Integrated analysis of the proteome and transcriptome in human ischemic cardiomyopathy

Kai Zhang
Department of Cardiovascular Surgery, Guangdong Provincial Key Laboratory of South China Structural Heart Disease, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology

Xianyu Qin
Department of Cardiovascular Surgery,Guangdong Provincial Key Laboratory of South China Structural Heart Disease, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology

Min Wu
Department of Cardiovascular Surgery, Guangdong Provincial Key Laboratory of South China Structural Heart Disease, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology

Pengju Wen
Department of Cardiovascular Surgery, Guangdong Provincial Key Laboratory of South China Structural Heart Disease, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology

Yueheng Wu
Department of Cardiovascular Surgery, Guangdong Provincial Key Laboratory of South China Structural Heart Disease, Guangdong Provincial People's Hospital & Guangdong Academy of Medical Sciences, School of Medicine, South China University of Technology

Jian Zhuang (zhuangjiangenetics@163.com)
Guangdong Cardiovascular Institute https://orcid.org/0000-0003-0142-4238

Research

Keywords: Ischemic cardiomyopathy (ICM), Data independent acquisition (DIA), RNA-seq, lncRNA

Posted Date: November 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-104553/v1

Abstract

Background: Ischemic cardiomyopathy (ICM) is the primary cause of heart failure, which leads to an unacceptable rate of mortality and morbidity. The molecular mechanisms involved in ICM remains incompletely understood. This study aimed to investigate the molecular mechanisms of ICM by integrated proteome and transcriptome analyses.

Methods and Results: Data independent acquisition (DIA) mass spectrometry and RNA-seq technologies were performed in left ventricular species from 5 ICM patients and 5 unused non-failing donors. A total of 546 differentially expressed proteins (DEPs) and 1080 mRNAs (DEGs) were identified in ICM compared with control, which were mainly involved in inflammatory/immune response, response to stress (such as hypoxia and reactive oxidative species), oxidative stress, and ECM (extracellular matrix) organization. Moreover, though the low correlation between transcriptome and proteome, 41 key genes were identified, which showed the same expression directions at mRNA and protein levels. Among them, HSP90AA1 occupied a central position in the PPI network. Furthermore, a differentially expressed lncRNA-mRNA-protein network was constructed, which consisted of 13, 11, and 11 differentially expressed lncRNAs, mRNAs, and proteins, respectively, and the expression of this network were validated by qRT-PCR.

Conclusion: This identified some key genes and a lncRNA-mRNA-protein regulatory network involved in ICM, which should provide a framework for an in-depth interrogation into the complex molecular mechanisms of ICM.

Background

Heart failure (HF) is one of the most common diseases with an unacceptable rate of mortality and morbidity, leading to an extremely serious medical and socio-economic burden[1, 2]. Among the various reasons for HF, ischemic cardiomyopathy (ICM) has emerged as the primary cause, resulting in more than 50% of HF patients[3]. ICM refers to the cardiac dysfunction caused by diffuse of severe coronary artery stenosis or even complete occlusion. The ischemic events trigger a series of cellular events, such as myocardial stunning, hibernation, or even death, accompanied by activation of fibroblast, which ultimately leads to cardiac remodeling or myocardial fibrosis, leading to adverse events such as heart failure[4]. Despite the advances in medical therapy that have led to a markedly improved prognosis, these patients still have a high rate of death. Therefore, in-depth and comprehensive understanding of the molecular mechanism, identification of new molecular markers and drug targets are particularly important for better management of ICM.

Increasing evidence has suggested the crucial roles of long non-coding RNA (lncRNA), more than 200 nucleotides in length, in many physiological and pathological processes of cardiovascular diseases, including ICM[5–7]. Although several lncRNAs have been identified to be involved in the pathogenesis of ICM[8, 9], the role of these transcripts is largely unknown and needed to be further explored.

With the development of the high-throughput techniques, numerous studies have tried to discover the molecular mechanism and potential biomarker of disease through those "omics" techniques from different levels, such as genomics, transcriptomics, proteomics, metabolomics, and so on[10]. Single biomarkers may not be sufficient to represent the complex biological process of a disease. In contrast, the omics techniques could capture thousands of variables and demonstrate the interrelation of these variables in the development of disease[10]. Compared with RNAs, the proteins are more reliable when reflecting the molecular status of disease for their relative stability. However, the proteomics technique may ignore the low abundant proteins like transcription factors[11]. Integration of distinct, complementary approaches, such as proteomics and transcriptomics, has the potential to yield novel and comprehensive insights into the complex biological process of diseases[12, 13]. However, not too many studies have systemically investigated the molecular mechanisms of ICM by integrated proteome and transcriptome analyses.

In the present study, human left ventricular tissues of ICM were systemically explored through transcriptomics and proteomics techniques to identify the relevant biomarkers and mechanisms. Moreover, the potential lncRNA-mRNA regulation pairs were identified by *antisense*-, *cis*-, and *trans*-predictions. These results might provide some important evidence to understand the mechanisms of ICM.

Materials And Methods

Ethics statement

This study was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of local Hospital. All participants provided informed written consent.

Tissue collection

The left ventricular tissue of five ICM patients with end-stage HF undergoing heart transplantation, as well as five abandoned non-failing donors were collected for analysis in this study. All the samples were dissected on ice surface and snap-frozen in liquid nitrogen or fixed in 4% paraformaldehyde for future analysis. The diagnosis of ICM was based on clinical history, echocardiography, and coronary angiography data. The clinical characteristics of the patients, such as clinical history and echocardiography data, are shown in Table 1.

Protein extraction and digestion

The samples were ground to powder in liquid nitrogen and dissolved in lysis buffer (7M Urea, 2% SDS, 1 Protease Inhibitor Cocktail (Roche Ltd. Basel, Switzerland)). After lysing on ice for 30 min, the samples were centrifuged at 15,000 rpm for 15 min at 4 °C. The supernatant was collected and precipitated with ice-cold acetone at -20 °C overnight. The precipitations were then cleaned with acetone three times and re-dissolved in 7M Urea by sonication on ice. The protein concentration of the supernatant was determined by using the BCA Protein Assay Kit. For protein digestion, 50 μg protein was hydrolyzed by sequencing-

grade trypsin (Promega, Madison, WI) at a substrate/enzyme ratio of 50:1 (W/W) at 37℃ for 16 hours. The digested peptides were then lyophilized under a vacuum.

High pH reverse-phase separation

The peptide mixture was re-diluted in buffer A (buffer A: 20mM ammonium formate in water, pH10.0, adjusted with ammonium hydroxide), and then fractionated by a linear gradient pH separation system, Ultimate 3000 system (ThermoFisher Scientific, MA, USA) connected to a reverse-phase column (XBridge C18 column, 4.6mm x 250 mm, 5μm, (Waters Corporation, MA, USA)). The linear gradient was 5% to 45% buffer B (20 mmol/L ammonium formate in 80% acetonitrile, pH 10.0, adjusted with ammonium hydroxide). After re-equilibration, the column worked at a flow rate of 1 ml/min at 30℃ for 40 min. Ten fractions were collected during the separation and freeze-dried in a vacuum concentrator.

Data-dependent acquisition (DDA) analysis

The dried fractions were re-dissolved in 30 μl buffer C (0.1% formic acid in water), and analyzed by online electrospray tandem mass spectrometry. The experiments were performed on EASY-nLC 1200 system connected to an Orbitrap Fusion Lumos Tribrid (Thermo Fisher Scientific, MA, USA). Three μl peptide was separated on an analytical column (Acclaim Pep Map C18, Thermo Fisher Scientific) at a flow rate of 200 nL/min, with a linear gradient of 5% to 35% buffer D (0.1% formic acid in acetonitrile) in 120 min. The electrospray voltage of 2 kV versus the inlet of the mass spectrometer was used.

The raw data were searched against the H. sapiens proteome database using Spectronaut X (Biognosys AG, Switzerland) with a default setting. Carbamidomethyl and oxidation were set as the fixed and variable modifications, respectively. False discovery rate (FDR) was set to 1% for peptide and protein identification.

Data-independent acquisition (DIA) analysis

DIA analysis was performed with the same mass spectrometer and LC system as DDA. The mass spectrometer was run under data independent acquisition mode, and automatically switched between MS and MS/MS mode. The raw data of DIA were also analyzed by using Spectronaut X with the default setting. iRT peptides were used to calibrate retention time. All results were filtered by FDR of 1%. Differentially expressed proteins (DEPs) were calculated using the t-test and Benjamini-Hochberg. And proteins with fold change > 1.2 or < 0.83 with adjusted p value < 0.05 were consider as DEPs.

RNA extraction, library preparation, and RNA sequencing

Total RNA was extracted using TRIzol reagent ((Invitrogen Life Technologies, USA). The concentration and quality of RNA were assessed using the NanoDrop1000 spectrophotometer (Thermo Fisher Scientific, USA) and Agilent 2100 Bioanalyzer (Agilent Technologies, USA). After removing the ribosomal RNA, the mRNA and ncRNA were digested into short fragments by Ribo-Zero™ rRNA Removal Kit (Epicentre Madison, USA), and were retro-transcribed into cDNA by SuperScript III Reverse Transcriptase kit (Life

Technologies, Thermo Fisher Scientific, USA). The cDNA was then purified with QiaQuick PCR extraction kit (Qiagen, Germany), followed by end polish, poly(A) addition, and Illumina sequencing adapters ligation. The complementary strand was digested, and the sequencing strands were amplified with 15 cycles of PCR reaction, followed by sequence on Illumina HiSeqTM 4000 platform.

RNA-seq Data processing

The fastp (version 0.18.0) was used to filter the raw reads to obtain high quality of clean reads. The reads mapped to the ribosomal RNA database by Bowtie2 (version 2.2.8) were removed. The genome reference profile was downloaded from the genome website, and the paired-end clean reads were aligned to the reference genome using HISTA2 (version 2.1.0). The reconstruction of the transcript was performed by StringTie (version 1.3.4). For lncRNA identification, the CNCI (version 2) and CPC (version 0.9-r2) were used to assess the protein-coding potential of novel transcripts. The overlap of the non-protein-coding potential results of the two software was defined as lncRNA. The fragment per kilobase of transcript per million mapped reads (FPKM) value was calculated to quantify the expression abundance and variations by StringTie. The differentially expressed mRNA (gene) (DEG) and lncRNA (DEL) were calculated by DESeq2 between two different groups. The mRNAs with adjusted p value < 0.05 and fold change > 1.5 or <0.67 were considered as DEGs, while lncRNAs with p value < 0.05 and fold change > 2 or < 0.5 were assigned as DELs.

Functional annotation of DEPs/DEGs

To obtain the overview of the characteristics and the mechanisms of the DEPs/DEGs involved in ICM, gene functional enrichment analysis, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway functions, were performed by using online databases of DAVID (https://david.ncifcrf.gov/) and Metascape (http://metascape.org/). P value < 0.05 was considered a statistically significant difference.

Protein-protein interaction (PPI) network construction

The PPI network was constructed by using an online database of STRING (http://string-db.org/). The interaction score > 0.4 (medium confidence) was considered as significance. The Cytoscape (version 3.6.1) was used to further analyze and visualize the network.

lncRNA-mRNA association analysis

To identify the *antisense*- characteristics of lncRNA, the software RNAplex (version 0.2) (http://www.tbi.univie.ac.at/RNA/RNAplex.1.html) was used to predict the complementary interaction of lncRNA and mRNA, basing on the calculation of minimum free energy through thermodynamics structure. Another function of lncRNAs is the *cis*-regulation of their neighboring genes on the same allele. Therefore, the lncRNAs were annotated again, and those in less than 100kb up/downstream of a gene were considered *cis*-regulators. Furthermore, the correlations of expression between lncRNAs and mRNAs were

calculated to reveal the *trans*- regulation function of lncRNAs. The Pearson's correlation of each pair of lncRNA and mRNA no less than 0.95 were considered as *trans*- regulation.

Validation by qRT-PCR

The procedure of the qRT-PCR was performed as described in previous studies[11, 14]. In brief, total RNA was extracted and reverse transcribed to cDNA. The SYBR-green qPCR kit (Life Technologies, USA) was used to detect the relative expression of the RNA. The primer sequences are shown in Table 2. The $2^{-\Delta\Delta Ct}$ method was used to calculate relative gene expression levels. Endogenous GAPDH was used for standardization.

Results

Identification of ICM-related DEPs

The proteomic analysis based on the DIA technique was performed to investigate the abundance of cardiac protein in end-stage ICM patients. As a result, a total of 28,724 unique peptides and 3419 protein groups were successfully identified. Compared with non-failing donors (control), a total of 546 proteins were differentially expressed in the ICM group, including 377 up-regulated DEPs and 169 down-regulated DEPs (Fig. 1A and 1B).

Functional enrichment analysis of ICM-related DEPs

To further understand the mechanisms involved in ICM, the GO and KEGG pathway analyses of the DEPs were explored. Through annotation on the DAVID database, a total of 149 terms of GO biological process (BP), 59 terms of GO cellular component (CC), 68 terms of molecular function (MF), and 34 terms of KEGG pathway were significantly enriched (Additional file 1). Regarding the GO-BP terms, the ICM-related DEPs were mainly involved in mitochondria function (GO:0070125, GO:0070126, GO:0032543, etc.), DNA activity (GO:0006336, GO:0006334, GO:0006352, GO:0000183, GO:0045815, etc.), muscle contraction (GO:0060048, GO:0002026, GO:0010881, etc.), response to stress, such as hypoxia and reactive oxygen species (GO:0098869, GO:0055114, GO:0071260, GO:0000302, etc.), and energy and lipid mentalism (GO:0006099, GO:0006635, GO:0046034, etc.) (Fig. 2A and Additional file 1). The variations in CC of DEPs were predominantly enriched in extracellular space (GO:0070062, GO:0031012, GO:0005615, etc.), mitochondrion (GO:0005739, GO:0005759, GO:0005743, etc.), ribosome (GO:0005840, GO:0005761, etc.), and cardiac structural framework (GO:0030018, GO:0031430, GO:0030315, etc.) (Fig. 2B and Additional file 1). The variations in MF were markedly enriched in banding activities, such as poly(A) RNA binding (GO:0044822), protein binding (GO:0005515), histone binding (GO:0042393), receptor binding (GO:0005102), etc. (Fig. 2C and Additional file 1). The KEGG pathway analysis results revealed that the ICM-related DEPs were mainly involved in autoimmune response, substance and energy metabolism, such as Systemic lupus erythematosus (hsa05322), Complement and coagulation cascades (hsa04610), Carbon metabolism (hsa01200), Fatty acid degradation (hsa00071), Citrate cycle (TCA cycle) (hsa00020), etc. (Fig. 2D and Additional file 1)

Furthermore, similar results were observed in the analysis of Metascape. As shown in Fig. 2E and 2F, the ICM-related DEPs were markedly enriched in cardiac muscle structure development/organization and contraction, oxidative stress, energy metabolism, lipid metabolism, responses to stress, nucleotide metabolism, ECM (extracellular matrix) organization, immune responses, and so on.

Identification of ICM-related DEGs and the common enriched functions between DEGs and DEPs

To investigate the transcriptomic variations of ICM, the RNA-seq technology was performed. After filtration, an average yield of 70 M clean reads was obtained, with an average alignment rate of 97.38% of each sample. After analysis of the expression level, a total of 1080 DEGs were identified, including 609 up-regulated and 471 down-regulated DEGs in ICM compared with control (Fig. 3A and 3B).

Moreover, the functions of these DEGs were investigated, and a total of 129 BP, 35 CC, 33 MF, and 17 KEGG pathways were enriched (Additional file 2). Similar to DEPs, the biological function of the DEGs were mainly involved in inflammatory/immune response, response to stress, and fibrosis, such as inflammatory response (GO:0006954), T cell activation (GO:0042110), response to hydrogen peroxide (GO:0042542), bone mineralization (GO:0030282), ECM-receptor interaction (hsa04512) (Additional file 2).

The common enriched terms of the DEGs and DEPs were shown in Fig. 3C to 3F. The common BP showed that the ICM-related DEGs or DEPs were involved in the processes of translation (GO:0006412), extracellular matrix organization (GO:0030198), muscle contraction (GO:0003009), and response to stimulators, such as calcium ion or ethanol (GO:0051592, GO:0045471, etc.) (Fig. 3C). The common variations in CC were mainly enriched in extracellular region (GO:0005576) and membrane, such as basement membrane (GO:0005604), cell-cell adherent junction (GO:0005913), apical plasma membrane (GO:0016324) (Fig. 3D). The common MF enrichment results were mainly involved in bind activities, such as protein binding (GO:0005515), calcium ion binding (GO:0005509), and actin filament binding (GO:0051015) (Figure 3E). The common enriched KEGG pathways were Ribosome (hsa03010), Complement and coagulation cascades (hsa04610), Systemic lupus erythematosus (hsa05322) (Fig. 3F).

Integrated analysis of proteome and transcriptome

The integrated analysis of transcriptome and proteome can provide a more comprehensive insight into the gene transcriptional profile and post-transcriptional regulation[15]. In the present study, the expression data of transcriptome and proteome were combined together, and the correlation between the two profiles was calculated. As shown in Fig. 4A, there was no significant correlation between the transcriptome and proteome ($r = 0.23$). According to the threshold of fold change, the integrated data could be divided into nine quadrants. The third and seventh quadrants represented the same trend for transcriptome and proteome expression, and a total of 28 significantly co-up regulated genes, as well as 13 significantly co-down regulated genes, were identified (Fig. 4A).

The functions of these co-regulated genes were mainly involved in response to hypoxia or oxidative stress, muscle contraction, and the complement cascade, such as HIF-1 signaling pathway, oxidation-reduction process, response to hydrogen peroxide, straited muscle contraction, and complement activation (Fig. 4B). The interactions of these co-regulated genes/proteins were shown the Fig. 4C, and only 18 interactions among 19 proteins were identified. The co-regulated genes with the top 10 degrees were shown in Fig. 4D, and HSP90AA1 showed a central position in the PPI network.

Identification of the DELs and differentially expressed lncRNA-mRNA-protein network construction

Increasing evidence has suggested the crucial role of lncRNA in regulating the transcriptional and post-transcriptional expression of coding genes through three main types, including *antisense*-, *cis*-, and *trans*-regulation[6, 16]. In the present study, a total of 1227 lncRNAs differentially expressed in ICM compared with control, including 626 up-regulated DELs and 601 down-regulated DELs (Fig. 5A and 5B). Moreover, 12 of lncRNA:*antisense*-mRNA pairs, 37 of lncRNA:*cis*-mRNA pairs, and 4283 of lncRNA:*trans*-mRNA pairs were identified between the DEGs and DELs. Additionally, the co-regulated genes between transcriptome and proteome with their regulatory lncRNA were extracted out, and their network was constructed. As shown in Fig. 5C, 13 DELs (DNM1P46-201, SNHG32-204, ZFAS1-204, CLASP1-205, PSMG3-AS1-203, LINC00881-202, GPAM-203, CHTOP-206, HLA-DPB1-217, AL513365.2-202, SNHG5-277, Z99127.4-201, SNHG5-283) were identified which regulated the expression of the 11 co-regulated genes (TF, CACYBP, FMOD, AOC3, TAGLN, PDK1, S100A13, S100A4, HMGN2, IGHA1, HLA-DPA1) by two *antisense*-, two *cis*-, and 13 *trans*- manners. For example, the decreased lncRNA of HLA-DPB1 might up-regulate the expression of the HLA-DPA1 mRNA via *antisense*- activity, which resulted in an increase of HLA-DPA1 protein (Fig. 5C).

Validation by qRT-PCR

To validate the expression of genes identified in the differentially expressed lncRNA-mRNA-protein network, the qRT-PCR was performed. As shown in Fig. 6, four lncRNAs (AL513365.2-202, DNM1P46-201, HLA-DPB1-217, and SNHG32-204) and one mRNA (S100A13) showed no statistical difference, while the other 9 lncRNAs and 10 mRNAs showed significant difference with the same directions as RNA-seq data.

Discussion

In the present study, an integrated analysis of transcriptome and proteome was performed to investigate the molecular mechanisms related to the pathogenesis of ICM. Each molecular feature represented partial aspects of the ICM, and the combination of lncRNA, mRNA, and protein information provided a more comprehensive insight into the cellular mechanisms related to ICM.

Compared with DNA and RNA, protein is the most directly functional molecule[17]. In this study, the DIA approach was used to identify the abundances as well as the changes of proteins in ICM, and a total of 546 proteins (including 377 up- and 169 down-regulated DEPs) were differentially expressed compared with NFD. Previous studies have revealed multiple cellular pathophysiologies involved in ICM, such as

oxidative stress, inflammation, mitochondrial dysfunction, apoptosis cascade, calcium overload, myocardial fibrosis, and so on[4, 18, 19]. Similarly, our results of the GO and pathway analysis also suggested that these DEPs participated in cardiac muscle structure development/organization and contraction, oxidative stress, energy metabolism, lipid metabolism, responses to stress and nucleotide metabolism, ECM organization, and immune responses.

The use of complementary approaches could attenuate some deficiencies and provide new perspectives compared with each omics method used in isolation. In this study, the RNA-seq technique was performed, and 1080 genes were identified to be changed in RNA level, including 609 up- and 471 down-regulated DEGs in ICM in contrasted to NFD. Similarly, these DEGs were also markedly involved in inflammatory/immune response, response to stress, oxidative stress, and fibrosis. The exact same terms enriched from DEPs and DEGs furtherly highlighted the key roles of the processes of ECM organization and response to stress, as well as the complement and coagulation cascades signal pathway, in the mechanisms of ICM, which were consistent with previous researches[19–21].

The integrated analysis of transcriptome and proteome has the potential to provide a systemic view on gene expression profile and complex biological functions related to the status of post-translational turnover and varying translation efficiencies[13, 15]. However, the correlation of gene expression profile at transcript and protein levels was poor overall, which was similar to other researches[15, 22]. Although the genetic information ultimately translated into amino acid information, the post-transcriptional and post-translational variations could result in an imperfect correlation between genetic programs and protein phenotypes[22, 23], that may partially interpret the poor correlation between transcripts and proteins. In this study, 41 key genes were identified, which showed the same expression directions at mRNA and protein levels. And these key genes were mainly associated with the process of response to hypoxia or oxidative stress, muscle contraction, as well as the complement cascade. Among the 41 genes, the HSP90AA1 occupied a central position in the PPI network, which was down-regulated with $\log_2$FC (fold change) of -1.41 and − 1.06 in mRNA and protein levels, respectively.

HSP90AA1 is a member of the HSP90 (heat shock protein) family, which is considered as a protein stabilizer binding to multiple receptors, such as epidermal growth factor receptor, transforming growth factor TGFβ receptor, and epithelial-mesenchymal transition factor, and thus protect the corresponding signal[24–26]. Previous researches have revealed the role of HSP90 in promoting fibrosis in different models, including renal fibrosis, liver fibrosis, and cardiac fibrosis in pressure overloaded rats[24, 27, 28]. It seems that HSP90 would act as a risk factor exacerbating the process of cardiac fibrosis and heart failure. However, studies also suggested the protective function of HSP90 in cardiomyocytes under ischemia and reperfusion conditions[29, 30]. It was demonstrated that over-expression of HSP90AA1 could alleviate the cardiomyocytes apoptosis, and knockdown of HSP90AA1 could enhance the apoptosis[30]. Therefore, the protective or aggravating role of HSP90AA1 in ICM or HF needs to be further investigated in the future.

LncRNAs have been suggested as crucial regulators of various cardiovascular diseases involved in response to stress, apoptosis, autophagy, proliferation, fibrosis, and so on[9, 31]. For example, overexpression of lncRNA Gm2691 could attenuate the apoptosis after myocardial infarction[32], while down-regulation of H19 could promote cell proliferation and inhibit cell apoptosis[33]. The mechanisms of lncRNAs in regulating the physiological or pathophysiological process are involved in a variety of approaches, especially binding miRNA as competing endogenous RNA, influencing neighboring genes as a *cis*-acting element, regulating remote genes as a *trans*-acting element, binding mRNA as *antisense*-regulator[16, 32, 34]. Thus, lncRNAs through those transcriptional and post-transcriptional modulation to regulate the relevant expressions of genes. In this study, 12 of lncRNA:*antisense*-mRNA pairs, 37 of lncRNA:*cis*-mRNA pairs, and 4283 of lncRNA:*trans*-mRNA pairs were identified between the DEGs and DELs. For the 41 key genes with the same expressive trends between mRNA and protein levels, only 11 genes (TF, CACYBP, FMOD, AOC3, TAGLN, PDK1, S100A13, S100A4, HMGN2, IGHA1, HLA-DPA1) were identified their related regulatory lncRNAs (DNM1P46-201, SNHG32-204, ZFAS1-204, CLASP1-205, PSMG3-AS1-203, LINC00881-202, GPAM-203, CHTOP-206, HLA-DPB1-217, AL513365.2-202, SNHG5-277, Z99127.4-201, SNHG5-283) through two *antisense*-, two *cis*-, and 13 *trans*- approaches. For example, the decrease of lncRNA HLA-DPB1-217 might up-regulate the expression of gene HLA-DPA1 through *antisense*-approach, while the increased lncRNA CLASP1-205 might up-regulate the expression of HLA-DPA1 through *trans*-approach. The two lncRNAs worked together, ultimately leading to the increase of HLA-DPA1 protein. Though most of the expressions of these lncRNAs were validated, their regulatory functions and their roles in the development of ICM should be further investigated in the future.

Conclusion

Taken together, we systematically investigated the transcriptomic and proteomic data of left ventricles collected from ICM as well as NFD. The results highlighted the processes of ECM organization, response to stress (such as oxidative stress), as well as the complement and coagulation cascades signal pathways in the mechanisms of ICM. Though the correlation between transcriptome and proteome was poor, 41 key genes were identified, which showed the same expressive directions at mRNA and protein levels. Moreover, the lncRNA regulatory network of the DEGs was investigated based on *antisense*-, *cis*-, and *trans*- regulation, and 13 lncRNAs were obtained, which regulated 11 of the 41 key genes. It is important to note that many of these lncRNAs and genes remain poorly understood in the pathogenesis of ICM and needs further exploration. This work should provide a framework for an in-depth interrogation into the complex molecular mechanisms of ICM.

Abbreviations

ICM, ischemic cardiomyopathy; DIA, data independent acquisition; DEPs, differentially expressed proteins; DEGs, differentially expressed genes; DEL, differentially expressed lncRNA; ECM, extracellular matrix; HF, heart failure; lncRNA, long non-coding RNA; PPI, protein-protein interaction; GO, Gene Ontology; BP,

biological process; CC, cellular component; MF, molecular function; NFD, non-failing donor; HF, heart failure.

Declarations

Ethics approval and consent to participate

The collection and research use of the human heart tissues was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of Guangdong People's Hospital (approval No. GDREC2016255H). All participants provided informed written consent.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The present study was supported by the Guangdong Provincial Science and Technology Planning Project, Grant/Award, Number: 2017A070701013, 2017B090904034, 2017B030314109.

Authors' contributions

KZ, MW, XQ, and PW collected the heart tissues and clinical data. KZ analyzed the data and draft the manuscript. YW, and JZ designed and revised the manuscript; All authors read and approved the final manuscript.

Acknowledgments

The authors would like to thank Guangzhou Genedenovo Biotechnology Co., Ltd for assisting in sequencing and bioinformatics analysis.

References

1. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, Chiuve SE, Cushman M, Delling FN, Deo R, et al: **Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association.** *Circulation* 2018, **137:**e67-e492.

2. Lu D, Xia Y, Chen Z, Chen A, Wu Y, Jia J, Sun A, Zou Y, Qian J, Ge J: **Cardiac Proteome Profiling in Ischemic and Dilated Cardiomyopathy Mouse Models.***Front Physiol* 2019, **10:**750.
3. Rosenbaum AN, Agre KE, Pereira NL: **Genetics of dilated cardiomyopathy: practical implications for heart failure management.***Nat Rev Cardiol* 2020, **17:**286-297.
4. Wu L, Li F, Zhao C, Ming Y, Zheng C, Li Y, Lei S, Chen C: **Effects and Mechanisms of Traditional Chinese Herbal Medicine in the Treatment of Ischemic Cardiomyopathy.***Pharmacol Res* 2020, **151:**104488.
5. Bär C, Chatterjee S, Thum T: **Long Noncoding RNAs in Cardiovascular Pathology, Diagnosis, and Therapy.***Circulation* 2016, **134:**1484-1499.
6. Chen S, Ma Q, Xue Y, Zhang J, Yang G, Wang T, Ma A, Bai L: **Comprehensive Analysis and Co-Expression Network of mRNAs and lncRNAs in Pressure Overload-Induced Heart Failure.***Front Genet* 2019, **10:**1271.
7. Lin F, Gong X, Yu P, Yue A, Meng Q, Zheng L, Chen T, Han L, Cao H, Cao J, et al: **Distinct Circulating Expression Profiles of Long Noncoding RNAs in Heart Failure Patients With Ischemic and Nonischemic Dilated Cardiomyopathy.***Front Genet* 2019, **10:**1116.
8. Zou Y, Wang Y, Rong Z, Wei B, Liu Y, Song Z, Li W, Hu E, Deng G, He Y, et al: **A susceptibility biomarker identification strategy based on significantly differentially expressed ceRNA triplets for ischemic cardiomyopathy.***Biosci Rep* 2020, **40**.
9. Saddic LA, Sigurdsson MI, Chang TW, Mazaika E, Heydarpour M, Shernan SK, Seidman CE, Seidman JG, Aranki SF, Body SC, Muehlschlegel JD: **The Long Noncoding RNA Landscape of the Ischemic Human Left Ventricle.***Circ Cardiovasc Genet* 2017, **10**.
10. Vernon ST, Hansen T, Kott KA, Yang JY, O'Sullivan JF, Figtree GA: **Utilizing state-of-the-art "omics" technology and bioinformatics to identify new biological mechanisms and biomarkers for coronary artery disease.***Microcirculation* 2019, **26:**e12488.
11. Chen L, Yang F, Chen X, Rao M, Zhang NN, Chen K, Deng H, Song JP, Hu SS: **Comprehensive Myocardial Proteogenomics Profiling Reveals C/EBPα as the Key Factor in the Lipid Storage of ARVC.***J Proteome Res* 2017, **16:**2863-2876.
12. Liu X, Qing W, Cui Y, Li X, Yang H: **In-depth transcriptomic and proteomic analyses of the hippocampus and cortex in a rat model after cerebral ischemic injury and repair by Shuxuetong (SXT) injection.***J Ethnopharmacol* 2020, **249:**112362.
13. Zhu X, Qiu Z, Ouyang W, Miao J, Xiong P, Mao D, Feng K, Li M, Luo M, Xiao H, Cao Y: **Hepatic transcriptome and proteome analyses provide new insights into the regulator mechanism of dietary avicularin in diabetic mice.***Food Res Int* 2019, **125:**108570.
14. Zheng W, Chen C, Chen S, Fan C, Ruan H: **Integrated analysis of long non-coding RNAs and mRNAs associated with peritendinous fibrosis.***J Adv Res* 2019, **15:**49-58.
15. Li X, Xie S, Qian L, Cai C, Bi H, Cui W: **Identification of genes related to skeletal muscle growth and development by integrated analysis of transcriptome and proteome in myostatin-edited Meishan pigs.***J Proteomics* 2020, **213:**103628.

16. Yang KC, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, Ewald GA, Mann DL, Nerbonne JM: **Deep RNA sequencing reveals dynamic regulation of myocardial noncoding RNAs in failing human heart and remodeling with mechanical circulatory support.** *Circulation* 2014, **129:**1009-1021.
17. Sajic T, Liu Y, Aebersold R: **Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications.** *Proteomics Clin Appl* 2015, **9:**307-321.
18. Zhou B, Tian R: **Mitochondrial dysfunction in pathophysiology of heart failure.** *J Clin Invest* 2018, **128:**3716-3726.
19. Li M, Parker BL, Pearson E, Hunter B, Cao J, Koay YC, Guneratne O, James DE, Yang J, Lal S, O'Sullivan JF: **Core functional nodes and sex-specific pathways in human ischaemic and dilated cardiomyopathy.** *Nat Commun* 2020, **11:**2843.
20. Kim GH, Uriel N, Burkhoff D: **Reverse remodelling and myocardial recovery in heart failure.** *Nat Rev Cardiol* 2018, **15:**83-96.
21. Suffritti C, Tobaldini E, Schiavon R, Strada S, Maggioni L, Mehta S, Sandrone G, Toschi-Dias E, Cicardi M, Montano N: **Complement and contact system activation in acute congestive heart failure patients.** *Clin Exp Immunol* 2017, **190:**251-257.
22. Lau E, Cao Q, Lam MPY, Wang J, Ng DCM, Bleakley BJ, Lee JM, Liem DA, Wang D, Hermjakob H, Ping P: **Integrated omics dissection of proteome dynamics during cardiac remodeling.** *Nat Commun* 2018, **9:**120.
23. Vogel C, Marcotte EM: **Insights into the regulation of protein abundance from proteomic and transcriptomic analyses.** *Nat Rev Genet* 2012, **13:**227-232.
24. García R, Merino D, Gómez JM, Nistal JF, Hurlé MA, Cortajarena AL, Villar AV: **Extracellular heat shock protein 90 binding to TGFβ receptor I participates in TGFβ-mediated collagen production in myocardial fibroblasts.** *Cell Signal* 2016, **28:**1563-1579.
25. Taipale M, Jarosz DF, Lindquist S: **HSP90 at the hub of protein homeostasis: emerging mechanistic insights.** *Nat Rev Mol Cell Biol* 2010, **11:**515-528.
26. Tomcik M, Zerr P, Pitkowski J, Palumbo-Zerr K, Avouac J, Distler O, Becvar R, Senolt L, Schett G, Distler JH: **Heat shock protein 90 (Hsp90) inhibition targets canonical TGF-β signalling to prevent fibrosis.** *Ann Rheum Dis* 2014, **73:**1215-1222.
27. Datta R, Bansal T, Rana S, Datta K, Chattopadhyay S, Chawla-Sarkar M, Sarkar S: **Hsp90/Cdc37 assembly modulates TGFβ receptor-II to act as a profibrotic regulator of TGFβ signaling during cardiac hypertrophy.** *Cell Signal* 2015, **27:**2410-2424.
28. Noh H, Kim HJ, Yu MR, Kim WY, Kim J, Ryu JH, Kwon SH, Jeon JS, Han DC, Ziyadeh F: **Heat shock protein 90 inhibitor attenuates renal fibrosis through degradation of transforming growth factor-β type II receptor.** *Lab Invest* 2012, **92:**1583-1596.
29. Budas GR, Churchill EN, Disatnik MH, Sun L, Mochly-Rosen D: **Mitochondrial import of PKCepsilon is mediated by HSP90: a role in cardioprotection from ischaemia and reperfusion injury.** *Cardiovasc Res* 2010, **88:**83-92.

30. Zhu WS, Guo W, Zhu JN, Tang CM, Fu YH, Lin QX, Tan N, Shan ZX: **Hsp90aa1: a novel target gene of miR-1 in cardiac ischemia/reperfusion injury.***Sci Rep* 2016, **6:**24498.
31. Li M, Duan L, Li Y, Liu B: **Long noncoding RNA/circular noncoding RNA-miRNA-mRNA axes in cardiovascular diseases.***Life Sci* 2019, **233:**116440.
32. Li T, Tian H, Li J, Zuo A, Chen J, Xu D, Guo Y, Gao H: **Overexpression of lncRNA Gm2691 attenuates apoptosis and inflammatory response after myocardial infarction through PI3K/Akt signaling pathway.***IUBMB Life* 2019, **71:**1561-1570.
33. Han Y, Xu H, Cheng J, Zhang Y, Gao C, Fan T, Peng B, Li B, Liu L, Cheng Z: **Downregulation of long non-coding RNA H19 promotes P19CL6 cells proliferation and inhibits apoptosis during late-stage cardiac differentiation via miR-19b-modulated Sox6.***Cell Biosci* 2016, **6:**58.
34. Tian W, Du Y, Ma Y, Gu L, Zhou J, Deng D: **MALAT1-miR663a negative feedback loop in colon cancer cell functions through direct miRNA-lncRNA binding.***Cell Death Dis* 2018, **9:**857.

Tables

Table 1. Clinical and echocardiographic characteristics of end-stage ICM patients.

	ICM 1	ICM 2	ICM 3	ICM 4	ICM 5
Age (years)	48	51	50	56	46
Male	1	1	1	1	0
Hypertension	1	1	0	1	0
Diabetes	1	1	0	0	0
NYHA	□	□	□	□	□
LVEF (%)	21	12	26	23	24
LVEDD (mm)	65	77	62	67	63
LVESD (mm)	56	69	54	60	54

NYHA, New York Heart Association; LVEF, left ventricular ejection fraction; LVEDD, left ventricular end diastolic diameter; LVESD, left ventricular end systolic diameter.

Table 2. Primer sequences used for qRT-PCR

Gene name	Forward	Reverse
HLA-DPA1	CCAGCGTTCCAACCACACTCAG	GAGGGCACAAAGGTCAGGTAATGG
HMGN2	GGTGAAGGACGAACCACAGAGAAG	CACTTGGCATCTCCAGCACCTTC
S100A4	TGGCGTGCCCTCTGGAGAAG	CCTCGTTGTCCCTGTTGCTGTC
S100A13	GCAGCAGAACCACTGACAGAGC	CCGAGTCCTGATTCACATCCAAGC
PDK1	TGACCGAGGAGGTGGCGTTC	CGTGGTTGGTGTTGTAATGCTTCC
TAGLN	TGAAGGCGGCTGAGGACTATGG	CTCTGTTGCTGCCCATCTGAAGG
AOC3	CCTGGTGAATGTGGTGCTGATCC	GCCCATGCCAAAGCCTCCATC
FMOD	GTACCTGGACCACAACAACCTGAC	GGACAGCCGCACATACAGCAG
CACYBP	TTGAAACCCATCTCTGTGGAAGGC	ATTCCGTGTCTCCTTTGGCTTGC
TF	TGCTGTTGCTGTGGTGAAGAAGG	AGACACTTGAAGGCTCCCGAGTAG
IGHA1	CATGGAACCATGGGAAGACCTTC	TCAGCGTCACCAGCTCGTTCA
ZFAS1-204	AAGTTCCCGTTTTGTGTGTGGT	CGTGGCTCCCAACCAATTTTAT
CLASP1-205	TATTTGGTACCAGGATTACTGAC	TTGGATGATGTAGAACTAGCAGA
GPAM-203	GACCTTGCTTATTCACCCACA	CAGGACTCAGCTATCAGTTTCG
SNHG32-204	AGAGCCATGCCTGTTCTCCTTC	CCCAATAAGGTGACAGCTGCTC
PSMG3-AS1-203	ACTTCATCTCAGCCACCCGGTT	ACCAGACTATCCGAGTTCCTGC
LINC00881-202	CTTCCCATGCACTCTTCCTGTT	TCCTTGTCTTCCATCCAGTGCA
CHTOP-206	GAGCACTGGCGTCTGTTTCCTT	GTTGAAGTTGCTGTTGGCTGCC
HLA-DPB1-217	AGACACAACTACGAGCTGGGC	ACCCCAGCTGTTTCCTCCTGT
AL513365.2-202	AGAATTTGAGCTGGCCCTTTAAC	CTGAACACAGTTCTCTGCTATTC
SNHG5-277	CGCACTTCGGTCTTTTACGT	CATCTTCAGAGCTGCTCCAC
Z99127.4-201	ACAGTTTCAGGCCCTCCTTT	GGCTCACTTATACCGAATCCA
SNHG5-283	TGAGTGTGGACGAGTAGCCAGT	TCTGGCCTCTATCAATGGGCAG
DNM1P46-201	TGAGACAAGAAACGTAAAACCC	GCCATGTAGGAGTCCACCAG

Figures

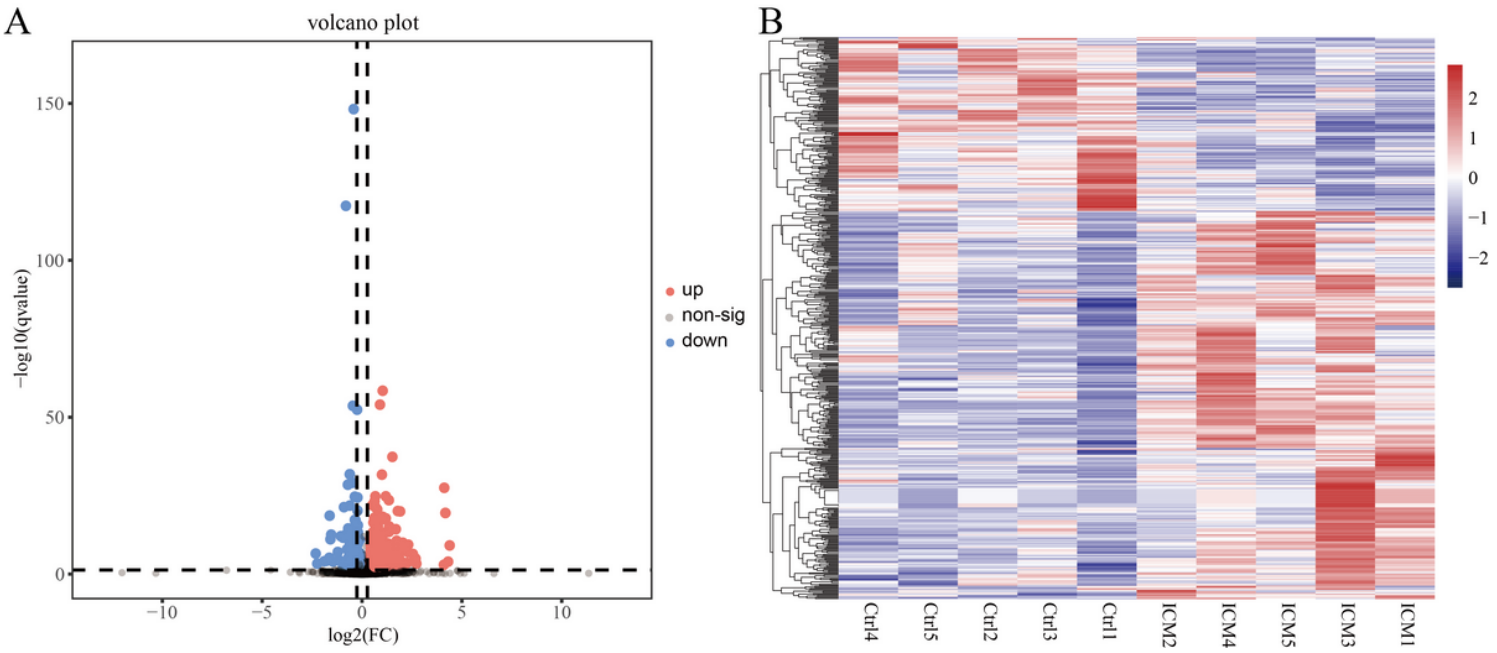

Figure 1

Identification of ICM-related DEPs (differentially expressed proteins). (A) The volcano plot of all the proteins. (B) The cluster heatmap of all the DEPs. Red indicates a higher expression level, while blue indicates a lower expression level.

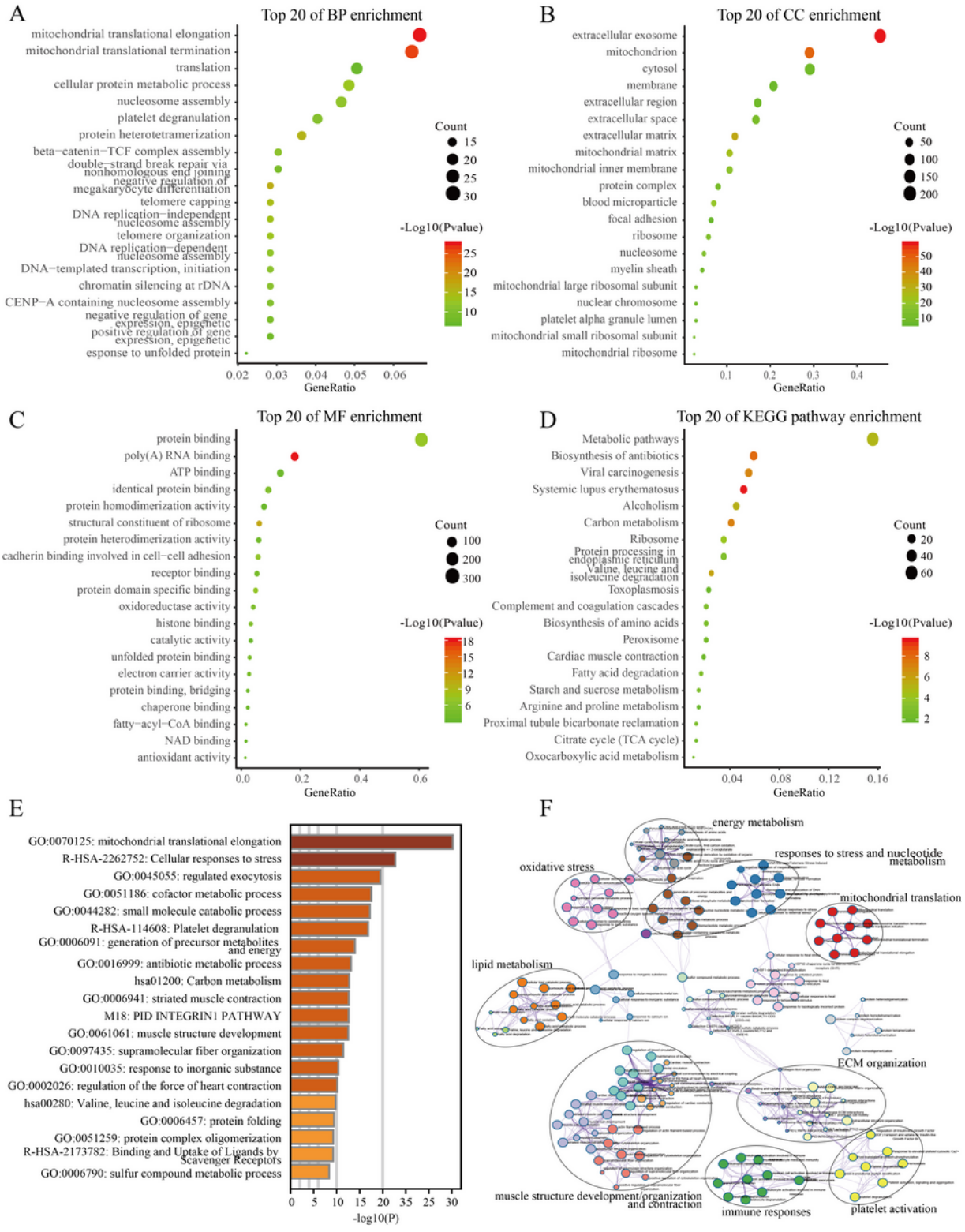

Figure 2

Functional enrichment analysis of ICM-related DEPs. The bubble diagram of the top 20 significant enriched terms of (A) biological process (BP), (B) cellular component (CC), (C) molecular function (MF), and (D) KEGG pathway via the DAVID database. In the bubble diagrams, dot sizes represent counts of enriched DEGs, and dot colors represent negative Log10 values (p values). (E) The bar diagram of

enriched pathways and biological processes, colored by p values, via the Metascape database. (F) Network of the enriched pathways and processes colored by cluster.

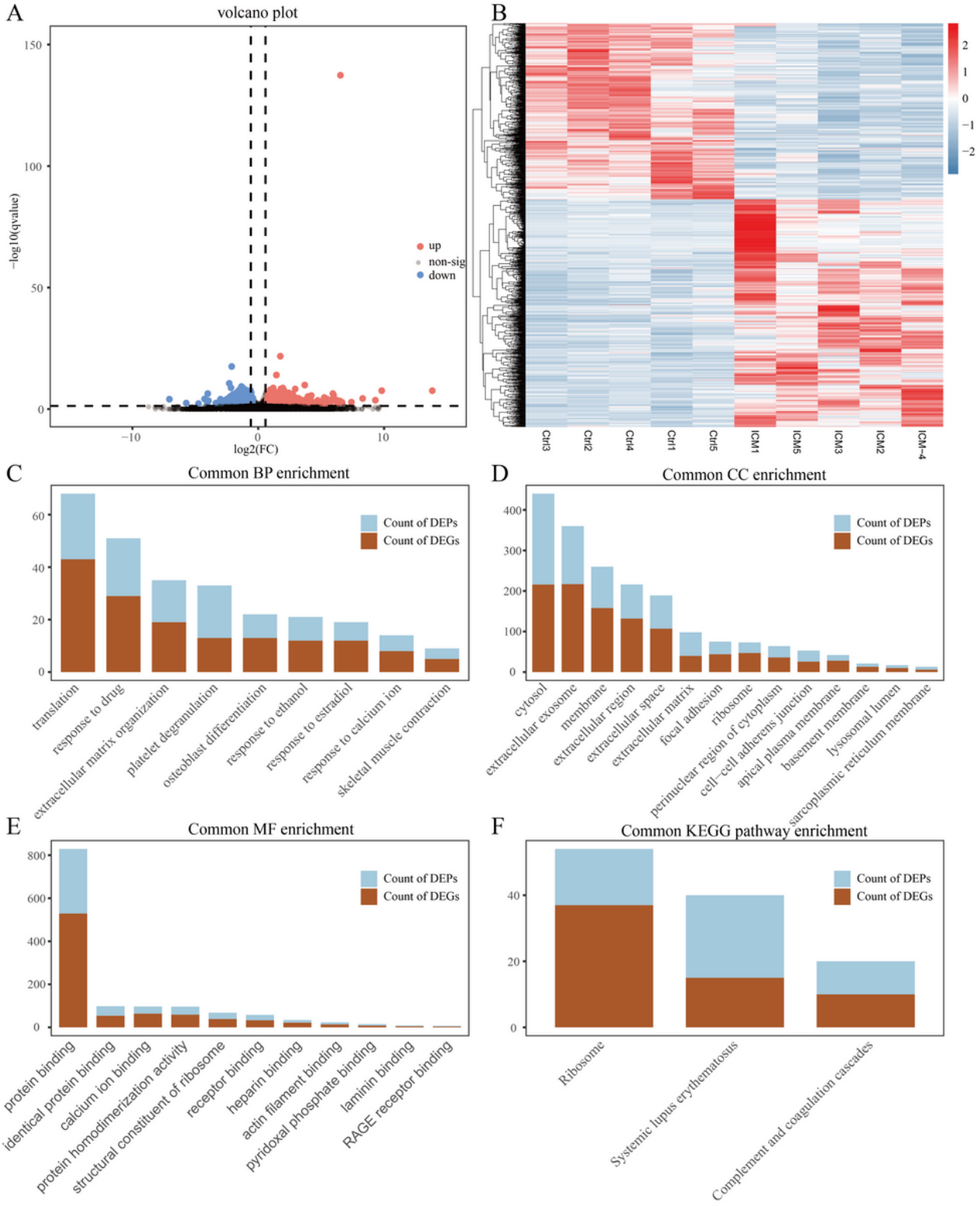

Figure 3

Identification of ICM-related DEGs (differentially expressed genes) and the common enriched functions between DEPs and DEPs. (A) The volcano plot of all the genes. (B) The cluster heatmap of all the DEGs. Red indicates a higher expression level, while blue indicates a lower expression level. (C-D) The bar

diagram of the common enriched functional terms of biological process (BP), cellular component (CC), molecular function (MF), and KEGG pathway via the DAVID database. The height of the bar represents the counts of the genes or proteins.

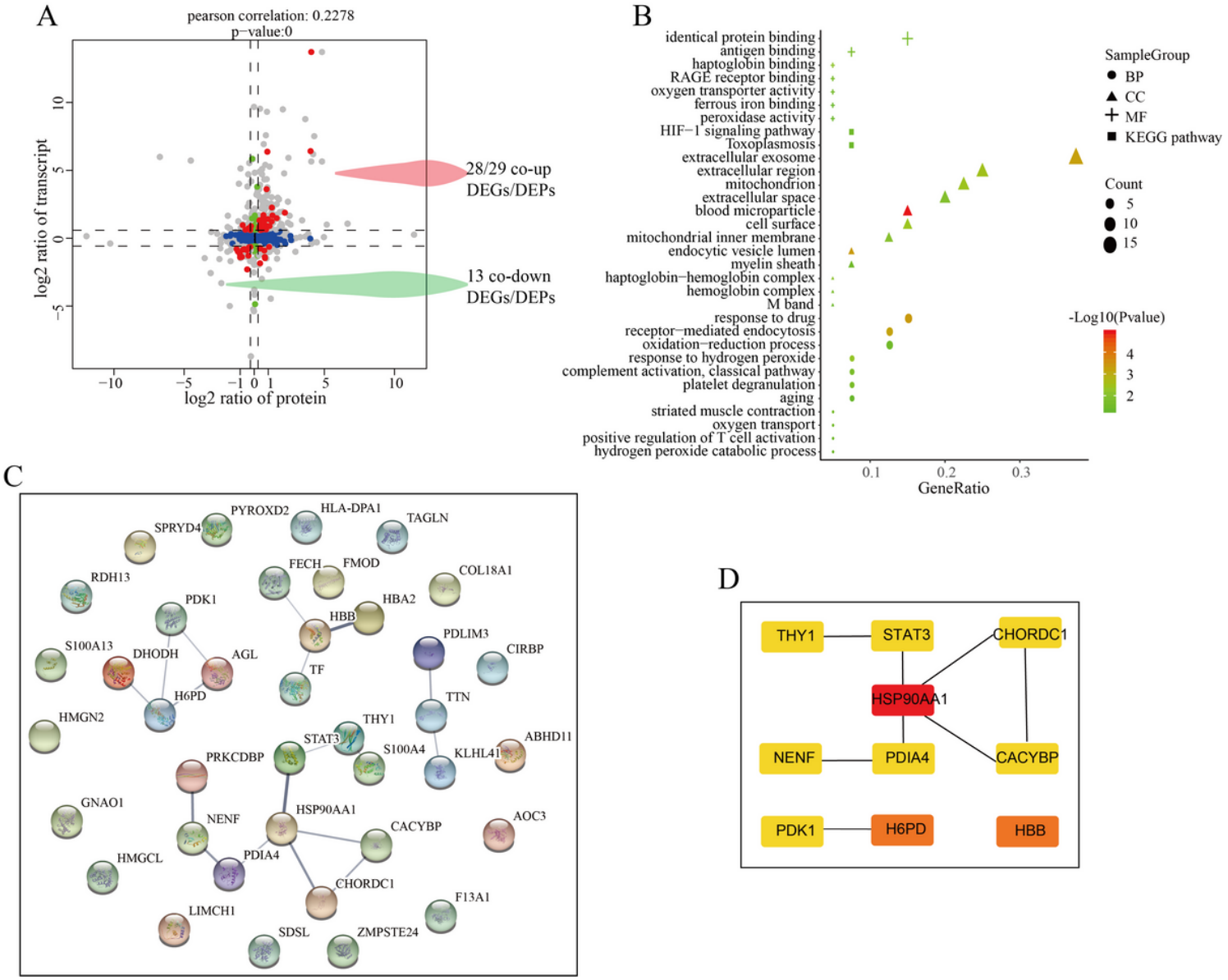

Figure 4

Integrated analysis of proteome and transcriptome. (A) The comparison of the gene expression levels between the transcriptome (fold change of FPKM) and proteome analysis (fold change of protein components). (B) The bubble diagram of the functional analysis of the co-up and co-down regulated genes. Dot sizes represent counts of enriched DEGs, and dot colors represent negative Log10 values (p values). (C) The protein-protein interaction (PPI) network of the co-up and co-down regulated genes. (D) Hub genes with top 10 degrees of the PPI network of the co-up and co-down regulated genes, colored by degree.

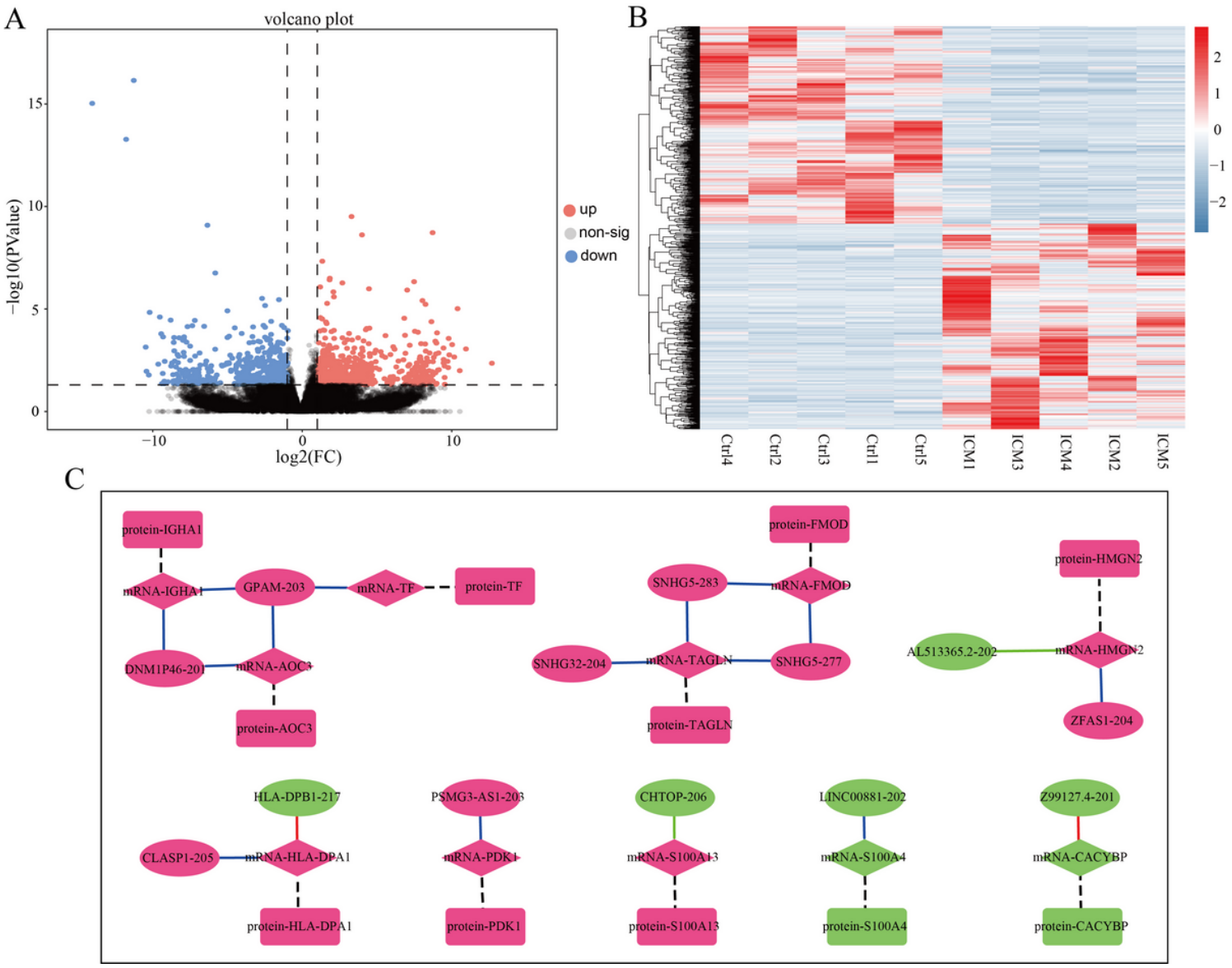

Figure 5

Identification of the DELs (differentially expressed lncRNAs) and differentially expressed lncRNA-mRNA-protein network construction. (A) The volcano plot of all the lncRNAs. (B) The cluster heatmap of all the DELs. Red indicates a higher expression level, while blue indicates a lower expression level. The differentially expressed lncRNA-mRNA-protein network. The ellipses represent the DELs, the diamonds represent the DEGs, and the rounded rectangles represent the DEPs. Red represents the upregulation and green represents the downregulation. The dotted lines represent the translation from mRNAs to proteins. The solid lines represent the regulation works of lncRNAs for mRNAs. The red, green, and blue color of the solid lines indicates the antisense-, cis-, and trans- regulation of lncRNAs.

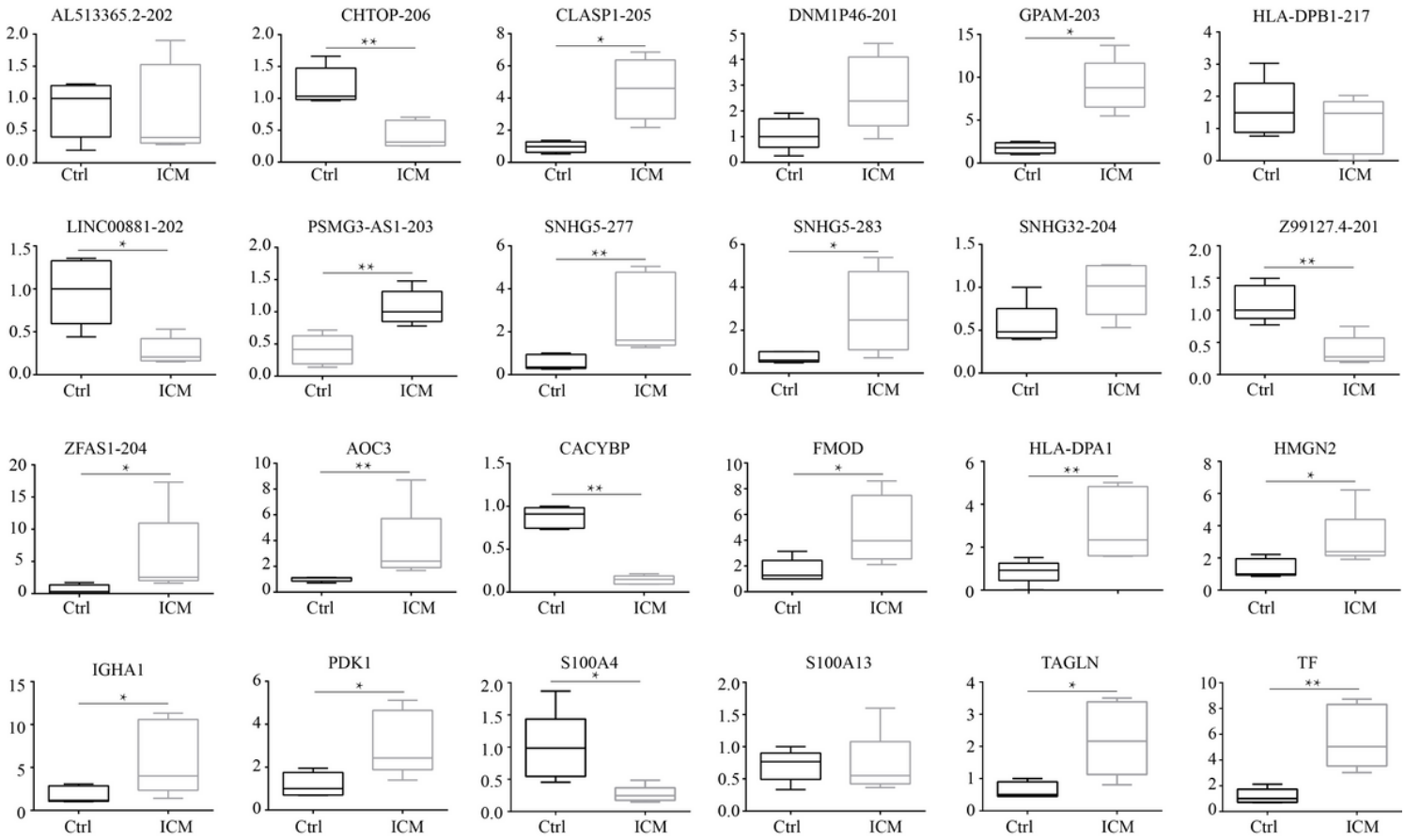

Figure 6

Validation by qRT-PCR for the differentially expressed lncRNA-mRNA-protein network. *p < 0.05, **p < 0.01

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xls
- Additionalfile2.xls