

# Integrative analysis reveals essential mRNA, long non-coding RNA (lncRNA), and circular RNA (circRNA) in paroxysmal and persistent atrial fibrillation patients

Haoliang Sun<sup>#</sup> , Junjie Zhang<sup>1, #</sup> , Yongfeng Shao 

Department of Cardiovascular Surgery, the First Affiliated Hospital of Nanjing Medical University; Jiangsu-China  
<sup>1</sup>Department of Cardiothoracic Surgery, Changzhou Wujin People's Hospital; Jiangsu-China

## ABSTRACT

**Objective:** This study aimed to investigate the functions of mRNA, long non-coding RNA (lncRNA), and circular RNA (circRNA) in paroxysmal and persistent atrial fibrillation (AF) patients.

**Methods:** A total of 9 left atrial appendage (LAA) tissues were collected from patients with AF (ParoAF patients = 3 and PersAF patients = 3) and donors (n=3). Genes and circRNAs were identified by per kilobase per million reads (RPKM) and number of circular reads/number of mapped reads/read length (SRPBM), respectively. Differentially expressed mRNAs (DE mRNAs), lncRNAs (DE lncRNAs), and circRNAs (DE circRNAs) were identified by  $|\log_2(\text{Fold Change})| \geq 2$  and p-value < 0.05. Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed. Protein-protein, mRNA-lncRNA, and circRNA-miRNA interaction networks were constructed. In addition, logistic analysis was conducted among AF and circRNAs.

**Results:** A total of 285 (116 up-regulated and 169 down-regulated) and 275 (110 up-regulated and 165 down-regulated) DE mRNAs, 575 (276 up-regulated and 299 down-regulated) and 583 (330 up-regulated and 253 down-regulated) DE lncRNAs, and 83 (48 up-regulated and 35 down-regulated) and 99 (58 up-regulated and 41 down-regulated) circRNAs were detected in ParoAF and PersAF, respectively, as compared with control. MAPK signal pathway as well as voltage-dependent, L type, and alpha 1C subunit calcium channel (*CACNA1C*) might participate in AF occurrence by preventing atrial parasympathetic remodeling. Collagen type I alpha 1 (*COL1A1*) and *COL1A2* mostly participated in the enriched GO and KEGG terms and connected with most of the DE mRNAs. The expression of chr10:69902697169948883 was a protective factor against PersAF after adjusting for age (p=0.022, 95% CI: 0.003–0.634).

**Conclusion:** We found that some mRNAs, lncRNAs, circRNAs, and pathways play essential roles in AF pathogenesis and development. Moreover, one protective factor against PersAF was detected.

**Keywords:** atrial fibrillation, RNA-Seq, pathway, mRNA, lncRNA, circRNA

**Cite this article as:** Sun H, Zhang J, Shao Y. Integrative analysis reveals essential mRNA, long non-coding RNA (lncRNA), and circular RNA (circRNA) in paroxysmal and persistent atrial fibrillation patients. *Anatol J Cardiol* 2021; 25: 414-28.

## Introduction

Atrial fibrillation (AF) is the most common type of arrhythmia, which affected 37.574 million individuals worldwide in 2017 (1). AF increases the probability of thromboembolism, ischemic stroke, congestive heart failure, psychological distress, and death (2). It impairs the quality of life of patients severely, especially patients who require open-heart surgery (3). There are three kinds of AF, which include

paroxysmal AF (ParoAF), persistent AF (PersAF), and permanent AF (4). PersAF appears to be more reliant on fibroblast proliferation and myocyte-fibroblast coupling when compared with ParoAF, which is mainly due to primary on pulmonary vein triggers (5). The mechanisms related to occurrence and development of ParoAF and PersAF have been studied continuously for a long time.

In recent years, next-generation sequencing (NGS) provided a more efficient way to study AF. There are growing evidences

<sup>#</sup>These authors contributed equally to this work.

**Address for Correspondence:** Yongfeng Shao, MD, Department of Cardiovascular Surgery, the First Affiliated Hospital of Nanjing Medical University, No. 300, Guangzhou Road, Nanjing/210029, Jiangsu-China  
 Phone: +86-13851536726 E-mail: shaoyongfeng\_2019@163.com

**Accepted Date:** 22.12.2020 **Available Online Date:** 21.05.2021

©Copyright 2021 by Turkish Society of Cardiology - Available online at [www.anatoljcardiol.com](http://www.anatoljcardiol.com)  
 DOI:10.14744/AnatolJCardiol.2020.57295



## HIGHLIGHTS

- Voltage-dependent, L type, and alpha 1C subunit calcium channel (*CACNA1C*) might involve in atrial fibrillation occurrence by preventing atrial parasympathetic remodeling.
- The expression of chr10:69902697|69948883 was a protective factor against persistent atrial fibrillation.
- Collagen type I alpha 1 (*COL1A1*) and *COL1A2* might be essential genes for atrial fibrillation.

that genome-wide association, microRNA (miRNA), mRNA, long non-coding RNA (lncRNA), and circular RNA (circRNA) play essential roles in pathogenesis and development of AF (6-11). For instance, miR29 was shown to play a crucial role in atrial fibrotic remodeling and may also be considered as a biomarker and/or therapeutic target (6). Sinner et al. (9) provided five novel loci for AF, which might be novel molecular targets for future biological and pharmacological investigation.

lncRNA longer than 200 nucleotides in length with non-protein coding function has attracted the attention of researchers (12). lncRNA plays a central role in many functions during heart development and various heart diseases, including cardiac hypertrophy, cardiac fibrosis (CF), AF, and heart failure (HF) (13-16). circRNA is a non-coding RNA that regulates gene expression and acts as miRNA "sponges" by competing with endogenous RNA (ceRNA) to suppress the activity of specific miRNA (17). Moreover, circRNAs serve as biomarkers for many diseases (18). In the study by Hu et al. (19), they detected several circRNAs that might be associated with inflammatory responses in AF. However, RNA-related studies on ParoAF and PersAF patients are still inadequate.

To identify the differentially expressed mRNAs (DE mRNAs), lncRNAs (DE lncRNAs), and circRNAs (DE circRNAs) on ParoAF and PersAF, left atrial appendages (LAA) in patients were investigated and studied through high-throughput RNA sequencing (RNA-Seq). We performed Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses for functional annotation of the DE mRNAs, mRNAs near the DE lncRNAs, and host genes of DE circRNAs were performed. Finally, the protein-protein, mRNA-lncRNA, and ceRNA interaction network was constructed. These results will provide valuable genes and pathways, which regulate the occurrence and development of AF.

## Methods

### Sample collection

The LAA of the donors was recruited. A total of 9 LAA tissues collected from the hearts of patients with AF (ParoAF patients=3 and PersAF patients=3) and donors (n=3) were used for NGS sequencing. All the healthy donors had sinus rhythm and were

in good condition. The ParoAF and PersAF LAA tissues were selected from the maze surgery of AF patients.

### RNA extraction

The LAA tissues were washed with ddH<sub>2</sub>O, cut into small pieces with a clean scalpel, and then snapped into liquid nitrogen for at least 4 h and finally stored at -80° until use. TRIzol Reagent (Invitrogen, Carlsbad, CA, USA) was used for total RNA extraction. The total amount was quantified and the quality was assessed. Samples with RNA amount >5 µg, RNA integrity number >7, and OD<sub>260/280</sub> >1.9 were employed for library preparation.

### Library preparation and sequencing

Ribosomal RNA (rRNA) in each RNA sample was deleted using Ribo-Minus kit (Thermo Fish, Waltham, MA, USA). Next, each RNA sample was quantified and applied for library preparation using TruSeq RNA Library Preparation Kit v2 (Illumina Inc, USA). All libraries were loaded into one lane on Illumina HiSeq X Ten (Illumina Inc, USA) platform, followed by 2 × 150 bp paired-end sequencing.

### Bioinformatics analysis of DE mRNAs, DE lncRNAs, and DE circRNAs

FastQC (v0.11.5, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was used to process raw sequencing data. Clean reads were mapped onto Homo sapiens (assembly GRCh38. P12) using TopHat (v2.0.12, <https://ccb.jhu.edu/software/tophat/index.shtml>), followed by assembly using Cufflinks (v2.2.1, <http://cufflinks.cbc.umd.edu>). Transcripts with reads per kilobase of exon per million read mapped (RPKM > 0) were retained for further analysis. For circRNA detection, RNA sequencing data were analyzed by CircRNA Identifier (CIRI) and an algorithm for *de novo* circRNA identification (20). Junction reads and circRNA candidates in SAM files were scanned twice by CIRI. CircRNAs that have junction reads ≥2 and appeared in at least two samples were regarded as high-confidence circRNAs. Finally, edgeR (<http://www.bioconductor.org/packages/release/bioc/html/edgeR.html>) was utilized to identify the DE mRNAs, DE lncRNAs, and DE circRNAs by pairwise comparisons, with the threshold of |Log<sub>2</sub> Fold Change (FC)| ≥2 and *p*-value < 0.05.

### GO and KEGG pathway enrichment analyses

We focused on the neighboring target genes of DE lncRNAs. We searched mRNAs nearest to DE lncRNAs both in upstream and downstream. Each circRNA was annotated to the linear host mRNA according to their position relationship on the chromosome. We used the linear host mRNA as the proxy of its related circRNAs.

GO (<http://geneontology.org/>) enrichment analysis was performed on all DE mRNAs, mRNAs nearby the DE lncRNAs, and host genes of DE circRNAs, with corrected *p*-value <0.05 as a threshold.

KEGG database (<https://www.kegg.jp/>) was applied for pathway mapping. KEGG pathway enrichment analysis identified

**Table 1. Primers designed for qRT-PCR validation of selected circRNAs**

CircRNAs	Forward primer (5'–3')	Reverse primer (5'–3')	Product length (bp)
chr10:69881195 69882097	CCGCTCGACTCCTCTCATTG	CCTCCCGGTTCACTCAAAA	208
chr10:69902697 69948883	GCAGGAATAGCGTCTGTGT	GAGACTCAAGCACTTCCGGG	200

CircRNA - circular RNA; qRT-PCR - Quantitative real-time PCR

metabolic pathways or signaling pathways significantly enriched in all DE mRNAs, mRNAs nearby the DE lncRNAs, and host genes of DE circRNAs in the whole genome using KOBAS, taking  $p$ -value  $<0.05$  as a threshold to identify the enrichment pathway. We further plotted the network of top ten KEGG pathways and mRNA in ParoAF and PersAF using Cytoscape software v3.6.1.

### Interaction networks

We constructed the protein-protein interaction network of DE mRNAs in ParoAF\_Control and PersAF\_Control. STRING database (<https://string-db.org/cgi/input.pl>) was used for the interaction correlation study and only those with correlation ( $r$ )  $>0.7$  were displayed. The DE mRNAs, which were also targeted by DE lncRNAs, were selected for the mRNA-lncRNA network study.

The common DE circRNAs in ParoAF\_Control and PersAF\_Control were selected for the circRNA-miRNA interaction network analysis. The miRNAs binding to circRNAs were identified in miRanda (v3.3a, <http://www.microrna.org>). CircRNAs with more miRNA binding sites were screened as candidate "sponge" circRNAs and used to construct the ceRNA interaction network. Cytoscape software v3.6.1 was used for the network drawing.

### Samples used for validation

Quantitative real-time polymerase chain reaction (qRT-PCR) was performed to validate the results of sequencing analysis on the DE lncRNAs and DE mRNAs in samples of Control, ParoAF, and PersAF. RNA samples were prepared as previously mentioned. The synthesis of the first-strand cDNA was compounded using PrimeScript<sup>TM</sup> RT reagent kit (Takara, Otsu, Shiga, Japan), followed by amplification using AceQ qPCR SYBR Green Master Mix (Vazyme Biotech Co., Nanjing, China) on ABI 7500 system RT-PCR instrument (Applied Biosystems, FosterCity, CA, USA). Expression levels of collagen type I alpha 1 (*COL1A1*), FK506 binding protein 5 (*FKBP5*), collagen type I alpha 2 (*COL1A2*), and RP11-442H21.2 were validated by qRT-PCR in the samples and ParoAF patients and 4 genes *COL1A1*, *COL1A2*, chemokine (C-C motif) ligand 5 (*CCL5*), and CTA-134P22.2 were detected in control samples and PersAF patients. The mRNAs and lncRNAs expression levels were calculated by the  $2^{-\Delta\Delta Ct}$  method.

Forty LAA samples isolated from the donors ( $n = 20$ , control) and patients with PersAF ( $n = 20$ ) were used for validation of circRNA expression. The primary clinical characteristics including, age, body mass index (BMI), left ventricular ejection fraction (LVEF), left atrium (LA), left ventricular diastolic dimension (LVDD), left ventricular systolic diameter (LVSD), and ventricular rate was collected from patients. RNA samples were extracted using previously reported methods. The expression levels of

circRNAs including, chr10: 69881195|69882097 and chr10: 69902697|69948883, were detected by qRT-PCR in heart samples from healthy donors and PersAF subjects. Primers were designed and presented in Table 1. Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) was used as the housekeeping gene. HiScript II one Step qRT-PCR Probe Kit (Vazyme Biotech Co., Nanjing, China) was used for the measurement. qRT-PCR was performed with ABI 7500 RT-PCR system (Thermo Fisher Scientific). The specific generation of the expressed PCR product was validated by melting curve analysis. The expression levels of circRNAs were calculated using the  $2^{-\Delta\Delta Ct}$  method.

### Statistical analysis

All statistical analyses were performed using IBM SPSS Statistics v22.0 (IBM Corp., Armonk, NY, USA) or GraphPad Prism version 8 (GraphPad Software, San Diego, CA). Comparison among three groups was performed by applying the non-parametric Kruskal–Wallis H test, followed by Mann–Whitney U-test. Data were presented as median (interquartile range). Comparison between two groups was performed using the non-parametric Mann–Whitney U-test, followed by Dunn's test. Normal data were presented as median (interquartile range). Logistics regression analysis was performed to determine the risk factors for AF after adjusting for age.  $P$ -value  $<0.05$  was considered as statistically significant.

## Results

### Clinical characteristics

There were no significant differences in age, BMI, LVDD, LVSD, systolic blood pressure, diastolic blood pressure, CHA2DS2-VASc score, HAS-BLED score, and EHRA across control donors, ParoAF, and PersAF patients ( $p > 0.05$ , Table 2).

### Differentially expressed mRNAs, lncRNAs, and circRNAs

A total of 285 (116 up-regulated and 169 down-regulated) and 275 (110 up-regulated and 165 down-regulated) DE mRNAs were detected in ParoAF and PersAF, respectively, when compared to control (Fig. 1a, 1b). A total of 63 common DE mRNAs were expressed in ParoAF\_Control and PersAF\_Control (Fig. 1c). The DE mRNAs of ParoAF\_Control and PersAF\_Control are provided in Supplementary Files 1, 2, respectively.

A total of 575 (276 up-regulated and 299 down-regulated) and 583 (330 up-regulated and 253 down-regulated) DE lncRNAs were detected in ParoAF and PersAF, respectively, when compared to the control (Fig. 1d, 1e). A total of 116 common DE lncRNAs were expressed in ParoAF\_Control and PersAF\_Control (Fig. 1f). The DE mRNAs of ParoAF\_Control and PersAF\_Control are provided in Supplementary Files 3, 4, respectively.

**Table 2. Clinical characteristics of subjects used for the sequencing**

Items	Control (n=3)	ParoAF (n=3)	PersAF (n=3)	P-value
Age (years)	39.00 (38.00–42.00)	54.00 (47.00–60.00)	57.00 (51.00–64.00)	0.061 <sup>a</sup>
Gender (F/M)	1/2	1/2	1/2	-
BMI (kg/m <sup>2</sup> )	21.70 (18.80–24.40)	23.07 (22.10–24.00)	25.17 (24.50–26.00)	0.141 <sup>a</sup>
LA (mm)	35.00 (34.00–38.00)	37.00 (30.00–42.00)	42.00 (40.00–51.00)	0.641 <sup>a</sup>
LVDD (mm)	46.00 (39.00–48.00)	46.00 (45.00–49.00)	44.00 (44.00–48.00)	0.219 <sup>a</sup>
LVSD (mm)	28.00 (24.00–30.00)	30.00 (30.00–36.00)	30.00 (28.00–31.00)	0.079 <sup>a</sup>
Ventricular rate	54.00 (50.00–64.00)	85.00 (81.00–108.0)	76.00 (65.00–80.00)	0.275 <sup>a</sup>
Systolic blood pressure	127.00 (118.00–130.00)	112.00 (101.00–129.00)	119.00 (111.00–134.00)	0.436 <sup>a</sup>
Diastolic blood pressure	78.00 (70.00–87.00)	84.00 (72.00–92.00)	84.00 (70.00–95.00)	0.297 <sup>a</sup>
CHA <sub>2</sub> DS <sub>2</sub> -VASc score	-	3.00 (2.00–5.000)	2.00 (1.00–3.00)	0.478 <sup>b</sup>
HASBLED score	-	4.00 (3.00–5.00)	2.00 (1.00–5.00)	0.988 <sup>b</sup>
EHRA	-	I(1)/II(2)	I(1)/II(2)	-

Values are expressed as median (interquartile range). P-value <0.05 was considered as statistically significant.  
(<sup>a</sup>) Differences were analyzed by non-parametric Kruskal–Wallis H test, followed by the Mann–Whitney U-test.  
(<sup>b</sup>) Differences were analyzed by non-parametric Mann–Whitney U-test, followed by Dunn's test.  
BMI - body mass index; LA- left atrium; LVDD- left ventricular diastolic dimension; LVEF- left ventricular ejection fraction; LVSD- left ventricular systolic diameter; CHA<sub>2</sub>DS<sub>2</sub>-VASc contains congestive heart failure/LV dysfunction, hypertension, age ≥75, diabetes mellitus, stroke/TIA/TE, Vascular disease, age =65–74, and sex category. HAS-BLED contains parameters of hypertension, abnormal renal function, abnormal liver function, previous stroke, bleeding history or predisposition, history of labile international normalized ratio, age ≥65, concomitant aspirin or nonsteroidal anti-inflammatory drug therapy, and substantial alcohol intake. EHRA - European Heart Rhythm Association; PersAF - persistent atrial fibrillation; ParoAF - paroxysmal atrial fibrillation

**Table 3. Clinical characteristics of Control donors and PersAF patients**

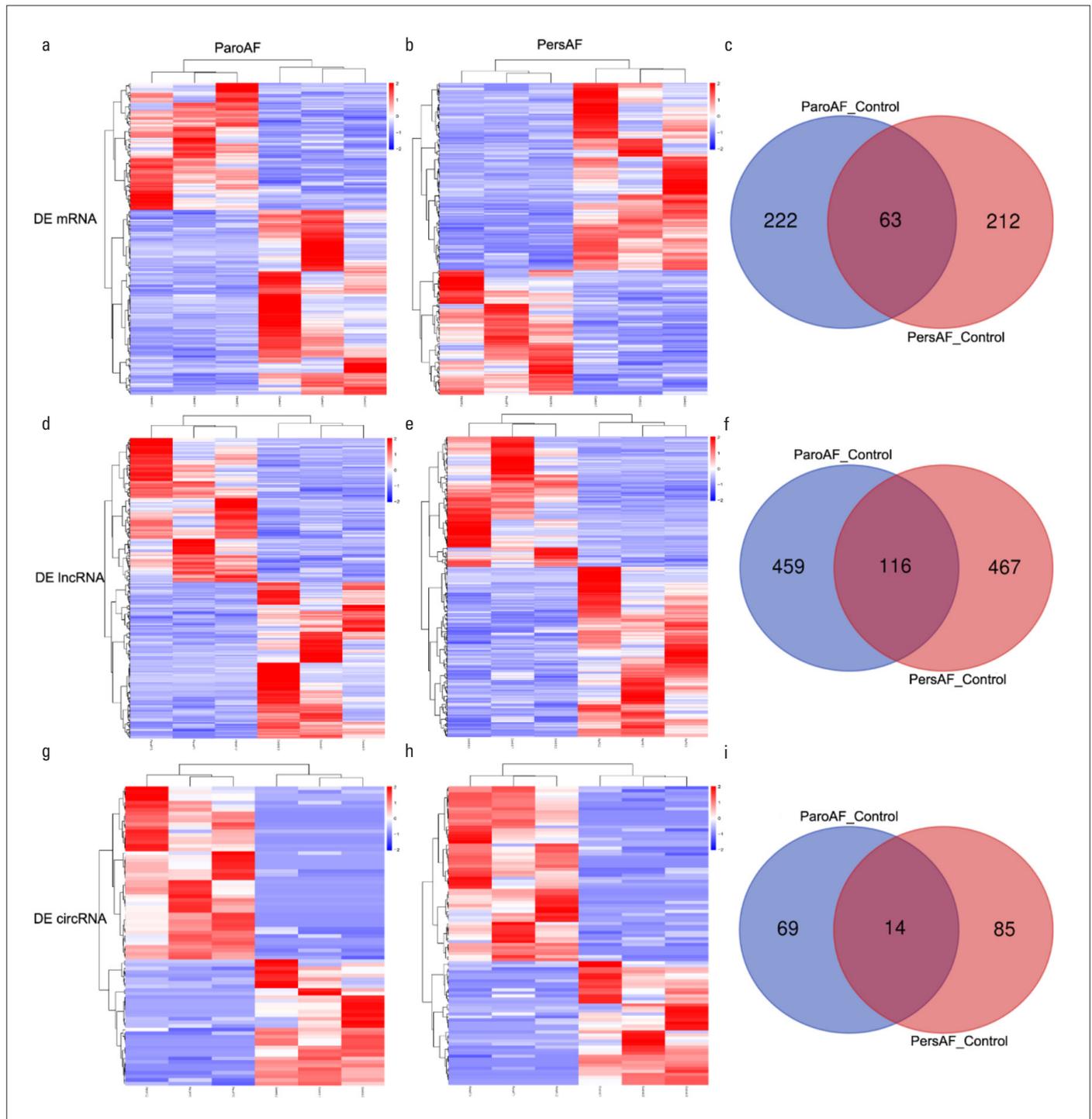
Items	Control (n=20)	PersAF (n=20)	P-value
Age (years)	31.00 (25.00–40.75)	65.00 (55.25–69.50)	<1.00E-04
BMI (kg/m <sup>2</sup> )	23.50 (20.85–25.32)	26.10 (24.12–27.87)	0.028
LA (mm)	35.00 (33.00–37.75)	42.50 (39.00–46.00)	<1.00E-04
LVDD (mm)	47.00 (45.25–48.00)	46.50 (45.00–50.00)	0.565
LVSD (mm)	30.00 (28.25–31.75)	30.00 (29.00–32.00)	0.841
Ventricular rate	69.00 (62.00–70.75)	84.50 (72.50–89.75)	0.024
Systolic blood pressure	115.50 (108.00–125.00)	122.00 (106.25–128.25)	0.665
Diastolic blood pressure	74.50 (72.00–78.00)	78.50 (74.25–83.50)	0.761
CHA <sub>2</sub> DS <sub>2</sub> -VASc score	-	2.50 (2.00–4.00)	-
HASBLED score	-	3.00 (1.25–3.00)	-
EHRA	-	I(3)/II(16)/III(1)	-
chr10:69881195169882097	0.30 (0.19–0.49)	0.025 (0.00–0.13)	0.004
chr10:69902697169948883	0.02 (0.00–0.09)	0.47 (0.23–0.77)	<1.00E-04

Values are expressed as median (interquartile range). Differences were analyzed by non-parametric Mann–Whitney U-test, followed by Dunn's test. P-value <0.05 was considered as statistically significant.  
BMI- body mass index; LA- left atrium; LVDD- left ventricular diastolic dimension; LVSD- left ventricular systolic diameter; CHA<sub>2</sub>DS<sub>2</sub>-VASc contains congestive heart failure/LV dysfunction, hypertension, age ≥75, diabetes mellitus, stroke/TIA/TE, Vascular disease, age =65–74, and sex category. HAS-BLED contains parameters of hypertension, abnormal renal function, abnormal liver function, previous stroke, bleeding history or predisposition, history of labile international normalized ratio, age ≥65, concomitant aspirin or nonsteroidal anti-inflammatory drug therapy, and substantial alcohol intake. EHRA - European Heart Rhythm Association; PersAF - persistent atrial fibrillation

A total of 83 (48 up-regulated, 35 down-regulated) and 99 (58 up-regulated, 41 down-regulated) circRNAs were DE in ParoAF and PersAF, respectively, when compared with control samples (ParoAF\_Control and PersAF\_Control), including 14 common circRNAs (Fig. 1g-1i). The total DE circRNAs of ParoAF\_Control and PersAF\_Control are listed in Supplementary Files 5 and 6.

### Alternative circularization of dysregulated circRNAs and logistics regression analysis

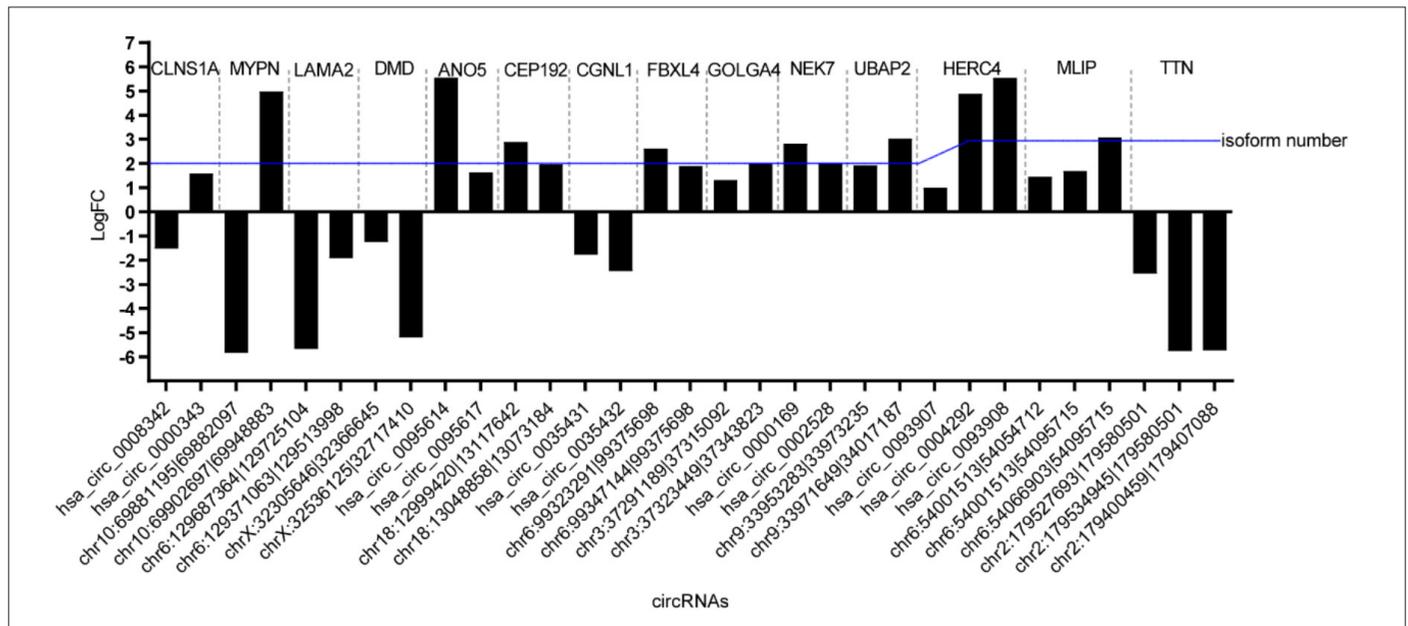
Among the host genes of DE circRNAs in ParoAF\_Control and PersAF\_Control, 14 host genes had more than two circRNA isoforms, of which 11 genes produced two circRNA isoforms and the other three genes produced three circRNA isoforms



**Figure 1.** Expression profiles of differentially expressed mRNAs, lncRNAs, and circRNAs of ParoAF and PersAF when compared to control. (a, b) represent the gene expression level of DE mRNAs in ParoAF\_Control and PersAF\_Control. (c) Shows the total and common number of DE mRNAs in ParoAF\_Control and PersAF\_Control. (d, e) represent the gene expression level of DE lncRNAs in ParoAF\_Control and PersAF\_Control. (f) shows the total and common number of DE lncRNAs in ParoAF\_Control and PersAF\_Control. (g, h) represent the gene expression level of DE circRNAs in ParoAF\_Control and PersAF\_Control. (i) shows the total and common number of DE circRNAs in ParoAF\_Control and PersAF\_Control. circRNA: circular RNA; DE - differentially expressed; lncRNA - long non-coding RNA; ParoAF - paroxysmal atrial fibrillation; PersAF - persistent atrial fibrillation

(Fig. 2). HECT and RLD domain containing E3 ubiquitin-protein ligase 4 (*HERC4*), muscular LMNA-interacting protein (*MLIP*), and titin (*TTN*) had three circRNA isoforms. It was worth noting that chloride channel nucleotide-sensitive 1A (*CLNS1A*) and

myopalladin (*MYPN*) produced circRNA isoforms with opposite expression patterns in PersAF patients when compared to control (Fig. 2). *CLNS1A* was the host gene of hsa\_circ\_0008342 and hsa\_circ\_0000343. *MYPN* was the host gene of



**Figure 2.** The number and expression level of circRNAs that have no less than two isoforms. The blue line represents the number of isoforms. Most of the genes have consistent isoform patterns, except *CLNS1A* and *MYPN*. *HERC4*, *MLIP* and *TTN* have three isoforms. *ANO5* - anoctamin 5; *CEP192* - centrosomal protein 192kDa; *CGNL1* - cingulin-like 1; *CLNS1A* - nucleotide-sensitive 1A; *DMD* - dystrophin; *FBXL4* - F-box and leucine-rich repeat protein 4; FC - fold change; *GOLGA4* - golgin A4; *HERC4* - RLD domain containing E3 ubiquitin-protein ligase 4; *LAMA2* - laminin, alpha 2; *MLIP* - muscular *LMNA*-interacting protein; *MYPN* - myopalladin; *NEK7* - *NIMA*-related kinase 7; *ParoAF* - paroxysmal atrial fibrillation; *PersAF* - persistent atrial fibrillation; *TTN* - titin; *UBAP2* - ubiquitin associated protein 2

**Table 4. Logistic analysis among AF and the expression of chr10:69881195|69882097 and chr10:69902697|69948883**

Items	$\beta$	OR	P-value	95% CI	
				Lower	Upper
chr10:69881195 69882097	-1.564	0.209	0.681	<0.0001	359.261
chr10:69902697 69948883	-3.102	0.045	0.022	0.003	0.634

AF - atrial fibrillation; CI - confidence interval; OR - odds ratios

chr10:69881195|69882097 and chr10:69902697|69948883. Among all the genes that had the same expression patterns, laminin, alpha 2 (*LAMA2*), dystrophin (*DMD*), anoctamin 5 (*ANO5*), and *HERC4* varied greatly on the level of expression (Fig. 2).

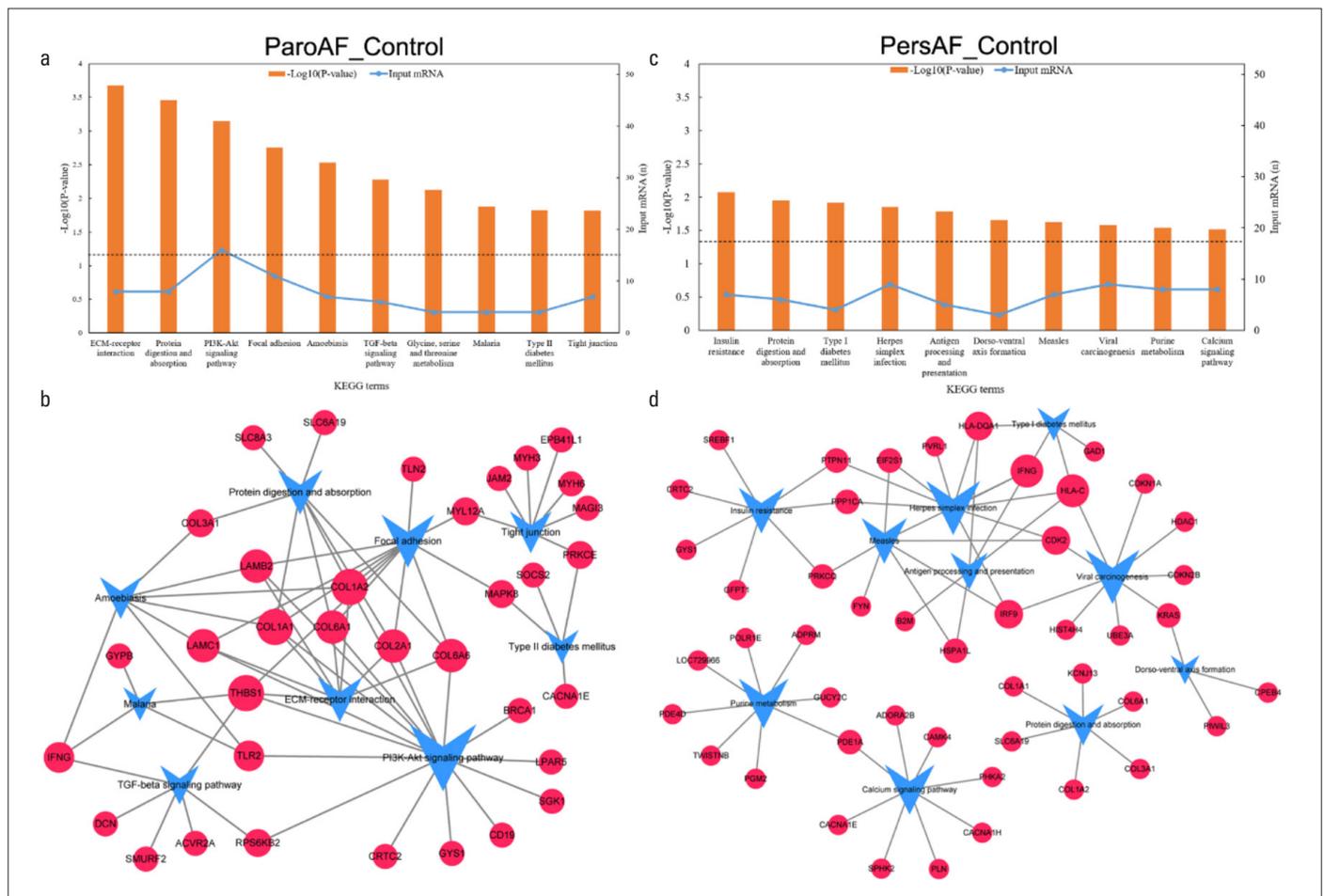
Since the association of *MYPN* with muscle diseases including hypertrophic cardiomyopathy, the two circRNA isoforms of it (chr10:69881195|69882097 and chr10:69902697|69948883) had opposite expression profiles in *PersAF* compared with control. In the validation cohort, there was significant difference in age ( $p < 0.0001$ ), BMI ( $p = 0.028$ ), LA ( $p < 0.0001$ ), ventricular rate ( $p = 0.024$ ), and expression levels of chr10:69881195|69882097 ( $p = 0.004$ ) and chr10:69902697|69948883 ( $p < 0.0001$ , Table 3) between control and *PersAF* samples. The expression level of chr10:69881195|69882097 and chr10:69902697|69948883 was significantly down-regulated and up-regulated, respectively, in patients with *PersAF* when compared with donors (Control). Logistics regression analysis showed that expression of chr10:69881195|69882097 was not associated with *PersAF* after adjusting for patients' age, while chr10:69902697|69948883 was a protective factor against *PersAF* (95% CI: 0.003–0.634,  $p = 0.022$ , Table 4).

### GO and KEGG pathway analysis on DE mRNAs

With the criterion of corrected  $p$ -value  $< 0.05$ , there were only two GO terms selected in *ParoAF*\_Control and they were "extracellular matrix" and "platelet-derived growth factor binding." *COL1A2*, *COL6A1*, *COL3A1*, and *COL2A1* participated in the two enriched GO terms. *COL1A2*, *COL6A1*, and *COL3A1* were up-regulated in *ParoAF* when compared to control. *COL2A1* was down-regulated in *ParoAF* when compared to control. With the same criterion, "collagen type I" and "platelet-derived growth factor binding" were enriched in *PersAF*\_Control, while *COL1A2* and *COL1A1* were enriched in both GO terms. Both *COL1A2* and *COL1A1* were up-regulated in *PersAF* when compared to control.

KEGG pathway enrichment analysis revealed that the top enriched pathways were "ECM-receptor interaction," "protein digestion and absorption," "PI3K-Akt signaling pathway Focal adhesion Amoebiasis," "TGF-beta signaling pathway," and "Amoebiasis" in *ParoAF* patients when compared to control donors (Fig. 3a). Up-regulated DE mRNAs, such as *COL1A1*, *COL1A2*, and laminin, beta 2 (*LAMB2*), participated in most of the top ten pathways (Fig. 3b). Down-regulated DE mRNA *COL2A1* was involved in most of the top ten pathways. DE mRNAs, decorin (*DCN*), SMAD specific E3 ubiquitin-protein ligase 2 (*SMURF2*), and interferon gamma (*IFNG*), participated in TGF- $\beta$  signaling pathway.

The top enriched pathways were "Insulin resistance," "Protein digestion and absorption," "Type I diabetes mellitus," "herpes simplex infection," and "Antigen processing and presentation Dorso-ventral axis formation" in *PersAF* patients when compared to control donors, with *IFNG* (up-regulated), major



**Figure 3.** KEGG analysis of DE mRNAs in ParoAF\_Control and PersAF\_Control. We listed the top ten enriched KEGG terms. (a, b) the most enriched KEGG terms and KEGG-mRNA network in ParoAF\_Control. (c, d) the most enriched KEGG terms and KEGG-mRNA network in PersAF\_Control. The blue lines at (a, c) represent the number of DE mRNAs involved in the KEGG terms. The black dotted lines represent  $-\log_{10}(P\text{-value}) = 0.05$ . (b, d) showed the network of DE mRNAs and KEGG terms. The “blue inverted triangles” represent the KEGG terms, and “red circles” represent the DE mRNAs. The size of the nodes was determined by the lines it connects. DE - differentially expressed.

histocompatibility complex class IC (*HLA-C*, up-regulated), interferon regulatory factor 9 (*IRF9*, down-regulated) being the mostly involved DE mRNAs (Fig. 3c, 3d).

### GO and KEGG functions on DE lncRNAs through cis

We searched mRNAs nearest to DE lncRNAs both upstream and downstream and analyzed their function through GO and KEGG. We listed the top ten GO terms both upstream and downstream in Table 5. The most enriched GO terms of upstream were “phosphoric diester hydrolase activity,” “regulation of nucleotide metabolic process,” and “cytoskeleton.” They were “sequence-specific DNA binding RNA polymerase II transcription factor activity,” “response to dietary excess,” and “positive regulation of gene expression” of downstream.

Otherwise, 23 and 47 KEGG pathways were also enriched in upstream and downstream mRNAs with the threshold of  $p$ -value  $< 0.05$ , respectively. The upstream mRNAs were mainly enriched in “Ras signaling pathway,” “MAPK signaling pathway,” and “PPAR signaling pathway” (Fig. 4a). Genes Mitogen-activated protein kinase 10 (*MAPK10*) and phospholipase A2 group IVA

(*PLA2G4A*) were related to more pathways than others in the upstream mRNAs (Fig. 4b). While the downstream mRNAs were involved in “Arrhythmogenic right ventricular cardiomyopathy (ARVC),” “Signaling pathways regulating pluripotency of stem cells,” as well as “Glycine, serine, and threonine metabolism” (Fig. 4c). For downstream mRNAs, calcium channel voltage-dependent L-type alpha 1C subunit (*CACNA1C*), ATPase  $\text{Ca}^{++}$  transporting cardiac muscle slow twitch 2 (*ATP2A2*), catenin (cadherin-associated protein) beta 1 88kDa (*CTNWB1*), phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*PIK3CA*), and protein kinase cAMP-dependent catalytic gamma (*PRKACG*) were involved in not less than four top ten pathways (Fig. 4d).

### GO and KEGG pathway analysis on host genes of DE circRNAs

GO analysis indicated that the host genes of DE circRNAs in ParoAF\_Control were mainly enriched in response to muscle stretch, I band, contractile fiber part, myofibril, and contractile fiber ( $p$ -value  $< 0.05$ ). In addition, the host genes of DE circRNAs in PersAF\_Control were mainly enriched in Z disc, I band, actinin

**Table 5. Top ten GO terms for cis target genes of upstream and downstream DE lncRNAs**

GO Term category	Term name	Corrected P-value
<b>Upstream</b>		
GO:0008081_MF	Phosphoric diester hydrolase activity	5.73E-06
GO:0006140_BP	Regulation of nucleotide metabolic process	1.60E-04
GO:0005856_CC	Cytoskeleton	3.90E-04
GO:0042995_CC	Cell projection	5.00E-04
GO:1900542_BP	Regulation of purine nucleotide metabolic process	5.00E-04
GO:0043005_CC	Neuron projection	5.90E-04
GO:1901701_BP	Cellular response to oxygen-containing compound	7.10E-04
GO:0008092_MF	Cytoskeletal protein binding	1.05E-03
GO:0045202_CC	Synapse	1.41E-03
GO:0048731_BP	System development	1.63E-03
<b>Downstream</b>		
GO:0000981_MF	Sequence-specific DNA binding RNA polymerase II transcription factor activity	3.12E-05
GO:0002021_BP	Response to dietary excess	1.00E-04
GO:0010628_BP	Positive regulation of gene expression	2.20E-04
GO:1902531_BP	Regulation of intracellular signal transduction	6.50E-04
GO:0045893_BP	Positive regulation of transcription, DNA-dependent	2.33E-03
GO:1902532_BP	Negative regulation of intracellular signal transduction	2.42E-03
GO:0000977_MF	RNA polymerase II regulatory region sequence-specific DNA binding	4.80E-03
GO:0001012_MF	RNA polymerase II regulatory region DNA binding	5.22E-03
GO:0045944_BP	Positive regulation of transcription from RNA polymerase II promoter	6.37E-03
GO:0001071_MF	Nucleic acid binding transcription factor activity	8.72E-03
BP - biological process; CC - cellular component; DE - differentially expressed; GO - Gene Ontology; MF - molecular function		

binding, and cardiac muscle cell differentiation, a muscle system process ( $p$ -value <0.05). Through GO analysis, we noted that the host genes were mainly related to fiber and muscle.

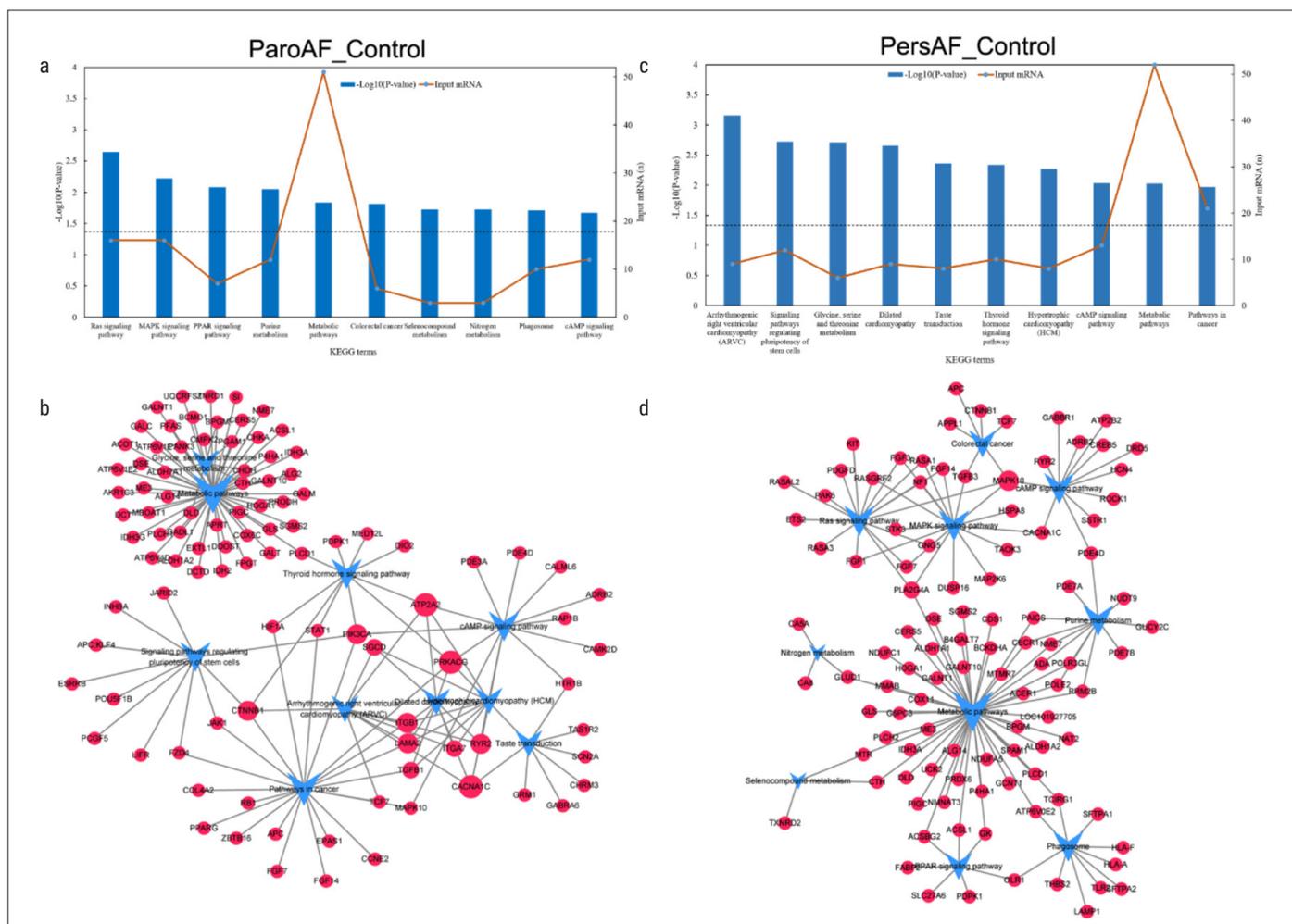
Pathways, such as dilated cardiomyopathy, hypertrophic cardiomyopathy, and arrhythmogenic right ventricular cardiomyopathy, were highly enriched in both ParoAF\_Control and PersAF\_Control, including host genes solute carrier family 8 (sodium/calcium exchanger) member 1 (*SLC8A1*), ryanodine receptor 2 (cardiac) (*RYR2*), integrin alpha 7 (*ITGA7*), among others. The calcium signaling pathway was only enriched in ParoAF\_Control. The top ten KEGG pathways and number of their input mRNAs of ParoAF\_Control and PersAF\_Control are shown in Figure 5a, 5c.

The network consisting of pathways and host genes were constructed (Fig. 5b, 5d). We found that TTN, epidermal growth factor (*EGF*), *ITGA7*, *RYR2*, and *SLC8A1* were hub genes in the network associated with ParoAF and PersAF. Moreover, *ITGA7*, *RYR2*, and *SLC8A1* were involved in more than five pathways in both ParoAF and PersAF. Among the host genes involved in more than four pathways, host genes adenylate cyclase 9 (*ADCY9*), phosphatidylinositol-4,5-bisphosphate 3-kinase, and catalytic subunits alpha (*PIK3CA*) were only involved in ParoAF, while *CACNA1C* was only related to PersAF.

### Interaction and co-expression network analysis

We studied the interactions between DE mRNAs of ParoAF\_Control, which are shown in Figure 6a, 6b. The up-regulated genes include *DCN*, *COL3A1*, *COL1A1*, secreted protein acidic cysteine-rich (*SPARC*), *COL6A1*, *COL1A2*, dynein cytoplasmic 1 heavy chain 1 (*DYNC1H1*), and cyclin B2 (*CCNB2*). The down-regulated genes, including breast cancer 1 early onset (*BRCA1*) and *COL2A1*, are the key genes that interacted with many other DE mRNAs in ParoAF\_Control network (Fig. 6a). In PersAF patients, up-regulated genes, including protein tyrosine phosphatase, non-receptor type 11 (*PTPN11*), major histocompatibility complex class II DQ alpha 1 (*HLA-DQA1*), cell division cycle 25A (*CDC25A*), beta-2-microglobulin (*B2M*), *COL6A1*, *COL3A1*, *SPARC*, *COL1A1*, and *COL1A2*, and down-regulated DE mRNAs, including cyclin-dependent kinase 2 (*CDK2*), kelch repeat and BTB (POZ) domain containing 13 (*KBTBD13*), histone deacetylase 1 (*HDAC1*), split hand/foot malformation (ectrodactyly) type 1 (*SHFM1*), F-box and leucine-rich repeat protein 19 (*FBXL19*), and ring finger protein 19A RBR E3 ubiquitin-protein ligase (*RNF19A*), are essential genes that interacted with other DE mRNAs (Fig. 6b).

Furthermore, we assessed the possible target genes of the DE lncRNAs (*Trans* prediction). A total of 709 and 1275 mRNAs were detected in ParoAF\_Control and PersAF\_Control, respec-



**Figure 4.** KEGG analysis of mRNAs nearby the DE lncRNAs. (a, b) the most enriched KEGG terms and KEGG-mRNA network in ParoAF\_Control. (c, d) the most enriched KEGG terms and KEGG-mRNA network in PersAF\_Control. We listed the top ten KEGG terms. The orange lines at (a, c) represent the number of mRNAs involved in the KEGG terms. The black dotted lines represent  $-\log_{10}(P\text{-value}) = 0.05$ . (b, d) showed the network of mRNAs and KEGG terms. The “blue inverted triangles” represent the KEGG terms, and “red circles” represent the mRNAs. The size of the nodes was determined by the lines it connects. DE - differentially expressed; KEGG - Kyoto Encyclopedia of Genes and Genomes; ParoAF - paroxysmal atrial fibrillation; PersAF - persistent atrial fibrillation

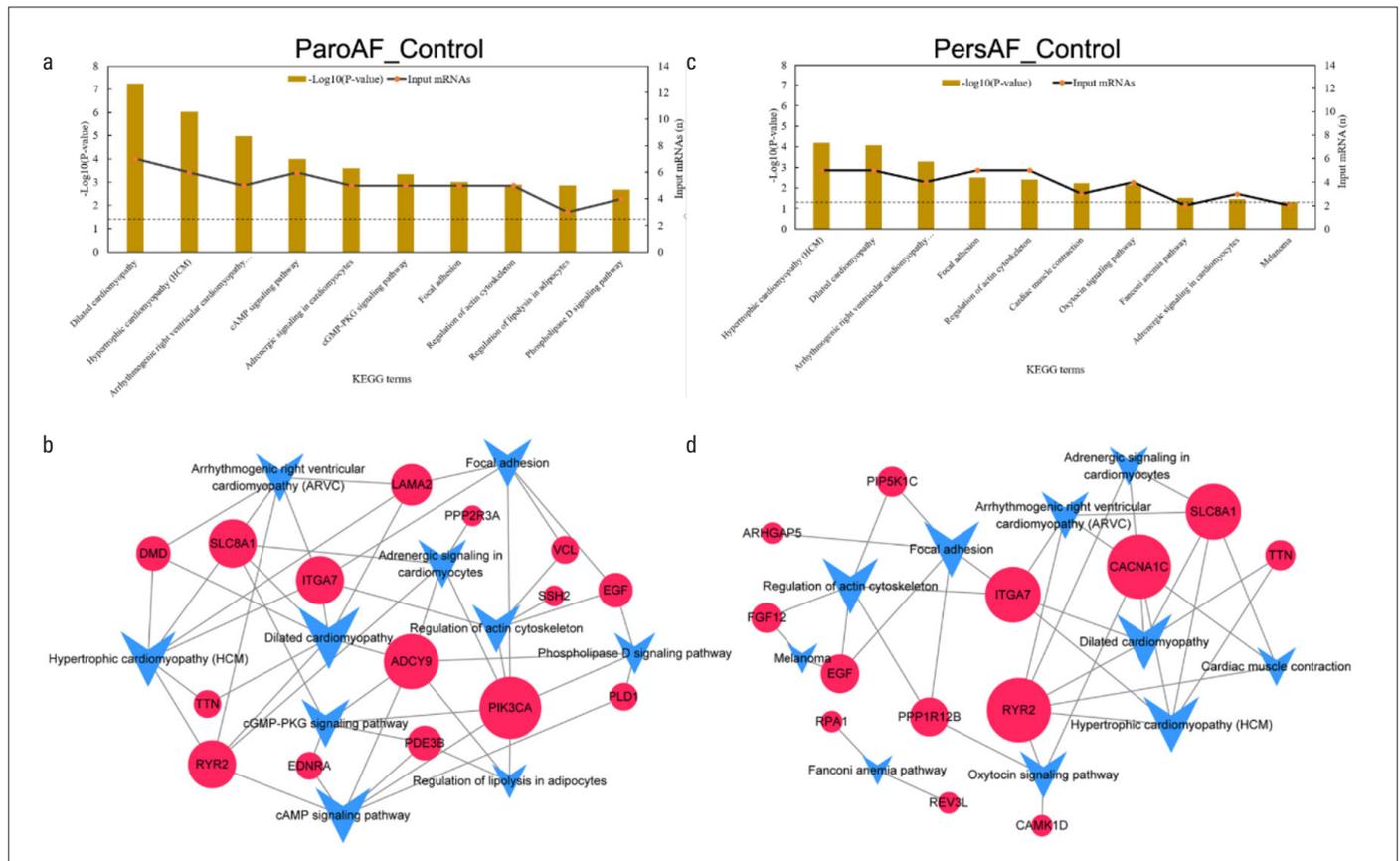
tively, with the threshold of  $r > 0.99$ . We selected DE mRNAs, which were also the target genes of DE lncRNAs, as shown in Figure 6c, 6d. There were 12 and 31 DE lncRNA and DE mRNA were co-expressed in ParoAF and PersAF patients, respectively, when compared to control.

The selected mRNAs were used for mRNA-lncRNA network analysis. The mRNA-lncRNA co-expression network of ParoAF\_Control revealed that anoctamin 3 (*ANO3*) had the co-down-expression trend with XLOC\_002352, XLOC\_110321, XLOC\_098657, and XLOC\_110333. Huntingtin interacting protein 1 related (*HIP1R*), fibrinogen C domain containing 1 (*FIBCD1*), tensin 4 (*TNS4*), and transcription factor Dp-2 (*TFDP2*) have a complex co-down-expression relationship with DE lncRNAs as XLOC\_027189, XLOC\_082023, XLOC\_095062, and XLOC\_077085. Translocated promoter region, nuclear basket protein, and contactin associated protein-like 2 (*CNTNAP2*) have co-up-expression with XLOC\_127021 (Fig. 6e). The mRNA-lncRNA co-expression network of PersAF\_Control also showed that *ANO3* co-expressed with XLOC\_002352, XLOC\_110321, XLOC\_098657, and

XLOC\_110333. XLOC\_130822 targeted both coiled-coil domain containing 63 (*CCDC63*) and solute carrier family 6 (neutral amino acid transporter) member 19 (*SLC6A19*, Fig. 6f).

### CircRNA-miRNA interaction network

We constructed the ceRNA network consisting of 14 common circRNAs DE between ParoAF\_Control and PersAF\_Control. A total of 3019 circRNA-miRNA interactions were predicted. Through the circRNA-miRNA interaction network, we found that hsa\_circ\_0124903, chr1:237890387|237951435, hsa\_circ\_0002142, and hsa\_circ\_0141424 acted as sponges of more than 300 miRNAs. These circRNAs have the host genes of mannosidase beta A (*MANBA*), lysosomal *RYS2*, DENN/MADD domain containing 2A (*DENND2A*), and zinc finger protein 709 (*ZNF709*), respectively. Conversely, there were 12 miRNAs (hsa\_miR-6735-5p, hsa\_miR-4739, hsa\_miR-4763-3p, hsa\_miR-3154, hsa\_miR-3916, hsa\_miR-4689, hsa\_miR-6511a-5p, hsa\_miR-665, hsa\_miR-671-5p, hsa\_miR-6797-5p, hsa\_miR-6847-5p, hsa\_miR-8089) that were sponged by not less than seven circRNAs (Fig. 7). We made



**Figure 5.** KEGG analysis of host genes of DE circRNAs. We listed the top ten KEGG terms. The brown chart represents the KEGG terms of host genes. (a, b) the most enriched KEGG terms and KEGG-mRNA network in ParoAF\_Control. (c, d) the most enriched KEGG terms and KEGG-mRNA network in PersAF\_Control. The solid black lines at (a, c) represent the number of host genes involved in the KEGG terms. The black dotted lines represent  $-\log_{10}(P\text{-value})=0.05$ . (b, d) showed the network of host genes and KEGG terms. The “blue inverted triangles” represent the KEGG terms, and “red circles” represent the host genes. The size of the nodes was determined by the lines it connects. DE - differentially expressed; KEGG - Kyoto Encyclopedia of Genes and Genomes; lncRNA - long non-coding RNA; ParoAF - paroxysmal atrial fibrillation; PersAF - persistent atrial fibrillation

the network with predicted miRNAs to have a less binding free energy ( $< -50$  kcal/mol) and found that hsa\_miR-671-5p, hsa\_miR-328-5p, and hsa\_miR-6782-5p sponged more circRNAs.

### Validation of differentially expressed lncRNAs and mRNAs with qRT-PCR

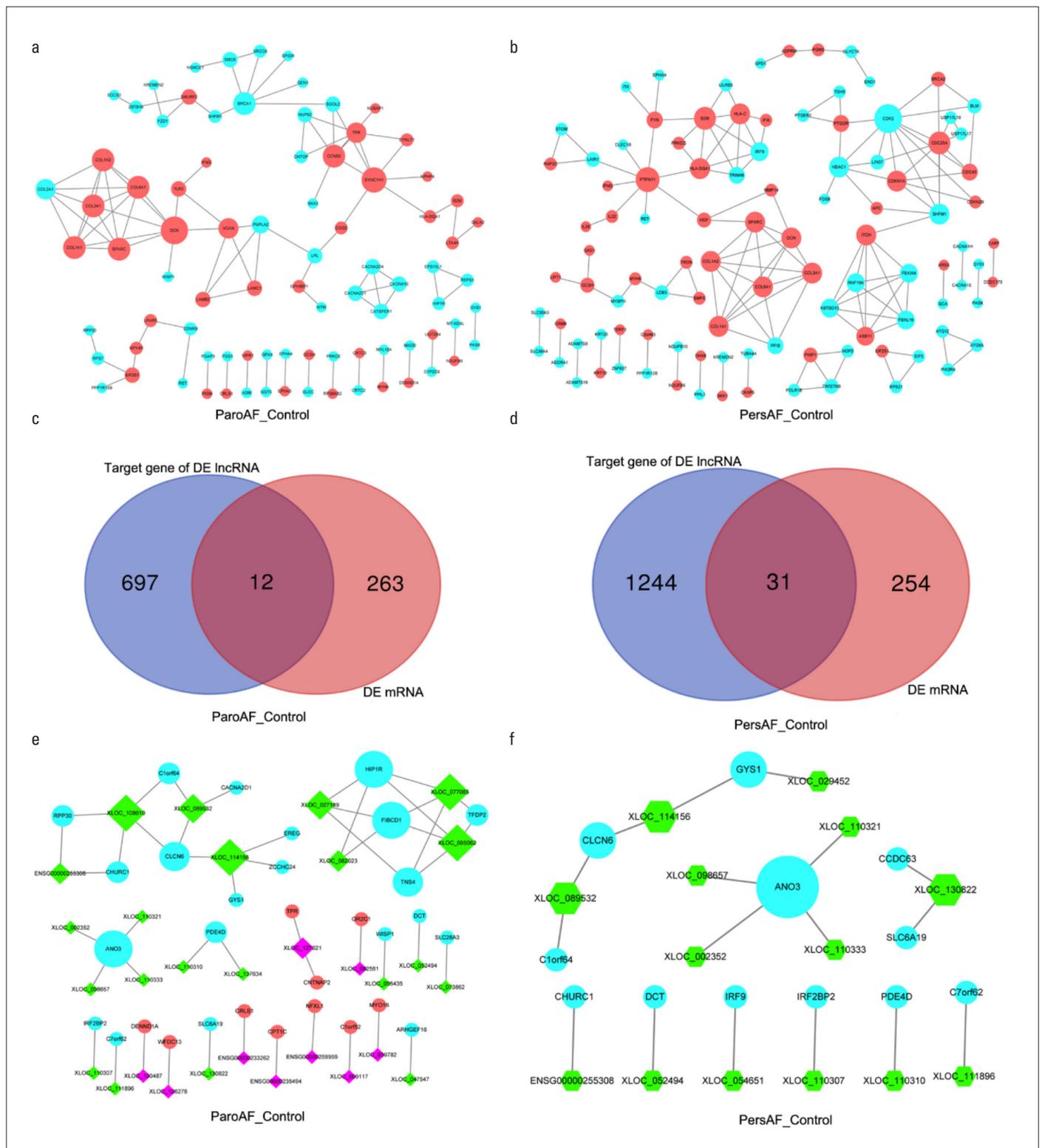
A total of four mRNAs and lncRNAs (*COL1A1*, *COL1A2*, *FKBP5*, and *RP11-442H21.2*) were validated by qRT-PCR in ParoAF and control samples (Fig. 8). On the other hand, a total of four mRNAs and lncRNAs (*COL1A1*, *COL1A2*, *CCL5*, and *CTA-134P22.2*) were validated by qRT-PCR in PersAF and control samples. The expression profiles of mRNAs and lncRNAs were consistent using either RNA-Seq or qRT-PCR, except *COL1A2* in PersAF\_Control, although deviation of exact fold change existed between the two techniques.

### Discussion

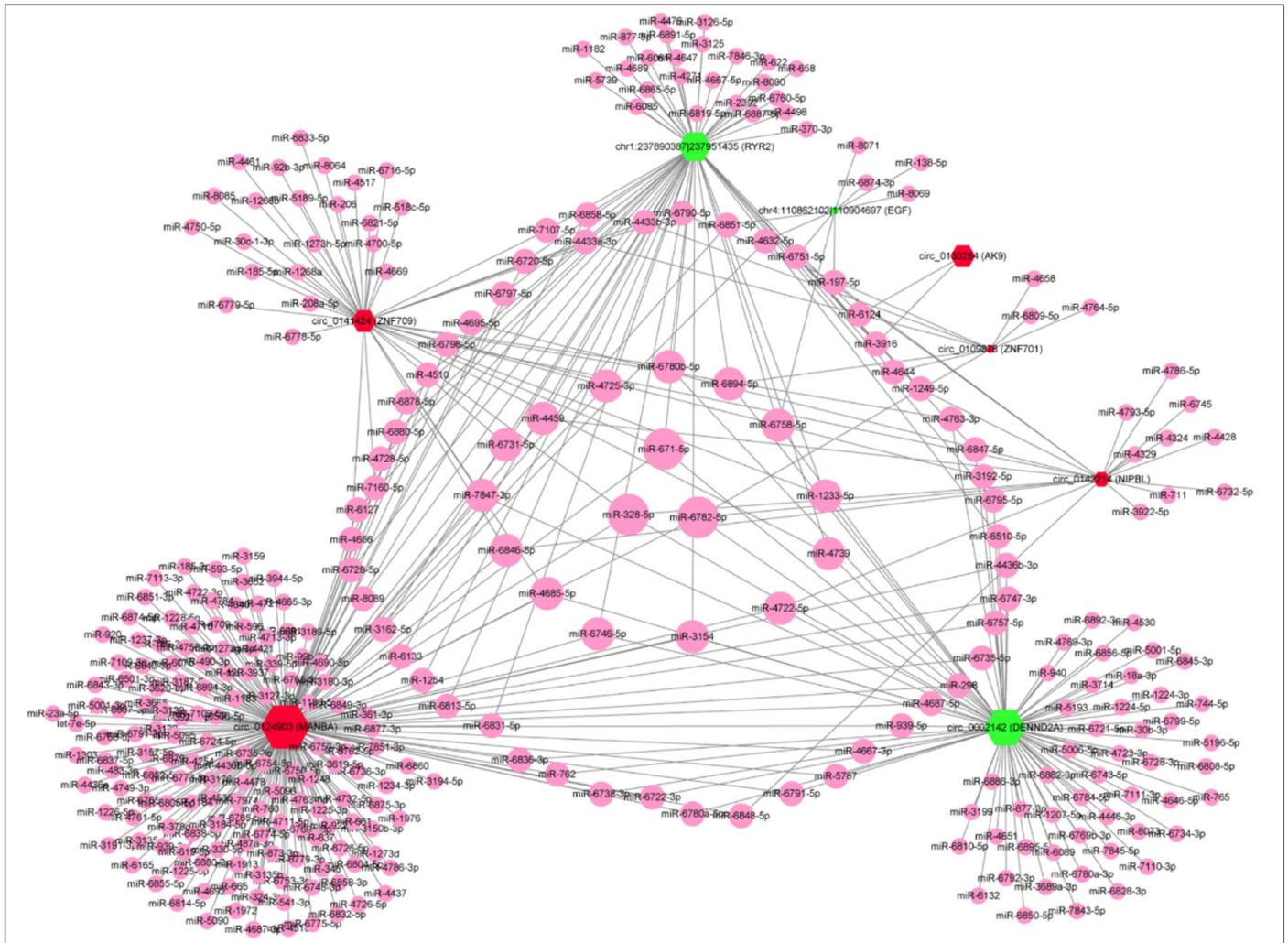
AF is the most prevalent heart disease worldwide, causing tremendous physical and psychological pain in individuals and a heavy burden to countries. In recent years, several researches have been channeled toward the mechanisms of AF RNAs, such

as mRNA, lncRNAs, and circRNAs, all play their various roles in regulating AF. In the present study, DE mRNAs, DE lncRNAs, and DE circRNAs in PersAF and ParoAF samples were clearly distinguished from control samples.

We focused on the relationship between circRNA isoforms and their host gene to examine whether circRNA isoforms originating from the same genes were dysregulated in AF. The circRNA isoforms of *CLNS1A* and *MYPN* were found to have opposite expression patterns in PersAF when compared with control, indicating that these expression-switched circRNA isoforms may have opposite functions in PersAF patients. We identified the two circRNA isoforms of *MYPN* (chr10:69881195|69882097 and chr10:69902697|69948883), which were significantly down-regulated and up-regulated, respectively, in PersAF patients when compared with control. In addition, logistics regression analysis showed that expression of chr10:69902697|69948883 was a protective factor for PersAF. Hu et al. (21) reported the significant up-regulation of circRNA02637 and downregulation of circRNA11156 in patients with PersAF when compared with control and *MYPN* and *TTN* was the host gene of circRNA02637 and circRNA11156, respectively. Hence, in our study, we found an obviously up-regulated circRNAs chr10:69902697|69948883



**Figure 6.** Interaction and co-expression network analysis of ParoAF and PersAF as compared to control. (a, b) represent the protein interaction network of ParoAF\_Control and PersAF\_Control, respectively. The red color represents the up-regulated DE mRNAs, while the blue color represents the down-regulated DE mRNAs. The Venn map in (c, d) showed the number of target genes of DE lncRNA and DE mRNAs of ParoAF\_Control and PersAF\_Control, respectively. (e, f) the co-expression network of DE lncRNA and DE mRNAs of ParoAF\_Control and PersAF\_Control, respectively. The “diamond nodes” represent DE lncRNAs, while the “circles nodes” represent mRNAs. The rose-red and green colors represent up-regulated and down-regulated DE lncRNAs, respectively. The red and blue color represent the up-regulated and down-regulated DE mRNAs, respectively. The size of the node in A, B, E, and F represent the DE lncRNAs or DE mRNAs they connect. DE - differentially expressed; lncRNA - long non-coding RNA; ParoAF - paroxysmal atrial fibrillation; PersAF - persistent atrial fibrillation



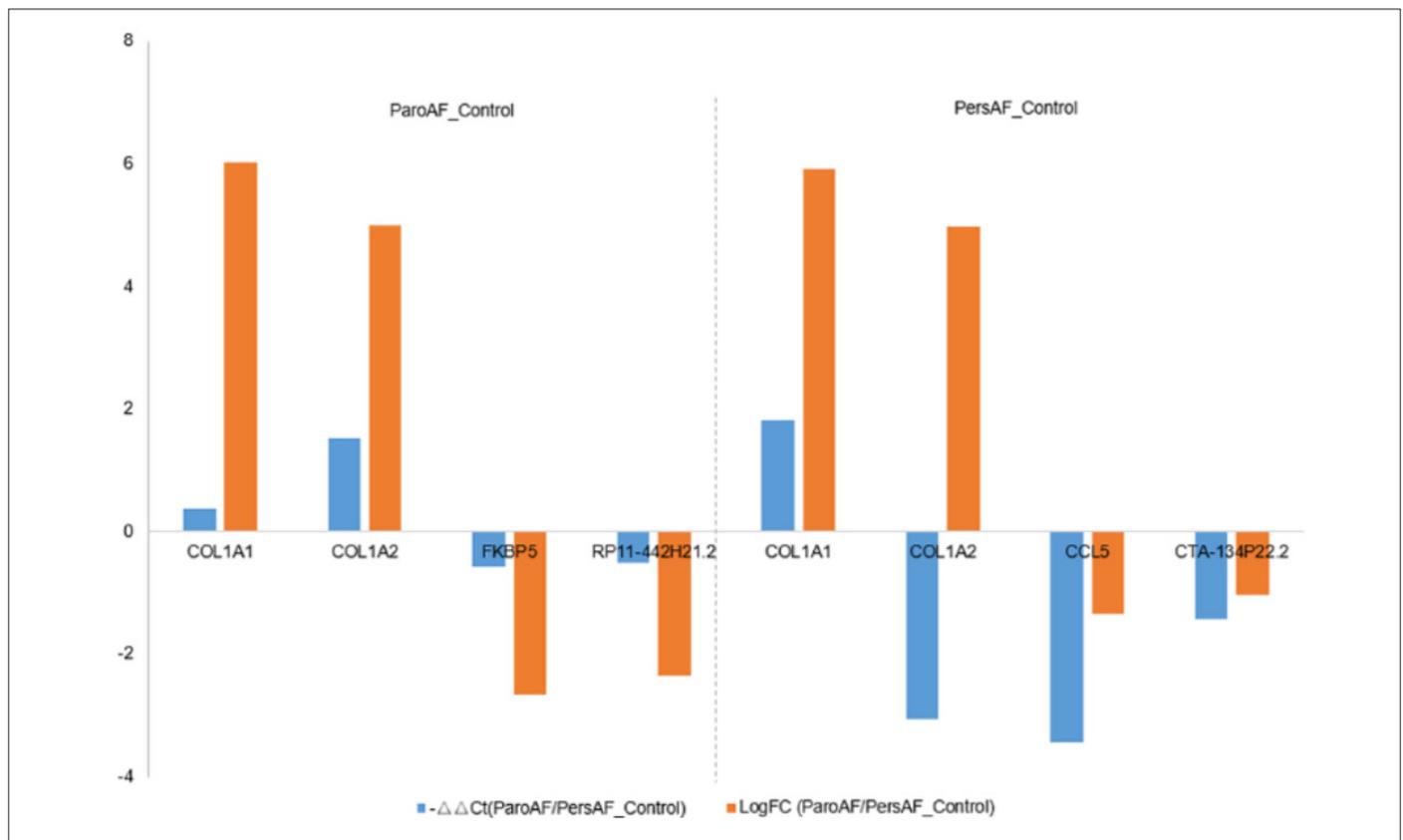
**Figure 7.** CeRNA network of selected circRNAs and miRNAs based on miRanda. Considering that the graphics cannot display the enormous amount of network information between 14 circRNAs and 3019 miRNAs, we selected circRNAs that had less binding free energy to make the network diagram using Cytoscape 3.6.1. “Hexagon node” represents circRNAs. Red and green colors represent up- and downregulation circRNAs, respectively. “Triangles node” represents as the miRNAs. The size of hexagon node represents the number of miRNAs that the circRNAs sponged. The size of circle node represents the number of sponging circRNAs. “hsa-” prefix omitted from all known circRNAs and miRNAs. CeRNA - competing endogenous RNA; circRNA - circular RNA

and down-regulated circRNAs chr10:69881195169882097 in patients with PersAF and ParoAF, respectively. Both of them were isoforms of *MYPN*, which play crucial roles in differentiation and development of muscle (22, 23). These results showed that their dysregulation might be of great value in the pathogenesis of AF.

In the present study, *COL1A2* and *COL1A1* were involved in platelet-derived growth factor binding in both ParoAF and PersAF samples. Moreover, *COL1A1*, *COL1A2*, and *COL2A1* also participated in most enriched KEGG pathways in ParoAF patients. In a study by Dawson et al. (6), it was revealed that *COL1A1* and *COL1A2* were target genes of miR-29, which might play a vital role in atrial fibrotic remodeling and considered as biomarkers or therapeutic targets. Gambini et al. (24) demonstrated that TGF-β1 could induce the up-regulation of *COL1A1* and *COL1A2* in human cardiac mesenchymal progenitor cells of AF. In one RNA-Seq analysis of ParoAF, the authors also found dysregulated *COL1A1* and *COL1A2* as in our study, which was not only validated by our

present study, but also demonstrated that these two genes participated in electrical and structural remodeling process of extracellular matrix/cardiac fibrosis (10). We then confirmed the expression of *COL1A1* and *COL1A2* through qRT-PCR. The expression level of *COL1A1* through qRT-PCR was consistent with that of RNA-Seq and was up-regulated in both ParoAF and PersAF samples. Moreover, in the present study, *COL1A1* and *COL1A2* act as essential genes that correlate with more DE mRNAs in the protein-protein network. These researches might suggest the crucial roles of these two genes on ParoAF and PersAF.

We focused on KEGG analysis of DE mRNAs, nearby mRNAs of DE lncRNAs, and host genes of DE circRNAs. Also, the pathways of TGF-β signaling, calcium signaling, MAPK signaling, and cAMP signaling were significantly enriched. The TGF-β signaling pathway was significantly enriched in ParoAF\_Control, with the involvement of *DCN*. The calcium signaling pathway was only enriched in PersAF samples, with the involvement of *PLN* enriched in it. For the DE lncRNA upstream and down-



**Figure 8.** qRT-PCR validation of DE mRNAs and DE lncRNAs. *CCL5* - chemokine (C-C motif) ligand 5; *COL1A1* - collagen type I alpha 1; *COL1A2* - collagen type I alpha 2; DE - differentially expressed; lncRNAs - long non-coding RNAs; ParoAF - paroxysmal atrial fibrillation; PersAF - persistent atrial fibrillation; qRT-PCR - quantitative real-time PCR

stream mRNAs, KEGG functional analysis revealed that *FGF14* and *CACNA1C* were enriched in the MAPK signaling pathway in the upstream mRNAs and *CACNA1C* and *HCN4* were enriched in the cAMP signaling pathway in the downstream mRNAs.

In a previous study of TGF- $\beta$  signal pathway of AF, it was revealed that the TGF- $\beta$  signal pathway might be involved in the course of atrial structural remodeling on AF and is related to the occurrence and maintenance of AF (25). In the study by Shen et al. (26), it was revealed that the activation of Gal-3 in AF could subsequently activate the TGF- $\beta$ 1/ $\alpha$ -SMA/Col I pathway in cardiac fibroblasts, which may strengthen atrial fibrosis. Suppressing the TGF- $\beta$  overexpression can prevent atrial remodeling (27-29). TGF- $\beta$  is also a promising therapeutic target for decreasing cardiac fibrosis (30). In the present study, the TGF- $\beta$  signaling pathway was differentially significant in ParoAF\_Control and the genes *SMURF2* and *DCN* were involved in it. *SMURF2* induces proteasomal degradation of Smad7, which promotes the TGF- $\beta$  signal pathway (29, 31), and was up-regulated in ParoAF\_Control in our study. *DCN* is a proteoglycan that binds to collagen fibrils in the ECM (extracellular matrix). *DCN* interacts with many growth factors, inhibiting the activity of TGF- $\beta$  (32), which is up-regulated in both ParoAF\_Control and PersAF\_Control in our study. Moreover, *DCN* was connected mostly in protein-protein networks with other DE mRNAs in the present study. Combined with these researches, the TGF- $\beta$  sig-

nal pathway and its involved gene *DCN* might play an essential role in the occurrence and maintenance of AF. Also, suppressing the overexpression of TGF- $\beta$  may prevent the occurrence of AF.

The calcium signaling pathway was DE in PersAF\_Control, including the gene *PLN* involved in it. Moreover, the host genes of DE circRNAs also participated in the calcium signaling pathway. The other research showed that calcium signaling pathway plays a vital role in the electrical remodeling of AF and might promote the recurrence of AF (12). Tan et al. (33) demonstrated that lncRNA *HOTAIR* is involved in the modulation of calcium homeostasis in human cardiomyocytes. They further confirmed that *HOTAIR* inhibited intracellular  $Ca^{2+}$  contents via regulation of L-type calcium channels. In a study regarding ParoAF and PersAF (34), it was revealed that PersAF mice had the phenomenon of enhanced diastolic  $Ca^{-\Delta\Delta Ct}$  release, marked conduction abnormalities, and atrial enlargement. *PLN* absence can increase  $Ca^{-\Delta\Delta Ct}$  transient amplitude and cause faster  $Ca^{-\Delta\Delta Ct}$  decay rate (35). Combining these findings with those of our study, we suggest that calcium signaling pathway may play an essential role in the processing of ParoAF to PersAF.

In the study of Zhang et al. (36), it was revealed that MAPKs/TGF- $\beta$ 1/TRAF6 signaling pathways participates in atrial fibrosis in patients with rheumatic heart disease, resulting to the occurrence of AF after cardiac surgery. MAPK is an important signal pathway that prevents atrial parasympathetic remodeling and

occurrence of AF via inhibiting MAPK pathway (37). The induction of AF and structural remodeling was related to MAPK expression and reduction of collagenase activity (12). *CACNA1C* was enriched in MAPK signal pathway, which is the direct target gene of miR-29a-3p. Combined with these results, we suggest that miR-29a-3p may be a potential therapeutic target in AF (38). Combining these findings with those of our study, MAPK signal pathway can be regarded as an important pathway that participates in AF occurrence by preventing atrial parasympathetic remodeling. The co-expression network revealed the correlation between lncRNAs and mRNAs, including transforming growth factor, calmodulin, and zinc finger protein. The co-expression results could broaden the understanding of the role of lncRNAs and mRNAs in AF.

Through KEGG pathway analysis, we found that DE circRNAs in PersAF and ParoAF were significantly associated with dilated cardiomyopathy, hypertrophic cardiomyopathy, and arrhythmogenic right ventricular cardiomyopathy, which suggest the crucial roles of these circRNAs in AF pathogenesis and development.

The network between pathways and host genes of DE circRNAs revealed that *ITGA7*, *RYR2*, *EGF*, *ADCY9*, *CACNA1C*, *PIN3CA*, and *SLC8A1* play an essential role in both ParoAF and PersAF. *EGF* and that *RYR2* host genes participate in most pathways in the two groups. In the study by Tang et al. (39), it was revealed that *RYR2* was down-regulated in AF, which is consistent with our study. In the study by Donald et al. (40), *RYR2* was identified as a potential therapeutic target in heart failure and AF. In another study on circRNAs of AF, *RYR2* acted as a host gene of some important circRNAs that were DE in AF when compared with control (41). Host genes, such as *ADCY9* and *PIK3CA*, only participate in ParoAF, while *CACNA1C* only participates in PersAF. These genes might play essential roles in the pathogenesis and development. Combined with the related studies, we boldly suggest that *RYR2*, *CACNA1C*, and other genes could serve as essential biomarkers for AF.

An interesting finding was the type of sample used for NGS. In the present study, GO enrichment analysis of host genes of common DE circRNAs between PersAF and ParoAF revealed that the main functions were correlated with muscle stretch and contractile fiber, as reported in Hu et al. (21). However, in a study, using lymphocytes from patients with AF (19), the host genes of DE circRNAs were mainly correlated with chromosome segregation, response to radiation, histone modification, among others. The circRNAs differed significantly in different kinds of samples. Therefore, it is especially important to choose proper materials for the experiment.

### Study limitations

There are some limitations in this study. First, the samples used in the present study were insufficient. Second, the validation of mRNA, lncRNA, circRNAs was not abundant.

### Conclusion

The present study analyzed the DE mRNA, DE lncRNAs, DE circRNAs, and their functions on ParoAF and PersAF patients.

Through this study, we found some pathways and genes, such as those related to TGF- $\beta$  signal pathway, calcium signaling pathway, *COL1A1*, *COL1A2*, among others, that might play essential roles in AF. Moreover, we also identified some circRNAs that could play an essential role in AF. CircRNAs, including chr1:237890387|237951435 and chr10:69902697|69948883, the host gene of *RYR2* and *MYPN*, respectively, might play essential roles in AF pathogenesis and development in ParoAF and PersAF via KEGG pathways associated with cardiomyopathy diseases. The expression of chr10:69902697|69948883 was identified to be a protective factor against PersAF. The host gene *CACNA1C* was only DE in PersAF and involved in most of the KEGG pathways associated with cardiomyopathy diseases, which might be crucial for the development from ParoAF to PersAF. Moreover, chr12:26661171|2721179, with the host gene of *CACNA1C*, was the sponge of miRNA-384/miRNA-328/miR-29a-3p, which might be a vital circRNA that participates in AF.

**Data availability:** The raw data were deposited on the NCBI Sequence Read Archive (SRA), with the SRA accession number of PRJNA531935.

**Conflict of interest:** None declared.

**Peer-review:** Externally peer-reviewed.

**Author contributions:** Concept – H.S., J.Z., Y.S.; Design – H.S., Y.S.; Supervision – H.S., J.Z., Y.S.; Fundings – None; Materials – H.S., J.Z.; Data collection &/or processing – J.Z.; Analysis &/or interpretation – H.S., J.Z., Y.S.; Literature search – H.S., J.Z., Y.S.; Writing – H.S.; Critical review – Y.S.

### References

1. Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *Int J Stroke* 2021; 16: 217-21. [\[Crossref\]](#)
2. Andrade J, Khairy P, Dobrev D, Nattel S. The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. *Circ Res* 2014; 114: 1453-68. [\[Crossref\]](#)
3. Yang S, Mei B, Feng K, Lin W, Chen G, Liang M, et al. Long-Term Results of Surgical Atrial Fibrillation Radiofrequency Ablation: Comparison of Two Methods. *Heart Lung Circ* 2018; 27: 621-8. [\[Crossref\]](#)
4. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC Jr, et al. 2019 AHA/ACC/HRS Focused Update of the 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2019; 74: 104-32. [\[Crossref\]](#)
5. Khurram IM, Habibi M, Gucuk Ipek E, Chrispin J, Yang E, Fukumoto K, et al. Left Atrial LGE and Arrhythmia Recurrence Following Pulmonary Vein Isolation for Paroxysmal and Persistent AF. *JACC Cardiovasc Imaging* 2016; 9: 142-8. [\[Crossref\]](#)
6. Dawson K, Wakili R, Ordög B, Clauss S, Chen Y, Iwasaki Y, et al. MicroRNA29: a mechanistic contributor and potential biomarker in atrial fibrillation. *Circulation* 2013; 127: 1466-75. [\[Crossref\]](#)

7. Wang Y, Cai H, Li H, Gao Z, Song K. Atrial overexpression of microRNA-27b attenuates angiotensin II-induced atrial fibrosis and fibrillation by targeting ALK5. *Hum Cell* 2018; 31: 251-60. [\[Crossref\]](#)
8. da Silva AMG, de Araújo JNG, de Oliveira KM, Novaes AEM, Lopes MB, de Sousa JCV, et al. Circulating miRNAs in acute new-onset atrial fibrillation and their target mRNA network. *J Cardiovasc Electrophysiol* 2018; 29: 1159-66. [\[Crossref\]](#)
9. Sinner MF, Tucker NR, Lunetta KL, Ozaki K, Smith JG, Trompet S, et al. Integrating genetic, transcriptional, and functional analyses to identify 5 novel genes for atrial fibrillation. *Circulation* 2014; 130: 1225-35. [\[Crossref\]](#)
10. Larupa Santos J, Rodríguez I, S Olesen M, Hjorth Bentzen B, Schmitt N. Investigating gene-microRNA networks in atrial fibrillation patients with mitral valve regurgitation. *PLoS One* 2020; 15: e0232719. [\[Crossref\]](#)
11. Zeng W, Jin J. The correlation of serum long non-coding RNA ANRIL with risk factors, functional outcome, and prognosis in atrial fibrillation patients with ischemic stroke. *J Clin Lab Anal* 2020; 34: e23352. [\[Crossref\]](#)
12. Ruan Z, Sun X, Sheng H, Zhu L. Long non-coding RNA expression profile in atrial fibrillation. *Int J Clin Exp Pathol* 2015; 8: 8402-10.
13. Song C, Zhang J, Liu Y, Pan H, Qi HP, Cao YG, et al. Construction and analysis of cardiac hypertrophy-associated lncRNA-mRNA network based on competitive endogenous RNA reveal functional lncRNAs in cardiac hypertrophy. *Oncotarget* 2016; 7: 10827-40. [\[Crossref\]](#)
14. Liang H, Pan Z, Zhao X, Liu L, Sun J, Su X, et al. lncRNA PFL contributes to cardiac fibrosis by acting as a competing endogenous RNA of let-7d. *Theranostics* 2018; 8: 1180-94. [\[Crossref\]](#)
15. Chen G, Guo H, Song Y, Chang H, Wang S, Zhang M, et al. Long non-coding RNA AK055347 is upregulated in patients with atrial fibrillation and regulates mitochondrial energy production in myocytes. *Mol Med Rep* 2016; 14: 5311-7. [\[Crossref\]](#)
16. Chen L, Yan KP, Liu XC, Wang W, Li C, Li M, et al. Valsartan regulates TGF- $\beta$ /Smads and TGF- $\beta$ /p38 pathways through lncRNA CHRFB to improve doxorubicin-induced heart failure. *Arch Pharm Res* 2018; 41: 101-9. [\[Crossref\]](#)
17. Barrett SP, Salzman J. Circular RNAs: analysis, expression and potential functions. *Development* 2016; 143: 1838-47. [\[Crossref\]](#)
18. Han B, Chao J, Yao H. Circular RNA and its mechanisms in disease: From the bench to the clinic. *Pharmacol Ther* 2018; 187: 31-44. [\[Crossref\]](#)
19. Hu X, Chen L, Wu S, Xu K, Jiang W, Qin M, et al. Integrative Analysis Reveals Key Circular RNA in Atrial Fibrillation. *Front Genet* 2019; 10: 108. [\[Crossref\]](#)
20. Gao Y, Wang J, Zhao F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol* 2015; 16: 4. [\[Crossref\]](#)
21. Hu M, Wei X, Li M, Tao L, Wei L, Zhang M, et al. Circular RNA expression profiles of persistent atrial fibrillation in patients with rheumatic heart disease. *Anatol J Cardiol* 2019; 21: 2-10. [\[Crossref\]](#)
22. Bagnall RD, Yeates L, Semsarian C. Analysis of the Z-disc genes PDLIM3 and MYPN in patients with hypertrophic cardiomyopathy. *Int J Cardiol* 2010; 145: 601-2. [\[Crossref\]](#)
23. Miyatake S, Mitsuhashi S, Hayashi YK, Purevjav E, Nishikawa A, Koshimizu E, et al. Biallelic Mutations in MYPN, Encoding Myopalladin, Are Associated with Childhood-Onset, Slowly Progressive Nemaaline Myopathy. *Am J Hum Genet* 2017; 100: 169-78. [\[Crossref\]](#)
24. Gambini E, Perrucci GL, Bassetti B, Spaltro G, Campostrini G, Lionetti MC, et al. Preferential myofibroblast differentiation of cardiac mesenchymal progenitor cells in the presence of atrial fibrillation. *Transl Res* 2018; 192: 54-67. [\[Crossref\]](#)
25. Mira YE, Muhuyati, Lu WH, He PY, Liu ZQ, Yang YC. TGF- $\beta$ 1 signal pathway in the regulation of inflammation in patients with atrial fibrillation. *Asian Pac J Trop Med* 2013; 6: 999-1003. [\[Crossref\]](#)
26. Shen H, Wang J, Min J, Xi W, Gao Y, Yin L, et al. Activation of TGF- $\beta$ 1/ $\alpha$ -SMA/Col I Profibrotic Pathway in Fibroblasts by Galectin-3 Contributes to Atrial Fibrosis in Experimental Models and Patients. *Cell Physiol Biochem* 2018; 47: 851-63. [\[Crossref\]](#)
27. Fu H, Li G, Liu C, Li J, Wang X, Cheng L, et al. Probucol prevents atrial remodeling by inhibiting oxidative stress and TNF- $\alpha$ /NF- $\kappa$ B/TGF- $\beta$  signal transduction pathway in alloxan-induced diabetic rabbits. *J Cardiovasc Electrophysiol* 2015; 26: 211-22. [\[Crossref\]](#)
28. Verheule S, Sato T, Everett T 4th, Engle SK, Otten D, Rubart-von der Lohe M, et al. Increased vulnerability to atrial fibrillation in transgenic mice with selective atrial fibrosis caused by overexpression of TGF- $\beta$ 1. *Circ Res* 2004; 94: 1458-65. [\[Crossref\]](#)
29. He X, Zhang K, Gao X, Li L, Tan H, Chen J, et al. Rapid atrial pacing induces myocardial fibrosis by down-regulating Smad7 via microRNA-21 in rabbit. *Heart Vessels* 2016; 31: 1696-708. [\[Crossref\]](#)
30. Xie J, Tu T, Zhou S, Liu Q. Transforming growth factor (TGF)- $\beta$ 1 signal pathway: A promising therapeutic target for attenuating cardiac fibrosis. *Int J Cardiol* 2017; 239: 9. [\[Crossref\]](#)
31. Cunnington RH, Nazari M, Dixon IM. c-Ski, Smurf2, and Arkadia as regulators of TGF- $\beta$  signaling: new targets for managing myofibroblast function and cardiac fibrosis. *Can J Physiol Pharmacol* 2009; 87: 764-72. [\[Crossref\]](#)
32. Järvinen TAH, Ruoslahti E. Generation of a multi-functional, target organ-specific, anti-fibrotic molecule by molecular engineering of the extracellular matrix protein, decorin. *Br J Pharmacol* 2019; 176: 16-25. [\[Crossref\]](#)
33. Tan WL, Xu M, Liu Z, Wu TY, Yang Y, Luo J, et al. HOTAIR inhibited intracellular Ca<sup>2+</sup> via regulation of Cav1.2 channel in human cardiomyocytes. *Cell Mol Biol (Noisy-le-grand)* 2015; 61: 79-83.
34. Li N, Chiang DY, Wang S, Wang Q, Sun L, Voigt N, et al. Ryanodine receptor-mediated calcium leak drives progressive development of an atrial fibrillation substrate in a transgenic mouse model. *Circulation* 2014; 129: 1276-85. [\[Crossref\]](#)
35. Baskin KK, Makarewicz CA, DeLeon SM, Ye W, Chen B, Beetz N, Schrewe H, Bassel-Duby R, Olson EN. MED12 regulates a transcriptional network of calcium-handling genes in the heart. *JCI Insight* 2017; 2: e91920. [\[Crossref\]](#)
36. Zhang D, Chen X, Wang Q, Wu S, Zheng Y, Liu X. Role of the MAPKs/TGF- $\beta$ 1/TRAF6 signaling pathway in postoperative atrial fibrillation. *PLoS One* 2017; 12: e0173759. [\[Crossref\]](#)
37. Liu L, Geng J, Zhao H, Yun F, Wang X, Yan S, et al. Valsartan Reduced Atrial Fibrillation Susceptibility by Inhibiting Atrial Parasympathetic Remodeling through MAPKs/Neurturin Pathway. *Cell Physiol Biochem* 2015; 36: 2039-50. [\[Crossref\]](#)
38. Zhao Y, Yuan Y, Qiu C. Underexpression of CACNA1C Caused by Overexpression of microRNA-29a Underlies the Pathogenesis of Atrial Fibrillation. *Med Sci Monit* 2016; 22: 2175-81. [\[Crossref\]](#)
39. Tang BP, Xu GJ, Shabiti Y, Yunus K, Abutirehemen M, Cheng ZH. [Alterations in gene expression of calcium handling proteins in patients with chronic atrial fibrillation]. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao* 2007; 29: 642-6.
40. Bers DM. Stabilizing ryanodine receptor gating quiets arrhythmic events in human heart failure and atrial fibrillation. *Heart Rhythm* 2017; 14: 420-1. [\[Crossref\]](#)
41. Costa MC, Cortez-Dias N, Gabriel A, de Sousa J, Fiúza M, Gallego J, et al. circRNA-miRNA cross-talk in the transition from paroxysmal to permanent atrial fibrillation. *Int J Cardiol* 2019; 290: 134-7. [\[Crossref\]](#)