doi: 10.4149/gpb_2022029

Integrated bioinformatics and machine learning strategies reveal *PRDX6* as the key ferroptosis-associated molecular biosignature of heart failure

Chenyang Jiang[1,2] and Weidong Jiang[3]

[1] *Department of Cardiology, The Second Affiliated Hospital of Nantong University, Nantong 226000, China*
[2] *The First Clinical Medical College of Guangxi Medical University, Nanning 530021, China*
[3] *Department of Cardiology, Nantong Hospital of Traditional Chinese Medicine, Nantong 226000, China*

Abstract. Heart failure (HF) is the leading cause of death and public health problems in the global population. This study aimed to identify and validate ferroptosis-related biomarkers associated with HF in clinical medicine using bioinformatics and machine learning strategies. Weighted co-expression network analysis (WGCNA) was applied to screen the module genes and analyze their biological functions and pathways. Ferroptosis-associated genes (FAG) in HF were determined and then machine learning algorithms were used for screening. Next, multiple external independent microarrays were used to verify molecular biosignature. Simultaneously, CIBERSORT was applied to estimate the immune infiltration landscape. Combined with the results of the WGCNA, 25 FAGs were determined and 6 FAMBs were selected by machine learning strategies. In addition, Peroxiredoxin 6 (*PRDX6*) was finally selected as the key ferroptosis-associated molecular biological feature based on multiple verifications of independent data sets. From the results of the infiltration and enrichment analysis, we believed that *PRDX6*, as a protective biomarker related to ferroptosis in HF, may help provide new ideas in the immunotherapy of HF.

Key words: Heart failure — Weighted co-expression network analysis — Gene expression omnibus — Bioinformatics — Immune cell

Abbreviations: *ATF4*, activating transcription factor 4; BP, biological processes; CC, cell component; DCM, dilated cardiomyopathy; FAG, ferroptosis-associated genes; FAMB, ferroptosis-associated molecular biosignature; GEO, Gene Expression Omnibus; GO, gene ontology; *GPX4*, glutathione peroxidase; GSEA, gene set enrichment analysis; HF, heart failure; ISCH, ischemic; LASSO, least absolute shrinkage and selection operator; MF, molecular function; *NF2*, neurofibromin 2; *NOX4*, NADPH oxidase 4; *PRDX6*, peroxiredoxin 6; *RGS4*, regulator of G protein signaling 4; *SLC40A1*, solute carrier family 40 member 1; SVM-RFE, support vector machine recursive feature elimination; *TLR4*, toll like receptor 4; TOM, topological overlap matrix; WGCNA, weighted co-expression network analysis.

Electronic supplementary material. The online version of this article (doi: 10.4149/gpb_2022029) contains Supplementary material.
Correspondence to: Weidong Jiang, Department of Cardiology, Nantong Hospital of Traditional Chinese Medicine, Nantong 226000, Jiangsu, China
E-mail: drjiangweidong@126.com

Introduction

Heart failure (HF) is a common factor in the death of cardiovascular disease; cardiovascular disease is a major threat to the health of global population. The reality is that due to insufficient heart output, the body cannot meet metabolic needs, and the quality-of-life survival for those suffering HF remains poor (Kim et al. 2021). As a global epidemic, HF affects 26 million adults worldwide. Studies have proved on the basis of available information that the cost of HF treatment in the next 10 years will increase by 127% (Mozaffarian et al. 2016). As of 2030, an estimated 8 million people in the United States would have HF. Concurrently; China has 4.2 million patients with HF with a prevalence of 1.3% (Hu et al. 2012). HF diagnostic markers such as natriuretic peptides provide the possibility to determine HF, but they are not stable enough at normal temperature. It is easily interfered by the age of the patient, a variety of diseases and drugs, and there are some shortcomings that cannot be ignored. In the era of big data, we still hope to discover new key molecular biological characteristics of HF. Access to new molecular markers of HF could bring benefits for faster and more efficient diagnosis of HF.

Ferroptosis is a new type of programmed cell death that is different from apoptosis, cell necrosis, and autophagy (Mou et al. 2019). Ferroptosis is characterized by iron dependence and reactive oxygen species (ROS) dependence, mainly focusing on cytological changes. Such changes include loss of the mitochondrial cristae, rupture of the mitochondrial membrane, and membrane condensation (Latunde 2017; Yu et al. 2017). In the process of biological changes, the activity of peroxidase is changed, and the antioxidant capacity of cells is gradually attenuated. When the accumulation of ROS causes cell oxidative death, ferroptosis occurs. In recent years, ferroptosis has not only made considerable progress in tumors but has also made discoveries in the cardiovascular field (Hu et al. 2021; Yu et al. 2021). Many investigations have proved that ferroptosis plays a critical role in cardiomyopathy, myocardial infarction, HF, atherosclerosis, etc. (Liu et al. 2018; Fang et al. 2019; Bai et al. 2020; Song et al. 2021). Studies have shown that the ferroptosis regulator Solute Carrier Family 7 Member 11 (*SLC7A11*), which is abundantly enriched in cardiomyocytes, can prevent HF caused by cardiac hypertrophy and cardiomyopathy (Fang et al. 2020; Zhang et al. 2021). In addition, studies have found that silencing toll like receptor 4 (*TLR4*) and NADPH oxidase 4 (*NOX4*) significantly delays ferroptosis in HF rats, suggesting that *TLR4-NOX4* may be a potential therapeutic target for HF by inhibiting ferroptosis (Chen et al. 2019). These studies help focus our attention on the close connection between ferroptosis and HF. This makes it more important to discover the potential molecular biomarkers between ferroptosis and HF.

The application of integrated bioinformatics methods to biological research helps to discover more sensitive and efficient diagnosis and treatment strategies. In this study, we analyzed data sets related to HF through bioinformatics and used machine learning strategies to screen and determine the key ferroptosis-associated molecular biomarkers. Immune cell infiltration is a novel method for assessing the abundance of immune cells in diseases. We used CIBERSORT to assess the infiltration of immune cells in HF. Combining the results of gene targets and immune cell infiltration will help understand the pathophysiological process of diseases and discover new ideas for immunomodulatory therapy. Our research results help to discover new ferroptosis targets for HF and provide new prevention and treatment strategies for individualized clinical treatment of HF.

Materials and Methods

HF microarray data and RNA-seq data preparation

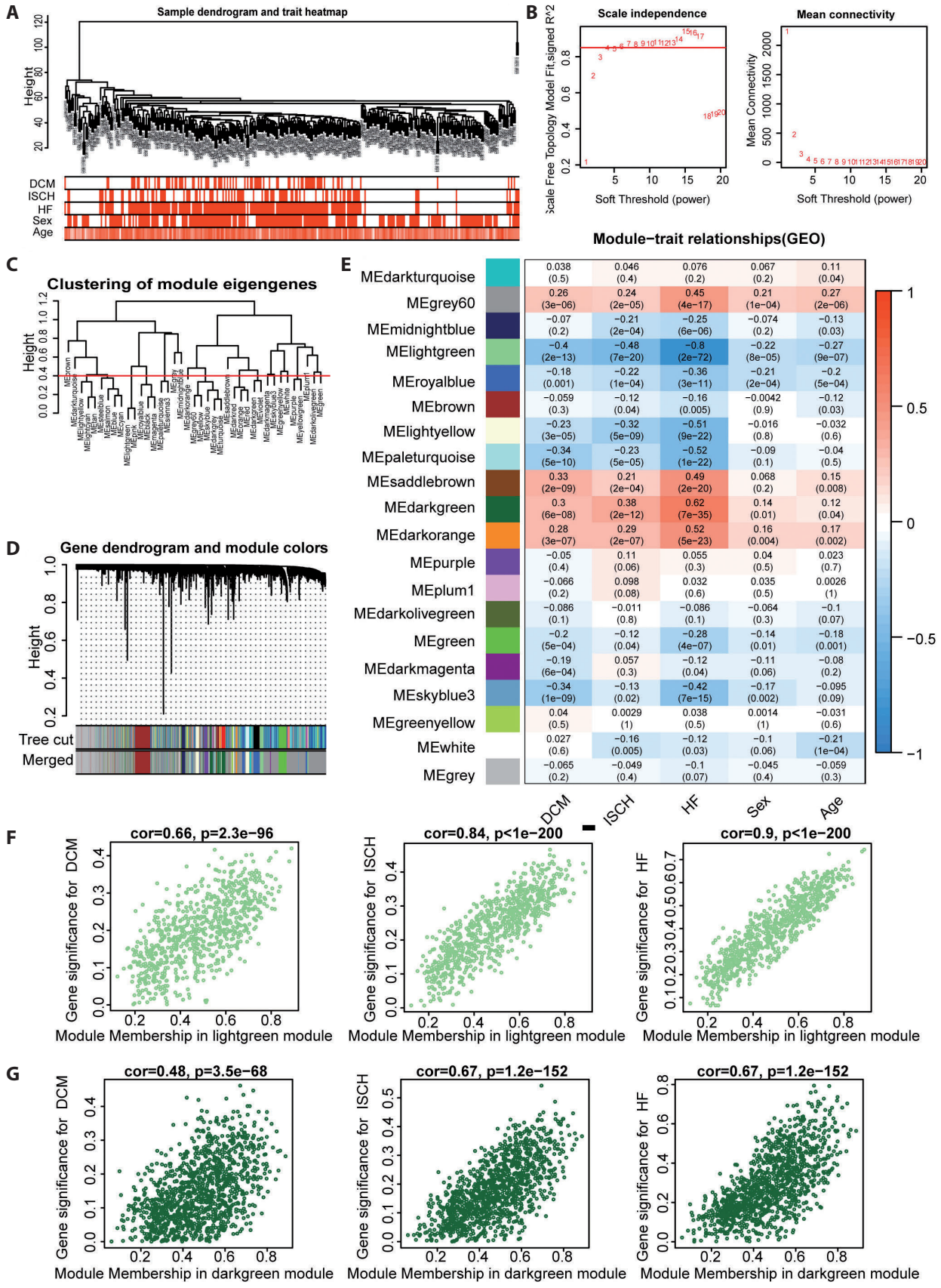
We screened the required HF data set from the Gene Expression Omnibus (GEO) database on the public platform. The selection criteria were as follows: 1. the data set excludes cancer samples; 2. the data set excludes complications such as diabetes, chronic kidney disease, and chronic obstructive pulmonary disease; and 3. the data set includes at least 6 samples. Based on the above standards, we obtained the GSE57338 human heart tissue sample microarray data sets, and the GSE71613, GSE161472, and GSE135055 human heart tissue high throughput sequencing data sets (Table S1 in Supplementary Material). GSE57338 included human left ventricular tissues from 136 normal controls and 177 HF samples. GSE71613 included 4 control samples and 4 left ventricular failure tissue samples. GSE161472 included 37 control samples and 47 HF samples with reduced ejection fractions. GSE135055 included the left ventricular myocardium of 9 normal controls and 21 patients with HF. In addition, we also obtained the C57BL/6 mouse microarray gene set GSE36074. This data set included 5 tissue samples in the Sham operation group, 7 tissue samples in the cardiac hypertrophy treatment group, and 7 tissue samples in the HF treatment group. The clinical information for GSE57338 includes the sex and age of the sample, and whether the patients have HF. The information of whether the cause of the patient's HF is dilated cardiomyopathy (DCM) or ischemic (ISCH) is also presented in GSE57338. GSE135055 clinical information included NYHA HF classification, including 2 cases of NYHA Class II, 7 cases of Class III, and 12 cases of Class IV. New York heart function classification (NYHA classification): I (Normal): no symptoms of HF; II: symptoms of HF with slight

activity; III: significantly limited activity. Heart failure symptoms with mild activity; IV: obvious symptoms of HF at rest. GSE36074 had corresponding mouse experimental characteristic data (left ventricular weight change, right ventricle weight, lung weight, tibia length, and ltindex). The ltindex is defined as the ratio of lung weight to tibia length. The large sample data set GSE57338 of 313 cases was used as the main analysis data set, whereas other external independent data sets GSE71613, GSE161472, GSE135055, and GSE36074 were used to verify the results. The required data were preprocessed, including background correction and log2 conversion, and the processed data set was used for subsequent analysis. The research design and analysis flowchart are presented in Figure 1.

Co-expression module formation

The weighted gene co-expression network analysis (WGCNA) is a method to find highly correlated gene modules. We used WGCNA to find highly relevant gene modules in the HF gene set to facilitate the identification of candidate biomarkers. We used the "*WGCNA*" package in R for the result analysis. First, the GSE57338 microarray data were checked for missing values, and then the samples were clustered and analyzed. The pickSoftThreshold function determined the soft threshold power β. Subsequently, we built a topological overlap matrix (TOM) to measure the average network connectivity of genes. Dynamic shearing divided the gene into multiple gene modules. After confirming that the cutting height was set to 0.4, similar modules were used to perform clustering and merging. Among the modules, the modules with the highest correlation with HF clinical information (HF, DCM, and ISCH) were considered the most critical. The characteristic genes in these modules were identified for further analysis. Finally, the saliency scatters plot shows the degree of correlation between the gene module membership and the gene significance correlation (GS).

Figure 1. Research design and analysis flow chart. GSE57338 is used for weighted gene co-expression network analysis of heart failure (HF) data set, enrichment analysis, and immune cell infiltration. Ferroptosis related genes are collected from the FerrDb database and other previous references. Use machine learning algorithms to determine Ferroptosis-associated molecular biomarkers (FAMB). Validate FAMB through the external data set screening of GSE161472, GSE71613 and GSE135055. The mouse data set GSE36074 is used for further verification of FAMB. Finally, the combination of FAMB and immune cell infiltration will help to discover new ideas for immunomodulatory therapy.


A
Sample dendrogram and trait heatmap
Height
DCM
ISCH
HF
Sex
Age
B
Scale independence
Scale Free Topology Model Fit,signed R^2
Soft Threshold (power)
Mean connectivity
Mean Connectivity
Soft Threshold (power)
C
Clustering of module eigengenes
Height
D
Gene dendrogram and module colors
Height
Tree cut
Merged
E
Module−trait relationships(GEO)
MEdarkturquoise 0.038 (0.5) 0.046 (0.4) 0.076 (0.2) 0.067 (0.2) 0.11 (0.04)
MEgrey60 0.26 (3e−06) 0.24 (2e−05) 0.45 (4e−17) 0.21 (1e−04) 0.27 (2e−06)
MEmidnightblue −0.07 (0.2) −0.21 (2e−04) −0.25 (6e−06) −0.074 (0.2) −0.13 (0.03)
MElightgreen −0.4 (2e−13) −0.48 (7e−20) −0.8 (2e−72) −0.22 (8e−05) −0.27 (9e−07)
MEroyalblue −0.18 (0.001) −0.22 (1e−04) −0.36 (3e−11) −0.21 (2e−04) −0.2 (5e−04)
MEbrown −0.059 (0.3) −0.12 (0.04) −0.16 (0.005) −0.0042 (0.9) −0.12 (0.03)
MElightyellow −0.23 (3e−05) −0.32 (5e−09) −0.51 (9e−22) −0.016 (0.8) −0.032 (0.6)
MEpaleturquoise −0.34 (5e−10) −0.23 (5e−05) −0.52 (1e−22) −0.09 (0.1) −0.04 (0.5)
MEsaddlebrown 0.33 (2e−09) 0.21 (2e−04) 0.49 (2e−20) 0.068 (0.2) 0.15 (0.008)
MEdarkgreen 0.3 (6e−08) 0.38 (2e−12) 0.62 (7e−35) 0.14 (0.01) 0.12 (0.04)
MEdarkorange 0.28 (3e−07) 0.29 (2e−07) 0.52 (5e−23) 0.16 (0.004) 0.17 (0.002)
MEpurple −0.05 (0.4) 0.11 (0.06) 0.055 (0.3) 0.04 (0.5) 0.023 (0.7)
MEplum1 −0.066 (0.2) 0.098 (0.08) 0.032 (0.6) 0.035 (0.5) 0.0026 (1)
MEdarkolivegreen −0.086 (0.1) −0.011 (0.8) −0.086 (0.1) −0.064 (0.3) −0.1 (0.07)
MEgreen −0.2 (5e−04) −0.12 (0.04) −0.28 (4e−07) −0.14 (0.01) −0.18 (0.001)
MEdarkmagenta −0.19 (6e−04) 0.057 (0.3) −0.12 (0.04) −0.11 (0.06) −0.08 (0.2)
MEskyblue3 −0.34 (1e−09) −0.13 (0.02) −0.42 (7e−15) −0.17 (0.002) −0.095 (0.09)
MEgreenyellow 0.04 (0.5) 0.0029 (1) 0.038 (0.5) 0.0014 (1) −0.031 (0.6)
MEwhite 0.027 (0.6) −0.16 (0.005) −0.12 (0.03) −0.1 (0.06) −0.21 (1e−04)
MEgrey −0.065 (0.2) −0.049 (0.4) −0.1 (0.07) −0.045 (0.4) −0.059 (0.3)
DCM
ISCH
HF
Sex
Age
F
cor=0.66, p=2.3e−96
Gene significance for DCM
Module Membership in lightgreen module
cor=0.84, p<1e−200
Gene significance for ISCH
Module Membership in lightgreen module
cor=0.9, p<1e−200
Gene significance for HF
Module Membership in lightgreen module
G
cor=0.48, p=3.5e−68
Gene significance for DCM
Module Membership in darkgreen module
cor=0.67, p=1.2e−152
Gene significance for ISCH
Module Membership in darkgreen module
cor=0.67, p=1.2e−152
Gene significance for HF
Module Membership in darkgreen module

Modular gene function enrichment

We selected the genes in the highly relevant gene modules obtained by WGCNA analysis for Gene Ontology (GO) enrichment analyses. An adjusted p-value of <0.05 was set as the critical standard for GO. The gene set enrichment analysis (GSEA) is an enrichment analysis method that maximizes the biological significance of genes by calculating enrichment scores. GSEA was performed to reveal the relevant pathways of interest. The R packages *"clusterProfiler," "org.Hs.eg.db,"* and *"org.Mm.eg.db"* were selected to perform GSEA feature enrichment. The critical standard was set to an adjusted p-value of <0.05 and an absolute value of NES of >1.

Machine learning algorithm

Overall, 259 ferroptosis-associated genes (FAG) were obtained after collating from the FerrDb database (http://zhounan.org/ferrdb) and previous literature (Stockwell et al. 2017; Bebber et al. 2020; Dai et al. 2020). All the ferroptosis-related genes are provided in Table S2. The clinical information related to gene and ferroptosis-associated gene information of the WGCNA module was integrated, and the FAG expression data collection in HF was obtained. In the era of smart data, we used more novel algorithms to screen FAG expression data sets to obtain ferroptosis-associated molecular biomarkers (FAMB). The machine learning algorithm Least Absolute Shrinkage and Selection Operator (LASSO) constructs a gene screening model based on the gene expression in GSE57338. Two algorithms, Logistic Regression (LR) and Random Forest (RF), perform 5-fold cross-validation on LASSO, and the receiver operating characteristic (ROC) curve is used to evaluate and screen efficiency. At the same time, Support Vector Machine Recursive Feature Elimination (SVM-RFE) was used to filter FAG. Finally, the results of LASSO and SVM-RFE were integrated to obtain the FAMB screened by the algorithm.

Immune cell infiltration estimation

A variety of immune cells are involved in the pathophysiological process of HF that plays an important role in cardiac hypertrophy and ventricular remodeling. We used immune cell infiltration to evaluate the association between immune cells and potential FAMB. The tissue immune infiltration landscape is implemented by CIBERSORT, and the proportion of 22 immune cells is estimated from gene expression data so that the abundance of multiple cell types in the mixed immune cell population can be obtained. After that, we performed correlation analysis on the estimated multiple immune cell types to find out whether there was a possible strong correlation between immune cells. Simultaneously, we explored whether immune cells significantly increase or decrease in disease. We also hope to explore the relationship between clinical information and immune cells. Furthermore, the relationship between immune cells and potential FAMB has also been studied.

External independent data set for FAMB verification

Multiple external independent data sets validate FAMB to obtain more accurate results. We validated FAMB expression in the RNA-seq dataset GSE161472 of 84 myocardial tissue samples and GSE71613 of 8 tissue samples. In addition, large-sample data set verification can improve accuracy. Genes with the same trend in the two data sets showed higher expression stability. Concurrently, the R package "pROC" was used to perform the diagnostic evaluation of the area under the ROC (AUC) of GSE161472. Furthermore, the FAMB was verified in the data set GSE135055 using the New York classification of HF. After aortic band treatment, the mouse HF data set GSE36074 contains experimental animal information, such as left ventricular weight change, right ventricle weight, lung weight, tibia length, and ltindex. Herein, left ventricular weight change is defined as ending left ventricular weight minus starting left ventricular weight, whereas ltindex is defined as the ratio of lung weight to tibia length. Immune cell subtypes, GSEA analysis, FAMB, and the correlation of experimental animal characteristics were analyzed together.

Statistical analysis

The statistical analysis in the article was performed using R software (version 4.0.2; https://www.r-project.org/). Use Spearman's correlation coefficient to analyze the association between continuous variables. Between the two groups, we used t-test (normal distribution) and Benjamini & Hochberg method correction or Mann-Whitney test (non-normal distribution) to evaluate. The p-value of < 0.05 was considered statistically significant.

◄ **Figure 2.** WGCNA identifies key modules. **A.** HF sample dendrogram with a heat map of traits. Each red vertical line represents a sample in a gene set. **B.** Scale-free fitting index graph and average connectivity graph. **C.** Suitable cut height dendrogram (cutting height was set to 0.4). **D.** Gene cluster dendrogram after merging. **E.** Modular feature correlation between modular features and clinical features. **F.** The saliency scatter plot of clinical features (DCM, ISCH, and HF) and lightgreen module. **G.** The saliency scatter plot of clinical features (DCM, ISCH, and HF) and darkgreen module.

Results

WGCNA in HF

WGCNA conducts genetic screening based on some clinical features of the GSE57338 microarray data. The clinical features include ISCH, DCM, HF, age, and sex (Fig. 2A). The final soft threshold power β was set to 6 (no scale = 0.85) (Fig. 2B). When the shearing height MEDissThres was 0.4 (Fig. 2C), similar modules were clustered and merged (Fig. 2D). According to the clinical characteristics, 20 modules were identified. Among the modules displayed by the heat map, the module feature gene possessing the highest correlation with clinical information was used for further analysis. The results show that the lightgreen and darkgreen modules have the highest correlation with clinical information (DCM, ISCH, and HF) (upregulation and downregulation) (Fig. 2E). The correlation between the lightgreen module and clinical information (for DCM, ISCH, and HF) was −0.4, −0.48, and −0.8. The correlation between the darkgreen module and clinical information (for DCM, ISCH, and HF) is 0.3, 0.38, and 0.62, respectively. Simultaneously, we also found that 20 modular genes have a low correlation with sex and age. The six saliency scatter plots show that the correlation between module membership and gene significance is relatively high. The specific correlation coefficients are as follows. Lightgreen module: DCM (*Cor* = 0.66, $p < 0.001$), ISCH (*Cor* = 0.84, $p < 0.001$), and HF (*Cor* = 0.90, $p < 0.001$) (Fig. 2F). Darkgreen module: DCM (*Cor* = 0.48, $p < 0.001$), ISCH (*Cor* = 0.67, $p < 0.001$), and HF (*Cor* = 0.67, $p < 0.001$) (Fig. 2G). Therefore, we choose the genes in these two modules to further identify ferroptosis-associated markers.

Enrichment analysis and biological significance

GO enrichment analysis divides gene function in disease into three potential parts: biological processes (BP), cell component (CC), and molecular function (MF). These three parts can evaluate the location of the gene, the biological mechanism that the gene participates in, and the type of reaction catalyzed at the molecular level of the gene. We selected genes in two related gene modules obtained through WGCNA analysis for GO enrichment analysis. GO enrichment analysis found that the most significant changes were mainly concentrated in various BP such as virus response, extracellular structure, and T cell regulation. The MF partly focuses on protein ribose binding and so on. The CC is concentrated in a protein complex, collagen trimer, and other parts (Fig. 3A). GSEA found that the MAPK signaling pathway, PI3K-Akt signaling pathway, B cell receptor signaling pathway, cellular senescence, phagosomes, and other pathways related to HF are involved (Fig. 3B).

Screening of FAMB in HF

The genes in the most significant upregulated and downregulated modules of WGCNA and 259 ferroptosis-associated genes were integrated to obtain 25 HF-related FAG sets. The screening of FAMB is generated from 25 FAGs. After LASSO dimensionality reduction screening, nine FAGs were obtained (Fig. 3C and D). Two algorithms, namely, LR and RF, were used to construct predictive models. It was verified that the AUC of the two algorithms reached 0.968 and 0.958, respectively (Fig. 3E), suggesting that LASSO screening is meaningful. Seven FAGs were more suitable after SVM-RFE screening (Fig. 3F). By integrating the FAG screened by two machine learning algorithms, we procured six FAMBs (Fig. 3G; Table S3). These six FAMBs, namely, STEP3 metalloreductase (*STEAP3*), activating transcription factor 4 (*ATF4*), neurofibromin 2 (*NF2*), peroxiredoxin 6 (*PRDX6*), regulator of G protein signaling 4 (*RGS4*), and solute carrier family 40 member 1 (*SLC40A1*), had significant potential analytical value.

HF-related tissue immune infiltration landscape

CIBERSORT was used to evaluate the immune cell landscape of the 313 samples data set GSE57338 and was able to estimate the relative distribution ratio of immune cell subtypes in HF and normal control tissues. Based on the clinical characteristics of GSE57338 and the difference in FAMB expressions, a heat map of 22 immune cells was drawn (Fig. 4A). Next, we analyzed the correlation between immune cells. We found that T cells CD4 memory resting has the most significant negative correlation with CD8 T cells, naïve B cells, and regulatory T cells (Tregs), whereas the most significant positive correlation was between Tregs and naïve B cells (Fig. 4B). Comparing 22 kinds of immune cells in HF and normal tissues, it was found that CD8 T cells ($p = 0.002$) and mast cells resting ($p < 0.001$) were significantly increased in the disease, macrophages M2 ($p < 0.001$), T cells CD4 memory resting ($p = 0.017$), Tregs ($p = 0.047$), and monocytes ($p < 0.001$) were significantly reduced in the disease group (Fig. 4C). Finally, we analyzed the correlation between clinical features of HF and immune cells and found that HF and nine immune cell subtypes (naïve CD4 T cells, CD8 T cells, macrophages M0, mast cells resting, neutrophils, monocytes, macrophages M2, T cells CD4 memory resting, and Tregs) are relevant. Mast cells resting has the most significant positive correlation with HF (*Cor* = 0.25, $p < 0.001$), and neutrophils have the most significant negative correlation with HF (*Cor* = −0.23, $p < 0.001$) (Fig. 4D).

External verification of FAMB

Six FAMBs (namely, *STEAP3*, *ATF4*, *NF2*, *PRDX6*, *RGS4*, and *SLC40A1*) were verified in an external independent data set. Compared with normal controls in GSE161472, *STEAP3* ($p <$

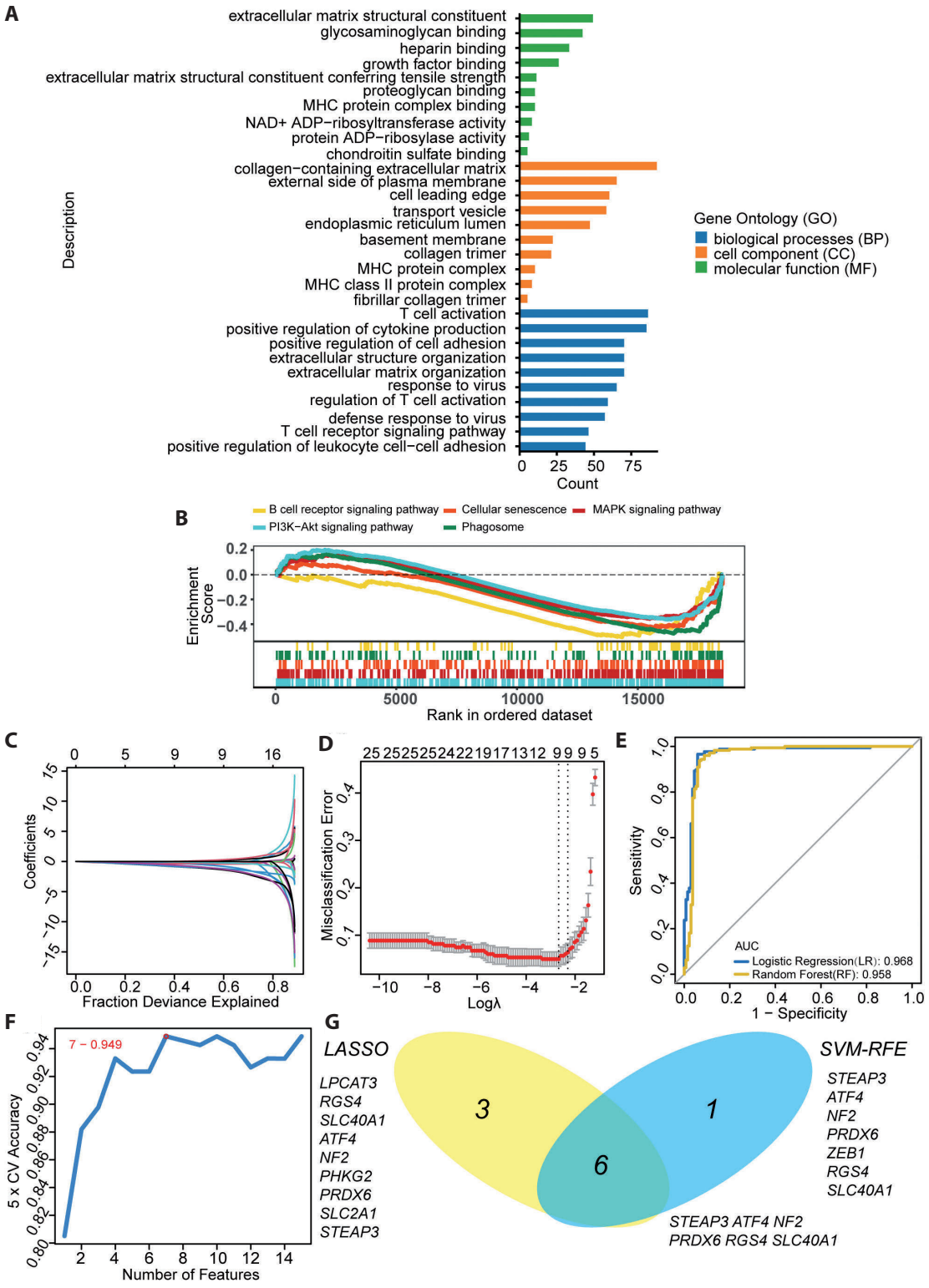


Figure 3. Pathway enrichment and FAMB screening. **A.** Biological significance in GO. Blue represents biological processes (BP), orange represents cell component (CC), and green represents molecular function (MF). **B.** Five pathways related to HF obtained from GSEA enrichment. **C.** Fraction deviance by LASSO. **D.** Appropriate min λ value (min λ = 9). **E.** Logistic regression (LR) and random forest (RF) are used to perform 5-fold cross-validation on the screening model constructed by LASSO. Receiver operating characteristic (ROC) curve shows verification efficiency. **F.** Seven FAMBs were screened by SVM-RFE. **G.** Six FAMBs after integration.

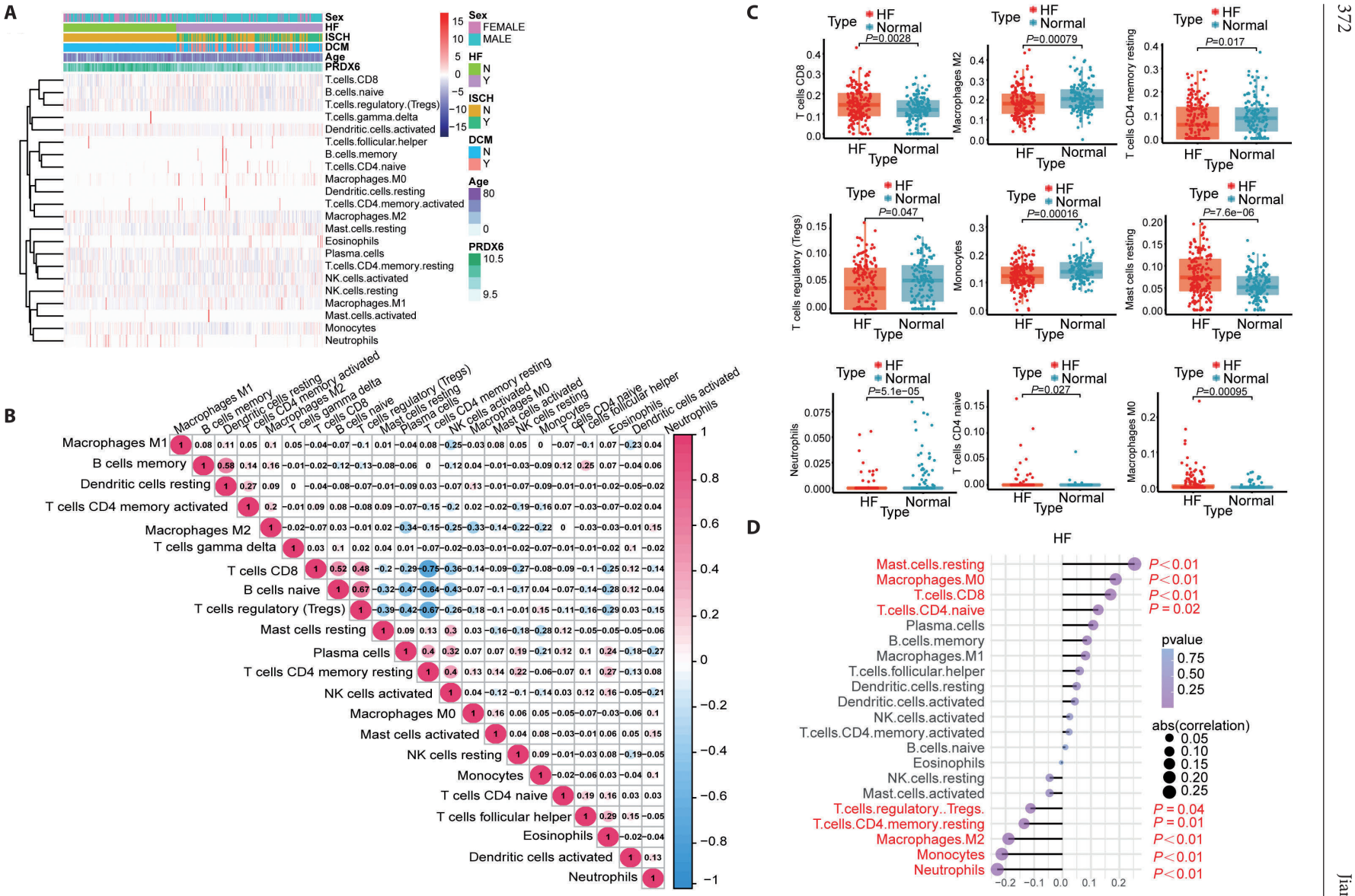


Figure 4. Immune cell landscape estimation. **A.** Heat map of immune cell subtypes of 313 samples. **B.** Correlation between immune cell subtypes. Blue represents negative correlation, and red represents positive correlation. The intensity of the color represents the strength of the correlation. **C.** Immune cell comparison between HF and normal control. Red represents HF samples, blue represents normal control samples. **D.** Correlation analysis of clinical feature (HF) and immune cells. Immune cells (statistically significant) are marked in red. N, no; Y, yes.

0.0001), *ATF4* ($p = 0.0006$), *PRDX6* ($p < 0.0001$), and *RGS4* ($p = 0.0022$) were statistically significant in HF disease (Fig. 5A). In GSE71613, compared with normal control samples, the differences in *STEAP3* ($p = 0.0121$), *SLC40A1* ($p = 0.0045$), and *PRDX6* ($p = 0.0424$) between the HF sample and the normal control sample were statistically significant (Fig. 5B). The verification results of *STEAP3* and *PRDX6* were better in the two data sets. Therefore, *STEAP3* and *PRDX6* were selected for further analysis. In GSE161472, ROC curve analysis can effectively distinguish FAMB between HF and normal control. In the results, we found that the AUC of *STEAP* and *PRDX6* were 0.781 and 0.765 (Fig. 5C). In addition, we verified the expression of *STEAP* and *PRDX6* in the data set GSE135055 with the New York classification of HF, and we found that the expression

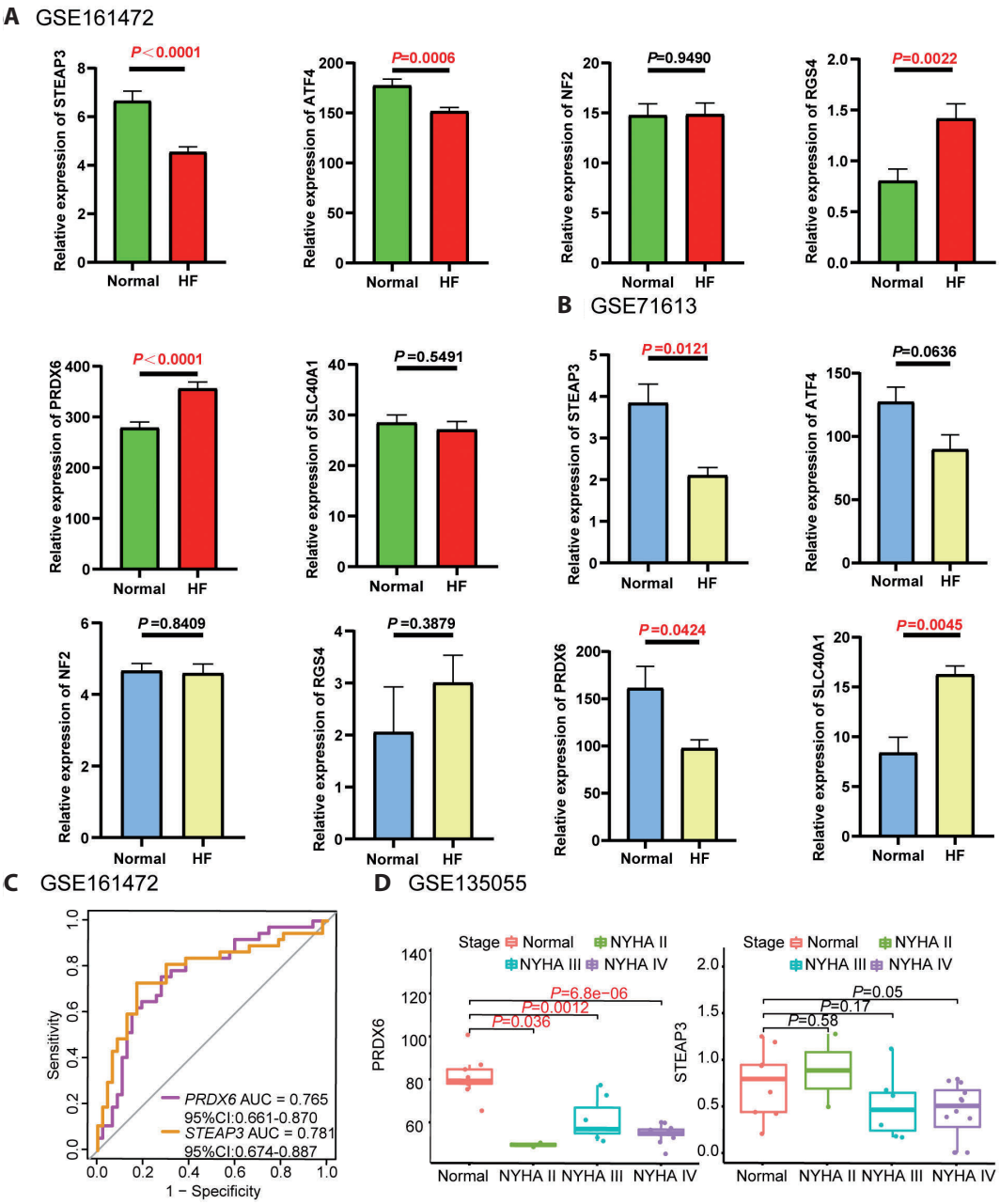


Figure 5. FAMB verification results. **A.** FAMB verification in GSE161472 data set. Green represents normal control samples, red represents HF samples. **B.** FAMB verification in GSE71613 data set. Blue represents normal control samples, yellow represents HF samples. **C.** FAMB ROC curve of GSE161472. **D.** FAMB verification in GSE135055. Red represents normal control, green represents NYHA II, blue represents NYHA III, and purple represents NYHA IV.

of *PRDX6* in each group of NYHA was lower than that in the normal control (Fig. 5D). There was no statistical significance between *STEAP3* HF and the normal control group. Therefore, combining the validation analysis of multiple data sets, we believe that *PRDX6* is more suitable for the subsequent analysis because it shows significant differences in multiple data sets.

FAMB verification in mice experiments

In the mouse data set, we found that the expression of *PRDX6* in the cardiac hypertrophy group and the HF group was significantly lower than that in the Sham group, which is consistent with our previous verification results (Fig. 6A). Left ventricular weight change, lung weight, and ltindex are positively correlated with the degree of cardiac hypertrophy and HF in mice. The pathophysiological process of HF is accompanied by an increase in left ventricular weight, lung weight, and the ratio of lung weight to tibia length. Analysis of the characteristics of *PRDX6* and C57BL/6 mice found that *PRDX6* and left ventricular weight change ($R = -0.72$, $p < 0.001$), lung weight ($R = -0.57$, $p = 0.011$), and ltindex ($R = -0.54$, $p = 0.016$) were negatively correlated, and there

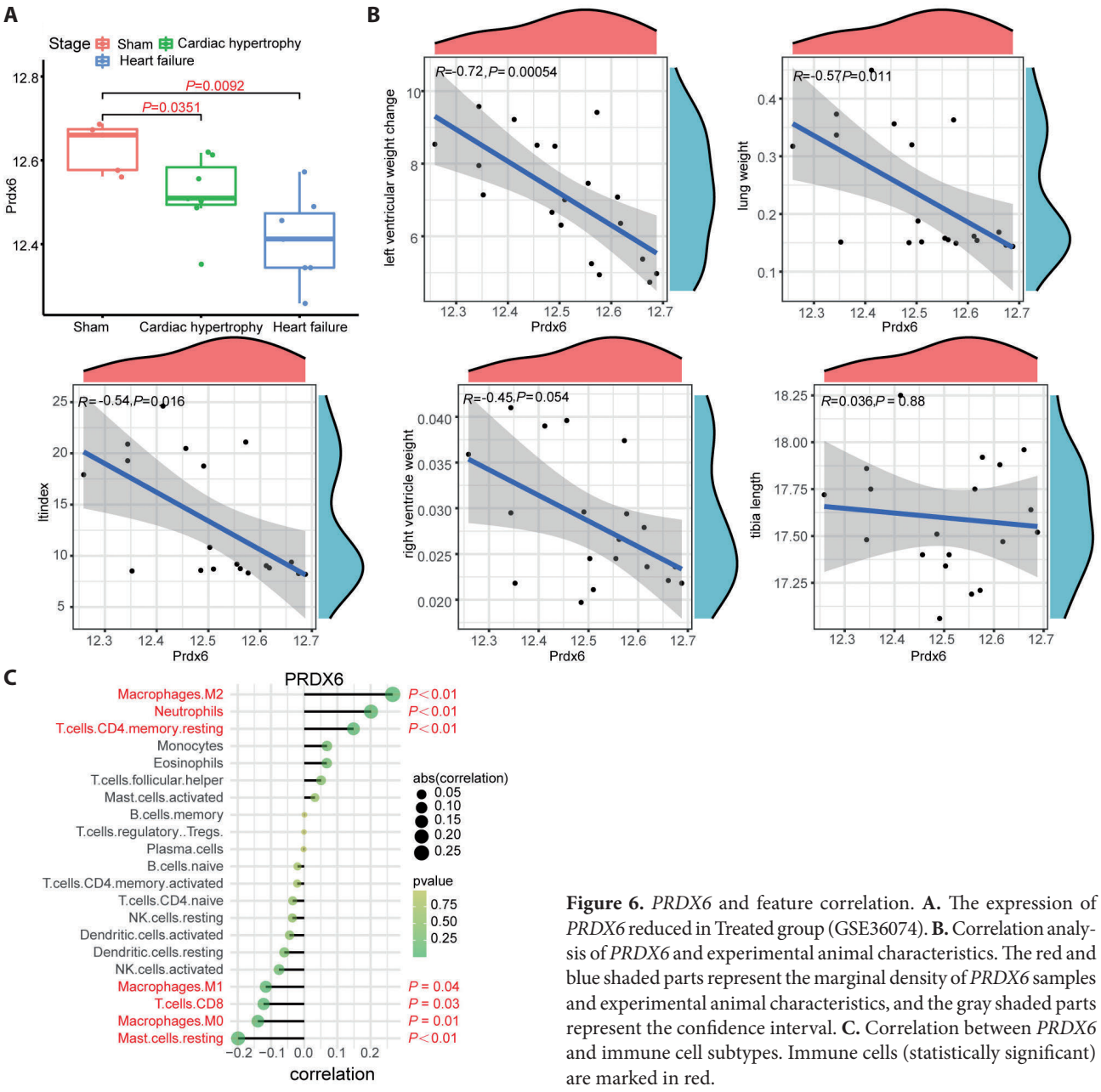


Figure 6. *PRDX6* and feature correlation. **A.** The expression of *PRDX6* reduced in Treated group (GSE36074). **B.** Correlation analysis of *PRDX6* and experimental animal characteristics. The red and blue shaded parts represent the marginal density of *PRDX6* samples and experimental animal characteristics, and the gray shaded parts represent the confidence interval. **C.** Correlation between *PRDX6* and immune cell subtypes. Immune cells (statistically significant) are marked in red.

was no statistically significant correlation with right ventricle weight ($R = -0.45$, $p = 0.054$) and tibia length ($R = 0.03$, $p = 0.880$) (Fig. 6B). It shows that *PRDX6* in mouse tissues is negatively correlated with the degree of HF.

In the left ventricular tissue samples of the HF data set, we found that *PRDX6* were correlated with 22 immune cell subtypes. *PRDX6*, which is considered the most suitable FAMB, has a greater correlation with CD8 T cells, macrophages M2, macrophages M0, macrophages M1, neutrophils, T cells CD4 memory resting, and mast cells resting. *PRDX6* is positively correlated with macrophages M2 (*Cor* = 0.26, $p < 0.001$), neutrophils (*Cor* = 0.20, $p < 0.001$), and T cells CD4 memory resting (*Cor* = 0.14, $p < 0.001$), whereas it is negatively correlated with macrophages M0 (*Cor* = −0.13, $p = 0.013$), macrophages M1 (*Cor* = −0.11, $p = 0.041$), mast cells resting (*Cor* = −0.19, $p < 0.001$), and CD8 T cells (*Cor* = −0.12, $p = 0.030$) (Fig. 6C).

CIBERSORT implements the tissue immune infiltration landscape and eliminates immune cell types with zero expression. The heat map shows the difference in the content of immune cells in samples classified according to the information of experimental animals and *PRDX6* expression (Fig. 7A). GSEA analysis found that chemical carcinogenesis−reactive oxygen species, MAPK signaling pathway, PI3K-Akt signaling pathway, oxidative phosphorylation, TGF-beta signaling pathway, dilated cardiomyopathy, and other pathways related to HF and ferroptosis were enriched (Fig. 7B). There are differences in activated NK cells ($p = 0.044$), monocytes ($p = 0.034$), macrophages M0 ($p = 0.005$), macrophages M2 ($p = 0.002$), and dendritic cells resting ($p < 0.001$) between the Sham group and the Treated group (Fig. 7C). Analysis of the correlation between *PRDX6* and macrophages found that *PRDX6* was negatively correlated with macrophages M0 ($R = -0.64$, $p = 0.002$) and positively correlated with macrophages M2 ($R = 0.58$, $p = 0.008$) (Fig. 7D).

Discussion

Heart failure, as the outcome of various cardiovascular diseases, has become one of the critical threats to the health of the aging population. Heart failure is one of the main causes of death among the elderly in the Western world (Florio et al. 2020; Triposkiadis et al. 2020). Patients with HF are often accompanied by severe symptoms, restricted activities, and reduced quality of life. Moreover, HF will continue to be one of the health problems that humans need to address in the future (Butler et al. 2020). In recent years, some biomarkers, echocardiography, and cardiac catheterization have improved the diagnosis and prognosis of HF (Henning et al. 2020). However, early discovery of new key molecular biological characteristics of HF can still provide significant support for the pathophysiological characteristics of HF, reduce mortality, and improve prognosis.

The new concept of programmed cell death, ferroptosis, proposed in 2012, has attracted our attention. This novel type of iron-dependent cell death cannot be simply explained by H_2O_2-dependent and iron-catalyzed ROS production (Dixon et al. 2012). At present, it is found in research that the main biochemical characteristics of ferroptosis are iron overload caused by iron homeostasis or lipid peroxidation caused by the accumulation of lipid ROS (Wu et al. 2021). The decrease in the activity of glutathione peroxidase (*GPX4*) breaks the metabolic reaction of lipid oxides. Moreover, Fe^{2+} produces ROS and promotes ferroptosis. In ferroptosis, a large number of cell metabolism processes are involved. In addition, mitochondria are considered to be the key sites for specific ferroptosis. A large number of studies have found that ferroptosis is closely related to the pathological process of tumors and ischemic diseases (Liang et al. 2019; Li K et al. 2021). Ferroptosis is related to acute kidney injury (Friedmann et al. 2021), stroke (Alim et al. 2019; Wan et al. 2019), Parkinson's disease (Guiney et al. 2017), diabetes (Li S et al. 2021), etc. Concurrently, the latest research also found that ferroptosis, as a new therapeutic target for cardiovascular disease, affects cardiomyopathy (Zhang et al. 2021), myocardial infarction (Lillo et al. 2021), HF (Ning et al. 2021), etc.

Advances in machine learning are key productivity tools in the modern world (Goecks et al. 2020). Machine learning uses certain modeling and prediction methods to identify patterns in high-throughput sequencing data sets and then learn to explore new areas of prediction. A large number of diagnostic models based on machine learning are used in the clinical field. Studies have reported that a joint diagnostic model based on random forest has been constructed to analyze the biomarkers of HF (Tian et al. 2020).

In the present study, WGCNA analysis combined with a machine learning dimensionality reduction method was used to identify potential ferroptosis-associated molecular biomarkers in HF. Based on independent external verification of multiple microarray data sets and RNA-seq data sets, *PRDX6* was selected for further analysis. *PRDX6* is closely related to clinical characteristics and HF characteristics in mice experiments, with the expression levels significantly reduced in the disease groups in different data sets. Subsequently, through CIBERSORT analysis of the immune infiltration landscape, we also found that *PRDX6* is related to a variety of immune cells, which may play a critical role in the pathophysiology of HF through immune-related pathways. This study determined the value of the ferroptosis-associated genes *PRDX6* in HF, indicating that it can be used as a report on the potential of ferroptosis-associated molecular biomarkers in HF. Although there are studies on the determination of HF biomarkers based on gene expression data from public databases, no studies

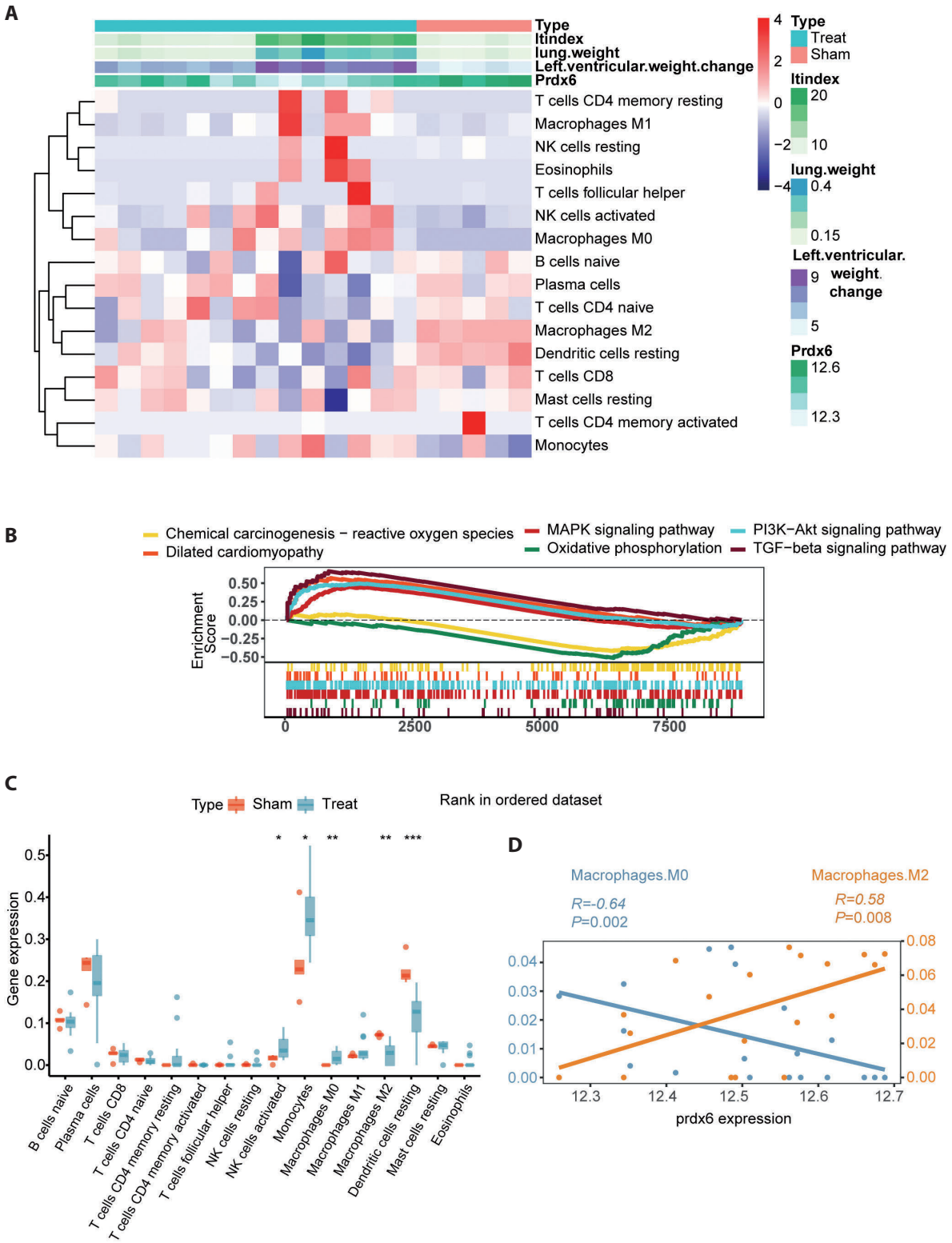


Figure 7. Immune infiltration and correlation verification. **A.** Results of immune infiltration in animal experiments. **B.** Six pathways enriched by GSEA. **C.** Immune subcellular difference between Sham group and Treated group. Red represents the Sham group, blue represents the Treated group, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **D.** The connection between *PRDX6* and immune cells obtained from GSE36074. *PRDX6* is positively correlated with macrophage M2 and negatively correlated with macrophage M0.

have been found to analyze ferroptosis-associated genes in HF using bioinformatics (Huang et al. 2018). Similarly, previous studies only used differentially expressed genes (DEG) to simply screen for biomarkers, which are possible confounding factors and not closely related to clinical features. Weighted gene co-expression network analysis is used to identify highly coordinated gene sets related to HF clinical information (Langfelder et al. 2018). As a widely used bioinformatics method, WGCNA can better remove the influence of confounding factors on biomarkers in HF. The enrichment analysis of the characteristic genes in the modules determined by WGCNA found that the virus response, extracellular structure, T cell regulation, protein ribose binding, protein complex, collagen trimer and other parts are more concentrated. GSEA found that the MAPK signaling pathway, PI3K-Akt signaling pathway, B cell receptor signaling pathway, cellular senescence, phagosomes, and other pathways related to HF are involved. Studies have reported the potential role of the MAPK signaling pathway and the PI3K-Akt signaling pathway in HF (Mutlak et al. 2021; Qin et al. 2021). The FAG collection related to HF was obtained by integrating the characteristic genes in the module with 259 ferroptosis-associated genes discovered in previous studies. To obtain FAMB, we used machine learning algorithms to screen the 25 FAGs obtained from the previous integration. By integrating the nine FAGs obtained after LASSO screening and the seven FAGs obtained after SVM-RFE screening, we identified six FAMBs (namely, *STEAP3*, *ATF4*, *NF2*, *PRDX6*, *RGS4*, and *SLC40A1*). Through CIBERSORT, our research analyzed the differences in immune infiltrating cells between HF tissues and normal control tissues. In addition, we analyzed the relationship between immune infiltrating cells and clinical features of HF. Comparison of HF and normal tissues revealed that a total of 9 immune cells are unbalanced (CD8 T cells, mast cells resting, macrophages M2, T cells CD4 memory resting, regulatory T cells, monocytes, neutrophils, naive CD4 T cells, and macrophages M0). The role of monocytes and macrophages in HF has received a great deal of attention (Glezeva et al. 2015a; Duncan et al. 2020). Research reports have found that macrophage M2 receptors are reduced in the peripheral blood of patients with HF and mouse models (Glezeva et al. 2015b; Zhang et al. 2021). There are also differences in T lymphocyte subsets (CD8T cells, T cells CD4 memory resting, T cells regulatory, and naive CD4 T cells) in the tissues of patients with HF. HF is associated with the accumulation of T lymphocytes and CD4 T cells in the heart tissue (Laroumanie et al. 2014; Youn et al. 2019). CD4 T cells can promote the transition from cardiac hypertrophy to HF under pressure overload. Similarly, the findings of the current study also support the correlation between clinical features of HF and macrophage subpopulations and T lymphocyte subpopulations. To explore the value of the six FAMBs, we further verified the RNA-seq data sets and finally believed that *PRDX6* was of excellent importance. This was verified on a dataset with NYHA classification and C57BL/6 experimental mouse information. Peroxiredoxin 6 (*PRDX6*) is a member of the nonselenium peroxidase family, and it can play a role in regulating phospholipid conversion and preventing oxidative damage. Studies have found that *PRDX6* can play a critical role in the repair of peroxidized cell membranes through two mechanisms (Fisher et al. 2017, 2018). Moreover, *PRDX6* and *GPX4* have similar enzymatic activities. Because of the core position of *GPX4* in ferroptosis, the role of *PRDX6* in ferroptosis is worth exploring (Lu et al. 2019). Previous studies have shown that peroxidase can be used as a potential target for cardiovascular disease (Jeong et al. 2021). A study in mice found that Peroxiredoxin 1 can prevent cardiac hypertrophy and HF caused by pressure overload by activating Nrf2/HO-1 signaling (Tang et al. 2020). Two recent basic studies on cardiovascular diseases have found that Peroxiredoxin 6 overexpression regulates doxorubicin-induced myocardial injury in rats (Guo et al. 2020). Overexpression of *PRDX6* can prevent Ang II-induced inflammation and endothelial dysfunction in human umbilical vein endothelial cells (Li et al. 2020). Similarly, in the subsequent correlation analysis, we found that *PRDX6* was significantly negatively correlated with experimental features related to HF in mice. These studies are consistent with the conclusions of the current study. *PRDX6*, as a ferroptosis-related gene, may have a protective effect on myocardial damage during the progression of HF. Subsequently, we conducted a correlation analysis between *PRDX6* and immune cell subtypes and found that it is also related to macrophage subpopulations and T lymphocyte subpopulations. In the subsequent estimation of immune infiltration, we found that activated NK cells, monocytes, macrophages M0, macrophages M2, and dendritic cells resting have significant differences between the mouse treatment group and the Sham group. Combined with the previous results, correlation analysis shows that *PRDX6* and macrophages M2 are positively correlated, whereas macrophages M0 are negatively correlated with *PRDX6*. Studies have found that more and more pathophysiological processes involve ferroptosis, accompanied by dysregulated immune responses (Chen et al. 2021). The relationship between ferroptosis and immune cells has also become a hot topic. For example, in atherosclerosis, ferroptosis can occur in a variety of cells, such as macrophages (Ouyang et al. 2021). *PRDX6* plays an important role in protecting cells from ferroptosis, and simultaneously, M2 in macrophages may have an inhibitory effect on lipid peroxidation produced by ferroptosis (Li and Huang 2021). Therefore, our research on ferroptosis-related biomarkers will help to discover new therapeutic targets, and when combined with immune cell analysis, it will help to provide new ideas for immunotherapy.

There are still several limitations in this study. First, our expression verification is performed on the tissues, and FAMB for serum or plasma verification and screening is still needed to be performed. Second, the functional study of *PRDX6* and ferroptosis in HF still needs experimental validation. Finally, a single biomarker lacks sufficient accuracy. Therefore, in future studies, we need to combine blood biochemical indicators to improve the verification of FAMB.

Conclusion

In this study, integrated bioinformatics and machine learning strategies identified *PRDX6* as a key molecular biological feature related to ferroptosis in HF. Ferroptosis-associated molecular markers combined with immune cell analysis can help provide new ideas for immunotherapy in HF.

Data availability. The data used in this study are available at the following links: Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions. WDJ proposed the idea and designed this study. CYJ wrote and improved the manuscript. CYJ carried out a literature survey. All authors read and approved the final manuscript.

Conflict of interests. The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments. We acknowledge the GEO database for providing their platforms and contributors for uploading their meaningful datasets. And we thank all participants involved in this study.

References

Alim I, Caulfield JT, Chen Y, Swarup V, Geschwind DH, Ivanova E, Seravalli J, Ai Y, Sansing LH, Ste Marie EJ, et al. (2019): Selenium drives a transcriptional adaptive program to block ferroptosis and treat stroke. Cell **177,** 1262-1279 https://doi.org/10.1016/j.cell.2019.03.032

Bai T, Li M, Liu Y, Qiao Z, Wang Z (2020): Inhibition of ferroptosis alleviates atherosclerosis through attenuating lipid peroxidation and endothelial dysfunction in mouse aortic endothelial cell. Free Radic. Biol. Med. **160,** 92-102 https://doi.org/10.1016/j.freeradbiomed.2020.07.026

Bebber CM, Müller F, Prieto Clemente L, Weber J, von Karstedt S (2020): Ferroptosis in cancer cell biology. Cancers **12,** 164 https://doi.org/10.3390/cancers12010164

Butler J, Khan MS, Mori C, Filippatos GS, Ponikowski P, Comin-Colet J, Roubert B, Spertus JA, Anker SD (2020): Minimal clinically important difference in quality of life scores for patients with heart failure and reduced ejection fraction. Eur. J. Heart. Fail. **22,** 999-1005 https://doi.org/10.1002/ejhf.1810

Chen X, Xu S, Zhao C, Liu B (2019): Role of TLR4/NADPH oxidase 4 pathway in promoting cell death through autophagy and ferroptosis during heart failure. Biochem. Biophys. Res. Commun. **516,** 37-43 https://doi.org/10.1016/j.bbrc.2019.06.015

Chen X, Kang R, Kroemer G, Tang D (2021): Ferroptosis in infection, inflammation, and immunity. J. Exp. Med. **218,** e20210518 https://doi.org/10.1084/jem.20210518

Dai C, Chen X, Li J, Comish P, Kang R, Tang D (2020): Transcription factors in ferroptotic cell death. Cancer Gene Ther. **27,** 645-656 https://doi.org/10.1038/s41417-020-0170-2

Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, Patel DN, Bauer AJ, Cantley AM, Yang WS, et al. (2012): Ferroptosis: An iron-dependent form of non-apoptotic cell death. Cell **149,** 1060-1072 https://doi.org/10.1016/j.cell.2012.03.042

Duncan SE, Gao S, Sarhene M, Coffie JW, Linhua D, Bao X, Jing Z, Li S, Guo R, Su J, et al (2020): Macrophage activities in myocardial infarction and heart failure. Cardiol. Res. Pract. **2020,** e4375127 https://doi.org/10.1155/2020/4375127

Fang X, Cai Z, Wang H, Han D, Cheng Q, Zhang P, Gao F, Yu Y, Song Z, Wu Q, et al. (2020): Loss of cardiac ferritin H facilitates cardiomyopathy via Slc7a11-mediated ferroptosis. Circ. Res. **127,** 486-501 https://doi.org/10.1161/CIRCRESAHA.120.316509

Fang X, Wang H, Han D, Xie E, Yang X, Wei J, Gu S, Gao F, Zhu N, Yin X, et al. (2019): Ferroptosis as a target for protection against cardiomyopathy. Proc. Natl. Acad. Sci. USA **116,** 2672-2680 https://doi.org/10.1073/pnas.1821022116

Fisher AB (2017): Peroxiredoxin 6 in the repair of peroxidized cell membranes and cell signaling. Arch. Biochem. Biophys. **617,** 68-83 https://doi.org/10.1016/j.abb.2016.12.003

Fisher AB, Vasquez-Medina JP, Dodia C, Sorokina EM, Tao J-Q, Feinstein SI (2018): Peroxiredoxin 6 phospholipid hydroperoxidase activity in the repair of peroxidized cell membranes. Redox Biol. **14,** 41-46 https://doi.org/10.1016/j.redox.2017.08.008

Florio MC, Magenta A, Beji S, Lakatta EG, Capogrossi MC (2020): Aging, microRNAs, and heart failure. Curr. Probl. Cardiol. **45,** 100406 https://doi.org/10.1016/j.cpcardiol.2018.12.003

Friedmann Angeli JP, Schneider M, Proneth B, Tyurina YY, Tyurin VA, Hammond VJ, Herbach N, Aichler M, Walch A, Eggenhofer E, et al. (2014): Inactivation of the ferroptosis regulator Gpx4 triggers acute renal failure in mice. Nat. Cell. Biol. **16,** 1180-1191 https://doi.org/10.1038/ncb3064

Glezeva N, Horgan S, Baugh JA (2015a). Monocyte and macrophage subsets along the continuum to heart failure: Misguided heroes or targetable villains? J. Mol. Cell. Cardiol. **89,** 136-145 https://doi.org/10.1016/j.yjmcc.2015.10.029

Glezeva N, Voon V, Watson C, Horgan S, McDonald K, Ledwidge M, Baugh J (2015b): Exaggerated inflammation and monocytosis associate with diastolic dysfunction in heart failure

with preserved ejection fraction: Evidence of M2 macrophage activation in disease pathogenesis. J. Card. Fail. **21,** 167-177 https://doi.org/10.1016/j.cardfail.2014.11.004

Goecks J, Jalili V, Heiser LM, Gray JW (2020): How machine learning will transform biomedicine. Cell **181,** 92-101 https://doi.org/10.1016/j.cell.2020.03.022

Guiney SJ, Adlard PA, Bush AI, Finkelstein DI, Ayton S (2017): Ferroptosis and cell death mechanisms in Parkinson's disease. Neurochem. Int. **104,** 34-48 https://doi.org/10.1016/j.neuint.2017.01.004

Guo J, Cao W, Chen C, Chen X (2020): Peroxiredoxin 6 overexpression regulates adriamycin-induced myocardial injury, oxidative stress and immune response in rats. Ann. Transl. Med. **8,** 1320 https://doi.org/10.21037/atm-20-6598

Henning RJ (2020): Diagnosis and treatment of heart failure with preserved left ventricular ejection fraction. World J. Cardiol. **12,** 7-25 https://doi.org/10.4330/wjc.v12.i1.7

Hu H, Chen Y, Jing L, Zhai C, Shen L (2021): The link between ferroptosis and cardiovascular diseases: A novel target for treatment. Front. Cardiovasc. Med. **8,** 710963 https://doi.org/10.3389/fcvm.2021.710963

Hu SS, Kong LZ, Gao RL, Zhu ML, Wang W, Wang YJ, Wu ZS, Chen WW, Liu MB, Editorial Board (2012): Outline of the report on cardiovascular disease in China, 2010. Biomed. Environ. Sci. **25,** 251-256

Huang H, Luo B, Wang B, Wu Q, Liang Y, He Y (2018): Identification of potential gene interactions in heart failure caused by idiopathic dilated cardiomyopathy. Med. Sci. Monit. **24,** 7697-7709 https://doi.org/10.12659/MSM.912984

Jeong S-J, Park J-G, Oh GT (2021): Peroxiredoxins as potential targets for cardiovascular disease. Antioxidants **10,** 1244 https://doi.org/10.3390/antiox10081244

Kim AH, Jang JE, Han J (2021): Current status on the therapeutic strategies for heart failure and diabetic cardiomyopathy. Biomed. Pharmacother. **145,** 112463 https://doi.org/10.1016/j.biopha.2021.112463

Langfelder P, Horvath S (2008): WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics **9,** 559 https://doi.org/10.1186/1471-2105-9-559

Laroumanie F, Douin-Echinard V, Pozzo J, Lairez O, Tortosa F, Vinel C, Delage C, Calise D, Dutaur M, Parini A, et al. (2014): CD4+ T cells promote the transition from hypertrophy to heart failure during chronic pressure overload. Circulation **129,** 2111-2124 https://doi.org/10.1161/CIRCULATIONAHA.113.007101

Latunde-Dada GO (2017): Ferroptosis: Role of lipid peroxidation, iron and ferritinophagy. Biochim. Biophys. Acta Gen. Subj. **1861,** 1893-1900 https://doi.org/10.1016/j.bbagen.2017.05.019

Li D-X, Chen W, Jiang Y-L, Ni J-Q, Lu L (2020): Antioxidant protein peroxiredoxin 6 suppresses the vascular inflammation, oxidative stress and endothelial dysfunction in angiotensin II-induced endotheliocyte. Gen. Physiol. Biophys. **39,** 545-555 https://doi.org/10.4149/gpb_2020029

Li K, Xu K, He Y, Lu L, Mao Y, Gao P, Liu G, Wu J, Zhang Y, Xiang Y, et al. (2021): Functionalized tumor-targeting nanosheets exhibiting Fe (II) overloading and GSH consumption for ferroptosis activation in liver tumor. Small **17,** e2102046 https://doi.org/10.1002/smll.202102046

Li S, Huang Y (2021): Ferroptosis: an iron-dependent cell death form linking metabolism, diseases, immune cell and targeted therapy. Clin. Transl. Oncol. **24,** 1-12 https://doi.org/10.1007/s12094-021-02669-8

Li S, Li Y, Wu Z, Wu Z, Fang H (2021): Diabetic ferroptosis plays an important role in triggering on inflammation in diabetic wound. Am. J. Physiol. Endocrinol. Metab. **321,** E509-E520 https://doi.org/10.1152/ajpendo.00042.2021

Liang C, Zhang X, Yang M, Dong X (2019): Recent progress in ferroptosis inducers for cancer therapy. Adv. Mater. **31,** e1904197 https://doi.org/10.1002/adma.201904197

Lillo-Moya J, Rojas-Solé C, Muñoz-Salamanca D, Panieri E, Saso L, Rodrigo R (2021): Targeting ferroptosis against ischemia/reperfusion cardiac injury. Antioxidants **10,** 667 https://doi.org/10.3390/antiox10050667

Liu B, Zhao C, Li H, Chen X, Ding Y, Xu S (2018): Puerarin protects against heart failure induced by pressure overload through mitigation of ferroptosis. Biochem. Biophys. Res. Commun. **497,** 233-240 https://doi.org/10.1016/j.bbrc.2018.02.061

Lu B, Chen X, Hong Y, Zhu H, He Q, Yang B, Ying M, Cao J (2019): Identification of PRDX6 as a regulator of ferroptosis. Acta. Pharmacol. Sin. **40,** 1334-1342 https://doi.org/10.1038/s41401-019-0233-9

Mou Y, Wang J, Wu J, He D, Zhang C, Duan C, Li B (2019): Ferroptosis, a new form of cell death: opportunities and challenges in cancer. J. Hematol. Oncol. **12,** 34 https://doi.org/10.1186/s13045-019-0720-y

Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Després J-P, et al. (2016): Heart disease and stroke statistics-2016 update: A report from the American Heart Association. Circulation **133,** e38-360 https://doi.org/10.1161/CIR.0000000000000366

Mutlak M, Kehat I (2021): Dual specific phosphatases (DUSPs) in cardiac hypertrophy and failure. Cell Signal. **84,** 110033 https://doi.org/10.1016/j.cellsig.2021.110033

Ning D, Yang X, Wang T, Jiang Q, Yu J, Wang D (2021): Atorvastatin treatment ameliorates cardiac function and remodeling induced by isoproterenol attack through mitigation of ferroptosis. Biochem. Biophys. Res. Commun. **574,** 39-47 https://doi.org/10.1016/j.bbrc.2021.08.017

Ouyang S, You J, Zhi C, Li P, Lin X, Tan X, Ma W, Li L, Xie W (2021): Ferroptosis: the potential value target in atherosclerosis. Cell Death Dis. **12,** 1-9 https://doi.org/10.1038/s41419-021-04054-3

Qin W, Cao L, Massey IY (2021): Role of PI3K/Akt signaling pathway in cardiac fibrosis. Mol. Cell. Biochem. **476,** 4045-4059 https://doi.org/10.1007/s11010-021-04219-w

Song Y, Wang B, Zhu X, Hu J, Sun J, Xuan J, Ge Z (2021): Human umbilical cord blood-derived MSCs exosome attenuate myocardial injury by inhibiting ferroptosis in acute myocardial infarction mice. Cell Biol. Toxicol. **37,** 51-64 https://doi.org/10.1007/s10565-020-09530-8

Stockwell BR, Friedmann Angeli JP, Bayir H, Bush AI, Conrad M, Dixon SJ, Fulda S, Gascón S, Hatzios SK, Kagan VE, et al. (2017): Ferroptosis: A regulated cell death nexus linking metabolism, redox biology, and disease. Cell **171,** 273-285

https://doi.org/10.1016/j.cell.2017.09.021
Tang C, Yin G, Huang C, Wang H, Gao J, Luo J, Zhang Z, Wang J, Hong J, Chai X (2020): Peroxiredoxin-1 ameliorates pressure overload-induced cardiac hypertrophy and fibrosis. Biomed. Pharmacother. **129,** 110357
https://doi.org/10.1016/j.biopha.2020.110357
Tian Y, Yang J, Lan M, Zou T (2020): Construction and analysis of a joint diagnosis model of random forest and artificial neural network for heart failure. Aging **12,** 26221-26235
https://doi.org/10.18632/aging.202405
Triposkiadis F, Xanthopoulos A, Parissis J, Butler J, Farmakis D (2020): Pathogenesis of chronic heart failure: cardiovascular aging, risk factors, comorbidities, and disease modifiers. Heart Fail. Rev. **27,** 337-344
https://doi.org/10.1007/s10741-020-09987-z
Wan J, Ren H, Wang J (2019): Iron toxicity, lipid peroxidation and ferroptosis after intracerebral haemorrhage. Stroke Vasc. Neurol. **4,** 93-95
https://doi.org/10.1136/svn-2018-000205
Wu X, Li Y, Zhang S, Zhou X (2021): Ferroptosis as a novel therapeutic target for cardiovascular disease. Theranostics **11,** 3052-3059
https://doi.org/10.7150/thno.54113
Youn J-C, Jung MK, Yu HT, Kwon J-S, Kwak J-E, Park S-H, Kim I-C, Park M-S, Lee SK, Choi S-W, et al. (2019): Increased frequency of CD4+CD57+ senescent T cells in patients with newly diagnosed acute heart failure: exploring new pathogenic mechanisms with clinical relevance. Sci. Rep. **9,** 12887
https://doi.org/10.1038/s41598-019-49332-5
Yu H, Guo P, Xie X, Wang Y, Chen G (2017): Ferroptosis, a new form of cell death, and its relationships with tumourous diseases. J. Cell Mol. Med. **21,** 648-657
https://doi.org/10.1111/jcmm.13008
Yu Y, Yan Y, Niu F, Wang Y, Chen X, Su G, Liu Y, Zhao X, Qian L, Liu P, et al. (2021): Ferroptosis: a cell death connecting oxidative stress, inflammation and cardiovascular diseases. Cell Death. Discov. **7,** 193
https://doi.org/10.1038/s41420-021-00579-w
Zhang H, Wang Z, Liu Z, Du K, Lu X (2021): protective effects of dexazoxane on rat ferroptosis in doxorubicin-induced cardiomyopathy through regulating HMGB1. Front. Cardiovasc. Med. **8,** 685434
https://doi.org/10.3389/fcvm.2021.685434
Zhang L, Chen J, Yan L, He Q, Xie H, Chen M (2021): Resveratrol ameliorates cardiac remodeling in a murine model of heart failure with preserved ejection fraction. Front. Pharmacol. **12,** 646240
https://doi.org/10.3389/fphar.2021.646240
Zhang X, Zheng C, Gao Z, Chen H, Li K, Wang L, Zheng Y, Li C, Zhang H, Gong M, et al. (2021): SLC7A11/xCT prevents cardiac hypertrophy by inhibiting ferroptosis. Cardiovasc. Drugs. Ther. **36,** 437-447
https://doi.org/10.1007/s10557-021-07220-z

Received: April 19, 2022
Final version accepted: May 30, 2022

doi: 10.4149/gpb_2022029

Supplementary Material

Integrated bioinformatics and machine learning strategies reveal *PRDX6* as the key ferroptosis-associated molecular biosignature of heart failure

Chenyang Jiang[1,2] and Weidong Jiang[3]

[1] *Department of Cardiology, The Second Affiliated Hospital of Nantong University, Nantong 226000, China*
[2] *The First Clinical Medical College of Guangxi Medical University, Nanning 530021, China*
[3] *Department of Cardiology, Nantong Hospital of Traditional Chinese Medicine, Nantong 226000, China*

Supplementary Table

Table S1. Gene set information

Gene Set	*Species*	*Data set type*	*Tissue data*	*Number of samples*	
				Disease group	Control group
GSE57338	Human	microarray data	heart ventricle	177	136
GSE71613	Human	High throughput data	heart ventricle	4	4
GSE161472	Human	High throughput data	heart ventricle	47	37
GSE135055	Human	High throughput data	heart ventricle	21	9
GSE36074	Mouse	microarray data	heart ventricle	14	5

Table S2. 259 ferroptosis-associated genes (FAG) were obtained after collating from the FerrDb database (http://zhounan.org/ferrdb) and previous literature

Gene Symbol	Role	Gene Symbol	Role	Gene Symbol	Role
RPL8	driver	*SCP2*	driver	*ACO1*	driver
ATP5MC3	driver	*TP53*	driver	*SLC38A1*	driver
CS	driver	*LPCAT3*	driver	*GLS2*	driver
EMC2	driver	*NRAS*	driver	*G6PDX*	driver
NOX1	driver	*KRAS*	driver	*ULK1*	driver
CYBB	driver	*HRAS*	driver	*ATG3*	driver
NOX3	driver	*TFR2*	driver	*ATG4D*	driver
NOX5	driver	*GOT1*	driver	*ATG5*	driver
DUOX1	driver	*CARS1*	driver	*BECN1*	driver
DUOX2	driver	*ATG7*	driver	*MAP1LC3A*	driver
G6PD	driver	*NCOA4*	driver	*GABARAPL2*	driver
PGD	driver	*ALOX12B*	driver	*GABARAPL1*	driver
VDAC2	driver	*ALOX15B*	driver	*ATG16L1*	driver
PIK3CA	driver	*ALOXE3*	driver	*WIPI1*	driver
FLT3	driver	*PHKG2*	driver	*WIPI2*	driver

(continued)

Table S2. (continued)

Gene Symbol	Role
SNX4	driver
ATG13	driver
ULK2	driver
SAT1	driver
EGFR	driver
MAPK3	driver
MAPK1	driver
BID	driver
ZEB1	driver
KEAP1	driver
DPP4	driver
ALOX15	driver
CDKN2A	driver
PEBP1	driver
SOCS1	driver
CDO1	driver
MYB	driver
MAPK8	driver
MAPK9	driver
SLC1A5	driver
CHAC1	driver
MAPK14	driver
LINC00472	driver
PRKAA2	driver
PRKAA1	driver
ABCC1	driver
MIR6852	driver
ACVR1B	driver
TGFBR1	driver
BAP1	driver
EPAS1	driver
HILPDA	driver
HIF1A	driver
ALOX12	driver
ACSL4	driver
HMOX1	driver
IFNG	driver
ANO6	driver
LPIN1	driver
HMGB1	driver
TNFAIP3	driver
TLR4	driver
NOX4	driver
ATF3	driver
ATM	driver
YY1AP1	driver
EGLN2	driver
MIOX	driver
TAZ	driver
MTDH	driver
IDH1	driver
SIRT1	driver
FBXW7	driver
PANX1	driver
DNAJB6	driver
BACH1	driver
LONP1	driver
PTGS2	marker
DUSP1	marker
NOS2	marker
NCF2	marker
MT3	marker
UBC	marker
ALB	marker
TXNRD1	marker
SRXN1	marker
GPX2	marker
BNIP3	marker
OXSR1	marker
SELENOS	marker
ANGPTL7	marker
DDIT4	marker
LOC284561	marker
ASNS	marker
TSC22D3	marker
DDIT3	marker
JDP2	marker
SESN2	marker
SLC1A4	marker
PCK2	marker
TXNIP	marker
VLDLR	marker
GPT2	marker
PSAT1	marker
LURAP1L	marker
SLC7A5	marker
HERPUD1	marker
XBP1	marker
CBS	marker
ZNF419	marker
KLHL24	marker
TRIB3	marker
ZFP69B	marker
ATP6V1G2	marker
VEGFA	marker
GDF15	marker
TUBE1	marker
ARRDC3	marker
CEBPG	marker
SNORA16A	marker
RGS4	marker
BLOC1S5-TXNDC5	marker
LOC390705	marker
KIM-1	marker
IL6	marker
CXCL2	marker
RELA	marker
HSD17B11	marker
AGPAT3	marker
SETD1B	marker
FTL	marker
MAFG	marker
IL33	marker
SLC40A1	marker
TF	marker
TFRC	marker
FTH1	marker
HAMP	marker
STEAP3	marker
DRD5	marker
DRD4	marker
MAP3K5	marker
SLC2A1	marker
SLC2A3	marker
SLC2A6	marker
SLC2A8	marker
SLC2A12	marker
GLUT13	marker
SLC2A14	marker
EIF2AK4	marker
EIF2S1	marker
ATF4	marker
ALOX5	marker
ACSF2	marker
IREB2	marker
GPX4	marker
NFE2L2	marker
ELAVL1	marker
SLC3A2	marker
TFAP2C	marker
SP1	marker
HBA1	marker
NNMT	marker
PLIN4	marker

(continued)

Table S2. (continued)

Gene Symbol	Role
HIC1	marker
STMN1	marker
RRM2	marker
CAPG	marker
HNF4A	marker
NGB	marker
YWHAE	marker
GABPB1	marker
AURKA	marker
MIR4715	marker
RIPK1	marker
PRDX1	marker
MIR30B	marker
SLC7A11	suppressor
AKR1C1	suppressor
AKR1C2	suppressor
AKR1C3	suppressor
RB1	suppressor
HSPB1	suppressor
HSF1	suppressor
GCLC	suppressor
SQSTM1	suppressor
NQO1	suppressor
MUC1	suppressor

Gene Symbol	Role
MT1G	suppressor
CISD1	suppressor
FANCD2	suppressor
FTMT	suppressor
HELLS	suppressor
FADS2	suppressor
SRC	suppressor
STAT3	suppressor
PML	suppressor
MTOR	suppressor
NFS1	suppressor
TP63	suppressor
CDKN1A	suppressor
MIR137	suppressor
ENPP2	suppressor
FH	suppressor
CISD2	suppressor
MIR9-1	suppressor
MIR9-2	suppressor
MIR9-3	suppressor
ISCU	suppressor
ACSL3	suppressor
OTUB1	suppressor
CD44	suppressor

Gene Symbol	Role
LINC00336	suppressor
BRD4	suppressor
PRDX6	suppressor
MIR17	suppressor
SCD	suppressor
NF2	suppressor
ARNTL	suppressor
JUN	suppressor
CA9	suppressor
TMBIM4	suppressor
HSPA5	suppressor
PLIN2	suppressor
MIR212	suppressor
Fer1HCH	suppressor
AIFM2	suppressor
LAMP2	suppressor
ZFP36	suppressor
PROM2	suppressor
CHMP5	suppressor
CHMP6	suppressor
CAV1	suppressor
GCH1	suppressor

Table S3. Results of LASSO and SVM-RFE

LASSO

1	LPCAT3
2	RGS4
3	SLC40A1
4	ATF4
5	NF2
6	PHKG2
7	PRDX6
8	SLC2A1
9	STEAP3

SVM-RFE

	FeatureName	FeatureID	AvgRank
1	STEAP3	24	2
2	ATF4	16	3.2
3	NF2	19	4.4
4	PRDX6	22	5.2
5	ZEB1	15	5.6
6	RGS4	13	6.2
7	SLC40A1	14	6.6
8	MYB	11	12
9	GPX2	18	12.6