

RESEARCH

Open Access



# Integrative analyses of biomarkers and pathways for heart failure

Shaowei Fan and Yuanhui Hu\*

## Abstract

**Background:** Heart failure (HF) is the most common potential cause of death, causing a huge health and economic burden all over the world. So far, some impressive progress has been made in the study of pathogenesis. However, the underlying molecular mechanisms leading to this disease remain to be fully elucidated.

**Methods:** The microarray data sets of GSE76701, GSE21610 and GSE8331 were retrieved from the gene expression comprehensive database (GEO). After merging all microarray data and adjusting batch effects, differentially expressed genes (DEG) were determined. Functional enrichment analysis was performed based on Gene Ontology (GO) resources, Kyoto Encyclopedia of Genes and Genomes (KEGG) resources, gene set enrichment analysis (GSEA), response pathway database and Disease Ontology (DO). Protein protein interaction (PPI) network was constructed using string database. Combined with the above important bioinformatics information, the potential key genes were selected. The comparative toxicological genomics database (CTD) is used to explore the interaction between potential key genes and HF.

**Results:** We identified 38 patients with heart failure and 16 normal controls. There were 315 DEGs among HF samples, including 278 up-regulated genes and 37 down-regulated genes. Pathway enrichment analysis showed that most DEGs were significantly enriched in BMP signal pathway, transmembrane receptor protein serine/threonine kinase signal pathway, extracellular matrix, basement membrane, glycosaminoglycan binding, sulfur compound binding and so on. Similarly, GSEA enrichment analysis showed that DEGs were mainly enriched in extracellular matrix and extracellular matrix related proteins. BBS9, CHRDL1, BMP4, MYH6, NPPA and CCL5 are central genes in PPI networks and modules.

**Conclusions:** The enrichment pathway of DEGs and GO may reveal the molecular mechanism of HF. Among them, target genes EIF1AY, RPS4Y1, USP9Y, KDM5D, DDX3Y, NPPA, HBB, TSIX, LOC28556 and XIST are expected to become new targets for heart failure. Our findings provide potential biomarkers or therapeutic targets for the further study of heart failure and contribute to the development of advanced prediction, diagnosis and treatment strategies.

**Keywords:** Potential key genes, Heart failure, Bioinformatics, Gene expression synthesis, Biomarkers

## Background

Cardiovascular disease is one of the main causes of human death, including coronary heart disease, hypertension, congenital heart disease, heart failure and other heart related diseases, as well as systemic vascular system related diseases such as atherosclerosis and lower extremity deep venous thrombosis. Among them, heart failure is the most common cardiovascular disease in clinic. It is estimated that about 64.3 million people worldwide

\*Correspondence: huiyuhui55@sohu.com  
Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

suffer from heart failure [1]. As everyone knows, heart failure often follows a variety of other diseases, such as coronary heart disease, hypertension, diabetes, etc. these diseases are characterized by an obvious age-dependent dependency. That is, the higher the risk factor for age, the greater the risk of diseases. However, research shows that the burden of heart failure in young people may be increasing, which means that more young people will enter the ranks of heart failure, and its prevalence tends to be younger [2].

In terms of the mechanism of heart failure, myocardial hypertrophy is the key to the occurrence and development of heart failure, but the exact reason for the transformation of myocardial hypertrophy into heart failure is not clear. In this process, it may play a role in the pathological increase of the circulating level of vasoactive substances (such as angiotensin II, catecholamine, endothelin, etc.), and the increase of the content of vasoactive substances will stimulate the myocardium and eventually lead to myocardial hypertrophy by stimulating their respective signal transduction pathways [3]. When hypertrophic cardiomyocytes are exposed to the environment with increased content of vasoactive substances for a long time, it will lead to subcellular defects, protease activation, metabolic disorders, abnormal calcium regulation, etc., which will increase the dysfunction of heart function, thus aggravating the occurrence and development process of heart failure [4]. Therefore, the cytokines and other molecular mechanisms involved in the occurrence and development of heart failure still need to be further studied, and its specific mechanism needs to be further proved at the molecular level [5].

Nowadays, the treatment cost of patients with heart failure has always been a large expenditure for any country. Therefore, early diagnosis of heart failure is of great significance for early intervention and treatment, which is equally important for individuals and countries. At present, the diagnosis of heart failure is mainly achieved by the results of BNP, NT-proBNP, echocardiography and other clinical symptoms, such as fatigue, dyspnea, low body position fluid retention. BNP and NT-proBNP are gold standard biomarkers for the diagnosis and prognosis of heart failure. Therefore, obtaining objective, accurate, reliable, noninvasive and biologically meaningful biomarkers of heart failure will greatly optimize the diagnosis, monitoring, treatment and prognosis of heart failure, which will also become the research focus for a long time in the future.

With the rapid development of high-throughput technology, various studies related to the pathophysiological process of heart failure continue to deepen, and more and more new biomarkers have been found, such as middle regional preatrial natriuretic peptide (MR-proANP),

middle regional adrenomedullin (MR-proADM), highly sensitive troponin, soluble ST2 (sST2) [6], growth differentiation factor (GDF)-15, Galectin-3 [7, 8], copeptin [8, 9], Cystatin C (Cys-C) [10] and Sirtuin (SIRT) [11], which have partially shown the potential to determine the diagnosis and prognosis of heart failure, but they are still insufficient in clinical evidence. In neurohumoral activities, pathological changes have taken place in the expression level of relevant biomarkers, signal molecules, cytokines and other substances often in the early stage of symptoms or even when there are no symptoms. Therefore, the cytokines and other molecular mechanisms involved in the occurrence and development of heart failure need to be further studied, in order to explore more biomarkers related to heart failure and improve their relevant clinical evidence, so as to improve the early diagnosis and prognosis management of heart failure and bring well-being to patients.

Heart failure is a major health problem in the world. It is necessary to explore the potential biomarkers and molecular mechanisms related to the occurrence and development of heart failure, so as to provide more targeted and effective treatment strategies for patients and bring benefits. At present, high-throughput technology has developed rapidly and is widely used in various fields. Integrated bioinformatics analysis is expected to become a key technology to clarify the etiology, pathogenesis and treatment of heart failure, which will benefit mankind in terms of human, material and financial resources.

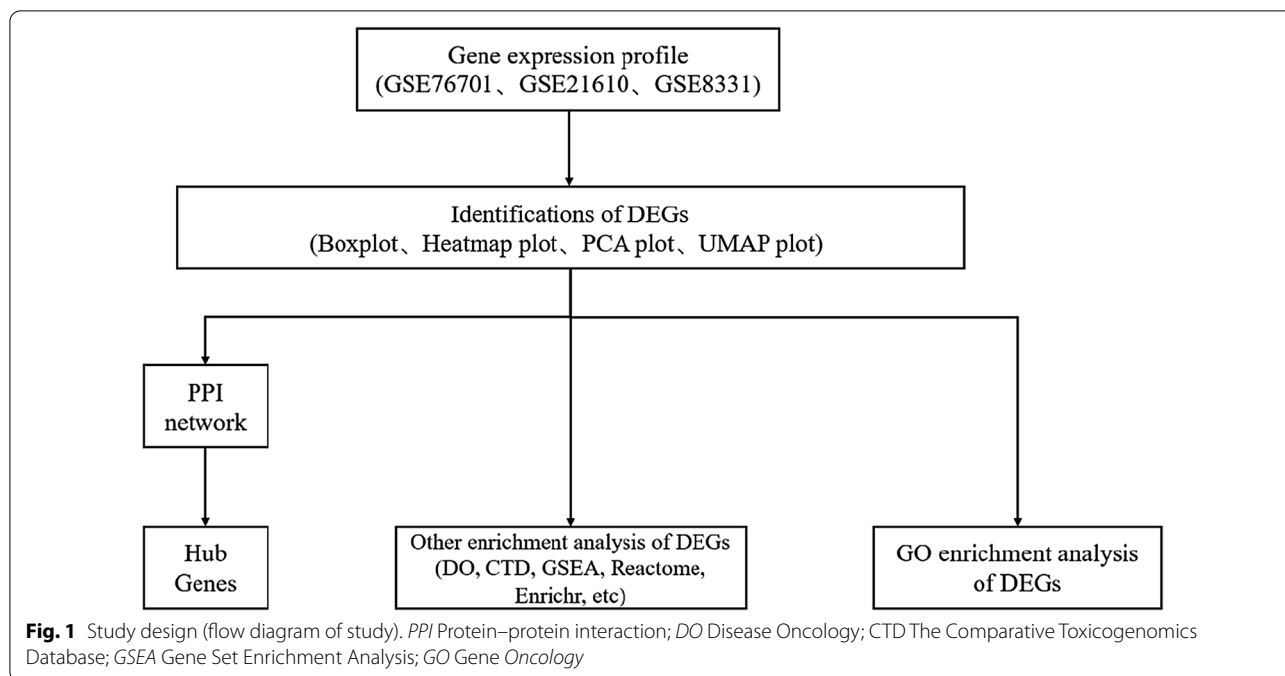
In this study, we analyzed three mRNA expression profiles (GSE76701, GSE21610 and GSE8331) of the same platform (GPL570) downloaded from GEO Database (<https://www.ncbi.nlm.nih.gov/geo/>) to determine the possible DEGs in the occurrence and development of heart failure, and analyzed their expression, function and interaction, so as to provide reference for exploring biomarkers or therapeutic targets of heart failure.

## Methods

Figure 1 clearly illustrates the flow chart of materials and methods. In this study, we integrated three datasets from GEO Database: GSE76701, GSE21610 and GSE8331. The characteristics of differential genes were displayed by Box plot, Heatmap plot, PCA plot and UMAP plot. Then, the construction of PPI network, the display of Hub gene, GO enrichment analysis and other enrichment analysis including DO, CTD, GSEA, Reactome and Enrichr were listed.

### Heart failure dataset

Download the heart failure dataset original files of three registered microarray datasets from NCBI GEO Database, including GSE76701, GSE21610 and



GSE8331 (Table 1). All these datasets are from the microarray platform of Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133\_Plus\_2]. In each dataset, human myocardial samples were selected only from HF and normal EF subjects, and finally 38 HF and 16 normal EF group samples were included for subsequent analysis.

**Data preprocessing**

The GSE76701 dataset contains 4 human HF samples and 4 human health samples, the GSE21610 dataset contains 30 human HF samples and 8 human health samples, and the GSE8331 dataset contains 4 human HF samples and 4 human health samples. A series of matrix text files for the dataset have been obtained. Subsequently, the limma R software package was used for background correction, quartile standardization and probe summary [12–15]. See Table 1 for details.

**Identification of DEGs between HF and healthy samples**

In this study, we used the limma R package (version 3.6.3; <https://www.r-project.org/>) [12]. The DEGs between HF samples and healthy samples were determined by the threshold standard of  $|\log_2(FC)| > 1$  and  $p_{adj} < 0.05$ .

**Function and pathway enrichment analysis of DEGs**

Gene Ontology (GO) resources (<http://geneontology.org/>, Accessed 20 Oct 2021) is a bioinformatics tool that provides a framework and a set of concepts to describe the function of all biological gene products [16]. Kyoto Encyclopedia of Genes and Genomes (KEGG) (<https://www.kegg.jp/>) is a database resource integrated the information of genomes, biological pathways, diseases and chemicals [17–19]. Reactome pathway database (<https://reactome.org/>) is a path annotation database that collects human biological paths and processes [20]. Enrichr (<https://maayanlab.cloud/Enrichr/>, Accessed 20 Oct 2021) is an online data processor [21–23]. Disease Ontology (DO) (<http://disease-ontology.org>, Accessed 20 Oct 2021), represents a

**Table 1** Characteristics of datasets in this study

GSE series	Platform	Total	HF	NG	Organism	Country	Contributors
GSE76701	GPL570	8	4	4	Myocardium	America	Kim EH, et al
GSE21610	GPL570	38	30	8	Myocardium	Germany	Schwientek P, et al
GSE8331	GPL570	8	4	4	Myocardium	Japan	Mano H

GSE, Gene Expression Omnibus; HF, Heart Failure; NG, Normal Group

comprehensive knowledge base of 8043 genetic, developmental and acquired human diseases [24]. Gene Set Enrichment Analysis (GSEA) (<http://software.broadinstitute.org/gsea/index.jsp>, Accessed 20 Oct 2021) is a computational method for interpreting gene expression data based on molecular signature database [25, 26]. Before enrichment analysis, the gene symbol code is converted into Entrez ID using the human genome annotation package "org. HS. eg.db". In order to better understand the biological functions and characteristics, the enrichment analysis is carried out with R software, using "Clusterprofiler" KEGG and GO enrichment analysis package, reactor pathway analysis "Reactomepa" package and do enrichment analysis "DOSE" package. The "GOplot" and "ggplot2" packages of R software are used for visual mapping. The relevant go biological function map is considered to be significantly rich if it meets the  $p$ . adjust value  $< 0.05$  and  $Q$  value  $< 0.05$ . For the important paths related to heart failure, the significance level, nominal  $p$  value and false discovery rate The cutoff value of (FDR)  $Q$  value is 0.05. For GSEA enrichment analysis, FDR  $Q$  value  $< 0.25$  and  $p$ .adjust value  $< 0.05$  are used as screening indexes.

#### Protein–protein interaction (PPI) network and potential key gene analysis

String database (<http://string-db.org/>, Accessed 20 Oct 2021) is used to construct PPI network to reveal the general organization principle of functional cell system and predict protein–protein interaction [27]. Through molecular complex detection (MCODE) of Cytoscape, the results of PPI network are modularized analyzed and visualized. Default parameters (degree cutoff)  $\geq 2$ , node score cut-off  $\geq 2$ , K nucleus  $\geq 2$ , maximum depth = 100). In order to select potential key genes, we synthesized the above important bioinformatics information for subsequent analysis. If DEG meets the inclusion criteria (adjusted  $p$  value  $< 0.05$  and  $|\log_2 FC| \geq 1$ ), it is considered as potential key genes. In addition, genes with connectivity greater than 5 in PPI network are also included.

#### Identification of potential key genes associated with heart failure

The Comparative Toxicogenomics Database (CTD, <http://ctdbase.org/>, Accessed 20 Oct 2021) synthesize information, including chemical gene/protein interactions, chemical diseases and gene disease relationships, to develop hypotheses related to disease mechanisms [28]. Use data in CTD to analyze the association between potential key genes and the risk of heart failure, atrial fibrillation, hypertension and sudden cardiac death.

## Results

### Dataset evaluation

The gene expression level of the combined GEO series with adjusted batch effect is standardized, and the results before and after standardization are shown in Fig. 2. Probes corresponding to 21,655 genes in GSE76701, GSE21610 and GSE8331 datasets were identified, and DEGs of heart failure were confirmed. The total number of filtered molecules was 21,655, of which 85 IDs met the threshold of  $|\log_2 (FC)| \geq 1$  &  $p$ .adj  $< 0.05$ . Under this threshold, 60 were highly expressed in HF group and 25 in normal group; 22 IDs met the threshold of  $|\log_2 (FC)| \geq 1.5$  &  $p$ .adj  $< 0.05$ . Under this threshold, 16 IDs were highly expressed in HF group and 6 IDs were highly expressed in normal group; There are 10 IDs that meet the threshold of  $|\log_2 (FC)| \geq 2$  &  $p$ .adj  $< 0.05$ . Under this threshold, there are 7 highly expressed IDs in HF group (EIF1AY, RPS4Y1, USP9Y, KDM5D, DDX3Y, NPPA and HBB) and 3 highly expressed IDs in normal group (TSIX, LOC28556 and XIST). See Table 2 for details. Figure 3 shows the Heatmap plot, Volcano plot, PAC plot and UMAP plot.

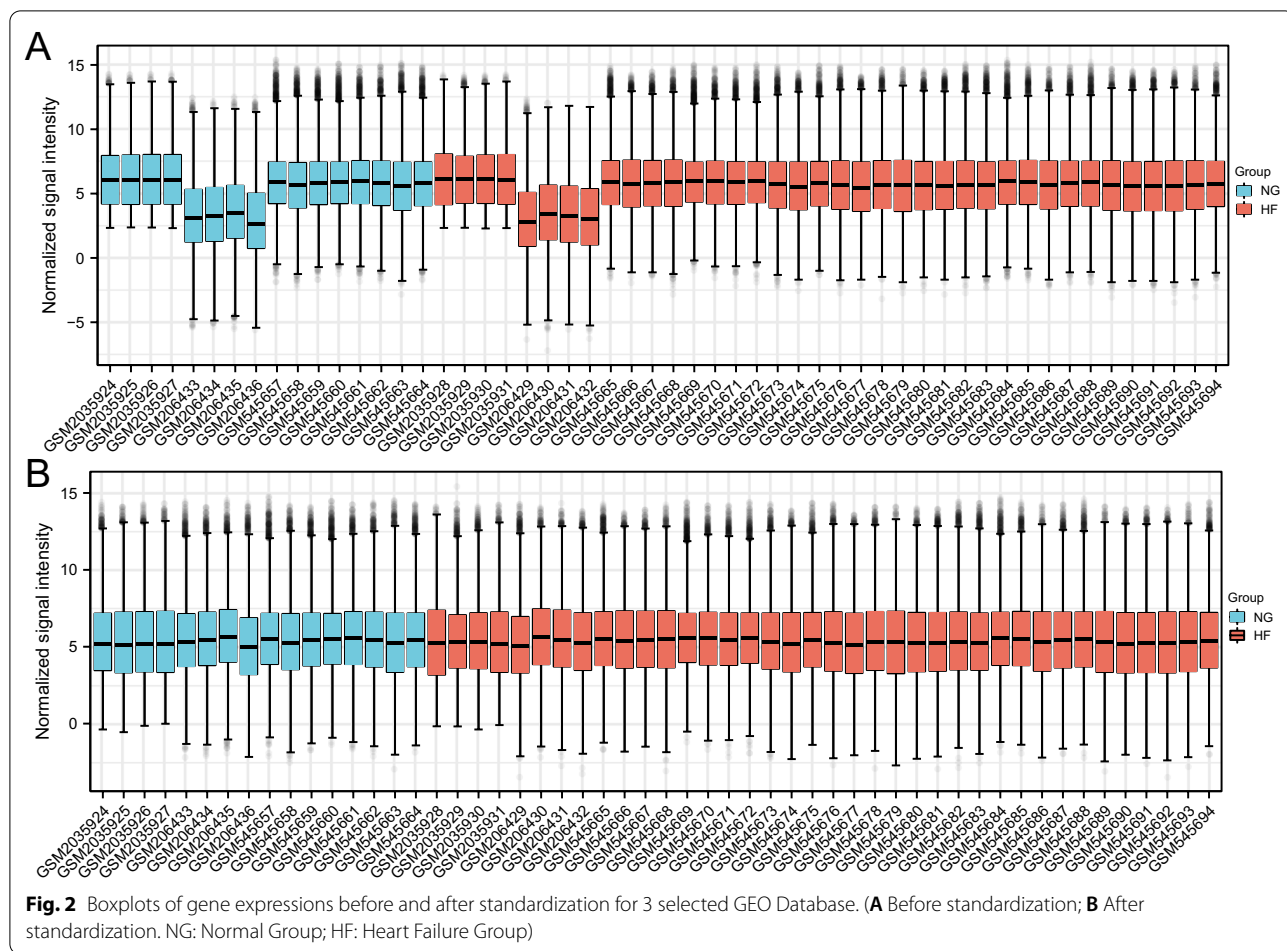
### Functional enrichment analysis of DEGs

#### GO enrichment analysis

In order to further study the biological functions of 10 DEGs, functional enrichment analysis was carried out, and the results are shown in Table 3. GO enrichment analysis showed that the functions of differentially expressed genes were mainly concentrated in the following 11 aspects: GO: 0030509 ~ BMP signaling pathway; GO: 0071772 ~ response to BMP; GO: 0071773 ~ cellular response to BMP stimulus; GO: 0003012 ~ muscle system process; GO: 0007178 ~ transmembrane receptor protein serine/threonine kinase signaling pathway; GO: 0062023 ~ collagen containing extracellular matrix; GO: 0005604 ~ basement membrane; GO: 0005614 ~ interstitial matrix; GO: 0008201 ~ heparin binding; GO: 0005539 ~ glycosaminoglycan binding; GO: 1901681 ~ sulfur compound binding. For the above GO meeting the requirements, the R language GOplot and ggplot 2 package are used for visual presentation (Fig. 4). According to the adjusted screening criteria of  $P$  value  $< 0.05$  and  $Q$  value  $< 0.05$ , there was no enrichment pathway in KEGG.

#### GSEA enrichment analysis

GSEA was used to test the combined GEO dataset to identify functional gene sets associated with heart failure. Finally, five groups of HF related expressions were determined (Table 4), of which only three data sets met FDR  $Q$  value  $< 0.25$  and  $p$ .adjust value  $< 0.05$ . These data sets include: (1) involved in encoding core extracellular



**Fig. 2** Boxplots of gene expressions before and after standardization for 3 selected GEO Database. (A Before standardization; B After standardization. NG: Normal Group; HF: Heart Failure Group)

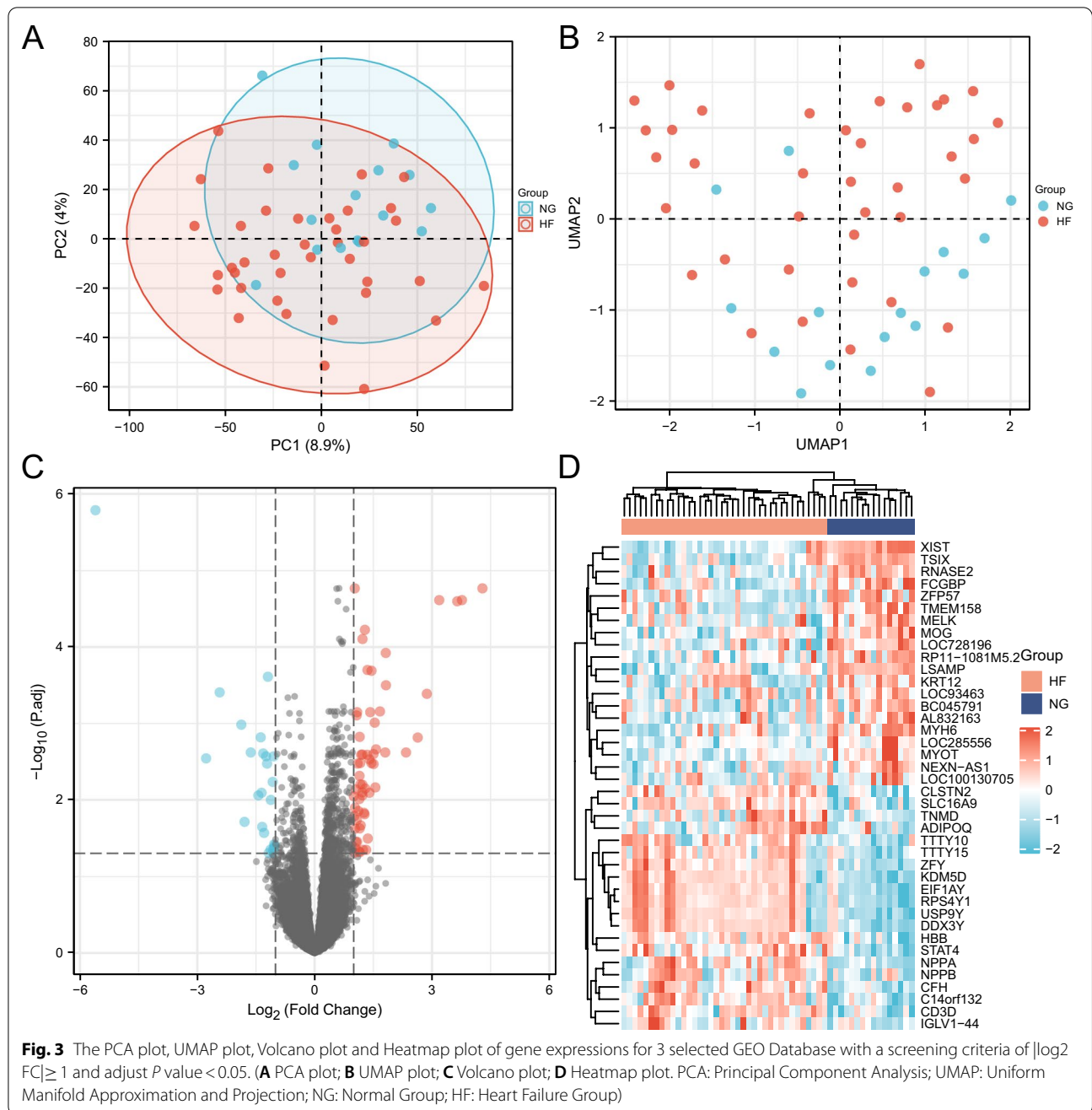
**Table 2** The DEGs of merged data set with the use of criteria of adjust *P* value < 0.05 and  $|\logFC| \geq 2$

Gene	LogFC	AveExpr	t	P value	Adjust P value	B
XIST	-5.60291	6.415566	-8.05301	7.55E-11	1.64E-06	14.17664
EIF1AY	4.290193	7.905949	7.068635	3.03E-09	1.71E-05	10.80004
RPS4Y1	3.765519	8.68866	6.835021	7.29E-09	2.43E-05	9.993851
USP9Y	3.646357	5.569263	6.73868	1.05E-08	2.52E-05	9.661358
KDM5D	3.185601	6.035742	6.814739	7.87E-09	2.43E-05	9.923851
DDX3Y	2.869435	6.737341	5.715082	4.73E-07	0.000409	6.15164
LOC285556	-2.77444	4.659843	-4.77551	1.39E-05	0.002863	3.038569
NPPA	2.635568	10.70313	5.101512	4.38E-06	0.00153	4.099873
TSIX	-2.42784	4.207833	-5.73785	4.35E-07	0.000392	6.228833
HBB	2.336796	10.47484	4.92105	8.32E-06	0.002397	3.509395

LogFC, log Fold Change; AveExpr, Average Expression

matrix (including ECM glycoprotein, collagen and proteoglycan), (2) involved in encoding structural ECM glycoprotein, and (3) involved in encoding extracellular matrix and extracellular matrix related proteins (Fig. 3). According to the visualization results of

GSEA, the eligible GSEA gene sets are as follows: (1) in NABA\_ CORE\_ Matrix gene set was significantly enriched (NES = 2.199; *p.adjust* = 0.037; FDR = 0.035); (2) At NABA\_ ECM\_ The glycoproteins gene set was significantly enriched (NES = 2.050; *p.adjust* = 0.037; FDR = 0.035) (Fig. 5).



**Enrichment analysis by enrichr and reactome**

Using Enrichr through online enrichment analysis, it was identified that 7 highly expressed IDs in HF group (EIF1AY, RPS4Y1, USP9Y, KDM5D, DDX3Y, NPPA and HBB) were related to Erythrocytes take up oxygen and release carbon dioxide, Physiological factors, Erythrocytes take up carbon dioxide and release oxygen, O<sub>2</sub>/CO<sub>2</sub> exchange in erythrocytes, HDMs demethylate histones, YAP1-and WWTR1 (TAZ)-stimulated gene expression; 7

highly expressed IDs in normal group (TSIX, LOC28556 and XIST) were related to Fatty acid metabolism (Fig. 6). The ten DEGs identified were enriched and analyzed by Reactome database (<https://reactome.org/>), and the related biological processes of HF were not enriched.

**PPI network construction and hub gene selection**

PPI analysis was performed on these DEGs using the string platform, and 42 nodes and 41 interactions were

**Table 3** Significant enriched GO terms and pathways of DEGs

ONTOLOGY	ID	Description	GeneRatio	BgRatio	p value	p.adjust	q value
BP	GO:0030509	BMP signaling pathway	6/66	157/18670	1.93e-05	0.016	0.014
BP	GO:0071772	Response to BMP	6/66	170/18670	3.02e-05	0.016	0.014
BP	GO:0071773	Cellular response to BMP stimulus	6/66	170/18670	3.02e-05	0.016	0.014
BP	GO:0003012	Muscle system process	8/66	465/18670	2.26e-04	0.067	0.058
BP	GO:0007178	Transmembrane receptor protein serine/threonine kinase signaling pathway	7/66	349/18670	2.27e-04	0.067	0.058
CC	GO:0062023	Collagen-containing extracellular matrix	11/69	406/19717	1.54e-07	2.37e-05	2.06e-05
CC	GO:0005604	Basement membrane	4/69	95/19717	3.44e-04	0.026	0.023
CC	GO:0005614	Interstitial matrix	2/69	12/19717	7.79e-04	0.040	0.035
MF	GO:0008201	Heparin binding	6/66	169/17697	3.92e-05	0.008	0.007
MF	GO:0005539	Glycosaminoglycan binding	6/66	229/17697	2.09e-04	0.021	0.019
MF	GO:1901681	Sulfur compound binding	6/66	250/17697	3.35e-04	0.022	0.020

GO, Gene Ontology; DEGs, Differentially Expressed Genes

finally determined (Fig. 7). In addition, an important module with 3 nodes and 3 edges is selected through MCODE. Two important modules with 6 nodes and 7 edges are selected through MCODE. BBS2, BBS7 and BBS9 are the hub nodes in module B, FRZB, CHRD, BMP4, MYH6, SLN and NPPA are the hub nodes in module C, and KLRB1, CD3D, CCL5, C3, CFH and FCN3 are the hub nodes in module D. Only BBS9, CHRD, BMP4, MYH6, NPPA and CCL5 were selected as hub genes. In addition, combined with the results of differential expression, enrichment analysis and PPI, BBS9, CHRD, BMP4, MYH6, NPPA and CCL5 were considered as Hub genes for further analysis.

#### Identification of potential key genes associated with heart failure

CTD is used to explore the interaction between potential key genes and heart failure. As shown in Fig. 8, potential key genes for heart failure, atrial fibrillation, hypertension, myocardial infarction, sudden cardiac death and myocarditis. The reasoning score in CTD reflects the association between chemicals, diseases and genes. The results of interaction showed that NPPA, HBB, DDX3Y and XIST had higher scores with heart failure.

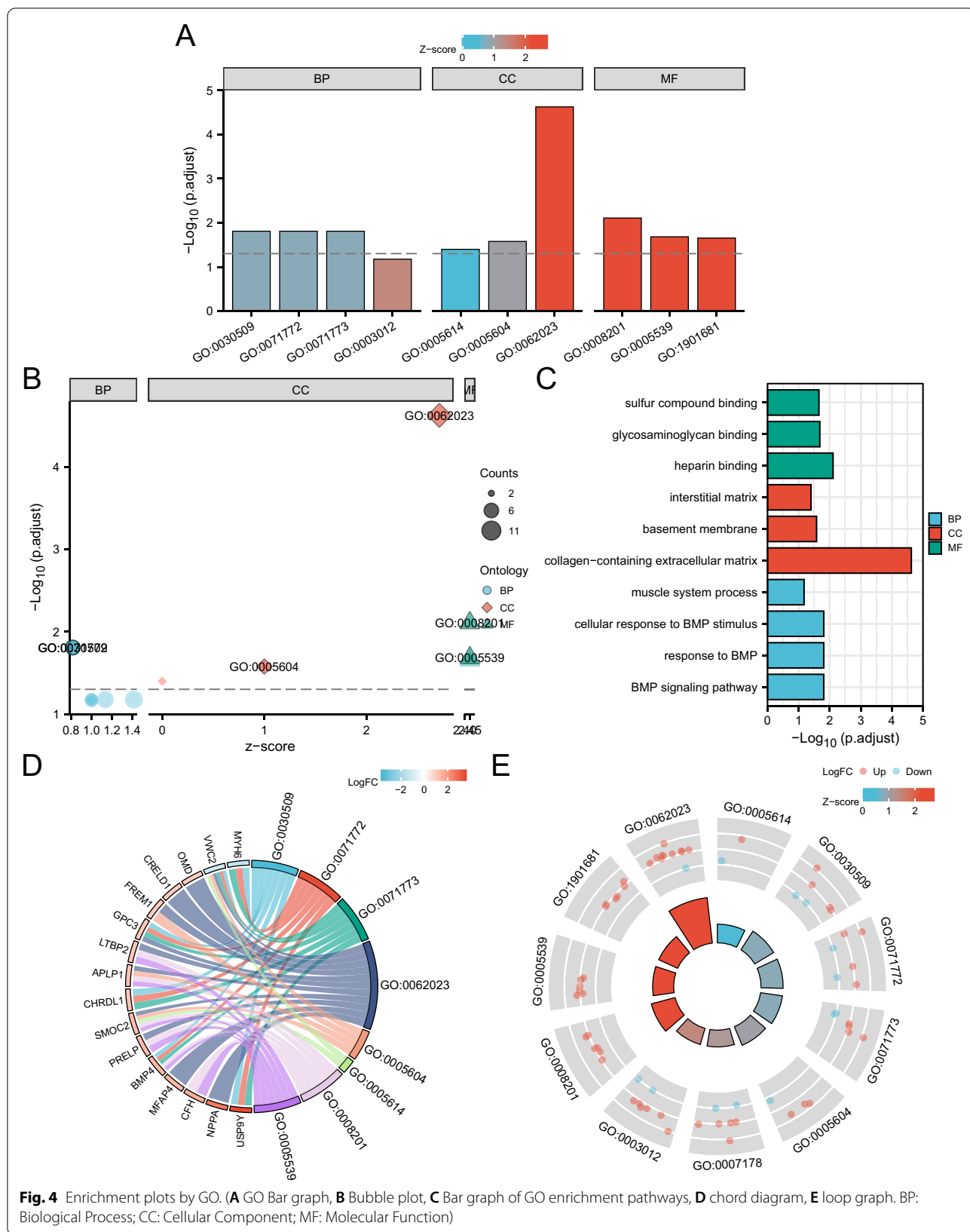
#### Discussion

At present, heart failure has made a great breakthrough in diagnosis, treatment and prevention compared with the past, and has achieved rich research results. In recent years, bioinformatics research such as biomarkers has made great progress and received more and more attention. Therefore, it is particularly important to study the biomarkers of heart failure for the diagnosis, treatment and prognosis of the disease. In this study, we integrated the gene expression profiles of 38 HF samples and 16

normal samples from 3 geo databases, and analyzed the data using bioinformatics tools. 85 IDs met the threshold of  $|\log_2(\text{FC})| \geq 1$  &  $p.\text{adj} < 0.05$ . Under this threshold, 60 IDs were highly expressed in HF group and 25 IDs were highly expressed in normal group; 22 IDs met the threshold of  $|\log_2(\text{FC})| \geq 1.5$  &  $p.\text{adj} < 0.05$ . Under this threshold, 16 IDs were highly expressed in HF group and 6 IDs were highly expressed in normal group; There are 10 IDs that meet the threshold of  $|\log_2(\text{FC})| \geq 2$  &  $p.\text{adj} < 0.05$ . Under this threshold, there are 7 highly expressed IDs in HF group (EIF1AY, RPS4Y1, USP9Y, KDM5D, DDX3Y, NPPA, HBB) and 3 highly expressed IDs in normal group (TSIX, LOC28556, XIST). These 10 potential key genes (EIF1AY, RPS4Y1, USP9Y, KDM5D, DDX3Y, NPPA, HBB, TSIX, LOC28556, XIST) and some important pathways related to the risk of heart failure have been identified, indicating that these may play an important role in the mechanism of the occurrence and development of heart failure.

#### Specific genes on the Y chromosome may be a risk factor for heart failure

EIF1AY, RPS4Y1, USP9Y, KDM5D [29] and DDX3Y, which are highly expressed in HF group, are located on Y chromosome. Considering that men account for a high proportion in the sample, such as data set GSE21610. However, gender is indeed a factor that cannot be ignored in cardiovascular diseases, especially in cardiovascular calcification [30]. In male sex hormones, there is a positive correlation between elevated testosterone and cardiovascular calcification, while in female sex hormones, the cardioprotective effect of estrogen is widely recognized. Therefore, when women enter menopause, the risk of cardiovascular disease will increase due to the decrease of estrogen level. When mineral components in

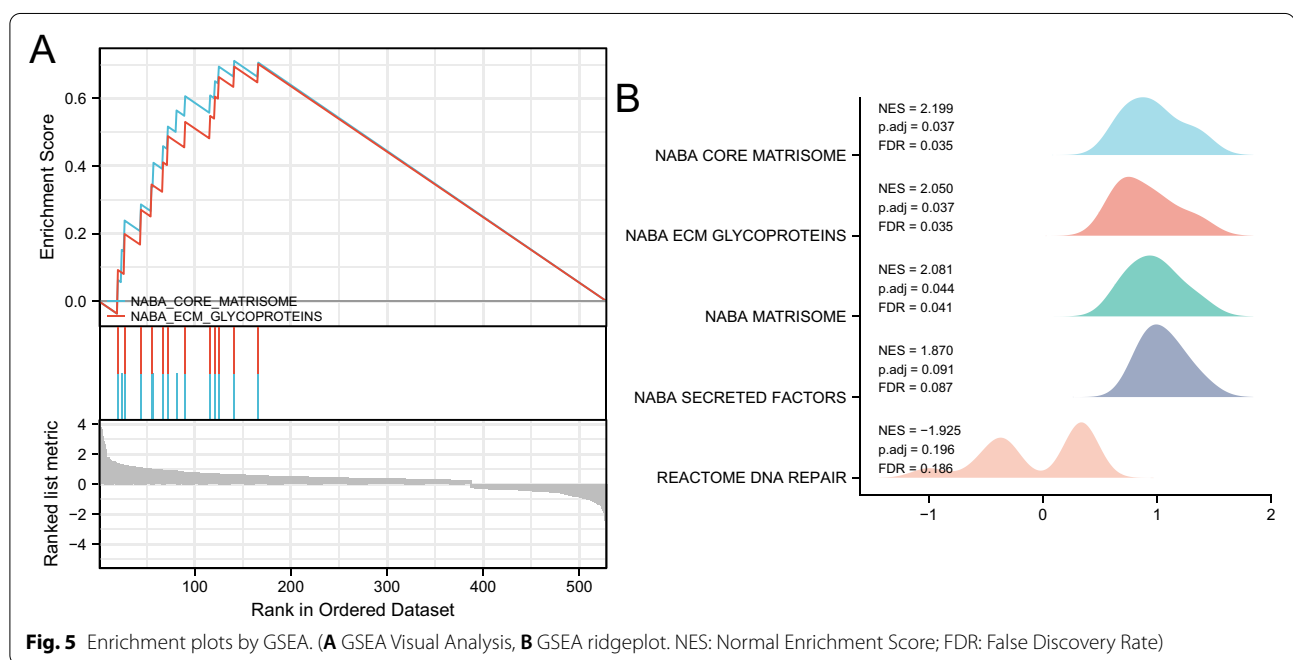




**Table 4** GSEA enrichment analysis

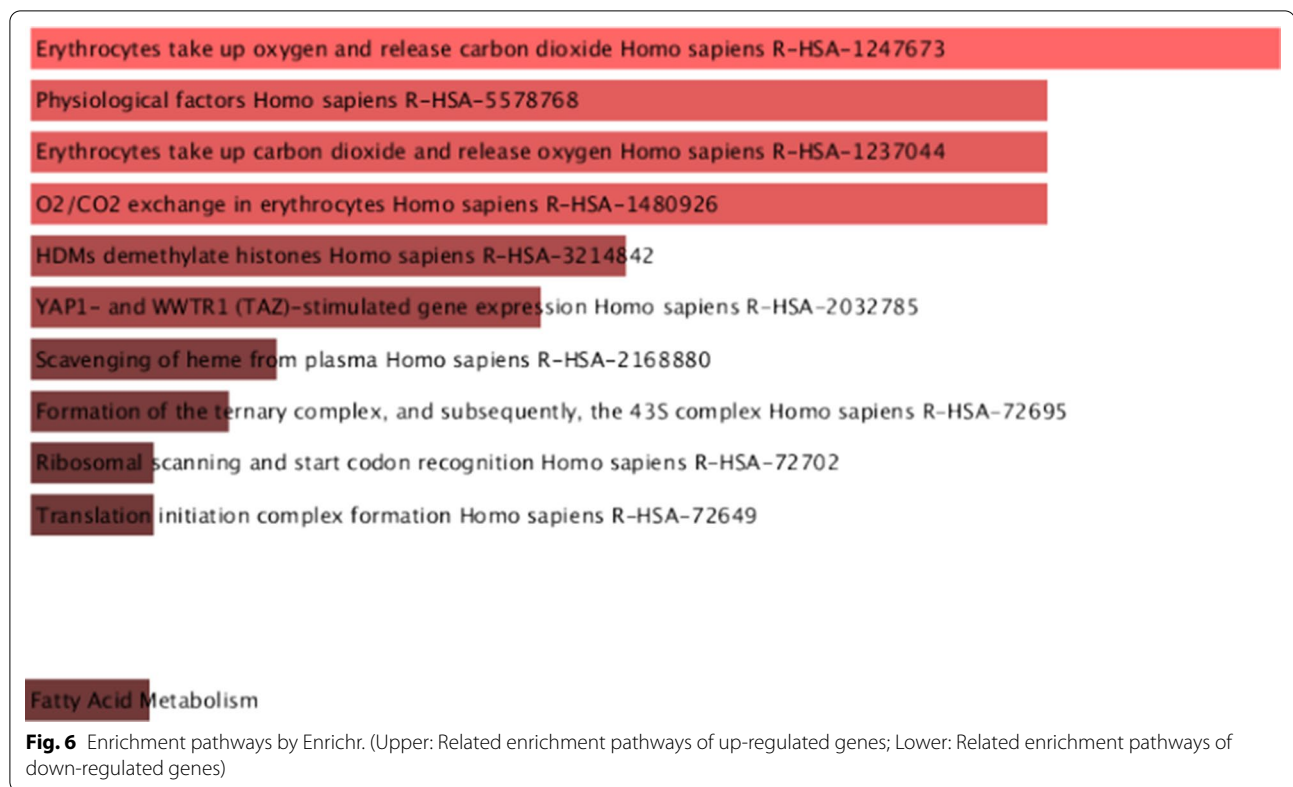
ID	Description	SetSize	EnrichmentScore	NES	p value	p.adjust	q values	Rank	Leading_edge
NABA_CORE_MATRISOME	Ensemble of genes encoding coreextracellular matrix including ECM glycol proteins, collagens and proteoglycans	15	0.7107363	2.199081	0.0014	0.03706	0.03519	141	Tags = 93%, List = 27%, Signal = 70%
NABA_ECM_GLYCOPROTEINS	Genes encoding structural ECM glycoproteins	12	0.7015504	2.049885	0.0015	0.03706	0.03519	166	Tags = 100%, List = 31%, Signal = 70%
NABA_MATRISOME	Ensemble of genes encoding extracellular matrix and extracellular matrix-associated proteins	44	0.5212222	2.08094	0.0026	0.04365	0.04144	141	Tags = 64%, List = 27%, Signal = 51%
NABA_SECRETED_FACTORS	Genes encoding secreted soluble factors	14	0.6127575	1.870019	0.0072	0.09133	0.08671	86	Tags = 64%, List = 16%, Signal = 55%
REACTOME_DNA_REPAIR	DNA Repair	13	0.543689	-1.9251	0.0192	0.19553	0.18564	249	Tags = 100%, List = 47%, Signal = 54%

NES, Normalized Enrichment score



the blood are deposited in blood vessels or heart valves, cardiovascular calcification can occur, including vascular calcification and heart valve calcification. According to its location, cardiovascular calcification can be divided into three types: atherosclerotic intimal vascular calcification, medial vascular calcification and aortic valve calcification [31]. Studies have shown that the occurrence of cardiovascular calcification has become a predictor

of cardiovascular disease-related risks [32]. The deposition of maladjusted calcium may lead to coronary atherosclerotic heart disease, aortic stenosis, hypertension and even heart failure, and become the first trigger to push down dominoes. Studies have shown that in vascular smooth muscle cells, estrogen receptors  $\alpha$  (ER $\alpha$ ). The expression of estrogen receptor was higher than that of estrogen receptor  $\beta$  (ER $\beta$ ). Estrogen is mainly through



ER $\alpha$  Inhibit RANKL signal, so as to reduce osteogenic differentiation and calcification of vascular smooth muscle cells by up regulating BMP and down regulating MGP [33]. In addition, estrogen and aromatase in blood vessels may also play the same role [34]. On the other hand, studies have shown that matrix vesicles (MVS) produced by the outer membrane of cardiovascular cells It also plays a role in cardiovascular calcification. When they are ingested by vascular smooth muscle receptors, they can cause changes in MAPK signal and calcium metabolism. The regulatory response process of calcium binding annexin is activated, while the expression of calcification inhibitors such as MGP and fetuin A is at a low level [35, 36].

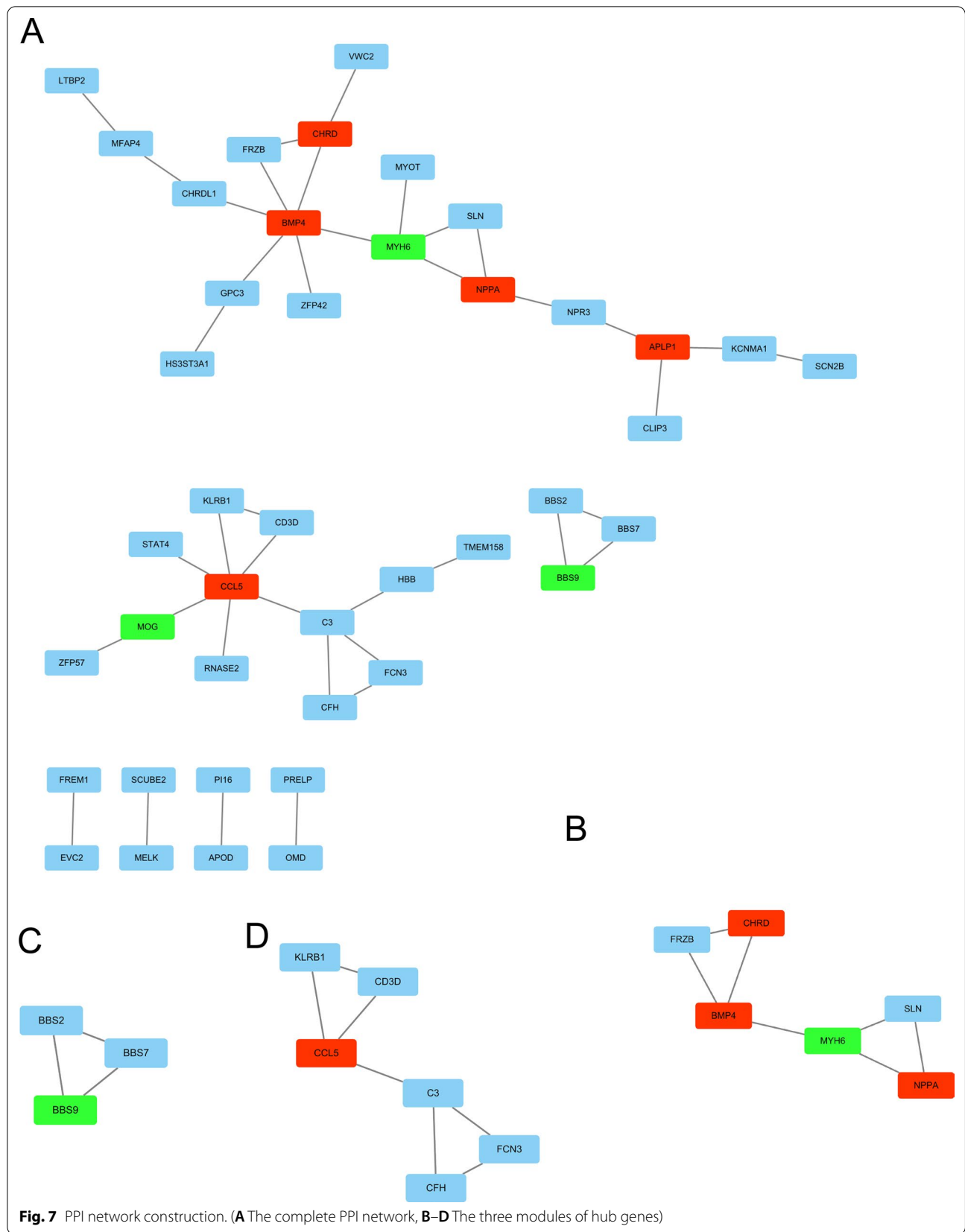
In conclusion, some key genes in this study may affect the cardiovascular system with the help of gender related factors, and finally play a promoting role in the occurrence and development of heart failure.

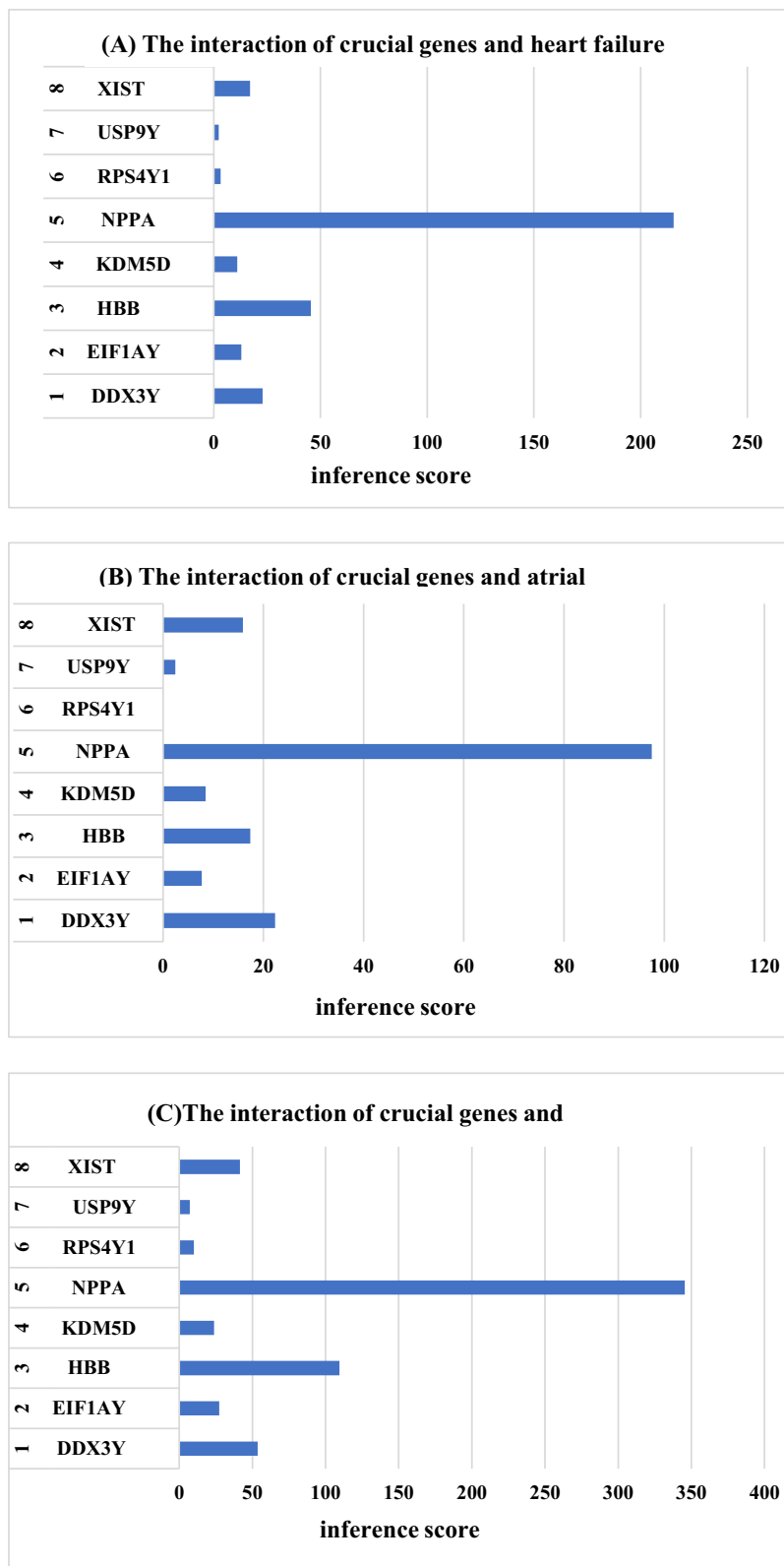
#### Other highly expressed genes related to the mechanism of heart failure

The highly expressed NPPA gene encodes atrial natriuretic peptide (ANP) protein, which is involved in regulating humoral and electrolyte homeostasis together with BNP and CNP. ANP produces intracellular cGMP by binding to guanylate cyclase receptor on cell membrane,

and binds to specific enzymes and ion channels, so as to play the biological function of natriuretic peptide [37, 38]. By studying the mixed muscle secretion phenotype of atrial cardiomyocytes, it is found that these cells can produce polypeptide hormone natriuretic peptide, so as to show the endocrine function of the heart. Studies have shown that NPPA is one of the earliest in situ gene expression during embryonic development, indicating that NPPA plays an important role in early cardiac development [39, 40]. Studies have shown that when acute heart failure occurs, the intermediate region sequence of pro-ANP (MR-proANP) may have greater advantages in judging the prognosis of patients [41]. In our results, compared with the normal group, NPPA was up-regulated in the heart tissue of heart failure, which was related to the traction stimulation of myocardium when heart failure occurred.

The highly expressed HBB gene encodes HBB protein, which is the key component of hemoglobin, so as to complete the important task of tissue oxygen supply. Studies have shown that patients with heart failure often have different degrees of hemodilution, and the changes of hemoglobin or hematocrit concentration can be used as an indirect marker to reflect congestion [42–45]. According to statistics, nearly one third of patients with heart failure are accompanied by anemia





**Fig. 8** The CTD analysis between potential key genes and diseases

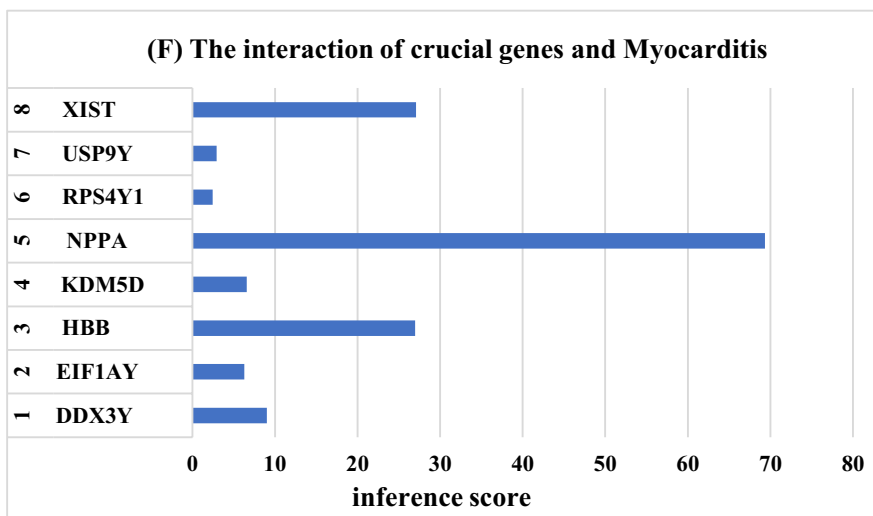
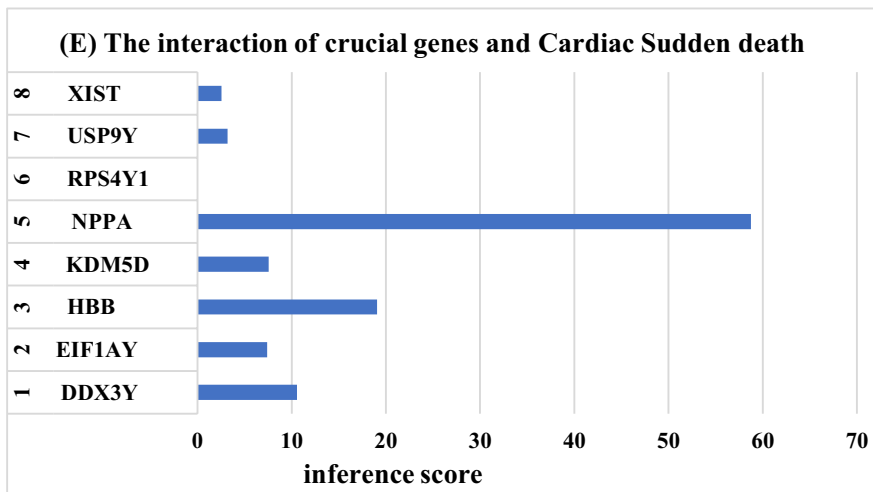
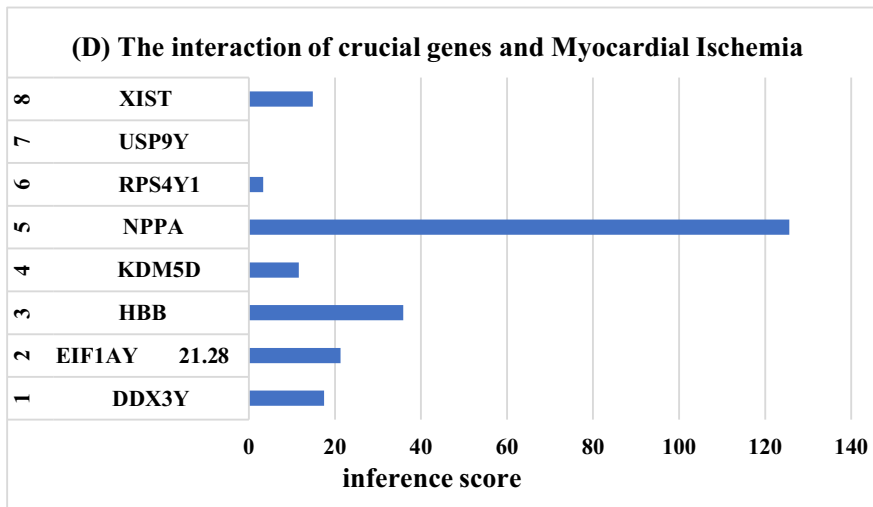


Fig. 8 continued

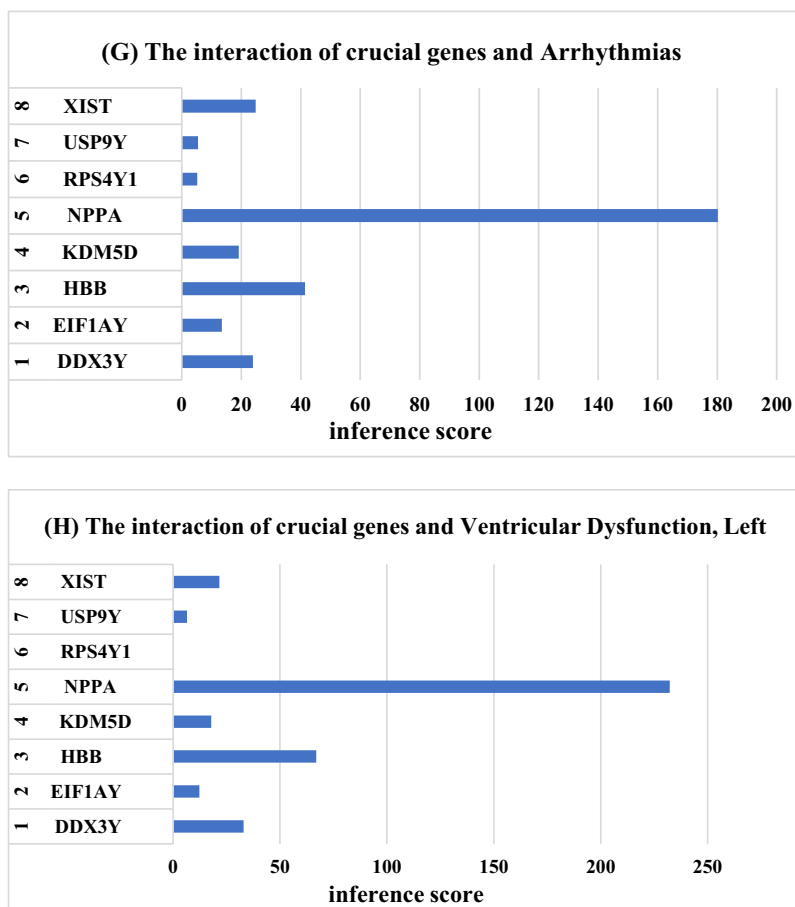


Fig. 8 continued

[46]. The existence of anemia may lead to more symptoms, increased hospitalization rate and increased mortality. Tissue and organ in patients with heart failure are often in an anoxic state. Studies show that erythropoietin levels in these patients are elevated and the overall concentration of erythropoietin is not significantly elevated due to retention of fluids [47, 48]. Erythropoietin is produced by the kidney and is the main stimulating factor for the production of red blood cells. Tissue hypoxia will promote the production of EPO, and the level of the latter is usually inversely proportional to the concentration of hemoglobin. Research shows that blocking  $\beta_2$ -adrenergic receptors can reduce hemoglobin levels [49]. When heart failure occurs, sympathetic nerve excitability increases, which can promote the increase of hemoglobin level. Therefore, according to our results, compared with the normal group, HBB is highly expressed, which may be related to the abnormal activation of sympathetic nervous system and the level of erythropoietin in heart failure.

**Possible relationship between BMP, extracellular matrix and heart failure**

Combining go enrichment analysis and PPI analysis, we pay attention to that BMP and extracellular matrix are relatively high-frequency keywords. Among them, XIST gene is also related to it.

Bone morphogenetic protein BMP is involved in the regulation of embryogenesis and organogenesis, and can participate in the development of cardiovascular structure and function in embryonic stage. When individuals mature, BMP can be used as an important endocrine regulator to participate in cardiovascular, metabolic and hematopoietic activities [50, 51]. Studies have shown that BMP is a transforming growth factor  $\beta$  (TGF- $\beta$ ), a member of the family carries out signal transmission with adjacent cells by means of paracrine or autocrine. Therefore, it plays a role in promoting the early development of organs. On the other hand, some BMP can also play a role in signal transmission in blood circulation, so as to affect distant tissues and organs, and finally complete the role of BMP in cardiovascular, metabolic and

hematopoietic functions [52, 53]. Studies have shown that BMP signaling pathway plays an important role in cardiovascular diseases, increasing its activity in vascular inflammation and atherosclerosis, while decreasing its activity in pulmonary hypertension and hereditary hemorrhagic telangiectasia [53]. BMP has TGF- $\beta$ , the commonness of the family is to bind to serine threonine kinase type I and type II receptors. Among them, the affinity for binding to type I receptors is higher, forming BMPRIA, also known as ALK3, BMPRII, also known as ALK6, ACVRL1, also known as ALK1, and ACVR1, also known as ALK2 heterotetramer complexes; binding with type II receptors to form BMPRII, ACTRIIA, and ACTRIIB heterotetramer complexes, which can be widely expressed in mesenchymal stem cells. In the tissues derived from, especially BMPRII is highly expressed in endothelial and endocardial tissues [54]. Studies have shown that when myocardial pathological hypertrophy is induced by pressure overload, the expression of BMP4 increases, but it does not increase in physiological hypertrophy induced by exercise [55, 56], and this effect can be inhibited by BMP inhibitors [56, 57]. BMP7 has been proved to inhibit cell apoptosis, myocardial fibrosis and anti calcification, and can improve the cardiac function of patients [58]. Studies have shown that BMP is closely related to the development of the heart during embryonic development. In the process of cardiac development, it needs to undergo endocardial to mesenchymal transition (EMT) and migrate as mesenchymal cells to fill the extracellular matrix separating the endocardial layer and the outer layer of myocardium [59, 60], and finally complete the development of the heart [61–63]. Therefore, the enriched pathways in this study are closely related to these mechanisms, which may provide some new evidence for these mechanisms [64].

Extracellular matrix (ECM) plays an important role in stabilizing the structure, transmitting signals and stress of cardiomyocytes, vascular cells and stromal cells. Therefore, the regulatory role of ECM is closely related to the occurrence and development of heart failure. The high expression of XIST gene in normal group is the main factor regulating the transcriptional silencing of X chromosome. At present, there are few studies on the role of XIST in cardiovascular diseases, mostly tumor related studies. It is reported that XIST and miR-101 can aggravate the occurrence of myocardial hypertrophy caused by excessive pressure load [65]. Studies have shown that silencing XIST can suck out mir-1277-5p through sponge action and inhibit the destruction of ECM [66, 67]. When cardiomyocytes are stimulated by external stimuli or their own dysfunction, it can affect the regulatory role of ECM and reduce stress. The increase of load can cause the transformation of cardiac fibroblasts

into myofibroblasts, that is, "interstitial fibrosis" and promote the synthesis of extracellular matrix, so as to reduce cardiac compliance, myocardial remodeling and accelerate the impairment of diastolic function [68]. Interstitial fibrosis is accompanied by the expansion of collagen area around cardiac microvascular adventitia, secondary to myocardial ischemia and hypoxia, aggravate the imbalance between supply and demand of blood oxygen under stress, and promote the occurrence and development of heart failure [69, 70]. Clinically, angiotensin converting enzyme inhibitors, angiotensin receptor inhibitors  $\beta$  [71, 72]. The application of adrenergic receptor antagonists and diuretics is to regulate the abnormal changes of ECM in patients with heart failure by reducing the load on the heart [73]. Among them, proteoglycan located in ECM plays a role in ECM, which belongs to glycosylated protein [74]. Research shows that fibroblasts' response to mechanical stress may include the following mechanisms: (1) After fibroblasts perceive the increase of mechanical load, they amplify the signal through intracellular cascade reaction and induce myocardial fibrosis by activating transcription factor myocardial related transcription factor (MRTF) [75]; (2) mechanical stress stimulates the key medium in the transformation of myofibroblasts—transforming growth factor TGF- $\beta$ . After treatment, it can promote matrix synthesis [76, 77]; (3) increased pressure load can directly activate renin angiotensin aldosterone system (RAAS), and stimulate fibroblast proliferation and ECM protein synthesis through angiotensin II type 1 receptor (AT1R) signal [78, 79]; (4) The increase of pressure negative charge can induce the expression of miRNA in fibroblasts, further activate MAPK signaling pathway, and finally promote the synthesis of matrix [80].

In the current research, the possible mechanisms of 10 potential key genes involved in the occurrence and development of heart failure have been discussed in the occurrence and development of heart failure. The results show that the above genes may become potential biomarkers and therapeutic targets of heart failure, hoping to provide some ideas for further exploration and research of heart failure. Yes, this study still has some limitations: (1) the included samples have limitations: in the included data set, the age, gender, race, nationality, region, living habits and family history of the samples can be called influencing factors. (2) the potential key factors obtained from the analysis need to be experimentally verified in clinical samples, such as RT-qPCR, Western blot, etc.

## Conclusions

Our study integrated relatively large sample size data from multiple geographic data sets and identified 10 potential key genes (EIF1AY, RPS4Y1, USP9Y, KDM5D, DDX3Y, NPPA, HBB, TSIX, LOC28556, XIST) by

bioinformatics analysis. The exploration of potential key genes of heart failure may provide some potential help for further identifying new biomarkers and useful therapeutic targets of heart failure susceptibility.

#### Acknowledgements

Not applicable.

#### Authors' contributions

Conception and design: S.F.; Data collection and analysis: S.F.; Manuscript writing: S.F.; Manuscript revising: Y.H. All authors read and approved the final manuscript.

#### Funding

This work was supported by the National Natural Science Foundation of China (No.: 81773444).

#### Availability of data and materials

The datasets analysed during the current study are available in the GEO Database repository, GSE76701 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE76701>. GSE21610 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE21610>. GSE8331 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8331>.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

Received: 29 October 2021 Accepted: 22 March 2022

Published online: 27 March 2022

#### References

- James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, Abbastabar H, Abd-Allah F, Abdela J, Abdelalim A, Abdollahpour I. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. 2018;392(10159):1789–858.
- Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. *Eur J Heart Fail*. 2020;22(8):1342–56.
- Dhalla NS, Saini-Chohan HK, Delfin RL, Vijayan E, Dent MR, Tappia PS. Subcellular remodelling may induce cardiac dysfunction in congestive heart failure. *Cardiovasc Res*. 2008;3:3.
- Shimizu I, Minamino T. Physiological and pathological cardiac hypertrophy. *J Mol Cell Cardiol*. 2016;97:245–62.
- Shah AK, Bhullar SK, Elimban V, Dhalla NS. Oxidative stress as a mechanism for functional alterations in cardiac hypertrophy and heart failure. *Antioxidants (Basel, Switzerland)*. 2021;10(6):931.
- Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Colvin MM, Drazner MH, Filippatos GS, Fonarow GC, Givertz MM. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *J Am Coll Cardiol*. 2016;68(13):1476–88.
- Gaggin HK, Januzzi JL Jr. Biomarkers and diagnostics in heart failure. *Biochim Biophys Acta*. 2013;1832(12):2442–50.
- Oikonomou E, Zografos T, Papamikroulis GA, Siasos G, Vogiatzi G, Theofilis P, Briasoulis A, Papaioannou S, Vavuranakis M, Gennimata V, Tousoulis D. Biomarkers in atrial fibrillation and heart failure. *Curr Med Chem*. 2019;26(5):873–87.
- Wang XY, Zhang F, Zhang C, Zheng LR, Yang J. The biomarkers for acute myocardial infarction and heart failure. *Biomed Res Int*. 2020;2020:2018035.
- Chen S, Tang Y, Zhou X. Cystatin C for predicting all-cause mortality and rehospitalization in patients with heart failure: a meta-analysis. *Biosci Rep*. 2019;39(2):BSR20181761.
- Akkafa F, Halil Altiparmak I, Erkus ME, Aksoy N, Kaya C, Ozer A, Sezen H, Oztuzcu S, Koyuncu I, Umurhan B. Reduced SIRT1 expression correlates with enhanced oxidative stress in compensated and decompensated heart failure. *Redox Biol*. 2015;6:169–73.
- Smyth GK. *limma: linear models for microarray data*. New York: Springer; 2005.
- Meltzer DPS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*. 2007;23(14):1846–7.
- Gu Z, Roland E, Matthias S. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847.
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: J Integr Biol*. 2012;16(5):284–7.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT. Gene Ontology: tool for the unification of biology. *Nat Genet*. 2000;25(11):25–9.
- Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res*. 2018;47(D1):D590–5.
- Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. 2019;28(11):1947–51.
- Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res*. 2021;49(D1):D545–d551.
- Jupe S, Fabregat A, Hermjakob H. Expression data analysis with reactome. *Curr Protoc Bioinform*. 2015;49(1):8–20.
- Chen EY, Tan CM, Kou Y, Duan Q, Ayan AM. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform*. 2013;14(1):128–128.
- Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Simon K, Jenkins SL, Jagodnik KM, Alexander L. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;W1:W90–7.
- Xie Z, Bailey A, Kuleshov MV, Clarke D, Ma'Ayan A. Gene set knowledge discovery with enrichr. *Curr Protoc*. 2021;1(3):e90.
- Schriml LM, Elvira M, James M, Becky T, Mike S, Lance N, Victor F, Linda J, Cynthia B, Richard L. Human disease ontology 2018 update: classification, content and workflow expansion. *Nucleic Acids Research*. 2018;D1:D1.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102(43):15545–50.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102(43):15545–50.
- Damian S, Gable AL, David L, Alexander J, Stefan W, Jaime HC, Milan S, Doncheva NT, Morris JH, Peer B. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2018;D1:D607.
- Davis AP, Grondin CJ, Johnson RJ, Sciaky D. The comparative toxicogenomics database: update 2019. *Nucleic Acids Res*. 2018;47(D1):D948–54.
- Li N, Dhar SS, Chen TY, Kan PY, Min GL. JARID1D is a suppressor and prognostic marker of prostate cancer invasion and metastasis. *Cancer Res*. 2016;76(4):831–43.
- Li Y, Jiang Q, Ding Z, Liu G, Yu P, Jiang G, Yu Z, Yang C, Qian J, Jiang H. Identification of a common different gene expression signature in ischemic cardiomyopathy. *Genes*. 2018;9(1):56.
- Woodward HJ, Zhu D, Hadoke PWF, MacRae VE. Regulatory role of sex hormones in cardiovascular calcification. *Int J Mol Sci*. 2021;22(9):4620.



32. Liu W, Zhang Y, Yu C-M, Ji Q-W, Cai M, Zhao Y-X, Zhou Y-J. Current understanding of coronary artery calcification. *J Geriatr Cardiol: JGC*. 2015;12(6):668–75.
33. Osako MK, Nakagami H, Koibuchi N, Shimizu H, Morishita R. Estrogen inhibits vascular calcification via vascular RANKL system. *Circ Res*. 2010;20(4):466–75.
34. Harada N, Sasano H, Murakami H, Ohkuma T, Nagura H, Takagi Y. Localized expression of aromatase in human vascular tissues. *Circ Res*. 1999;84(11):1285.
35. Praneet C, Chen NX, Kalisha O, McClintock JN, Moe SM, Chandra JS, Van WA. Differential miRNA expression in cells and matrix vesicles in vascular smooth muscle cells from rats with kidney disease. *PLoS ONE*. 2015;10(6):e0131589.
36. Chen NX, O'Neill KD, Moe SM. Matrix vesicles induce calcification of recipient vascular smooth muscle cells through multiple signaling pathways. *Kidney Int*. 2018;93(2):343–54.
37. Tulassay T, Seri I, Rascher W. Atrial natriuretic peptide and extracellular volume contraction after birth. *Acta Paediatr Scand*. 1987;76(3):444–6.
38. Samson WK. Atrial natriuretic factor inhibits dehydration and hemorrhage-induced vasopressin release. *Neuroendocrinology*. 1985;40(3):277–9.
39. Goetze JP, Bruneau BG, Ramos HR, Ogawa T, de Bold MK, de Bold AJ. Cardiac natriuretic peptides. *Nat Rev Cardiol*. 2020;17(11):698–717.
40. Zeller R, Bloch KD, Williams BS, Arceci RJ, Seidman CE. Localized expression of the atrial natriuretic factor gene during cardiac embryogenesis. *Genes Dev*. 1987;1(7):693–8.
41. Seronde MF, Gayat E, Logeart D, Lassus J, Laribi S, Boukef R, Sibellas F, Launay JM, Manivet P, Sadoune M, Nouira S, Solal AC, Mebazaa A. Comparison of the diagnostic and prognostic values of B-type and atrial-type natriuretic peptides in acute heart failure. *Int J Cardiol*. 2013;168(4):3404–11.
42. Testani JM, Chen J, McCauley BD, Kimmel SE, Shannon RP. Potential effects of aggressive decongestion during the treatment of decompensated heart failure on renal function and survival. *Circulation*. 2010;122(3):265–72.
43. Greene SJ, Gheorghide M, Vaduganathan M, Ambrosy AP, Mentz RJ, Subacius H, Maggioni AP, Nodari S, Konstam MA, Butler J, Filippatos G. Haemoconcentration, renal function, and post-discharge outcomes among patients hospitalized for heart failure with reduced ejection fraction: insights from the EVEREST trial. *Eur J Heart Fail*. 2013;15(12):1401–11.
44. van der Meer P, Postmus D, Ponikowski P, Cleland JG, O'Connor CM, Cotter G, Metra M, Davison BA, Givertz MM, Mansoor GA, Teerlink JR, Massie BM, Hillege HL, Voors AA. The predictive value of short-term changes in hemoglobin concentration in patients presenting with acute decompensated heart failure. *J Am Coll Cardiol*. 2013;61(19):1973–81.
45. Kobayashi M, Girerd N, Duarte K, Chouihed T, Chikamori T, Pitt B, Zannad F, Rossignol P. Estimated plasma volume status in heart failure: clinical implications and future directions. *Clin Res Cardiol*. 2021;110(8):1159–72.
46. van Veldhuisen DJ, Anker SD, Ponikowski P, Macdougall IC. Anemia and iron deficiency in heart failure: mechanisms and therapeutic approaches. *Nat Rev Cardiol*. 2011;8(9):485–93.
47. van der Meer P, Voors AA, Lipsic E, Smilde TD, van Gilst WH, van Veldhuisen DJ. Prognostic value of plasma erythropoietin on mortality in patients with chronic heart failure. *J Am Coll Cardiol*. 2004;44(1):63–7.
48. Grote Beverborg N, van der Wal HH, Klip IT, Voors AA, de Boer RA, van Gilst WH, van Veldhuisen DJ, Gansvoort RT, Hillege HL, van der Harst P, Bakker SJ, van der Meer P. High serum erythropoietin levels are related to heart failure development in subjects from the general population with albuminuria: data from PREVEND. *Eur J Heart Fail*. 2016;18(7):814–21.
49. Komajda M, Anker SD, Charlesworth A, Okonko D, Metra M, Di Lenarda A, Remme W, Moullet C, Swedberg K, Cleland JG, Poole-Wilson PA. The impact of new onset anaemia on morbidity and mortality in chronic heart failure: results from COMET. *Eur Heart J*. 2006;27(12):1440–6.
50. David L, Mallet C, Keramidas M, Lamande N, Gasc JM, Dupuis-Girod S, Plauchu H, Feige JJ, Bailly S. Bone morphogenetic protein-9 is a circulating vascular quiescence factor. *Circ Res*. 2008;102(8):914–22.
51. Vukicevic S, Grgurevic L. BMP-6 and mesenchymal stem cell differentiation. *Cytokine Growth Factor Rev*. 2009;20(5–6):441–8.
52. Laux DW, Young S, Donovan JP, Mansfield CJ, Upton PD, Roman BL. Circulating Bmp10 acts through endothelial Alk1 to mediate flow-dependent arterial quiescence. *Development*. 2013;140(16):3403–12.
53. Herrera B, Inman GJ. A rapid and sensitive bioassay for the simultaneous measurement of multiple bone morphogenetic proteins. Identification and quantification of BMP4, BMP6 and BMP9 in bovine and human serum. *BMC Cell Biol*. 2009;10(1):20.
54. Wu X, Sagave J, Rutkovskiy A, Haugen F, Baysa A, Nygård S, Czibik G, Dahl CP, Gullestad L, Vaage J. Expression of bone morphogenetic protein 4 and its receptors in the remodeling heart. *Life Sci*. 2014;97(2):145–54.
55. Sun B, Sheng Y, Huo R, Hu CW, Lu J, Li SL, Liu X, Wang YC, Dong DL. Bone morphogenetic protein-4 contributes to the down-regulation of Kv4.3 K<sup>+</sup> channels in pathological cardiac hypertrophy. *Biochem Biophys Res Commun*. 2013;436(4):591–4.
56. Sun B, Rong H, Yue S, Li Y, Dong DL. Bone morphogenetic protein-4 mediates cardiac hypertrophy, apoptosis, and fibrosis in experimentally pathological cardiac hypertrophy. *Hypertension*. 2013;61(2):352.
57. Pachori AS, Custer L, Hansen D, Clapp S, Kempa E, Klingensmith J. Bone morphogenetic protein 4 mediates myocardial ischemic injury through JNK-dependent signaling pathway. *J Mol Cell Cardiol*. 2010;48(6):1255–65.
58. Aluganti Narasimulu C, Singla DK. The role of bone morphogenetic protein 7 (BMP-7) in inflammation in heart diseases. *Cells*. 2020;9(2):280.
59. Selleri L, Zappavigna V, Ferretti E. "Building a perfect body": control of vertebrate organogenesis by PBX-dependent regulatory networks. *Genes Dev*. 2019;33(5–6):258–75.
60. Tam PP, Parameswaran M, Kinder SJ, Weinberger RP. The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. *Development*. 1997;124(9):1631–42.
61. Waller BR, Wessels A. Cardiac morphogenesis and dysmorphogenesis. Totowa, NJ: Humana Press; 2000.
62. Srivastava D, Olson EN. A genetic blueprint for cardiac development. *Nature*. 2000;407(6801):221–6.
63. Kruthof B, Duim SN, Moerkamp AT, Goumans MJ. TGFβ and BMP signaling in cardiac cushion formation: lessons from mice and chicken. *Differentiation*. 2012;84(1):89–102.
64. Morrell NW, Bloch DB, Dijke PT, Goumans MJ, Bloch KD. Targeting BMP signalling in cardiovascular disease and anaemia. *Nat Rev Cardiol*. 2016;13(2):106–20.
65. Xiao L, Gu Y, Sun Y, Chen J, Wang X, Zhang Y, Gao L, Li L. The long noncoding RNA XIST regulates cardiac hypertrophy by targeting miR-101. *J Cell Physiol*. 2019;234(8):13680–92.
66. Akbari Dilmaghani N, Shoorei H, Sharifi G, Mohaqiq M, Majidpoor J, Dinger ME, Taheri M, Ghafouri-Fard S. Non-coding RNAs modulate function of extracellular matrix proteins. *Biomed Pharmacother*. 2021;136:111240.
67. Zhou J, Zhou Y, Wang CX. LncRNA-MIAT regulates fibrosis in hypertrophic cardiomyopathy (HCM) by mediating the expression of miR-29a-3p. *J Cell Biochem*. 2018;120(5):7265–75.
68. Ytrehus K, Hulot J-S, Perrino C, Schiattarella GG, Madonna R. Perivascular fibrosis and the microvasculature of the heart: Still hidden secrets of pathophysiology? *Vasc Pharmacol*. 2018;107:78–83.
69. Dai Z, Aoki T, Fukumoto Y, Shimokawa H. Coronary perivascular fibrosis is associated with impairment of coronary blood flow in patients with non-ischemic heart failure. *J Cardiol*. 2012;60(5–6):416–21.
70. Varagic J, Frohlich ED, Diez J, Susic D, Ahn J, Gonzalez A, Lopez B. Myocardial fibrosis, impaired coronary hemodynamics, and biventricular dysfunction in salt-loaded SHR. *Am J Physiol Heart Circ Physiol*. 2006;290(4):H1503.
71. Kate H, Ida L, Andrew MC, Geir C. The soft- and hard-heartedness of cardiac fibroblasts: mechanotransduction signaling pathways in fibrosis of the heart. *J Clin Med*. 2017;6(5):53.
72. Rienks M, Papageorgiou AP, Frangogiannis NG, Heymans S. Myocardial extracellular matrix an ever-changing and diverse entity. *Circ Res*. 2014;114(5):872–88.
73. Christensen G, Herum KM, Lunde IG. Sweet, yet underappreciated: proteoglycans and extracellular matrix remodeling in heart disease. *Matrix Biol*. 2018;75:286–99.
74. Karamanos NK. Matrix pathobiology—central roles for proteoglycans and heparanase in health and disease. *FEBS J*. 2017;284(1):7–9.
75. Kennedy L, Xu SW, Ca Rter DE, Abraham DJ, Leask A. Fibroblast adhesion results in the induction of a matrix remodeling gene expression program. *Matrix Biol*. 2008;27(4):274–81.

76. Lighthouse JK, Small EM. Transcriptional control of cardiac fibroblast plasticity. *J Mol Cell Cardiol.* 2016;91:52–60.
77. Zhao XH, Laschinger C, Arora P, Szász K, Kapus A, Mcculloch CA, Zhao XH, Laschinger C, Arora P, et al. Force activates smooth muscle alpha-actin promoter activity through the Rho signaling pathway. *J Cell Sci.* 2007;120(Pt 10):1801–9.
78. Vincent S, Anne K, Chow ML, Elena Z, Li CX, Hideyuki K, Caldarone CA, Boris H. Integrins  $\alpha\beta 5$  and  $\alpha\beta 3$  promote latent TGF- $\beta 1$  activation by human cardiac fibroblast contraction. *Cardiovasc Res.* 2014;3:407–17.
79. Ping K, Shinde AV, Su Y, Russo I, Frangogiannis NG. Opposing actions of fibroblast and cardiomyocyte Smad3 signaling in the infarcted myocardium. *Circulation.* 2017;137(7):707–24.
80. Frangogiannis NG. The extracellular matrix in ischemic and nonischemic heart failure. *Circ Res.* 2019;125(1):117–46.

### Publisher's Note

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

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

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

At BMC, research is always in progress.

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

