Bioinformatics Analysis and Identification of Genes and Pathways in Ischemic Cardiomyopathy

Jing Cao
Central South University First Hospital: Xiangya Hospital Central South University

Zhaoya Liu
Central South University Third Xiangya Hospital

Jie Liu
Central South University First Hospital: Xiangya Hospital Central South University

Chan Li
Central South University First Hospital: Xiangya Hospital Central South University

Guogang Zhang
Central South University Third Xiangya Hospital

Ruizheng Shi (✉ xyshiruizheng@csu.edu.cn)
Xiangya Hospital Central South University  https://orcid.org/0000-0003-1554-0846

Research Article

Keywords: ischemic cardiomyopathy, differential expressed genes, bioinformatics analysis

DOI: https://doi.org/10.21203/rs.3.rs-404922/v1

License: ©  This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background

Ischemic cardiomyopathy (ICM) is considered to be the common cause of heart failure, which has high prevalence and mortality. This study aimed to investigate the different expressed genes (DEGs) and pathways in the pathogenesis of ICM using bioinformatics analysis.

Methods

The control and ICM datasets GSE116250, GSE46224 and GSE5406 were collected from the gene expression omnibus (GEO) database. DEGs were identified using limma package of R software and co-expressed genes were identified with Venn diagrams. Then, the gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed to explored the biological functions and signaling pathways. Protein-protein interaction (PPI) networks were assembled with Cytoscape software to identify hub genes related to the pathogenesis of ICM.

Results

A total of 844 DEGs were screened from GSE116250, of which 447 up-regulated and 397 down-regulated genes respectively. A total of 99 DEGs were singled out from GSE46224, of which 58 up-regulated and 41 down-regulated genes respectively. 30 DEGs were screened from GSE5406, including 10 genes with up-regulated expression and 20 genes with down-regulated expression. 5 up-regulated and 3 down-regulated co-expressed DEGs were intersected in three datasets. GO and KEGG pathway analyses revealed that DEGs mainly enriched in collagen fibril organization, protein digestion and absorption, AGE-RAGE signaling pathway and other related pathways. Collagen alpha-1(III) chain (COL3A1), collagen alpha-2(I) chain (COL1A2) and lumican (LUM) are the three hub genes in all three datasets through PPI network analysis. The expression of 5 DEGs (SERPINA3, FCN3, COL3A1, HBB, MXRA5) in heart tissues by qRT-PCR results were consistent with our GEO analysis, while expression of 3 DEGs (ASPN, LUM, COL1A2) were opposite with GEO analysis.

Conclusions

These findings from this bioinformatics network analysis investigated key hub genes, which contributed to better understand the mechanism and new therapeutic targets of ICM.

Introduction

Heart failure (HF), a leading cause of high morbidity and mortality of cardiovascular disease and a global health care problem, has an estimated global prevalence of approximately 26 million (1, 2). Ischemic cardiomyopathy (ICM) accounts for more than 40%-60% of systolic HF, which characterized prior myocardial revascularization or myocardial infarction (MI) and severe coronary heart disease (3, 4). The severity and duration of ischemia that contributes to myocardium injury and irreversible myocardium necrosis is common in earlier phase of ICM pathogenesis, which results in cardiac fibrosis and remodeling and eventual HF (5, 6). To date, the detailed pathogenic mechanism in the progression of ICM have remained poorly understood and
efforts in drug development for ICM are currently focused on the symptomatic treatment, such as angiotensin-converting enzyme inhibitor (ACEI), β receptor blocker etc(7, 8).

The dysregulated hub genes, proteins and key pathways are representative of genomic mechanism of the process of diseases(9, 10). Therefore, the mechanism of genes and proteins correlation and key pathways in the pathogenesis of ICM are necessary to improve the prevention, diagnosis and treatment. Recently, with the continuous development of genomic technologies, microarray technology and RNA sequence have increased the ability of experts to investigate the genes regulation of cardiovascular disease(11). A few studies have reported that unique genes contribute to the pathogenesis of heart failure. A study of myocardial tissues in 8 ICM patients and 8 non-failing humans revealed 160 different expressed microRNA and 679 long non-coding RNA, which suggest an important role for expression signature of miRNA and IncRNA in the pathogenesis of ICM(12). Previous study of 196 failing and 16 non-failing human hearts demonstrated that MEF2, NKX, GATA, and NF-AT transcription factors play a crucial role in progression of HF(13). Although these databases had revealed different expressed genes (DEGs) in ICM, co-expressed DEGs by the integrated analyses of multiple datasets have not been fully investigate.

In the present study, original data from microarray analyses and RNA sequence conducted on ICM patients heart samples were downloaded from the Gene Expression Omnibus (GEO) database and comprehensive analysis was implemented.

Then, the gene ontology (GO) and Kyoto Encyclopedia of Genes /Genomes (KEGG) pathway enrichment analyses and protein-protein interactions (PPI) were performed to explore key genes and the molecular mechanism involved in the pathogenesis of clinical ICM, which may provide therapeutic targets for treatment and prevention.

Materials And Methods

Acquisition of data of gene expression profiles

The datasets of gene expression profiles with sequence number of GSE116250,GSE46224 and GSE5406 were downloaded from GEO which performed microarray or RNA sequence and matched normal tissues. Totally 14 non-failing donors and 14 ICM patients left ventricular tissues were consisted in GSE116250; in GSE46224, there were 8 non-failing control and 8 ICM patients heart tissues; in GSE5406, the number of cases from ICM patients and normal human was 108 and 16 respectively.

Screening of differential expressed genes

R language was applied to analyzed the initial data to identify DEGs. We utilized the Affy package to perform the data normalization, including conversion or raw data and background correction. Then the data were subjected to DEGs analysis using the Limma package. \( P<0.05 \) and \( |\log_2(FC)|>1 \) were set as the threshold and the genes that met the criteria were considered as DEGs. Finally, the “heatmap” package of “R” software was applied to construct heat maps, and the regions in which the differential genes were mainly concentrated were highlighted.

Gene ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis
Annotation of cellular components, biological processes and molecular functions of DEGs were determined by using Gene Ontology (GO) enrichment analysis. The route of gene cluster and related functions was determined using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. R clusterProfiler package was used to explore GO enrichment and KEGG pathway analysis with a cut-off criterion of adjusted $P < 0.05$.

Protein-protein interaction (PPI) network analysis

In our study, to determine functional interactions between DEGs, protein-protein interaction (PPI) networks were created using Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) online Tool. PPI networks were visualized by Cytoscape software. In co-expressed network, Cytohubba, a plugin in the Cytoscape software, was adopted to calculate the degree of each protein node. In our study, the top ten genes were selected as hub genes.

RNA extraction and cDNA synthesis

Total RNA was extracted from heart tissues of ICM patients and the control using RNA isolation plus reagent (Takara, Osaka, Japan). Then the concentration and quality of total RNA were examined by Thermo Fisher NanoDrop, the PrimerScript™ RT Master Mix (Takara, Osaka, Japan) was used to reversely transcript the RNA into cDNA.

Real-time quantitative PCR

All the protocols and the use of human heart tissues were approved by the Xiangya Hospital of Central South University Institutional Review Board. 4 patients diagnosed as ICM and 2 controls were enrolled in this study. The failing hearts were harvested from patients who diagnosed with ICM and accepted heart transplantation. The normal hearts were obtained from gift of hope and were used as controls. The detailed characteristics of the patients were listed in Additional file 5: Table S5. cDNA was generated with oligo(dT) primers and the Superscript first-strand synthesis kit (Invitrogen) from 1ug of total RNA as the manufacturer’s instructions. Then cDNA was diluted and used for RT-PCR using SYBR Green following the application instructions. The Primer sequences are summarized as in the Additional file 6: Table S6.

Statistical analysis

Results are expressed as Mean±SD. t-tests or nonparametric tests were used to analyze the scores of two different groups. A two-tailed value of $P < 0.05$ was considered statistically significant. Statistics were done using GraphPad Prism 8.0 and SPSS 20.0.

Results

Identification of DEGs in ICM samples compared with not-failing samples

The basic information of the datasets related to ICM was shown in Additional file 1: Table S1. A total of 844 differential genes screened out between the ICM and not-failing groups from GSE116250, of which 447 up-regulated and 397 down-regulated genes respectively ($p < 0.05$, fold change > 2). The volcano plots and heat map displayed the DEGs from GSE116250 are shown in Fig. 1a and Fig. 1d. A total of 99 DEGs screened out between the ICM and not-failing groups from GSE46224, of which 58 up-regulated and 41 down-regulated
genes respectively (p < 0.05, fold change > 2). As shown in Fig. 1b and Fig. 1e, the volcano plots and heat map displayed the DEGs from GSE46224. A total of 30 differential genes screened out between the ICM and control group from GSE5406, of which 10 up-regulated and 20 down-regulated genes respectively (p < 0.05, fold change > 2). The volcano plots and heat map displayed the DEGs from GSE5406 are shown in Fig. 1c and Fig. 1f. The top 10 DEGs of three datasets respectively are shown in Table 1. Subsequently, we mapped the Venn diagram of the differentially expressed genes in the three datasets, and identified 8 co-expressed differentially expressed genes (co-DEGs) (Fig. 2a). A total of 8 co-DEGs are shown in Table 2.

Table 1
The top 10 differential expressed genes (DEGs) in the left ventricular myocardial tissue of patients with ischemic cardiomyopathy in GSE116250, GSE46224, GSE5406.

| Gene               | Log2 FC | adjust P | Gene | Log2 FC | adjust P | Gene | Log2 FC | adjust P |
|--------------------|---------|-----------|------|---------|-----------|------|---------|-----------|
| MTATP6P1           | -16.93972 | 0.00065   | HBB  | 5.03998  | 0.00078   | ASPN | 2.44397  | 0.00000   |
| MTATP8P1           | -16.50207 | 0.00055   | NPPB | 4.67375  | 0.02280   | MYOT | -2.27140 | 0.00000   |
| GCSHP3             | 13.86860   | 0.01736   | HBA1 | 4.66989  | 0.00149   | LUM  | 2.06227  | 0.00000   |
| NDUFS1             | 11.21495   | 0.02041   | HBA2 | 4.45526  | 0.00100   | NPPA | 1.76952  | 0.00000   |
| RN7SKP232          | 10.83579   | 0.00000   | MYH6 | -2.9162  | 0.00100   | MXRA5 | 1.63951  | 0.00000   |
| TRBV5-4            | 10.08304   | 0.00000   | TUBA3E | -2.8583 | 0.00078  | HBB   | 1.50740  | 0.00004   |
| RP11-777B9.5       | -9.87759   | 0.00029   | COMP | 2.65490  | 0.01899   | FKBP5 | -1.49216 | 0.00000   |
| CUX1               | 9.40378    | 0.04378   | FMOD | 2.63632  | 0.00896   | NRAP  | -1.49200 | 0.00462   |
| ACTA1              | 7.81456    | 0.00576   | LUM  | 2.41298  | 0.00710   | HOPX  | -1.48882 | 0.00000   |
| RP5-857K21.11      | -7.33930   | 0.03388   | SFRP4 | 2.40590 | 0.01113   | RPS4Y1 | 1.42053  | 0.00121   |

Abbreviation: Log2 FC, Log2 fold change; adjust P, adjust P value.
Table 2
co-expressed differential expressed genes (co-DEGs).

| Gene  | GSE116250 Log2 FC | adjust P value | GSE46224 Log2 FC | adjust P value | GSE5406 Log2 FC | adjust P value |
|-------|------------------|---------------|------------------|---------------|------------------|---------------|
| LUM   | 2.06461371       | 0.00040205    | 2.41298228       | 0.00710086    | 2.06226548       | 3.48E-15      |
| FCN3  | -1.65780841      | 0.00115544    | -1.58296821      | 0.01053148    | -1.33651814      | 8.47E-16      |
| MXRA5 | 1.91960666       | 0.00116672    | 1.61809866       | 0.00455141    | 1.63950658       | 9.47E-09      |
| HBB   | 2.67938588       | 0.00132138    | 5.03998253       | 0.00077615    | 1.50739663       | 3.77E-05      |
| SERPINA3 | -2.15000927   | 0.00367814    | -1.79533725      | 0.02799646    | -1.29892016      | 2.03E-14      |
| ASPN  | 1.54437120       | 0.01250974    | 2.09911174       | 0.00339544    | 2.44397193       | 1.20E-14      |
| COL1A2 | 1.31286048      | 0.01716398    | 1.56899182       | 0.02778214    | 1.23113629       | 0.00150730    |
| COL3A1 | 1.38917275      | 0.03254463    | 1.52824442       | 0.03899465    | 1.22361165       | 0.002995064   |

Abbreviation: Log2 FC, Log2 fold change.

Gene Ontology function enrichment analysis

GO analyses were performed on the detected 8 co-DEGs to examine their biological functions in detail. The enrichment analysis related to biological processes were mostly enriched in collagen fibril organization ($P < 0.0001$), response to transforming growth factor beta ($P < 0.0001$), and platelet activation ($P < 0.0001$), whereas those relating to cellular components were mainly enriched in collagen-containing extracellular matrix ($P < 0.0001$), extracellular matrix ($P < 0.0001$), and the fibrillar collagen trimer ($P < 0.0001$). In addition, molecular function analysis showed that co-DEGs were involved in extracellular matrix structural constituent ($P < 0.0001$), platelet-derived growth factor binding ($P < 0.0001$), and extracellular matrix structural constituent conferring compression resistance ($P < 0.0001$) (Fig. 2b). In the next step, we conducted separate GO analyses on these three databases. The DEGs of GSE116250 datasets related to biological processes were mostly enriched in extracellular structure organization ($P < 0.0001$), extracellular matrix organization ($P < 0.0001$) and collagen fibril organization ($P < 0.0001$). Homoplastically, the enrichment analysis related to cellular components were mostly enriched in collagen-containing extracellular matrix ($P < 0.0001$), extracellular matrix ($P < 0.0001$), and basement membrane ($P < 0.0001$), whereas those relating to molecular function analysis were mostly enriched in extracellular matrix structural constituent ($P < 0.0001$), glycosaminoglycan binding ($P < 0.0001$) and heparin binding ($P < 0.0001$) (Fig. 3a). The DEGs of GSE46224 datasets related to biological processes were mostly enriched in extracellular structure organization ($P < 0.0001$), extracellular matrix organization ($P < 0.0001$), and collagen fibril organization ($P < 0.0001$). The enrichment analysis related to cellular components showed that DEGs were involved in collagen-containing extracellular matrix ($P < 0.0001$), extracellular matrix ($P < 0.0001$), and fibrillar collagen trimer ($P < 0.0001$). In addition, molecular function analysis was mostly enriched in extracellular matrix structural constituent ($P < 0.0001$), heparin binding ($P < 0.0001$) and glycosaminoglycan binding ($P < 0.0001$) (Fig. 3b). The DEGs of GSE46224 datasets with regarded to biological processes were
mostly enriched in extracellular structure organization \((P<0.0001)\), collagen fibril organization \((P<0.0001)\) and muscle system process \((P<0.0001)\), and related to cellular components were mostly enriched in extracellular matrix \((P<0.0001)\), collagen-containing extracellular matrix \((P<0.0001)\), and fibrillar collagen trimer \((P<0.0001)\). The molecular function analysis of DEGs were mainly focused on extracellular matrix structural constituent \((P<0.0001)\), extracellular matrix structural constituent conferring compression resistance \((P<0.0001)\) and platelet-derived growth factor binding \((P<0.0001)\) (Fig. 3c).

**Kyoto Encyclopedia of Genes and Genomes pathway analysis**

To get in more information about the momentous pathways of DEGs, a Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was exerted, and the top 12 significant pathways are shown in Fig. 4.

A total of ten key pathways were found through the KEGG pathway analysis of co-DEGs, which were mostly enriched in protein digestion and absorption \((P=0.0008)\), AGE-RAGE signaling pathway in diabetic complications \((P=0.0009)\) and amoebiasis \((P=0.0009)\) (Fig. 4a). Respectively, the top 12 significant pathways of GSE116250 were mainly enriched in Rheumatoid arthritis \((P<0.0001)\), Graft-versus-host disease \((P<0.0001)\) and Human infection \((P<0.0001)\) (Fig. 4b). And the top 12 significant pathways of GSE46224 were mainly enriched in Malaria \((P<0.0001)\), ECM-receptor interaction \((P<0.0001)\) and Phagosome \((P=0.0001)\). Only one pathway was enriched to the GSE5406 which was related to AGE-RAGE signaling pathway in diabetic complications \((P=0.0008)\) (Fig. 4d).

**PPI Network Analysis**

A PPI network analysis of DEGs were conducted with strings database. As shown in Fig. 5a, protein Collagen alpha-1(III) chain (COL3A1), Collagen alpha-2(I) chain (COL1A2), Lumican (LUM), Asporin (ASPN) and Matrix-remodeling-associated protein 5 (MXRA5) are interconnected among the 8 co-DEGs, while the remaining 3 proteins had no remarkable interactions. 449 nodes, 69 nodes and 16 nodes were found to be interconnected in GSE11250, GSE46224 and GSE5406 severally. Collagen alpha-1(I) chain (COL1A1), COL3A1, COL1A2, matrix metalloproteinase-2 (MMP2), biglycan (BGN), collagen alpha-1(V) chain (COL5A1), connective tissue growth factor (CTGF), LUM, fibromodulin (FMOD) and proenkephalin A (PENK) were the top 10 hub genes in the GSE116250 dataset (Fig. 5b, Additional file2: Table S2). And COL1A1, COL3A1, COL1A2, BGN, LUM, periostin (POSTN), thrombospondin-2 (THBS2), FMD0, CTGF and Collagen alpha-1(XVI) chain (COL16A1) were the top 10 hub genes in the GSE46224 dataset in which there were 7 hub genes the same as in GSE116250 (Fig. 5c, Additional file3: Table S3). In GSE5406 dataset, COL3A1, COL1A2, osteoglycin (OGN), LUM, ASPN, Filamin-C (FLNC), myotilin (MYOT), myosin-6 (MYH6), plasminogen activator inhibitor 1 (SERPINE1) and pentraxin 3 (PTX3) were the top 10 hub genes in the GSE5406dataset which have 3 hub genes the same as both GSE116250 and GSE46224 datasets (Fig. 5d, Additional file4: Table S4).

**Validation of the 8 co-DEGs in myocardial tissue of ICM patients**

8 co-DEGs validated in GSE116250, GSE46224 and GSE5406 were chosen for qRT-PCR analysis (Table 2). As shown in Fig. 6, 5 co-DEGs were down-regulated: SERPINA3 (NC 0.82 ± 0.25 vs ICM 0.05 ± 0.06, \(p<0.01\)), FCN3 (NC 1.35 ± 0.49 vs ICM 0.25 ± 0.16, \(p<0.05\)), LUM (NC 0.86 ± 0.20 vs ICM 0.31 ± 0.18, \(p<0.05\)), COL1A2 (NC 0.74 ± 0.37 vs ICM 0.30 ± 0.26, \(p=0.158\)) and ASPN (NC 0.63 ± 0.52 vs ICM 0.47 ± 0.32, \(p=0.533\)). In addition, gene HBB (NC 0.55 ± 0.64 vs ICM 9.83 ± 12.27, \(p=0.133\)) and COL3A1 (NC 0.81 ± 0.27 vs ICM 0.97 ±
were up-regulated while MXRA5 (NC 0.99 ± 0.21 vs ICM 1.01 ± 0.58, \( p = 0.800 \)) was no significant difference between the two groups.

Discussion

Ischemic cardiomyopathy (ICM) is the main cause of HF with reduced ejection fraction (EF), accounting for approximately two thirds of all cases, which is correlated with a worse prognosis(14). Although some new therapies such as stem cell transplantation and new drugs had emerged, recent years ICM remains high incidence and mortality(15). Elucidating the underlying mechanism of ICM may contribute to the early diagnosis and effective treatment HF patients. Bioinformatics analysis provided information for the identification of different genes and protein to the progress of ICM. In the present study, 8 co-DEGs were identified in three datasets (GSE116250, GSE46224 and GSE5406). Analysis of GO enrichment and KEGG pathway were performed and PPI network was constructed to further analyze molecular process underlying of ICM progression. We selected SERPINA3, FCN3, LUM, COL1A2, COL3A1, ASPN, HBB, MXRA5 to verify their expression in ICM. Expression of 5 DEGs (SERPINA3, FCN3, COL3A1, HBB, MXRA5) in heart tissues by qRT-PCR results were consistent with our GEO analysis, while expression of 3 DEGs (ASPN, LUM, COL1A2) were opposite with GEO analysis. These results may contribute to better understanding of the mechanism of ICM development.

Collagen types I and III (Col I and Col III) are the major fibrillar collagens produced fibroblasts, including cardiac fibroblasts of the heart(16). The triple-helical collagen molecules are assembled by independently coded \( \alpha_1(I) \), \( \alpha_2(I) \), and \( \alpha_1(III) \) chains(17). Previous studies demonstrated that Col I and Col III are involved in extracellular matrix protein (ECM) during the progression of cardiac remodeling in the failing heart(18, 19). Several lines of evidence indicate that expression of COL3A1 and COL1A2 coding for the two collagens mainly increased in cardiovascular diseases, such as ICM and dilated cardiomyopathy(20–22). In our study, our results showed that COL3A1 was up-regulated in heart tissues of ICM compared with normal tissues, which showed correlation with cardiac remodeling with ICM.

Lumican (LUM) and Asporin (ASPN) belong to the small leucine-rich proteoglycans (SLRP) family consisting of a core protein with leucine-rich repeats and one or more attached glycosaminoglycan chains, which are considered to control the assembly of collagen fibers in the ECM(23–25). LUM is an ECM localized proteoglycan associated with inflammatory conditions and known to bind collagens(26, 27). Previous study has indicated that LUM protein and mRNA levels are increased in cardiac tissues of patients and mice with HF compared with control hearts(28). Moreover, moderate lack of LUM attenuated cardiac fibrosis and improved diastolic function following pressure overload in mice(29). ASPN, firstly identified as an extracellular secreted protein in 2001, is distinct from other class 1 SLRP family members because of its unique aspartate residues named the D-repeat(30). It has been demonstrated that ASPN in the myocardium of the db/db mice was significantly higher than that in the control mice(31). Thus, ASPN could play an important role in pathological cardiac remodeling. The findings suggest that LUM and ASPN may contribute to fibrinogenesis of cardiac remodeling of ICM.

MXRA5, a member of the MXRA protein family, plays an important role in cell adhesion and ECM remodeling(32). The protein is found in humans that is encoded by the MXRA5 gene, which contains 7 leucine-rich repeats and 13 immunoglobulin-like C2-type domains related to perlecan with a molecular weight of 312KD
and is reported to express in primates but not in rat or mouse\(33, 34\). MXRA5 mRNA is increased in human chronic ischemic myocardium and in primary fibroblast isolated from aging patients. Our results further confirmed that MXRA5 mRNA levels is enhanced in heart tissues of ICM patients, although there is no statistical significance compared with normal hearts. Recently, MXRA5 is found to be closely associated with the mitogen-activated protein kinases (MAPKs) pathways, which plays a key role in the regulation of cell proliferation, survival, differentiation and apoptosis\(35\). However, the role of MXRA5 in the progression of ICM remains unclear.

As with most bioinformatic network analysis studies of human diseases, this study had several limitations. The decreased expression of COL1A2, LUM and ASPN and no statistical difference of up-regulated expression of COL3A1 and MXRA5 were not consistent with the GEO analysis, which may due to the limited number of heart tissues \(n = 4 \text{ ICM, } n = 2 \text{ normal hearts}\) for qRT-PCR confirmation and demands further investigation to increase the number of each group to confirm these results. In addition, the sample of GSE46224 was relatively small and the sampling method did not eliminate the effect of gender and medications that may alter the gene expression. Although we detected the hub genes expression in cardiac tissues, the related pathways and mechanisms were not confirmed in vitro studies or other functional studies which need further investigation.

**Conclusion**

The present study was intended to identify DEGs and pathways with comprehensive bioinformatic analysis to find the potential biomarkers and predict the progression of ICM. We further confirmed COL3A1, COL1A2, LUM, ASPN and MXRA5 might be exploited as diagnostic and prognostic indicators for ICM. Finally, AGE-RAGE signaling pathway, ECM-receptor interaction and protein digestion and absorption may be essential signaling pathways in the development of ICM. Therefore, our findings provided a potential basis to understand the cause and mechanism of ICM and therapeutic targets for clinic treatment.

**Abbreviations**

ICM: ischemic cardiomyopathy

DEGs: different expressed genes

GEO: gene expression omnibus

GO: gene otology

KEGG: Kyoto Encyclopedia of Genes and Genomes

PPI: Protein-protein interaction

HF: Heart failure

MI: myocardial infarction

COL3A1: protein Collagen alpha-1(III) chain
COL1A2: Collagen alpha-2(I) chain

Declarations

Author contribution

JC, ZY-L and R-ZS conceived and designed the study. JC, ZY-L, JL performed the majority the work described here. JC and CL performed the bioinformatics analysis JCâ–ZYL and JL performed the experimental validation. RZS and GGZ revised the manuscript. All the authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Funding

The work was supported by grants from the National Program on Key Basic Research Project of China (Grant No. 2019YFF0216304 to S.R.Z.), Outstanding Youth Foundation Project of Hunan Natural Science Foundation (Grant No. 2019JJ20036 to S.R.Z.), Chinese Cardiovascular Association V.G foundation (Grant No.2017-CCA-VG-005 to S.R.Z), National Nature Science Foundation of China (Grant No. 81873479 to Z.G.G), National Nature Science Foundation of China (Grant No. 82000301 to L.Z.Y).

Availability of data and materials

All relevant data and materials during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

This study was reviewed and approved by the Committee at Xiangya hospital of Central South University (IRB No. 2017121009). All participants provided written informed consent prior to research participation.

Competing interest

The authors declare that they have no competing interests.

References

1. Roger VL. Epidemiology of heart failure. Circ Res. 2013;113(6):646-59.
2. Malagoli A, Rossi L, Bursi F, Zanni A, Sticozzi C, Piepoli MF, et al. Left Atrial Function Predicts Cardiovascular Events in Patients With Chronic Heart Failure With Reduced Ejection Fraction. J Am Soc Echocardiogr. 2019;32(2):248-56.
3. Schuster A, Morton G, Chiribiri A, Perera D, Vanoverschelde JL, Nagel E. Imaging in the management of ischemic cardiomyopathy: special focus on magnetic resonance. J Am Coll Cardiol. 2012;59(4):359-70.
4. Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. Nat Rev Cardiol. 2011;8(1):30-41.
5. Nayak A, Liu C, Mehta A, Ko YA, Tahhan AS, Dhindsa DS, et al. N8-Acetylspermidine: A Polyamine Biomarker in Ischemic Cardiomyopathy With Reduced Ejection Fraction. J Am Heart Assoc. 2020;9(11):e016055.

6. Shinde AV, Frangogiannis NG. Fibroblasts in myocardial infarction: a role in inflammation and repair. J Mol Cell Cardiol. 2014;70:74-82.

7. Kong P, Christia P, Frangogiannis NG. The pathogenesis of cardiac fibrosis. Cell Mol Life Sci. 2014;71(4):549-74.

8. Wu L, Li F, Zhao C, Ming Y, Zheng C, Li Y, et al. Effects and Mechanisms of Traditional Chinese Herbal Medicine in the Treatment of Ischemic Cardiomyopathy. Pharmacol Res. 2020;151:104488.

9. Xiong Y, Mi BB, Liu MF, Xue H, Wu QP, Liu GH. Bioinformatics Analysis and Identification of Genes and Molecular Pathways Involved in Synovial Inflammation in Rheumatoid Arthritis. Med Sci Monit. 2019;25:2246-56.

10. Franco D, Bonet F, Hernandez-Torres F, Lozano-Velasco E, Esteban FJ, Aranega AE. Analysis of microRNA Microarrays in Cardiogenesis. Methods Mol Biol. 2016;1375:207-21.

11. Kang S, Song J. Robust gene selection methods using weighting schemes for microarray data analysis. BMC Bioinformatics. 2017;18(1):389.

12. Yang KC, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, et al. Deep RNA sequencing reveals dynamic regulation of myocardial noncoding RNAs in failing human heart and remodeling with mechanical circulatory support. Circulation. 2014;129(9):1009-21.

13. Hannenhalli S, Putt ME, Gilmore JM, Wang J, Parmacek MS, Epstein JA, et al. Transcriptional genomics associates FOX transcription factors with human heart failure. Circulation. 2006;114(12):1269-76.

14. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, et al. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. Circulation. 2017;135(10):e146-e603.

15. Florea V, Rieger AC, DiFede DL, El-Khorazaty J, Natsumeda M, Banerjee MN, et al. Dose Comparison Study of Allogeneic Mesenchymal Stem Cells in Patients With Ischemic Cardiomyopathy (The TRIDENT Study). Circ Res. 2017;121(11):1279-90.

16. Namba T, Tsutsui H, Tagawa H, Takahashi M, Saito K, Kozai T, et al. Regulation of fibrillar collagen gene expression and protein accumulation in volume-overloaded cardiac hypertrophy. Circulation. 1997;95(10):2448-54.

17. Ghosh AK. Factors involved in the regulation of type I collagen gene expression: implication in fibrosis. Exp Biol Med (Maywood). 2002;227(5):301-14.

18. Myllyharju J, Kivirikko KI. Collagens and collagen-related diseases. Ann Med. 2001;33(1):7-21.

19. Talman V, Ruskoaho H. Cardiac fibrosis in myocardial infarction-from repair and remodeling to regeneration. Cell Tissue Res. 2016;365(3):563-81.

20. Frangogiannis NG. The Extracellular Matrix in Ischemic and Nonischemic Heart Failure. Circ Res. 2019;125(1):117-46.

21. Echegaray K, Andreu I, Lazcano A, Villanueva I, Saenz A, Elizalde MR, et al. Role of Myocardial Collagen in Severe Aortic Stenosis With Preserved Ejection Fraction and Symptoms of Heart Failure. Rev Esp Cardiol (Engl Ed). 2017;70(10):832-40.
22. Yamamoto K, Masuyama T, Sakata Y, Nishikawa N, Mano T, Yoshida J, et al. Myocardial stiffness is determined by ventricular fibrosis, but not by compensatory or excessive hypertrophy in hypertensive heart. Cardiovasc Res. 2002;55(1):76-82.

23. Chen X, Li X, Hu X, Jiang F, Shen Y, Xu R, et al. LUM Expression and Its Prognostic Significance in Gastric Cancer. Front Oncol. 2020;10:605.

24. Brezillon S, Untereiner V, Mohamed HT, Ahallal E, Proult I, Nizet P, et al. Label-Free Infrared Spectral Histology of Skin Tissue Part II: Impact of a Lumican-Derived Peptide on Melanoma Growth. Front Cell Dev Biol. 2020;8:377.

25. Zhang Z, Li H, Zhao Y, Guo Q, Yu Y, Zhu S, et al. Asporin promotes cell proliferation via interacting with PSMD2 in gastric cancer. Front Biosci (Landmark Ed). 2019;24:1178-89.

26. Mohammadzadeh N, Melleby AO, Palmero S, Sjaastad I, Chakravarti S, Engebretsen KVT, et al. Moderate Loss of the Extracellular Matrix Proteoglycan Lumican Attenuates Cardiac Fibrosis in Mice Subjected to Pressure Overload. Cardiology. 2020;145(3):187-98.

27. Dupuis LE, Berger MG, Feldman S, Doucette L, Fowlkes V, Chakravarti S, et al. Lumican deficiency results in cardiomyocyte hypertrophy with altered collagen assembly. J Mol Cell Cardiol. 2015;84:70-80.

28. Mohammadzadeh N, Lunde IG, Andenaes K, Strand ME, Aronsen JM, Skrbic B, et al. The extracellular matrix proteoglycan lumican improves survival and counteracts cardiac dilatation and failure in mice subjected to pressure overload. Sci Rep. 2019;9(1):9206.

29. Chen SW, Tung YC, Jung SM, Chu Y, Lin PJ, Kao WW, et al. Lumican-null mice are susceptible to aging and isoproterenol-induced myocardial fibrosis. Biochem Biophys Res Commun. 2017;482(4):1304-11.

30. Li H, Zhang Z, Chen L, Sun X, Zhao Y, Guo Q, et al. Cytoplasmic Asporin promotes cell migration by regulating TGF-beta/Smad2/3 pathway and indicates a poor prognosis in colorectal cancer. Cell Death Dis. 2019;10(2):109.

31. Li YL, Yu F, Li BY, Fu CL, Yu X, Xu M, et al. The protective effects of grape seed procyanidin B2 against asporin mediates glycated low-density lipoprotein induced-cardiomyocyte apoptosis and fibrosis. Cell Biol Int. 2019.

32. Xiao H, Jiang Y, He W, Xu D, Chen P, Liu D, et al. Identification and functional activity of matrix-remodeling associated 5 (MXRA5) in benign hyperplastic prostate. Aging (Albany NY). 2020;12(9):8605-21.

33. Ding L, Li S, Zhang Y, Gai J, Kou J. MXRA5 is decreased in preeclampsia and affects trophoblast cell invasion through the MAPK pathway. Mol Cell Endocrinol. 2018;461:248-55.

34. Xiong D, Li G, Li K, Xu Q, Pan Z, Ding F, et al. Exome sequencing identifies MXRA5 as a novel cancer gene frequently mutated in non-small cell lung carcinoma from Chinese patients. Carcinogenesis. 2012;33(9):1797-805.

35. Gabrielsen A, Lawler PR, Yongzhong W, Steinbruchel D, Blagoja D, Paulsson-Berne G, et al. Gene expression signals involved in ischemic injury, extracellular matrix composition and fibrosis defined by global mRNA profiling of the human left ventricular myocardium. J Mol Cell Cardiol. 2007;42(4):870-83.

Figures
Figure 1

Volcano plot and heat map of all the differentially expressed genes (DEGs) between Non-failure patients and heart failure patients in cardiac tissues. A, volcano plot of DEGs in GSE116250 dataset; B, volcano plot of DEGs in GSE46224 dataset; C, volcano plot of DEGs in GSE5406 dataset; Green dot represent DEGs with fold change <1, red dot represents DEGs with fold change >1, adjust P <0.05. D, heat map of DEGs in GSE116250 dataset; E, heat map of DEGs in GSE46224 dataset; F, heat map of DEGs in GSE5406 dataset; The gradient colour change from blue to red represents the changing process from downregulation to upregulation.
Figure 2

Functional enrichment analysis of common differentially expressed genes (co-DEGs). A, Venn diagram of co-DEGs; B, Gene Ontology (GO) analysis of co-DEGs regard to biological process (BP), cellular component (CC), and molecular function (MF).
Figure 3

Gene Ontology (GO) analysis of all DEGs. A, the top 10 enriched GO terms in biological process (BP), cellular component (CC), and molecular function (MF) of GSE116250 dataset; B, the top 10 enriched GO terms in biological process (BP), cellular component (CC), and molecular function (MF) of GSE462224 dataset; C, the top 10 enriched GO terms in biological process (BP), cellular component (CC), and molecular function (MF) of GSE5406 dataset.
Figure 4

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis of all DEGs. A, the enriched KEGG pathways in co-DEGs; B, the enriched KEGG pathways of DEGs in GSE116250 dataset; C, the enriched KEGG pathways of DEG in GSE46224 dataset; D, the enriched KEGG pathways in GSE5406 dataset.
**Figure 5**

Protein–protein interaction (PPI) network analysis. A, PPI network of co-DEGs identified by Cytoscape; B and C, PPI network and hub gene of DEGs in GSE116250; D and E, PPI network and hub gene of DEGs in GSE46224; G and H, PPI network and hub gene of DEGs in GSE5406.
Figure 6

Relative gene expression verification of 8 co-DEGs. The expression of SERPINA3 (A), FCN3 (B), LUM (C), COL1A2 (D), ASPN (E), HBB (F), COL3A1 (G), MXRA5 (H) were measured by real time-qPCR. *P < 0.05; **P < 0.01.

Supplementary Files

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

- Additionalfile1TableS1.docx
- Additionalfile2TableS2.docx
- Additionalfile3TableS3.docx
- Additionalfile4TableS4.docx
- Additionalfile5tableS5.docx
- Additionalfile6tableS6.docx