Int J Clin Exp Pathol 2016;9(10):9867-9878
www.ijcep.com /ISSN:1936-2625/IJCEP0037886

Original Article

Integrated analysis of dysregulated lncRNA and mRNA expression profiles in pulmonary vein sleeves from patients with long-standing persistent atrial fibrillation

Zhi Li*, Yuji Zhang*, Yu Liu*, Jinsong Han, Yan Zhu, Zongtao Yin, Jian Zhang, Huishan Wang

*Department of Cardiovascular Surgery, Shenyang Northern Hospital, Shenyang, Liaoning, China. *Equal contributors and co-first authors.*

Received August 14, 2016; Accepted August 24, 2016; Epub October 1, 2016; Published October 15, 2016

Abstract: Atrial fibrillation (AF) is currently the most common arrhythmia in the world, which is an important contributor to cardiac morbidity and mortality, and classified to five types: first diagnosed AF, paroxysmal AF, persistent AF, long-standing persistent AF (LSP-AF), and permanent AF. LncRNAs, emerging as a novel therapeutic in cardiovascular medicine, attracted widespread attention, although an association LSP-AF has not been reported. An integrated analysis of dysregulated lncRNA and mRNA expression profiles were conducted in Pulmonary vein sleeves (PVS) between the patients (n=6) with mitral valve disease (MVD) and dialed left atria who develop LSP-AF and the patients (n=6) with MVD and dialed left atria who were in normal sinus rhythm (NSR), focusing specifically on the identification and characterization of lncRNAs and mRNA potentially involving in initiation of LSP-AF by a second generation lncRNA microarray. LncRNA microarray analysis shows 579 lncRNAs (393 higher, 186 lower), and 349 mRNAs (278 higher, 71 lower) was found between the LSP-AF and NSR. GO categories, pathway analyses, and interaction network showed a consistent result that differentially expressed genes contribute to the pathogenesis of LSP-AF. To conclude, this dysregulation of the lncRNA and mRNA profile provides a novel insight into the etiology of AF and furthermore, illustrates the intricate relationship between coding and ncRNA transcripts in LSP-AF. These data may offer a background/reference resource for future functional studies of lncRNAs and mRNA related to LSP-AF.

Keywords: Gender, atrial fibrillation, remodeling, mitral valve disease

Introduction

Atrial fibrillation (AF) is one of the most prevalent arrhythmias in clinical practice and is related to increasing cardiovascular morbidity and mortality [1-3]. Clinically, it is reasonable to distinguish five types of AF based on the presentation and duration of the arrhythmia: first diagnosed, paroxysmal, persistent, long-standing persistent and permanent AF [4]. Unfortunately, the precise mechanisms of AF were not well elucidated, leading to the demand of investigating the exact mechanisms of the disease and developing treatments. Identification of novel biomarker influencing the development of AF is critical to the understanding and future prevention of the disease.

Recent clinical data implicate lncRNAs have been implicated in cardiovascular function and disease, such as AF, cardiac development, heart failure, cardiac regeneration and coronary disease [5-7]. However, the association between lncRNA and LSP-AF has not been reported.

In our study, we performed genome-wide lncRNA and mRNA expression profiling using PVS from the patients with MVD and a history with and without LSP-AF. As AF associated with MVD is often accompanied by left atrial enlargement [8], which is well known to be an independent risk factor for fibrotic remodeling and AF [9], we only included LSP-AF or NON- AF patients with a similar degree of atrial dilation in order to exclude any experimental bias.

Materials and methods

Human pulmonary vein sleeves

Pulmonary vein sleeves were obtained from a biorepository of human atrial tissues from 12

Table 1. Clinical characteristics of study population

Classification	NSR (n=6)	LSP-AF (n=6)	*P*-value
Age, y	61.5±6.5	60.3±3.3	0.70
Gender Male/Female	3/3	3/3	1/1
Clinical diagnosis (n%)			
MS+MR	3 (50%)	5 (83.33%)	0.22
MR	3 (50%)	1 (16.67%)	0.22
AVD	2 (33.33%)	1 (16.67%)	0.51
TVD	3 (50%)	2 (33.33%)	0.56
CHD	2 (33.33%)	1 (16.67%)	0.51
Cerebral infarction	1 (16.67%)	1 (16.67%)	1
Pulmonary hypertension	1 (16.67%)	0 (0%)	0.30
Hypertension	1 (16.67%) (III)	0 (0%)	0.30
LAD (mm)	51.50±3.83	47.00±5.10	0.11
LVEF	0.60±0.03	0.61±0.05	0.69
LVFS	0.28±0.02	0.30±0.02	0.24
LVEDD (mm)	55.83±5.64	55.83±4.71	1

NSR, normal sinus rhythm; LSP-AF, long standing persistent atrial fibrillation; MS, mitral stenosis; MR, mitral regurgitation; AVD, aortic valve disease; TVD, tricuspid valve disease; LAD, left atrial diameter; LVEF, left ventricular ejection fraction; LVFS, left ventricular fractional shortening; LVEDD, left ventricular end diastolic dimension. *P*-value (NSR vs. LSP-AF).

MVD patients with dilated left atria who underwent cardiac surgery at the Thoracic Department of Shenyang Northern Hospital and divided into two groups by AF status (NSR and LSP-AF) (Supplementary Figure 1). The Chinese Ethics Committee of Registering Clinical Trials approved this study, and the investigation complied with the principles that govern the use of human tissues outlined in the Declaration of Helsinki. All patients gave informed consent before participating in the study. Approved NO. of ethic committee: ChiECRCT-20150025. There was no industry involvement in the design or performance of the study or in the analysis of the data. Samples were snap-frozen in liquid nitrogen and kept at -80°C until RNA extraction. Patients' characteristics are listed in the **Table 1**. The dilated left atria were defined according to ESC guidelines as follows [10]: male left atrial diameter (LAD) > 40 mm, female LAD > 38 mm. The diagnosis of LSP-AF was made according to AHA/ACC/ESC/HRS guidelines [11]. The LSP-AF patients had a history of AF for 1 yr. after which a rhythm control strategy is often adopted [4]. The NSR patients in normal sinus rhythm had no history of any type of AF or of using antiarrhythmic drugs.

RNA extraction

To extract RNA, frozen tissues were ground into powder with mortar and pestle and resuspended in TRIzol reagent (Invitrogen, Carlsbad, CA, USA). RNA was isolated from the homogenate following the manufacturer's protocol. A260/A280 in all samples was more than 1.8 and there were no degradation fragments in electrophoresis images (data not shown).

Microarray analysis

Purified RNA was amplified and transcribed into cRNA utilizing a random priming method and cDNA was labeled and hybridized to the Affymetrix Human Transcriptome Array 2.0 arrays. Totally, 70524 transcripts both coding and non-coding were covered. The arrays were scanned by GeneChip® Command Console® Software (AGCC) and all CEL files from the chips were normalized using Expression console (Affymetrix, Santa Clare, CA, USA). Due to the small sample size, we adopted the random variance model t-test to filter the differentially expressed transcripts between control and infected groups. After the significance analysis and false discovery rate (FDR) analysis, differentially expressed genes were selected according to their *p* value threshold [12]. Differentially expressed genes with statistical significance were identified through Heat maps and Volcano Plot filtering. The threshold used to screen up-or down-regulated genes was $P < 0.05$ and $FDR < 0.7$. The microarray data discussed in this study have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE76899 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE76899).

Validation by qRT-PCR

Total RNA extraction and cDNA transcription were conducted as above. For real-time qRT-PCR, we added 1 μl of cDNA to 12.5 μl of SYBR-Green Gene Expression Master Mix (Applied Biosystems, USA), 10.5 μl of DEPC-treated water and 0.5 μl of reverse and forward primers. cDNA was amplified for 40 cycles on the ABI 7500 Real-Time PCR system (Applied Biosystems, USA). Two lncRNAs (IFNG-AS1 and ENST00000401123) were randomly selected to prove the consistency of microarray and qPCR. The primers sequences used are listed

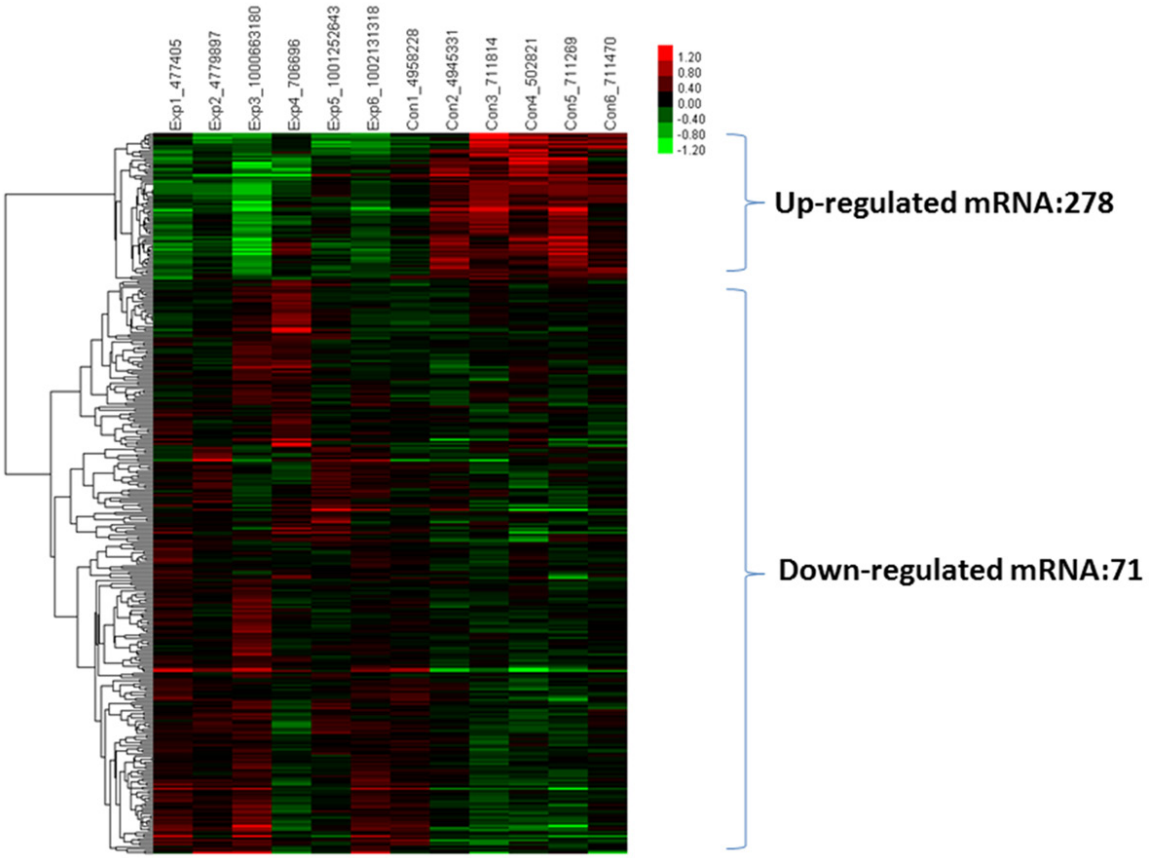


Figure 1. Heat map and hierarchical clustering of mRNA profile comparison between the LSP-AF (n=6) and NSR (n=6) samples. Total mRNA is 349: up-regulated mRNA, 278; Down-regulated mRNA, 71. The relative gene log2 expression changes are expressed by a color gradient intensity scale, as shown in the top left corner. Green color indicates down-regulation, and red color indicates up-regulation of mRNA expression. Each row represents a separate sample and each column represents a single mRNA. Left spectrum represents NSR samples and right spectrum represents LSP-AF samples. The differentially expressed mRNAs are clearly self- segregated into clusters.

in Supplementary Table 1. β-actin was used as a reference to obtain the relative expression of target lncRNAs which was determined with the comparative cycle threshold (CT) ($2^{-\Delta CT}$) method, in which $\Delta CT=CT_{\text{target lncRNA}}-CT_{GAPDH}$.

Gene ontology analysis and pathway analysis of LSP-AF related gene

Gene ontology (GO) analysis was applied to analyze the main function of the differential expression genes according to the Gene ontology project. Fisher's exact test and χ^2 test were used to classify the GO category, and the false discovery rate (FDR) was calculated to correct *P* value. The standard of difference screening was $P < 0.05$. Similarly, pathway analysis was used to find the significant pathway of the differential gene according to KEGG, Biocarta and Reatome. Fisher's exact test and χ^2 tests were used to select the significant pathway, and the threshold of significance was defined by *P* value and FDR [13-15].

Signal-net network diagram

Gene-gene interaction network was constructed based on the data of differentially expressed mRNAs. For instance, if there is confirmative evidence that two genes interact with each other, an interaction edge is assigned between the two genes. The considered evidence is the source of the interaction database from KEGG. Networks are stored and presented as graphs, where nodes are main genes (protein, compound, etc.) and edges represent relation types between the nodes, e.g. activation or phosphorylation.

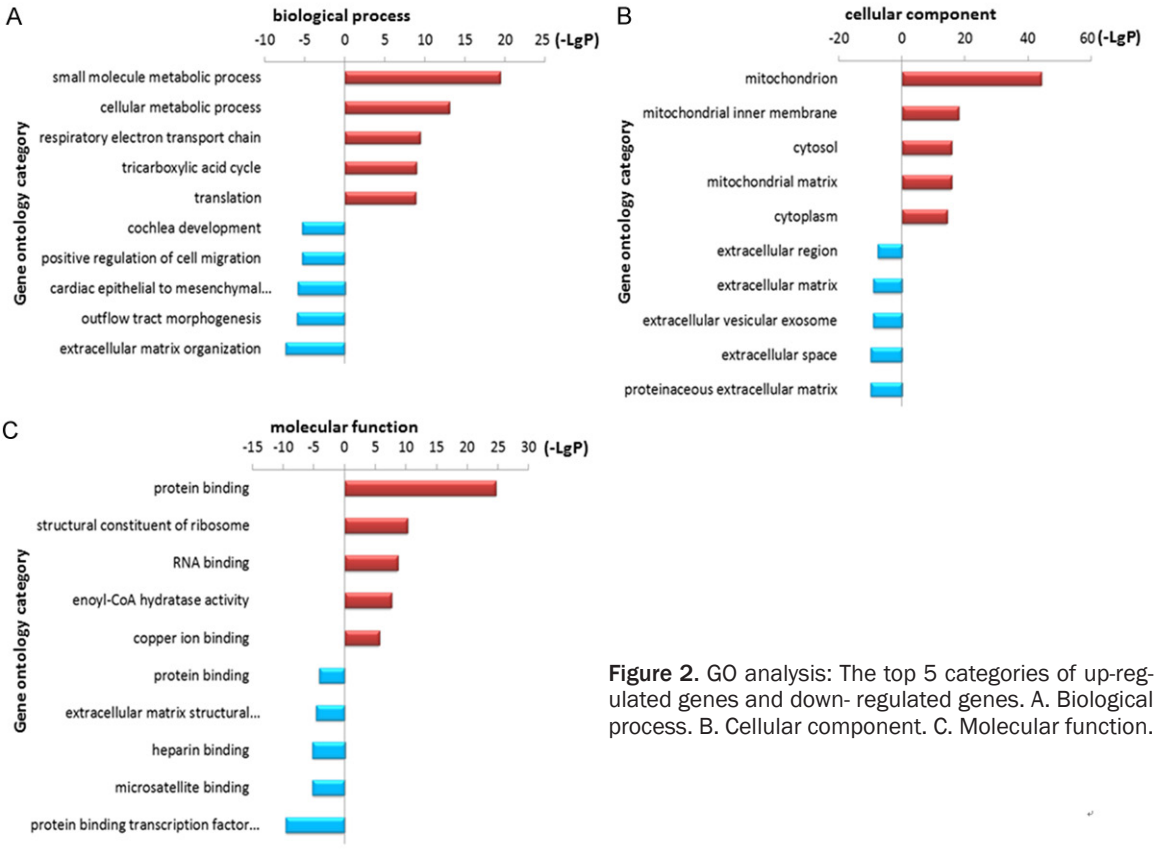


Figure 2. GO analysis: The top 5 categories of up-regulated genes and down- regulated genes. A. Biological process. B. Cellular component. C. Molecular function.

LncRNA-mRNA-network analysis

LncRNA-mRNA interaction network was networks were built according to the normalized signal intensity of specific expression genes. We constructed the network adjacency between 2 genes, i and j, defined as a power of the Pearson correlation between the corresponding gene expressions profiles xi and xj. By computing them, we obtained the gene adjacency matrix M (i, j). The adjacency matrix M (i, j) was visualized as a graph, and the topological properties of this graph were examined [16]. To make a visual representation, only the strongest correlations (0.96 or greater) were drawn in these renderings.

Statistical analysis

All statistical analyses were performed using the Student's t-test with PASW Statistics 18. A *P* value less than 0.05 was considered statistically significant, and all the statistical tests were two-sided. False discovery rate (FDR) was calculated to correct the *P* value.

Results

Patients characteristics

There were no gender-related differences in LAD, age, clinical diagnosis of etiology, and accompanied disease between AF and NAF (**Table 1**).

Dysregulated mRNA expression profiles of PVS from patients with LSP-AF

Hierarchical clustering and Volcano plots showed systematic variations in the expression of mRNAs. A total of 349 mRNAs were identified to be differentially expressed. Among them, 278 and 71 mRNAs were upregulated and downregulated in AF tissues compared with controls, respectively (**Figure 1**). These results indicate that mRNAs are involved in the pathogenesis of AF.

GO analysis

The GO project is a collaborative effort to construct and use ontologies to facilitate the bio-

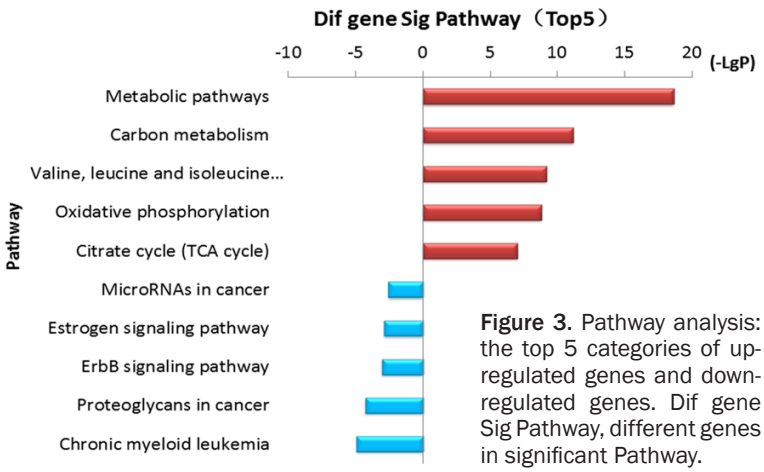


Figure 3. Pathway analysis: the top 5 categories of up-regulated genes and down-regulated genes. Dif gene Sig Pathway, different genes in significant Pathway.

logically meaningful annotation of genes and their products in a wide variety of organisms and is the key functional classification system of the NCBI.

We performed GO analysis for the functions of differential expressed (both up-regulated and down-regulated) mRNA in LSP-AF as shown in **Figure 2**. Differentially expressed transcripts in biological process were involved as follows (**Figure 2A**): up-regulated mRNA categories (Top5), (1) small molecule metabolic process; (2) cellular metabolic process; (3) respiratory electron transport chain; (4) tricarboxylic acid cycle; (5) translation; down-regulated mRNA categories (Top5), (1) extracellular matrix organization; (2) outflow tract morphogenesis; (3) cardiac epithelial to mesenchymal transition; (4) positive regulation of cell migration; (5) cochlea development. Differentially expressed transcripts in cellular component were involved as follows (**Figure 2B**): up-regulated mRNA categories (Top5), (1) mitochondrion; (2) mitochondrial inner membrane; (3) cytosol; (4) mitochondrial matrix; (5) cytoplasm; down-regulated mRNA categories (Top5), (1) proteinaceous extracellular matrix; (2) extracellular space; (3) extracellular vesicular exosome; (4) extracellular matrix; (5) extracellular region. Differentially expressed transcripts in molecular function were involved as follows (**Figure 2C**): up-regulated mRNA categories (Top5), (1) protein binding; (2) structural constituent of ribosome; (3) RNA binding; (4) enoyl-CoA hydratase activity; (5) copper ion binding; down-regulated mRNA categories (Top5), (1) protein binding transcription factor activity; (2) microsatellite binding; (3) heparin binding; (4) extracellular matrix structural constituent; (5) protein binding.

Pathway analysis

Ingenuity Pathway Analysis (IPA) was used to identify pathways and gene networks represented among the sets of protein-coding mRNAs identified in the AF gene expression signature. We performed KEGG pathway enrichment analysis as shown in **Figure 3**. The most significant pathways were involved as follows (**Figure 3**): up-regulated mRNA categories (Top5), (1) Metabolic pathways; (2) Carbon metabolism; (3) Valine, leucine and isoleucine degradation; (4) Oxidative phosphorylation; (5) Citrate cycle (TCA cycle); down-regulated mRNA categories (Top5), (1) Chronic myeloid leukemia; (2) Proteoglycans in cancer; (3) ErbB signaling pathway; (4) Estrogen signaling pathway; (5) MicroRNAs in cancer.

Signal-net analysis of AF-related genes

To investigate the key genes involved in LSP-AF, signal-net analysis of differential genes associated with LSP-AF exposure showed that 59 genes have a direct or indirect interaction relationship (**Figure 4**). The largest direct connected subnetwork includes four up-regulated genes (SDHD, SDHC, SUCLG2, and SUCLA2) on the Metabolic related pathway. These findings agree with current knowledge of the involvement of the Metabolic related pathway in AF [17].

Dysregulated lncRNA expression profiles of PVS from patients with LSP-AF

Hierarchical clustering and Volcano plots showed systematic variations in the expression of lncRNAs. A total of 579 lncRNAs were identified to be differentially expressed. Among them, 393 and 186 lncRNAs were upregulated and downregulated in AF tissues compared with controls, respectively (**Figure 5**). These results indicate that lncRNAs are involved in the pathogenesis of AF.

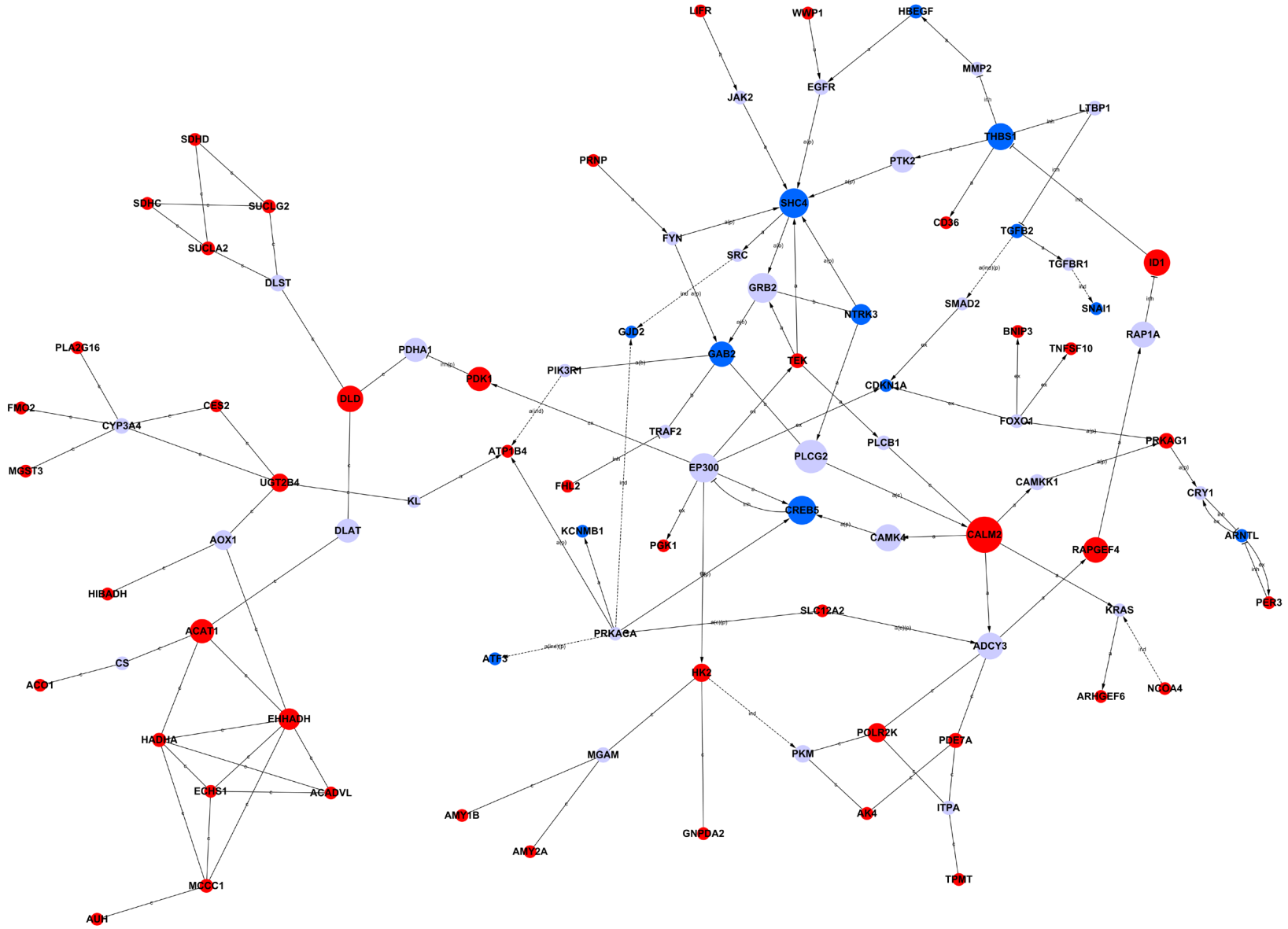

LIFR
WNT1
HBEGF
MMP2
JAK2
EGFR
LTBP1
THBS1
SDHD
PRNP
PTK2
SDHC
SUCLG2
SHC4
CD36
SUCLA2
FYN
TGFB2
SRC
TGFBR1
ID1
DLST
GRB2
SMAD2
SNAI1
NTRK3
GAB2
BNIP3
RAP1A
PLA2G16
PDHA1
PIK3R1
TEK
TNFSF10
DLD
PDK1
CDKN1A
FMO2
CES2
CYP3A4
FOXO1
PRKAG1
TRAF2
PLCB1
ATP1B4
EP300
PLCG2
MGST3
UGT2B4
CAMKK1
CRY1
FHL2
KL
CREB5
KCNMB1
CAMK4
CALM2
ARNTL
AOX1
DLAT
PGK1
RAPGEF4
HIBADH
PER3
SLC12A2
KRAS
ACAT1
PRKACA
ADCY3
CS
ATF3
ACO1
HK2
ARHGEF6
NCOA4
EHHADH
PDE7A
HADHA
MGAM
PKM
PDE2K
ECHS1
ACADVL
AK4
ITPA
AMY1B
GNPDA2
AMY2A
TPMT
MCCC1
AUH

Figure 4. Signal-net analysis of differential AF-related genes in LSP-AF. The molecular networks were constructed, where nodes are main genes and edges represent relation types between the nodes. a(b), activation (binding/association); a, activation; c, compound; ex, expression; inh, inhibition; a(p), activation (phosphorylation); ind, indirect effect; inh(p), inhibition (phosphorylation); a(ind), activation (indirect effect); a(c), activation (compound); p, phosphorylation; a(ind)(p), activation (indirect effect) (phosphorylation). a(c)(p), activation (compound) (phosphorylation). Red denotes the up-regulated genes. Blue denotes the down-regulated genes. Light purple denotes Joining gene according to the database.

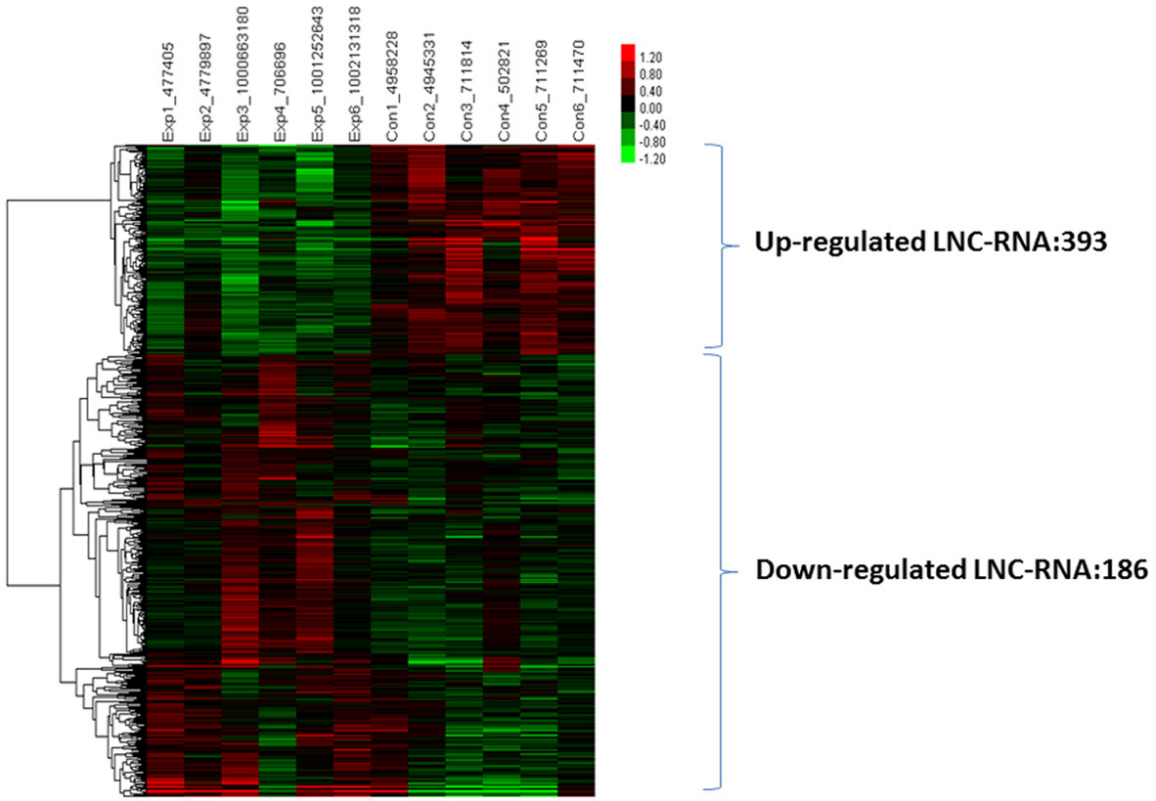


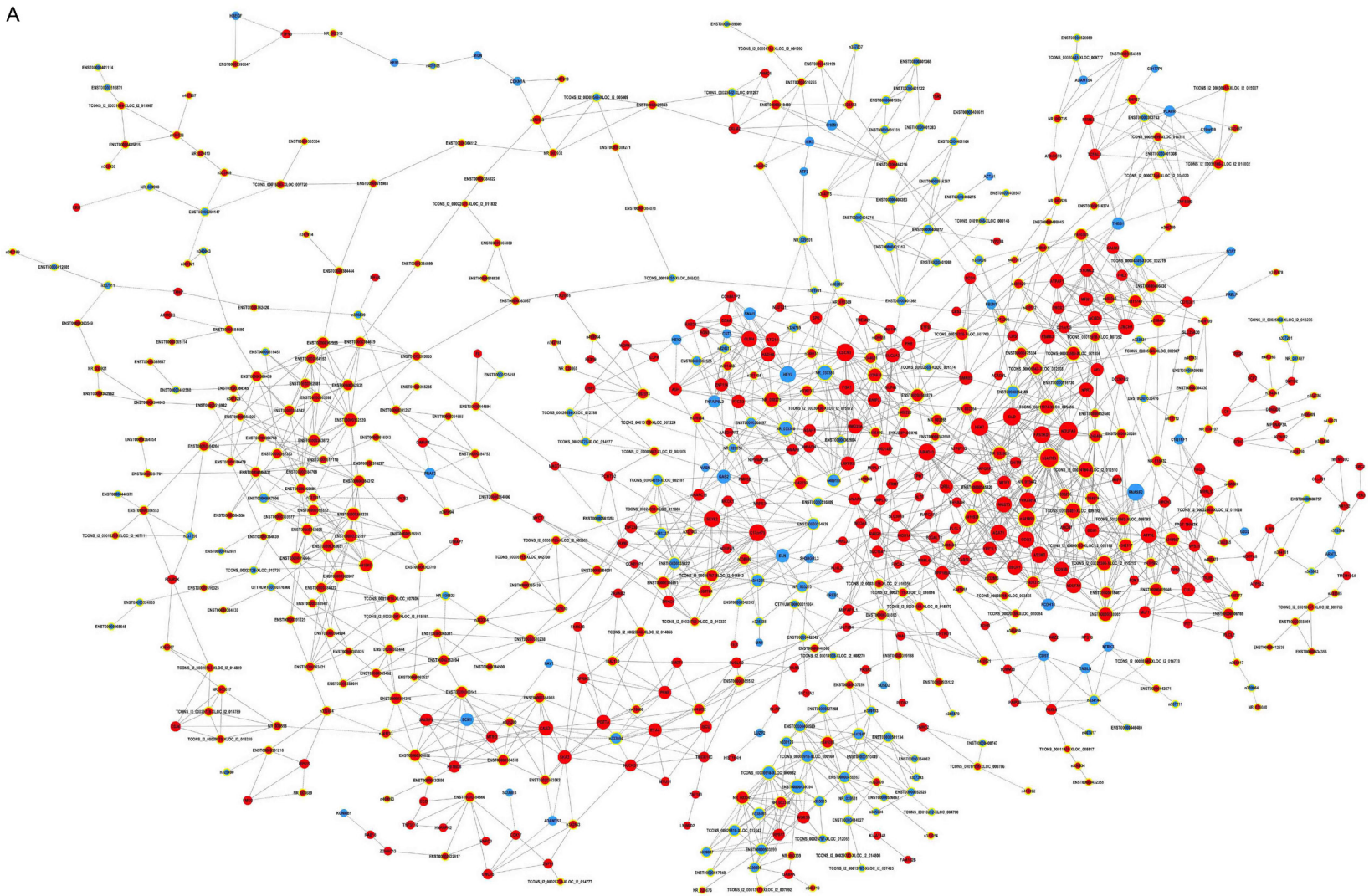
Figure 5. Heat map and hierarchical clustering of LNC-RNA profile comparison between the LSP-AF (n=6) and NSR (n=6) samples. Total mRNA is 579: up-regulated LNC-RNA, 393; Down-regulated LNC-RNA, 186. The relative gene log2 expression changes are expressed by a color gradient intensity scale, as shown in the top left corner. Green color indicates down-regulation, and red color indicates up-regulation of LNC-RNA expression. Each row represents a separate sample and each column represents a single LNC-RNA. Left spectrum represents NSR samples and right spectrum represents LSP-AF samples. The differentially expressed LNC-RNAs are clearly self-segregated into clusters.

Validation of differentially expressed lncRNAs

Two lncRNAs (IFNG-AS1 and ENST0000040-1123) were randomly selected to prove the consistency of microarray and qPCR. As expected, the expression of lncRNA IFNG-AS1 was upregulated and ENST00000401123 was down-regulated in the AF samples versus control samples (Supplementary Figure 2), consistent with the microarray results.

Construction of lncRNA-mRNA co-expression network

To explore which lncRNAs and mRNAs play critical roles in AF progression, we constructed a co-expression network of the differentially expressed correlated lncRNAs and mRNAs in LSP-AF (**Figure 6B**) and NSR group (**Figure 6A**), respectively, according to the similarity of gene expression patterns in each group. The nodes in the figures stand for the genes, and the lines between nodes stand for the relationship between the genes. The node size is on behalf of the gene's degree. The greater the degree is, the more interactions between the target gene and other genes. The clustering coefficient is on behalf of the density of an area of the network. Genes with greater clustering coefficient indicates the region of this gene is more important for the regulation of whole network. The


A

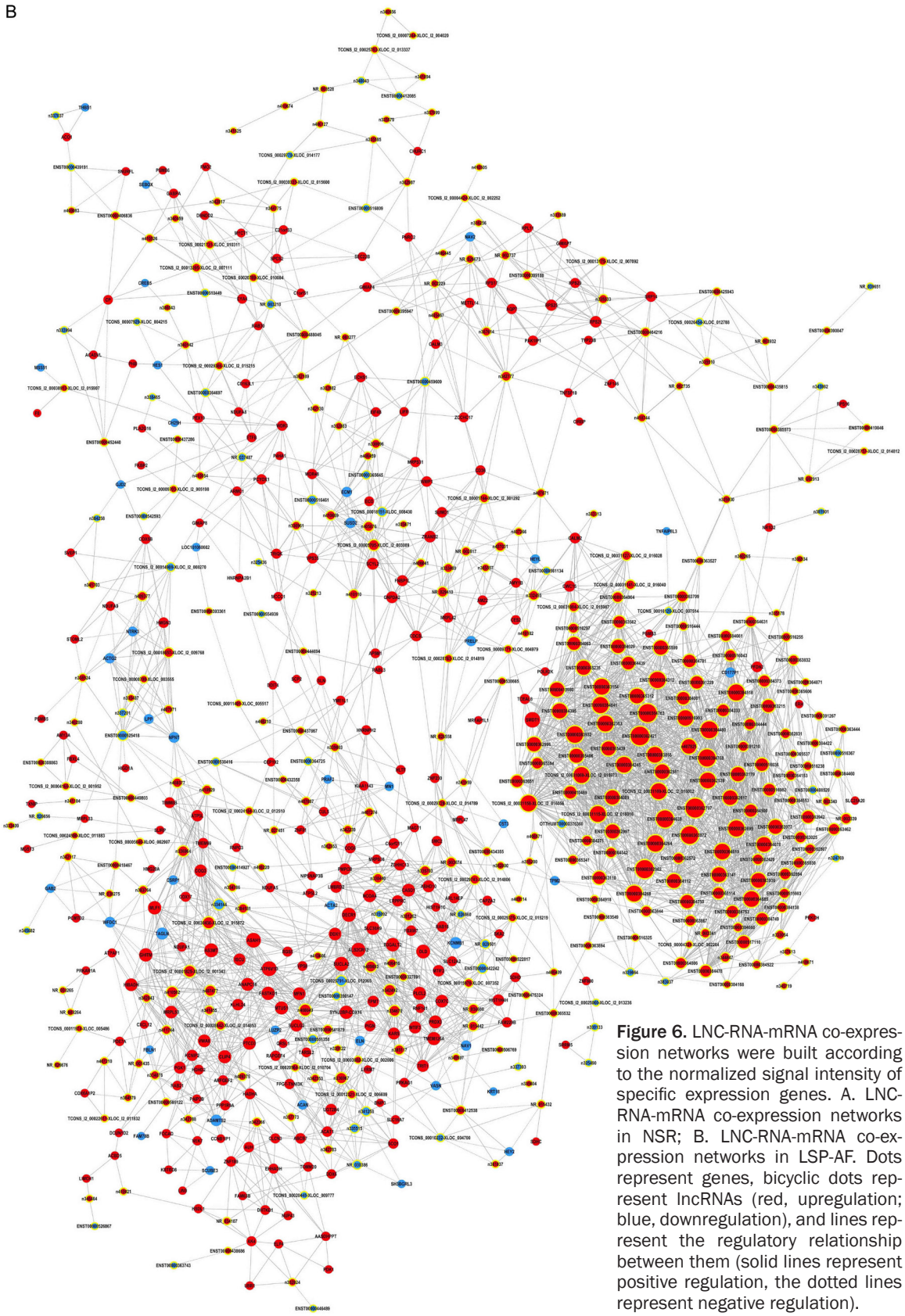


Figure 6. LNC-RNA-mRNA co-expression networks were built according to the normalized signal intensity of specific expression genes. A. LNC-RNA-mRNA co-expression networks in NSR; B. LNC-RNA-mRNA co-expression networks in LSP-AF. Dots represent genes, bicyclic dots represent lncRNAs (red, upregulation; blue, downregulation), and lines represent the regulatory relationship between them (solid lines represent positive regulation, the dotted lines represent negative regulation).

network structure of the LSP-AF (**Figure 6B**) tissues and corresponding NSR (**Figure 6A**) tissues samples was markedly different.

Genes with a large degree value and clustering coefficient in NSR and LSP-AF samples (clustering coefficient > 0.6, Degree 10) were listed in Supplementary Tables 2 and 3, respectively.

The critical change of co-expression status of a gene between the co-expression network of LSP-AF and NSR suggests that the expression of the gene very possibly undertakes a pivotal change in the process of LSP-AF pathogenesis Gene like this may play an important role in the LSP-AF's genesis and development, so away to find these genes is to select the genes that have the most pivotal changes in the co-expression network of the LSP-AF compared with that of the NSR. Through analyzing the changes of co-expression status of genes between the LSP-AF and NSR samples, we have got a list of genes (|diffK| > 0.6) that may participate in the pathogenesis of LSP-AF (Supplementary Table 4), such as TCONS_00024590-XLOC_011883 may be participate in the regulation of LSP-AF in biological processes (small molecule metabolic process; oxidation-reduction process; signal transduction; lipid metabolic process; xenobiotic metabolic process), cellular component (extracellular vesicular exosome; membrane; integral to membrane; intracellular membrane-bounded organelle; endoplasmic reticulum membrane) and several pathways (Drug metabolism-cytochrome P450), by interacting with MGST3, according to GO and Pathway categories and deserves for further study.

Discussion

AF is one of the most prevalent arrhythmias in clinical practice and is related to increasing cardiovascular morbidity and mortality [1-3]. Recent clinical data implicate lncRNAs have been implicated in cardiovascular function and disease, such as cardiac development, heart failure, cardiac regeneration coronary disease and AF. Braveheart (Bvht) was first as identified heart-associated lncRNA with potential regulatory function in cardiovascular development [18]. The lncRNA ANRIL has been shown to participate in the molecular mechanisms underpinning heart failure [19]. The MIAT was identified as an lncRNA associated with increased risk of myocardial infarction [20]. In addition, Prof. Zhu and Jiang found a lot of lncRNA are closely related with AF in Blood [7] and left atrial appendage [6]. However, the association between lncRNA and LSP-AF has not been reported.

Pulmonary vein (PV) sleeves play a critical role in the initiation and maintenance of AF [21]. In our study, we performed genome-wide lncRNA and mRNA expression profiling in PVS from the patients with and without LSP-AF using microarray and preliminary explore the role of lncRNA and mRNA in the pathogenesis of AF. We identified 579 dysregulated lncRNAs (393 higher, 186 lower), and 349 dysregulated mRNAs (278 higher, 71 lower) in LSP-AF. Our results illustrated significant changes of lncRNAs and mRNA expression in atrial tissues of AF patients and dysregulated lncRNAs and mRNA may play regulatory roles in the mechanism of AF, which may further provide potential therapeutic targets for prophylaxis and treatment of AF.

However, perhaps the most important current challenge is that the knowledge embedded in pathways regarding how various genes interact with each other is not currently exploited. Microarray technology makes it possible to measure the expression levels of almost all the human genes and therefore facilitates the identification of genes and pathways that are related to disease initiation and development. Gene ontology (GO) and Pathway Analysis were applied to classify differentially expressed AF related genes. We found that the most obvious significant up-regulated gene category is metabolism-related (small molecule metabolic process) and Oxidative stress-related (mitochondrion); the most obvious significant down-regulated gene category is fibrosis remodeling-related (extracellular matrix organization, proteinaceous extracellular matrix). The mechanism of atrial fibrillation is mainly electrical remodeling and structural remodeling [22]. The metabolism-related categories mainly involved in AF by promoting oxidative stress and inflammation [23], while oxidative stress-related categories mainly involved in AF by promoting fibrosis remodeling [24]. Therefore, our screened significant gene categories mainly involved in structural remodeling, which plays an important role in LSP-AF.

Signal-net analysis showed that 59 AF-related genes have a direct or indirect interaction relationship (**Figure 4**). The largest direct connect-

ed subnetwork includes four up-regulated genes (SDHD, SDHC, SUCLG2, and SUCLA2) on the Metabolic pathway. These findings agree with current knowledge of the involvement of the Metabolic pathway in AF [17, 23].

It has been found that 80% of noncoding RNAs are lncRNAs, which are involved in gene expression and function regulations. Not like mRNA, the functions of most lncRNAs are still not determined. In order to investigate the function of lncRNAs in AF, we combined differentially expressed lncRNAs and mRNA to construct a co-expression network. We found that many lncRNAs were significantly correlated with multiple protein-coding genes. Although we have identified some differentially expressed mRNA and lncRNAs in LSP-AF as TCONS_00024590-XLOC_011883 may be participate in the regulation of LSP-AF in biological processes (small molecule metabolic process; oxidation-reduction process; signal transduction; lipid metabolic process; xenobiotic metabolic process), cellular component (extracellular vesicular exosome; membrane; integral to membrane; intracellular membrane-bounded organelle; endoplasmic reticulum membrane) and several pathways (Drug metabolism-cytochrome P450), by interacting with MGST3, it is too early for us to confirm their relationship. Therefore, subgroup analysis of mRNA and lncRNAs should be performed to explore this relationship in the future. In addition, most differentially expressed lncRNAs have no official Human Genome Nomenclature Committee symbol and their function is still unclear. Further functional studies are required to elucidate their roles in LSP-AF.

In conclusion, this dysregulation of the lncRNA and mRNA profile provides a novel insight into the etiology of AF and furthermore, illustrates the intricate relationship between coding and ncRNA transcripts in LSP-AF. These data may offer a background/reference resource for future functional studies of lncRNAs and mRNA related to AF.

Acknowledgements

We thank Q. Ming of Genminix Informatics Co., Ltd. for bioinformatics assistance. This study was supported in part by a Grant-in-Aid from the Liaoning Province, Science and Technology public relations project (Grant No. 20140200-65 and 201601402).

Disclosure of conflict of interest

None.

Address correspondence to: Huishan Wang, Department of Cardiovascular Surgery, Shenyang Northern Hospital, 83 Wenhua Road, Shenhe District, Shenyang 110016, Liaoning, China. Tel: +86-24-28897-382; Fax: +86-24-23912376; E-mail: huishanw@126.com

References

[1] Schnabel RB, Yin X, Gona P, Larson MG, Beiser AS, McManus DD, Newton-Cheh C, Lubitz SA, Magnani JW, Ellinor PT, Seshadri S, Wolf PA, Vasan RS, Benjamin EJ, Levy D. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the framingham heart study: A cohort study. Lancet 2015; 386: 154-162.

[2] Menezes AR, Lavie CJ, DiNicolantonio JJ, O'Keefe J, Morin DP, Khatib S, Milani RV. Atrial fibrillation in the 21st century: A current understanding of risk factors and primary prevention strategies. Mayo Clin Proc 2013; 88: 394-409.

[3] Xu J, Cui G, Esmailian F, Plunkett M, Marelli D, Ardehali A, Odim J, Laks H, Sen L. Atrial extracellular matrix remodeling and the maintenance of atrial fibrillation. Circulation 2004; 109: 363-368.

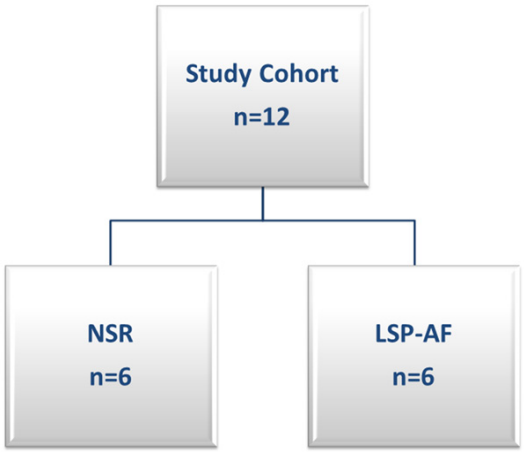
[4] European Heart Rhythm Association; European Association for Cardio-Thoracic Surgery, Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, Van Gelder IC, Al-Attar N, Hindricks G, Prendergast B, Heidbuchel H, Alfieri O, Angelini A, Atar D, Colonna P, De Caterina R, De Sutter J, Goette A, Gorenek B, Heldal M, Hohloser SH, Kolh P, Le Heuzey JY, Ponikowski P, Rutten FH. Guidelines for the management of atrial fibrillation: The task force for the management of atrial fibrillation of the european society of cardiology (esc). Eur Heart J 2010; 31: 2369-2429.

[5] Kataoka M, Wang DZ. Non-coding rnas including mirnas and lncrnas in cardiovascular biology and disease. Cells 2014; 3: 883-898.

[6] Ruan Z, Sun X, Sheng H, Zhu L. Long non-coding rna expression profile in atrial fibrillation. Int J Clin Exp Pathol 2015; 8: 8402-8410.

[7] Xu Y, Huang R, Gu J, Jiang W. Identification of long non-coding rnas as novel biomarker and potential therapeutic target for atrial fibrillation in old adults. Oncotarget 2016; 7: 10803-10811.

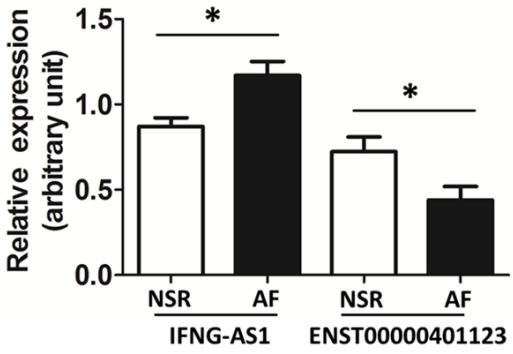
[8] Proietti M, Raparelli V, Basili S, Olshansky B, Lip GY. Relation of female sex to left atrial diameter and cardiovascular death in atrial fibrillation: The affirm trial. Int J Cardiol 2016; 207: 258-263.
[9] Broughton ST, O'Neal WT, Salahuddin T, Soliman EZ. The influence of left atrial enlargement on the relationship between atrial fibrillation and stroke. J Stroke Cerebrovasc Dis 2016; 25: 1396-402.
[10] Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, Picard MH, Roman MJ, Seward J, Shanewise J, Solomon S, Spencer KT, St John Sutton M, Stewart W; American Society of Echocardiography's Nomenclature and Standards Committee; Task Force on Chamber Quantification; American College of Cardiology Echocardiography Committee; American Heart Association; European Association of Echocardiography, European Society of Cardiology. Eur J Echocardiogr 2006; 7: 79-108.
[11] European Heart Rhythm Association; Heart Rhythm Society, Fuster V, Rydén LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, Halperin JL, Le Heuzey JY, Kay GN, Lowe JE, Olsson SB, Prystowsky EN, Tamargo JL, Wann S, Smith SC Jr, Jacobs AK, Adams CD, Anderson JL, Antman EM, Hunt SA, Nishimura R, Ornato JP, Page RL, Riegel B, Priori SG, Blanc JJ, Budaj A, Camm AJ, Dean V, Deckers JW, Despres C, Dickstein K, Lekakis J, McGregor K, Metra M, Morais J, Osterspey A, Zamorano JL; American College of Cardiology; American Heart Association Task Force on Practice Guidelines; European Society of Cardiology Committee for Practice Guidelines; Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation. Acc/aha/esc 2006 guidelines for the management of patients with atrial fibrillation--executive summary: A report of the american college of cardiology/american heart association task force on practice guidelines and the european society of cardiology committee for practice guidelines (writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation). J Am Coll Cardiol 2006; 48: 854-906.
[12] Bai W, Yang J, Yang G, Niu P, Tian L, Gao A. Long non-coding rna nr_045623 and nr_028291 involved in benzene hematotoxicity in occupationally benzene-exposed workers. Exp Mol pathol 2014; 96: 354-360.
[13] Dupuy D, Bertin N, Hidalgo CA, Venkatesan K, Tu D, Lee D, Rosenberg J, Svrzikapa N, Blanc A, Carnec A, Carvunis AR, Pulak R, Shingles J, Reece-Hoyes J, Hunt-Newbury R, Viveiros R, Mohler WA, Tasan M, Roth FP, Le Peuch C, Hope IA, Johnsen R, Moerman DG, Barabasi AL, Baillie D, Vidal M. Genome-scale analysis of in vivo spatiotemporal promoter activity in caenorhabditis elegans. Nat Biotechnol 2007; 25: 663-668.
[14] Draghici S, Khatri P, Tarca AL, Amin K, Done A, Voichita C, Georgescu C, Romero R. A systems biology approach for pathway level analysis. Genome Res 2007; 17: 1537-1545.
[15] Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The kegg resource for deciphering the genome. Nucleic Acids Res 2004; 32: D277-280.
[16] Prieto C, Risueno A, Fontanillo C, De las Rivas J. Human gene coexpression landscape: Confident network derived from tissue transcriptomic profiles. PLoS One 2008; 3: e3911.
[17] Hajhosseiny R, Matthews GK, Lip GY. Metabolic syndrome, atrial fibrillation, and stroke: Tackling an emerging epidemic. Heart Rhythm 2015; 12: 2332-2343.
[18] Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, Ding H, Butty VL, Torrey L, Haas S, Abo R, Tabebordbar M, Lee RT, Burge CB, Boyer LA. Braveheart, a long noncoding rna required for cardiovascular lineage commitment. Cell 2013; 152: 570-583.
[19] Congrains A, Kamide K, Oguro R, Yasuda O, Miyata K, Yamamoto E, Kawai T, Kusunoki H, Yamamoto H, Takeya Y, Yamamoto K, Onishi M, Sugimoto K, Katsuya T, Awata N, Ikebe K, Gondo Y, Oike Y, Ohishi M, Rakugi H. Genetic variants at the 9p21 locus contribute to atherosclerosis through modulation of anril and cdkn2a/b. Atherosclerosis 2012; 220: 449-455.
[20] Ishii N, Ozaki K, Sato H, Mizuno H, Saito S, Takahashi A, Miyamoto Y, Ikegawa S, Kamatani N, Hori M, Saito S, Nakamura Y, Tanaka T. Identification of a novel non-coding rna, miat, that confers risk of myocardial infarction. J Hum Genet 2006; 51: 1087-1099.
[21] Klimek-Piotrowska W, Holda MK, Piatek K, Koziej M, Holda J. Normal distal pulmonary vein anatomy. PeerJ 2016; 4: e1579.
[22] Jalife J. Novel upstream approaches to prevent atrial fibrillation perpetuation. Cardiol Clin 2014; 32: 637-650.
[23] Bell DS, O'Keefe JH. Metabolic syndrome and postoperative atrial fibrillation (poaf). Eur Heart J 2009; 30: 1167-1168.
[24] Kottkamp H. Human atrial fibrillation substrate: Towards a specific fibrotic atrial cardiomyopathy. Eur Heart J 2013; 34: 2731-2738.



Supplementary Figure 1. Group of selected patients.

Supplementary Table 1. Primers for RT-PCR

	Forward sequence (5'-3')	Reverse sequence (5'-3')
IFNG-AS1	GCTGATGATGGTGGTGGCAATCT	TTAGCAGTTGGTGGGCTTCT
ENST00000401123	GGTCCGAGTGTATTGGTGTTTAGA	AGGCTAGTCAAGTGAAGCAATGC
β-actin	GCTCAGGAGGAGCAATGATCTTG	GTACGCCAACACAGTGCTGTC



Supplementary Figure 2. Validation of differentially expressed lncRNAs in NSR (n=6) and LSP-AF (n=6) by qRT-PCR. The bars represent standard errors. The validation results of the lncRNAs (IFNG-AS1 and ENST00000401123) indicated that the qPCR results correlated well with the microarray data. The results are presented as mean ± SEM. *P < 0.05 vs. corresponding controls.

Supplementary Table 2. Genes with a large degree value and clustering coefficient in NSR samples (clustering coefficient > 0.6, Degree 10)

Gene/ncRNA	Type	Style	Degree	Clustering coefficient
DECR1	Coding	Up	14	0.648351648
n410929	Noncoding	Up	14	0.626373626
ENST00000545520	Noncoding	Up	13	0.628205128
NR_003341	Noncoding	Up	12	0.606060606
n409375	Noncoding	Up	11	0.672727273
YME1L1	Coding	Up	11	0.672727273
PPP1R9A	Coding	Up	11	0.654545455
ENST00000364189	Noncoding	Down	11	0.636363636
ECM1	Coding	Down	10	0.644444444
TCONS_00005647-XLOC_002908	Noncoding	Up	10	0.644444444

NSR, normal sinus rhythm.

Supplementary Table 3. Genes with a large degree value and clustering coefficient in LSP-AF samples (clustering coefficient > 0.6, Degree 10)

Gene/ncRNA	Type	Style	Degree	Clustering coefficient
ENST00000383932	Noncoding	Up	48	0.624113475
ENST00000384041	Noncoding	Up	46	0.64057971
ENST00000384312	Noncoding	Up	46	0.602898551
ENST00000362353	Noncoding	Up	45	0.688888889
TCONS_l2_00031069-XLOC_l2_015973	Noncoding	Up	45	0.635353535
TCONS_l2_00031115-XLOC_l2_016016	Noncoding	Up	45	0.635353535
TCONS_l2_00031158-XLOC_l2_016056	Noncoding	Up	45	0.635353535
TCONS_l2_00031100-XLOC_l2_016002	Noncoding	Up	45	0.628282828
ENST00000363154	Noncoding	Up	45	0.623232323
ENST00000384345	Noncoding	Up	44	0.694503171
ENST00000362539	Noncoding	Up	44	0.634249471
ENST00000383855	Noncoding	Up	44	0.604651163
ENST00000516993	Noncoding	Up	43	0.61683278
ENST00000384395	Noncoding	Up	41	0.735365854
ENST00000365312	Noncoding	Up	40	0.719230769
ENST00000362881	Noncoding	Up	39	0.690958165
n407825	Noncoding	Up	39	0.607287449
ENST00000384089	Noncoding	Up	38	0.783783784
ENST00000362996	Noncoding	Up	38	0.698435277
ENST00000365599	Noncoding	Up	38	0.628733997
ENST00000365235	Noncoding	Up	38	0.608819346
ENST00000410980	Noncoding	Up	38	0.608819346
ENST00000365439	Noncoding	Up	37	0.749249249
ENST00000362997	Noncoding	Up	37	0.725225225
ENST00000365406	Noncoding	Up	37	0.698198198
ENST00000363977	Noncoding	Up	37	0.624624625
ENST00000364112	Noncoding	Up	37	0.618618619
ENST00000384029	Noncoding	Up	34	0.604278075
SMDT1	Coding	Up	32	0.651209677
ENST00000384063	Noncoding	Up	32	0.629032258

ENST00000364439	Noncoding	Up	31	0.632258065
ENST00000516036	Noncoding	Up	31	0.627956989
ENST00000363110	Noncoding	Up	31	0.615053763
ENST00000363215	Noncoding	Up	30	0.650574713
ENST00000384001	Noncoding	Up	30	0.627586207
ENST00000410489	Noncoding	Up	29	0.655172414
ENST00000363651	Noncoding	Up	27	0.749287749
ENST00000384749	Noncoding	Up	27	0.60968661
ENST00000384333	Noncoding	Up	26	0.670769231
ENST00000363562	Noncoding	Up	26	0.64
ENST00000384373	Noncoding	Up	24	0.634057971
ENST00000365384	Noncoding	Up	22	0.74025974
ENST00000516043	Noncoding	Up	21	0.652380952
ENST00000362894	Noncoding	Up	21	0.619047619
ENST00000363462	Noncoding	Up	21	0.614285714
ENST00000384661	Noncoding	Up	20	0.726315789
ENST00000363709	Noncoding	Up	20	0.631578947
ENST00000364904	Noncoding	Up	18	0.699346405
MTIF3	Coding	Up	18	0.666666667
KARS	Coding	Up	17	0.647058824
ENST00000364631	Noncoding	Up	17	0.632352941
ENST00000384271	Noncoding	Up	16	0.883333333
ENST00000516297	Noncoding	Up	15	0.761904762
ENST00000515983	Noncoding	Up	15	0.752380952
ENST00000384138	Noncoding	Up	15	0.723809524
ENST00000363832	Noncoding	Up	15	0.695238095
ENST00000362807	Noncoding	Up	15	0.647619048
TCONS_l2_00031084-XLOC_l2_015987	Noncoding	Up	14	0.78021978
ENST00000364153	Noncoding	Up	14	0.648351648
ENST00000365341	Noncoding	Up	13	0.871794872
TCONS_l2_00031127-XLOC_l2_016028	Noncoding	Up	13	0.820512821
TCONS_l2_00031141-XLOC_l2_016040	Noncoding	Up	13	0.820512821
CWC15	Coding	Up	13	0.653846154
NR_015442	Noncoding	Up	13	0.615384615
CALM2	Coding	Up	12	0.666666667
NAV1	Coding	Down	10	0.777777778
SCO1	Coding	Up	10	0.666666667
ARFGEF2	Coding	Up	10	0.644444444
CD177P1	Coding	Down	10	0.622222222
FAM229B	Coding	Up	10	0.622222222
SMC2	Coding	Up	10	0.622222222

LSP-AF, long-standing persistent atrial fibrillation.

Supplementary Table 4. Co-expression status of genes between the LSP-AF and NSR samples

Gene/ncRNA	Type	Style	Exp_Degree	|Diff K|
n342783	Noncoding	Up	6	0.896551724
ENST00000362962	Noncoding	Up	54	0.844077961
ENST00000383932	Noncoding	Up	48	0.827586207
ENST00000363154	Noncoding	Up	45	0.775862069
ENST00000384268	Noncoding	Up	45	0.775862069
ENST00000364639	Noncoding	Up	50	0.775112444
n341959	Noncoding	Up	3	0.774362819
ENST00000364763	Noncoding	Up	54	0.757121439
NDUFA9	Coding	Up	7	0.748875562
ENST00000362421	Noncoding	Up	56	0.748125937
TCONS_00020461-XLOC_009792	Noncoding	Up	0	0.739130435
C17orf75	Coding	Up	0	0.739130435
RNASE2	Coding	Down	0	0.739130435
ENST00000383936	Noncoding	Up	40	0.689655172
NEK7	Coding	Up	6	0.67916042
n407825	Noncoding	Up	39	0.672413793
ENST00000362617	Noncoding	Up	39	0.672413793
ENST00000384480	Noncoding	Up	49	0.670914543
ENST00000384041	Noncoding	Up	46	0.662668666
ENST00000410980	Noncoding	Up	38	0.655172414
NR_033360	Noncoding	Down	0	0.652173913
ENST00000362997	Noncoding	Up	37	0.637931034
ENST00000363179	Noncoding	Up	37	0.637931034
TCONS_l2_00024184-XLOC_l2_012510	Noncoding	Up	6	0.635682159
MGST3	Coding	Up	1	0.634932534
ENST00000384345	Noncoding	Up	44	0.628185907
ENST00000364685	Noncoding	Up	44	0.628185907
CLCN3	Coding	Up	9	0.627436282
LIMCH1	Coding	Up	2	0.617691154
TCONS_00011974-XLOC_005486	Noncoding	Up	2	0.617691154
ENST00000365235	Noncoding	Up	38	0.611694153
NR_038275	Noncoding	Up	5	0.609445277
ENST00000439394	Noncoding	Down	0	0.608695652
TCONS_00003580-XLOC_001356	Noncoding	Up	0	0.608695652
ENST00000384078	Noncoding	Up	40	0.602698651
ENST00000384768	Noncoding	Up	50	0.6011994
PRKAR1A	Coding	Up	3	0.600449775

NSR, normal sinus rhythm. LSP-AF, long-standing persistent atrial fibrillation.