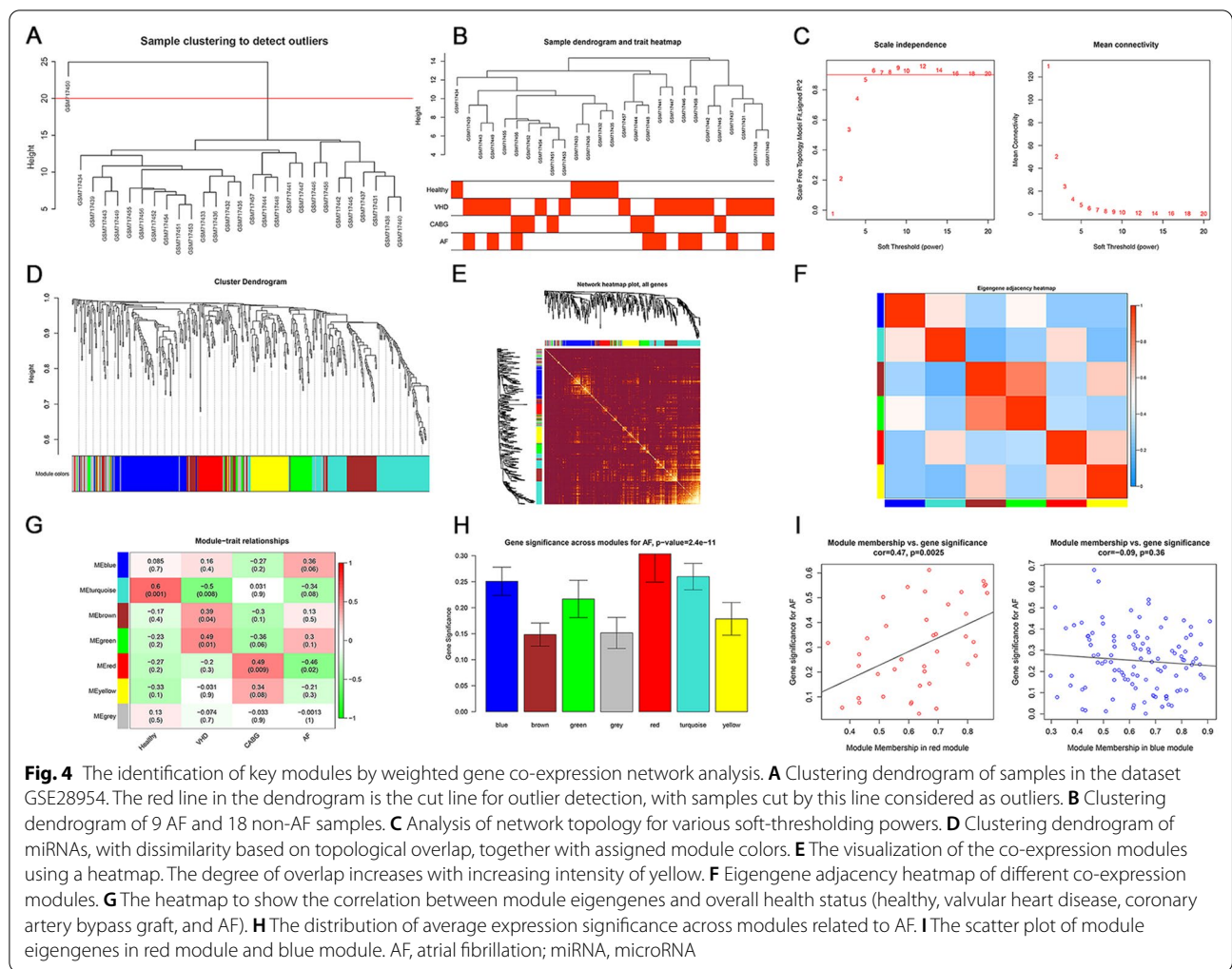
**Fig. 2** (See legend on previous page.)



Furthermore, the  $|MM|$  and  $|GS|$  values of hsa-miR-146b-5p were 0.48 and 0.62, respectively, whereas the numbers separately reached 0.46 and 0.68 in hsa-miR-378a-5p. Therefore, hsa-miR-146b-5p and hsa-miR-378a-5p were further regarded as hub miRNAs.

### Target gene prediction and key gene validation

A total of 2935 target genes and 3000 miRNA-gene pairs were obtained after miRNA-gene mapping by electronic databases (i.e. ENOCRI, miRTarBase, Tarbase). The association between AF-related hub miRNAs and genes



was further verified by overlapping target genes and the DEGs. Finally, 38 genes were regarded as key genes.

### Analysis of the miRNA-gene interaction network

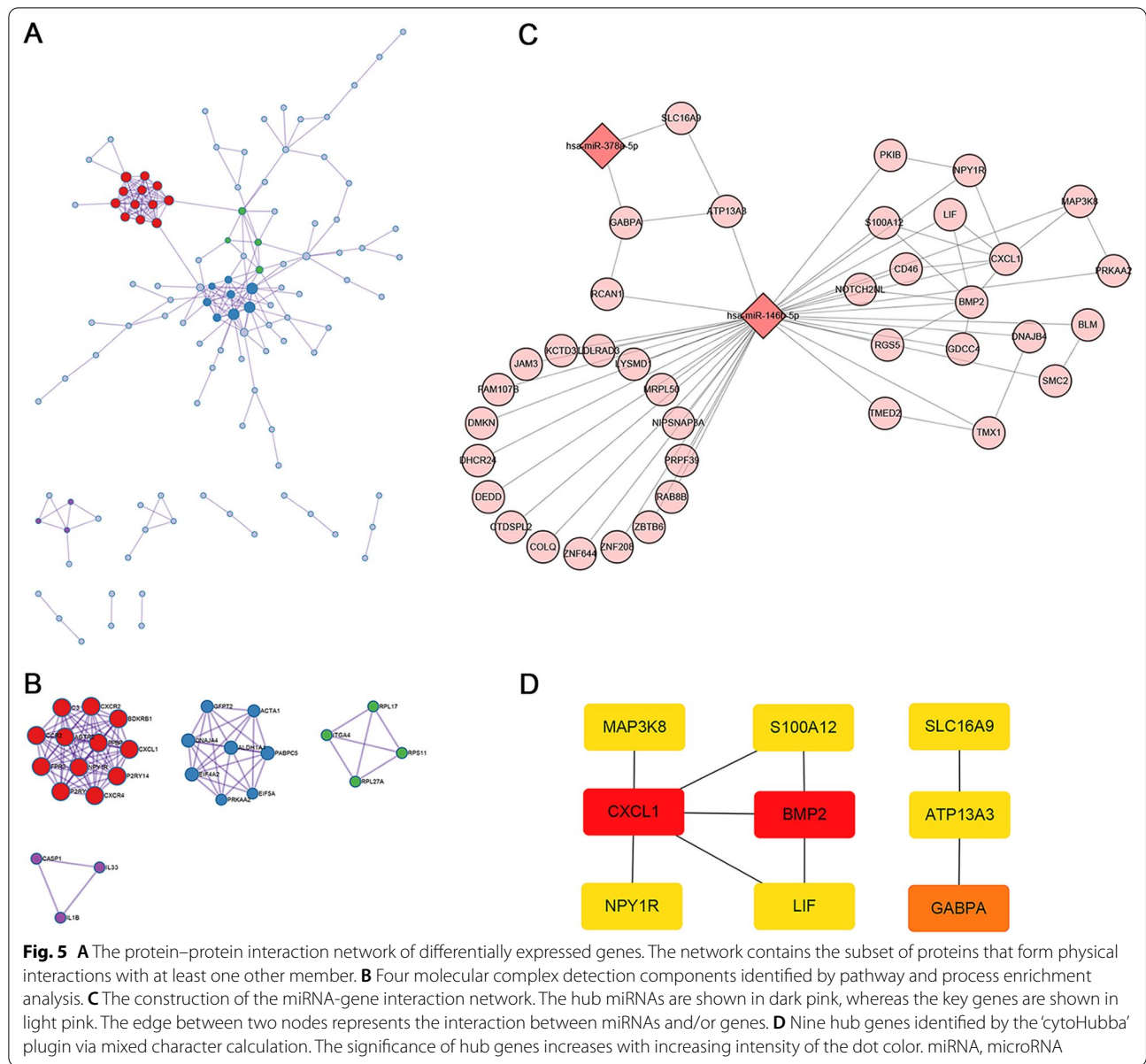
To better understand the interaction between the DEGs, we used Metascape to perform PPI enrichment analysis and illustrate a list of important gene components (Fig. 5A). Then, we performed pathway and process enrichment analysis and identified 4 molecular complex detection (MCODE) components (Fig. 5B). The results showed that biological functions of MCODE components were mainly involved in G protein-coupled receptor activity, generation of precursor metabolites and energy, protein translation process, and positive regulation of cytokine secretion.

Moreover, we constructed a miRNA-gene interaction network to determine hub genes related to the regulation of key signaling pathways in the development of AF. As shown in Fig. 5C, 40 nodes (2 miRNAs and 38 genes) and 57 pairs of interaction relationship were observed in the

network. The hsa-miR-146b-5p was the top hub node in the network with a degree of 36. Further, the ‘cytoHubba’ plugin detected 9 hub genes based on topological parameters of the network, including *ATP13A3*, *BMP2*, *CXCL1*, *GABPA*, *LIF*, *MAP3K8*, *NPY1R*, *S100A12*, and *SLC16A2* (Fig. 5D).

### Discussion

AF is the most common sustained cardiac arrhythmias, characterized by rapid and uncoordinated atrial activity, which leads to an increased risk of complications (e.g. stroke, heart failure, premature death) and constitutes a substantial economic burden worldwide. However, the mechanisms underlying AF are intricate and have not been fully understood; there remains an unmet medical need in the treatment of AF to date. In the present study, we used miRNA and gene expression data to identify predictive and prognostic biomarkers in AF management. Furthermore, we performed integrated analysis to construct a miRNA-gene interaction network for further validation.



Our study performed WGCNA to identify miRNA co-expression modules and the MTRs related to AF; the results showed a significant correlation of red module and blue module with the development of AF. The differential expression analysis showed that 3 miRNAs and 38 key genes were differentially expressed in atrial tissue samples from AF patients. Overall, we identified 2 hub miRNAs (hsa-miR-146b-5p and hsa-miR-378a-5) and constructed a miRNA-gene interaction network containing 9 hub genes (*ATP13A3*, *BMP2*, *CXCL1*, *GABPA*, *LIF*, *MAP3K8*, *NPY1R*, *S100A12*, and *SLC16A2*), of which 7 (*ATP13A3*, *BMP2*, *CXCL1*, *LIF*, *MAP3K8*, *NPY1R*, and *S100A12*) were linked to

hsa-miR-146b-5p. Functional enrichment analysis, including biological functions of MCODE components, showed that the DEGs were mainly involved in inflammation-related items, highlighting a central role of inflammation in the development of AF.

New-onset or exacerbation of AF results from the interaction between trigger and substrate conceptually [6]; moreover, the maintenance of AF requires a vulnerable substrate, largely attributable to genetic predisposition, cardiovascular neurohormonal regulation, and disease-related cardiac remodeling [44, 45]. Cardiac electrophysiology remodeling represents the change in properties of ion channels, whereas structural remodeling



represents the change in organization structure (including atrial dilatation and fibrosis) [46]. While the exact mechanism underlying the development of AF has not been clearly established, atrial dilatation and fibrosis are known to influence a variety of biological processes, primarily through regulation of angiotensin II and related mediators (e.g. transforming growth factor [TGF]- $\beta$ 1, platelet derived growth factor [PDGF], connective tissue growth factor [CTGF]) [47]. AF can stimulate the release of pro-inflammatory cytokines and chemokines related to cardiovascular disease and tissue injuries such as angiotensin II, interleukin (IL)-6, IL-8, and tumor necrosis factor (TNF)- $\alpha$  [48]. In addition, accumulating evidence has demonstrated that mechanical stretch and injuries stimulate leukocyte activation and the release of pro-inflammatory mediators, including nicotinamide adenine dinucleotide phosphate oxidase-derived reactive oxygen species, fibroblast growth factors, and regulatory hormones [49].

Inflammation has been associated with a clinically relevant effect on the increase of cardiac/serum inflammatory biomarkers, and was shown to regulate expression patterns of specific miRNAs and genes in AF patients, thereby potentiating the fibrosis of an established AF substrate [50–53]. AF, largely driven by inflammation, also has been reported to initiate additional inflammatory responses, aggravate atrial remodeling, and deteriorate cardiac function [54]. Such interactions may constitute a key inducement of the transformation from fibroblasts into myofibroblasts and attain significance for the mechanism underlying AF. Other inflammation-related factors such as oxidative stress, apoptosis, and thrombogenesis from the bioactivation of coagulation cascade, were also generally associated with the development of AF [50]. Inflammation, particularly those related to atrial fibrosis, represents an important mechanism of AF. Herein, we summarize a list of miRNAs and genes potentially associated with the inflammatory process in the presence of AF.

BMPs, bone morphogenetic proteins, constitute the largest branch of TGF- $\beta$  family ligands. In addition to enhancing bone regeneration, BMPs also exert anti-inflammatory activity by inhibiting inflammatory cell infiltration in patients with cardiovascular disease [55, 56]. BMP2 and BMP4 have been reported to participate in the development of cardiac function, and so their interruption or discontinuation can lead to cardiac function abnormalities in adulthood [57, 58]. Howden et al. [59] demonstrated that QRS duration and ST interval were significantly reduced in BMP2/BMP4 knockout mice compared with wide-type group, suggesting a potential role for BMP2 and BMP4 in the development of short QT interval syndrome and following initiation of AF. BMP7, an inhibitor of apoptosis, fibrosis, and

calcification, has been shown to stimulate the differentiation from pro-inflammatory infiltrated monocytes to anti-inflammatory M2 macrophages and reduce the progression of cardiac dysfunction in multiple cardiovascular diseases [60]. Our results indicated that BMPs were down-regulated in subjects with AF than in subjects with sinus rhythm, with its expression significantly correlating with the presence of AF.

CXCL1, C-X-C motif chemokine ligand 1, is a member of the alpha chemokine subfamily involved in enhanced neutrophil chemotaxis and phagocytosis during inflammation. It has been suggested that the level of CXCL1 was in keep with the extent of myocardial inflammation and the percentage of CD14<sup>++</sup>DN16<sup>-</sup> monocytes [61]. Moreover, CXCL1-induced myocardial or systemic inflammation in patients with acute cardiomyopathy was significantly associated with monocyte adhesion, macrophage infiltration into the myocardium, and the release of inflammatory cells from bone marrow. Fan et al. [62] also demonstrated that the level of CXCL1 was significantly higher in the AF group than in the control group (by stimulation of cytokines), which was consistent with that seen in our study. Overall, molecular mechanisms of action of CXCL1 behind AF have not been fully elucidated, and remain to be confirmed.

LIF, leukemia inhibitory factor, is a secreted glycoprotein of interleukin-6-type cytokine family. LIF has been shown to transmit biological information through a heterodimer receptor complex comprising LIF receptor and Interleukin 6 signal transducer in a variety of inflammatory processes, including acute phase reaction, hematopoiesis, bone metabolism, and cancer progression [63]. Our study indicated that LIF was down-expressed in atrial tissue of AF patients, suggesting that inflammation-related atrial fibrosis was of particular importance for the development of AF. However, an earlier animal study in adult pigs found no significant difference between the AF group and the control group, which was conflicting with that seen in previous reports [64]. Additional studies would help to determine the position of LIF in the development of AF.

Consistent with the results of recent reports, our study further demonstrated that hsa-miR-146b-5p was over-expressed in AF patients [53, 65]. has-miR-146b-5p has been reported to function as an intracellular mediator in maladaptive remodeling of atrial fibrosis by breaking the balance of MMP/TIMP axis and increasing the level of collagen content of cardiac fibroblasts in AF patients [53, 66]. Emerging evidence suggests that vascular smooth muscle cells (VSMCs) can achieve perceived plasticity and exert non-professional phagocytic activity to maintain inflammatory or senescent condition [67]. Notably, MBNL1 was significantly down-regulated via its

interaction with hsa-miR-146b-5p, which was expected to influence VSMC proliferation and differentiation [68, 69]. In human monocytes and murine cardiomyocytes, has-miR-146b-5p regulated the expression of several pro-inflammatory mediators, including tumor necrosis factor-associated factor 6 (TRAF6), Interleukin-1 receptor (IL-1R)-associated kinase (IRAK1), IL-6, and signal transducer and activator of transcription 3 (STAT3) [70, 71]. Studies in mouse models showed that has-miR-146b-5p down-regulated TRAF6 expression (known to act upstream of the NF- $\kappa$ B pathway) under hypoxia conditions, thereby inhibiting cardiac fibrosis and preventing cardiac dysfunction in patients with heart failure [71]. In addition to hub genes described above, *ATP13A3*, *MAP3K8*, *NPY1R*, and *S100A12* were regulated by hsa-miR-146b-5p; their exact mechanisms of action underlying AF remain elusive, and therefore, further molecular biology studies are warranted in the future. In general, hsa-miR-146b-5p may be a central mediator in the development of AF.

hsa-miR-378a-5p was another hub miRNA that could influence the development of AF, primarily through reversible inhibition of *GABPA* and *SLC16A2*. Our results showed that they were down-regulated, in contrast to hsa-miR-378a-5p that was over-expressed in the presence of AF. hsa-miR-146b-5p, targeting anti-proliferative protein TOB2, has been shown to positively regulate the cell cycle and the following angiogenic process, thereby inducing or enhancing the development of AF by inflammation or fibrosis [72–74]. Despite the lack of studies on the association of *GABPA* with AF, *GABPA* has been reported to down-regulate genes involved in the inflammatory response and oxidative stress at transcriptional level [75]. Additional research is needed to fully elucidate the potential cardioprotective effects, if any, of *SLC16A2* in the development of AF.

The main strength of this study is the use of atrial tissue samples from AF patients instead of commonly reported murine aneurysmal model or human VSMCs in miRNA studies, whose results need additional testing and analysis. Another strength is to provide an overview and perspective from the interaction between miRNAs and genes for AF management, despite requiring more in-depth investigation and further validation.

Our study has several limitations. First, although 2 independent datasets were used for integrated analysis, input data were insufficient to accurately identify and validate predictive and prognostic biomarkers in AF management. Second, given that a substantial proportion of AF samples from the dataset GSE28954 were strongly correlated with dilated atria due to valvular heart disease, our findings were expected to be outmost applicable to patients with AF and concurrent valvular heart disease,

with a controversial generalizability to the whole population. Third, our study merely focused on detecting hub miRNAs and genes related to the development of AF, without considering epigenetic mechanisms behind multifactorial nature of disease, which are generally stable and are often affected by environmental factors. Fourth, despite advances in the pathogenesis of AF over the past decade, there remains a paucity of functional studies on hub molecules identified in our study, making it difficult to understand their implication for potential mechanisms of AF. Fifth, the paucity of confirmatory experiments is another noteworthy limitation; further results from molecular biology studies would be beneficial.

## Conclusion

WGCNA, in combination with a miRNA-gene interaction network, identified 2 hub miRNAs and 9 hub genes in the occurrence and progression of AF. This study may improve the understanding of molecular mechanisms and reveal potential therapeutic targets against AF. Further confirmatory experiments are warranted to validate these findings and elucidate underlying mechanisms.

## Abbreviations

AF: Atrial fibrillation; WGCNA: Weighted gene co-expression network analysis; DEM: Differentially expressed miRNA; DEG: Differentially expressed gene; FDR: False discovery rate; GO: Gene ontology; GS: Gene significance; KEGG: Kyoto encyclopedia of genes and genomes; MCODE: Molecular complex detection; ME: Module eigengene; MM: Module membership; MS: Module significance; MTR: Module-trait relationship; PPI: Protein–protein interaction.

## Acknowledgements

None.

## Authors' contributions

QQ, J-YS, Z-YZ and R-XW developed the concept of the study; QQ, J-YS and Z-YZ designed this study and carried out data analysis; QQ wrote the manuscript with the help from J-YS and Z-YZ; FL, YS, S-SL, and R-XW provided critical reviews of the paper. All authors have read and approved the final manuscript.

## Funding

This work was supported by the National Natural Science Foundation of China [No. 81770331].

## Availability of data and materials

All expression data were acquired from the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) database with the accession numbers GSE28954 and GSE79768.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

**Author details**

<sup>1</sup>Department of Cardiology, The Affiliated Wuxi People's Hospital of Nanjing Medical University, No. 299, Qingyang Road, Wuxi 214023, China. <sup>2</sup>Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China.

Received: 6 July 2021 Accepted: 8 November 2021

Published online: 15 November 2021

**References**

- Zulkifly H, Lip GH, Lane DA. Epidemiology of atrial fibrillation. *Int J Clin Pract*. 2018;72(3):e13070.
- Rahman F, Kwan GF, Benjamin EJ. Global epidemiology of atrial fibrillation. *Nat Rev Cardiol*. 2014;11(11):639–54.
- Lernfelt G, Mandalenakis Z, Hornestam B, Lernfelt B, Rosengren A, Sundh V, et al. Atrial fibrillation in the elderly general population: a 30-year follow-up from 70 to 100 years of age. *Scand Cardiovasc J SCL*. 2020;54(4):232–8.
- Johansson C, Dahlqvist E, Andersson J, Jansson JH, Johansson L. Incidence, type of atrial fibrillation and risk factors for stroke: a population-based cohort study. *Clin Epidemiol*. 2017;9:53–62.
- Chen LY, Chung MK, Allen LA, Ezekowitz M, Furie KL, McCabe P, et al. Atrial fibrillation burden: moving beyond atrial fibrillation as a binary entity: a scientific statement from the American Heart Association. *Circulation*. 2018;137(20):e623–44.
- Huang X, Li Y, Zhang J, Wang X, Li Z, Li G. The molecular genetic basis of atrial fibrillation. *Hum Genet*. 2020;139(12):1485–98.
- Ni H, Iseppa AF, Giles WR, Narayan SM, Zhang H, Edwards AG, et al. Populations of in silico myocytes and tissues reveal synergy of multi-atrial-predominant K(+) -current block in atrial fibrillation. *Br J Pharmacol*. 2020;177(19):4497–515.
- Nattel S, Heijman J, Zhou L, Dobrev D. Molecular basis of atrial fibrillation pathophysiology and therapy: a translational perspective. *Circ Res*. 2020;127(1):51–72.
- Fakuade FE, Steckmeister V, Seibert F, Gronwald J, Kestel S, Menzel J, et al. Altered atrial cytosolic calcium handling contributes to the development of postoperative atrial fibrillation. *Cardiovasc Res*. 2020;2020:cvaa162.
- Van Wagoner DR, Chung MK. Inflammation, inflammasome activation, and atrial fibrillation. *Circulation*. 2018;138(20):2243–6.
- Yang Y, Zhao J, Qiu J, Li J, Liang X, Zhang Z, et al. Xanthine oxidase inhibitor allopurinol prevents oxidative stress-mediated atrial remodeling in Alloxan-induced diabetes mellitus rabbits. *J Am Heart Assoc*. 2018;7(10):e008807.
- Yu X, Wang M, Li L, Zhang L, Chan MTV, Wu WKK. MicroRNAs in atopic dermatitis: a systematic review. *J Cell Mol Med*. 2020;24(11):5966–72.
- Ghafari-Fard S, Vafaee R, Shoorei H, Taheri M. MicroRNAs in gastric cancer: biomarkers and therapeutic targets. *Gene*. 2020;757:144937.
- Titze-de-Almeida SS, Soto-Sánchez C, Fernandez E, Koprlich JB, Brotchie JM, Titze-de-Almeida R. The promise and challenges of developing miRNA-based therapeutics for Parkinson's disease. *Cells*. 2020;9(4):841.
- Galenko O, Jacobs V, Knight S, Taylor M, Cutler MJ, Muhlestein JB, et al. The role of microRNAs in the development, regulation, and treatment of atrial fibrillation. *J Interv Cardiac Electrophysiol Int J Arrhythm Pacing*. 2019;55(3):297–305.
- Binas S, Knyrim M, Hupfeld J, Kloeckner U, Rabe S, Mildnerberger S, et al. miR-221 and -222 target CACNA1C and KCNJ5 leading to altered cardiac ion channel expression and current density. *Cell Mol Life Sci CMLS*. 2020;77(5):903–18.
- Hu J, Gao C, Wei C, Xue Y, Shao C, Hao Y, et al. RBfox2-miR-34a-Jph2 axis contributes to cardiac decompensation during heart failure. *Proc Natl Acad Sci USA*. 2019;116(13):6172–80.
- Cheng WL, Kao YH, Chao TF, Lin YK, Chen SA, Chen YJ. MicroRNA-133 suppresses ZFX3-dependent atrial remodeling and arrhythmia. *Acta Physiol (Oxford, England)*. 2019;227(3):e13322.
- Tremblay BL, Guénard F, Lamarche B, Pérusse L, Vohl MC. Weighted gene co-expression network analysis to explain the relationship between plasma total carotenoids and lipid profile. *Genes Nutr*. 2019;14:16.
- Liang JW, Fang ZY, Huang Y, Liuyang ZY, Zhang XL, Wang JL, et al. Application of weighted gene co-expression network analysis to explore the key genes in Alzheimer's disease. *J Alzheimer's Dis JAD*. 2018;65(4):1353–64.
- Sahu J, Panda D, Baruah G, Patar L, Sen P, Borah BK, et al. Revealing shared differential co-expression profiles in rice infected by virus from reoviridae and sequiviridae group. *Gene*. 2019;698:82–91.
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2013;41:D991–d5.
- Cooley N, Cowley MJ, Lin RC, Marasco S, Wong C, Kaye DM, et al. Influence of atrial fibrillation on microRNA expression profiles in left and right atria from patients with valvular heart disease. *Physiol Genomics*. 2012;44(3):211–9.
- Tsai FC, Lin YC, Chang SH, Chang GJ, Hsu YJ, Lin YM, et al. Differential left-to-right atria gene expression ratio in human sinus rhythm and atrial fibrillation: Implications for arrhythmogenesis and thrombogenesis. *Int J Cardiol*. 2016;222:104–12.
- Niu X, Zhang J, Zhang L, Hou Y, Pu S, Chu A, et al. Weighted gene co-expression network analysis identifies critical genes in the development of heart failure after acute myocardial infarction. *Front Genet*. 2019;10:1214.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003;19(2):185–93.
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innov NY*. 2021;2(3):100141.
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Gene Ontol Consort Nat Genet*. 2000;25(1):25–9.
- Gene Ontology Consortium. The gene ontology resource: enriching a GOld mine. *Nucleic Acids Res*. 2021;49(D1):D325–d34.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
- Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci Publ Protein Soc*. 2019;28(11):1947–51.
- Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res*. 2021;49(D1):D545–51.
- Storey JD, Bass AJ, Dabney A, Robinson D. qvalue: Q-value estimation for false discovery rate control [R package version 2.24.0]. 2021.
- Na KJ, Choi H. Tumor metabolic features identified by (18)F-FDG PET correlate with gene networks of immune cell microenvironment in head and neck cancer. *J Nuclear Med Off Publ Soc Nuclear Med*. 2018;59(1):31–7.
- Zhou XG, Huang XL, Liang SY, Tang SM, Wu SK, Huang TT, et al. Identifying miRNA and gene modules of colon cancer associated with pathological stage by weighted gene co-expression network analysis. *Oncotargets Ther*. 2018;11:2815–30.
- Li W, Wang L, Wu Y, Yuan Z, Zhou J. Weighted gene co-expression network analysis to identify key modules and hub genes associated with atrial fibrillation. *Int J Mol Med*. 2020;45(2):401–16.
- Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v20: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014;42:D92–d7.
- Yang JH, Li JH, Shao P, Zhou H, Chen YQ, Qu LH. starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. *Nucleic Acids Res*. 2011;39:D202–d9.
- Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res*. 2018;46(D1):D296–d302.
- Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellis I, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res*. 2018;46(D1):D239–45.

43. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):D607–13.
44. Lip GY, Fauchier L, Freedman SB, Van Gelder I, Natale A, Gianni C, et al. Atrial fibrillation. *Nat Rev Dis Primers.* 2016;2:16016.
45. Sánchez J, Gomez JF, Martínez-Mateu L, Romero L, Saiz J, Trenor B. Heterogeneous effects of fibroblast-myocyte coupling in different regions of the human atria under conditions of atrial fibrillation. *Front Physiol.* 2019;10:847.
46. Wijesurendra RS, Casadei B. Mechanisms of atrial fibrillation. *Heart (Br Cardiac Soc).* 2019;105(24):1860–7.
47. Burstein B, Nattel S. Atrial fibrosis: mechanisms and clinical relevance in atrial fibrillation. *J Am Coll Cardiol.* 2008;51(8):802–9.
48. Guo Y, Lip GY, Apostolakis S. Inflammation in atrial fibrillation. *J Am Coll Cardiol.* 2012;60(22):2263–70.
49. Dikalov SI, Nazarewicz RR. Angiotensin II-induced production of mitochondrial reactive oxygen species: potential mechanisms and relevance for cardiovascular disease. *Antioxid Redox Signal.* 2013;19(10):1085–94.
50. Harada M, Van Wagoner DR, Nattel S. Role of inflammation in atrial fibrillation pathophysiology and management. *Circ J.* 2015;79(3):495–502.
51. Tran KV, Majka J, Sanghai S, Sardana M, Lessard D, Milstone Z, et al. Micro-RNAs are related to epicardial adipose tissue in participants with atrial fibrillation: data from the MiRhythm study. *Front Cardiovasc Med.* 2019;6:115.
52. Zhang H, Liu L, Hu J, Song L. MicroRNA regulatory network revealing the mechanism of inflammation in atrial fibrillation. *Med Sci Monit Int Med J Exp Clin Res.* 2015;21:3505–13.
53. Wang J, Song S, Xie C, Han J, Li Y, Shi J, et al. MicroRNA profiling in the left atrium in patients with non-valvular paroxysmal atrial fibrillation. *BMC Cardiovasc Disord.* 2015;15:97.
54. Hu YF, Chen YJ, Lin YJ, Chen SA. Inflammation and the pathogenesis of atrial fibrillation. *Nat Rev Cardiol.* 2015;12(4):230–43.
55. Lukač N, Katavić V, Novak S, Šučur A, Filipović M, Kalajzić I, et al. What do we know about bone morphogenetic proteins and osteochondrogenesis in inflammatory conditions? *Bone.* 2020;137:115403.
56. Nguyen V, Meyers CA, Yan N, Agarwal S, Levi B, James AW. BMP-2-induced bone formation and neural inflammation. *J Orthop.* 2017;14(2):252–6.
57. Moazzen H, Wu Y, Engineer A, Lu X, Aulakh S, Feng Q. NOX2 is critical to endocardial to mesenchymal transition and heart development. *Oxid Med Cell Longev.* 2020;2020:1679045.
58. Felin JE, Mayo JL, Loos TJ, Jensen JD, Sperry DK, Gaufin SL, et al. Nuclear variants of bone morphogenetic proteins. *BMC Cell Biol.* 2010;11:20.
59. Howden R, Cooley I, Van Dodewaard C, Arthur S, Cividanes S, Leamy L, et al. Cardiac responses to 24 hrs hyperoxia in Bmp2 and Bmp4 heterozygous mice. *Inhalation Toxicol.* 2013;25(9):509–16.
60. Aluganti Narasimhulu C, Singla DK. The role of bone morphogenetic protein 7 (BMP-7) in inflammation in heart diseases. *Cells.* 2020;9(2):280.
61. Scally C, Abbas H, Ahearn T, Srinivasan J, Mezincescu A, Rudd A, et al. Myocardial and systemic inflammation in acute stress-induced (takotsubo) cardiomyopathy. *Circulation.* 2019;139(13):1581–92.
62. Fan G, Wei J. Identification of potential novel biomarkers and therapeutic targets involved in human atrial fibrillation based on bioinformatics analysis. *Kardiol Pol.* 2020;78(7–8):694–702.
63. Jalvy S, Veschambre P, Fédou S, Rezvani HR, Thézé N, Thiébaud P. Leukemia inhibitory factor signaling in *Xenopus* embryo: Insights from gain of function analysis and dominant negative mutant of the receptor. *Dev Biol.* 2019;447(2):200–13.
64. Tsai CT, Lin JL, Lai LP, Lin CS, Huang SK. Membrane translocation of small GTPase Rac1 and activation of STAT1 and STAT3 in pacing-induced sustained atrial fibrillation. *Heart Rhythm.* 2008;5(9):1285–93.
65. Wang JG, Meng X, Han J, Li Y, Luo TG, Wang J, et al. Differential expressions of miRNAs in patients with nonvalvular atrial fibrillation. *Zhonghua Yi Xue Za Zhi.* 2012;92(26):1816–9.
66. Wang J, Wang Y, Han J, Li Y, Xie C, Xie L, et al. Integrated analysis of microRNA and mRNA expression profiles in the left atrium of patients with nonvalvular paroxysmal atrial fibrillation: role of miR-146b-5p in atrial fibrosis. *Heart Rhythm.* 2015;12(5):1018–26.
67. Sorokin V, Vickneson K, Kofidis T, Woo CC, Lin XY, Foo R, et al. Role of vascular smooth muscle cell plasticity and interactions in vessel wall inflammation. *Front Immunol.* 2020;11:599415.
68. Lim YH, Ryu J, Kook H, Kim YK. Identification of long noncoding RNAs involved in differentiation and survival of vascular smooth muscle cells. *Mol Ther Nucleic Acids.* 2020;22:209–21.
69. Woo CC, Liu W, Lin XY, Dorajoo R, Lee KW, Richards AM, et al. The interaction between 30b–5p miRNA and MBNL1 mRNA is involved in vascular smooth muscle cell differentiation in patients with coronary atherosclerosis. *Int J Mol Sci.* 2019;21(1):11.
70. Taganov KD, Boldin MP, Chang KJ, Baltimore D. NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc Natl Acad Sci USA.* 2006;103(33):12481–6.
71. Chouvarine P, Legchenko E, Geldner J, Riehle C, Hansmann G. Hypoxia drives cardiac miRNAs and inflammation in the right and left ventricle. *J Mol Med (Berl).* 2019;97(10):1427–38.
72. Feng M, Li Z, Aau M, Wong CH, Yang X, Yu Q. Myc/miR-378/T0B2/cyclin D1 functional module regulates oncogenic transformation. *Oncogene.* 2011;30(19):2242–51.
73. Krist B, Florczyk U, Pietraszek-Gremplewicz K, Józkowicz A, Dulak J. The role of miR-378a in metabolism, angiogenesis, and muscle biology. *Int J Endocrinol.* 2015;2015:281756.
74. Berntsson J, Smith JG, Johnson LSB, Söderholm M, Borné Y, Melander O, et al. Increased vascular endothelial growth factor D is associated with atrial fibrillation and ischaemic stroke. *Heart (Br Cardiac Soc).* 2019;105(7):553–8.
75. Pinheiro DML, de Oliveira AHS, Coutinho LG, Fontes FL, de Medeiros Oliveira RK, Oliveira TT, et al. Resveratrol decreases the expression of genes involved in inflammation through transcriptional regulation. *Free Radical Biol Med.* 2019;130:8–22.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

