Heliyon 10 (2024) e27049

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification of circadian rhythm-related gene classification patterns and immune infiltration analysis in heart failure based on machine learning

Xuefu Wang [a,1], Jin Rao [b,1], Li Zhang [c,1], Xuwen Liu [c], Yufeng Zhang [a,b,*]

[a] School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai, China
[b] Department of Cardiothoracic Surgery, Shanghai Changzheng Hospital, Naval Medical University, Shanghai, China
[c] Guangxi University, Nanning, China

ARTICLE INFO

Keywords:
Heart failure
Circadian rhythm
Machine learning
Immune infiltration
Unsupervised clustering
Bioinformatics

ABSTRACT

Background: Circadian rhythms play a key role in the failing heart, but the exact molecular mechanisms linking changes in the expression of circadian rhythm-related genes to heart failure (HF) remain unclear.

Methods: By intersecting differentially expressed genes (DEGs) between normal and HF samples in the Gene Expression Omnibus (GEO) database with circadian rhythm-related genes (CRGs), differentially expressed circadian rhythm-related genes (DE-CRGs) were obtained. Machine learning algorithms were used to screen for feature genes, and diagnostic models were constructed based on these feature genes. Subsequently, consensus clustering algorithms and non-negative matrix factorization (NMF) algorithms were used for clustering analysis of HF samples. On this basis, immune infiltration analysis was used to score the immune infiltration status between HF and normal samples as well as among different subclusters. Gene Set Variation Analysis (GSVA) evaluated the biological functional differences among subclusters.

Results: 13 CRGs showed differential expression between HF patients and normal samples. Nine feature genes were obtained through cross-referencing results from four distinct machine learning algorithms. Multivariate LASSO regression and external dataset validation were performed to select five key genes with diagnostic value, including NAMPT, SERPINA3, MAPK10, NPPA, and SLC2A1. Moreover, consensus clustering analysis could divide HF patients into two distinct clusters, which exhibited different biological functions and immune characteristics. Additionally, two subgroups were distinguished using the NMF algorithm based on circadian rhythm associated differentially expressed genes. Studies on immune infiltration showed marked variances in levels of immune infiltration between these subgroups. Subgroup A had higher immune scores and more widespread immune infiltration. Finally, the Weighted Gene Co-expression Network Analysis (WGCNA) method was utilized to discern the modules that had the closest association with the two observed subgroups, and hub genes were pinpointed via protein-protein interaction (PPI) networks. GRIN2A, DLG1, ERBB4, LRRC7, and NRG1 were circadian rhythm-related hub genes closely associated with HF.

Conclusion: This study provides valuable references for further elucidating the pathogenesis of HF and offers beneficial insights for targeting circadian rhythm mechanisms to regulate immune

* Corresponding author. Department of Cardiothoracic Surgery, Shanghai Changzheng Hospital, Naval Medical University, Shanghai, China.
E-mail address: zhyf19810824@163.com (Y. Zhang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e27049
Received 25 May 2023; Received in revised form 17 December 2023; Accepted 22 February 2024
Available online 9 March 2024

responses and energy metabolism in HF treatment. Five genes identified by us as diagnostic features could be potential targets for therapy for HF.

1. Introduction

Heart failure (HF) represents a complex clinical syndrome and is among the primary causes of mortality globally [1]. At present, there are limited treatment options for heart failure (HF), resulting in a poor prognosis and a 5-year survival rate of approximately 50% [2,3]. From a clinical viewpoint, heart failure represents the final outcome of numerous cardiac disorders in which the heart fails to pump blood efficiently enough to match the oxygen and nutrient requirements of tissues and organs, ultimately leading to impaired bodily function [4].While some progress has been made in HF research in recent decades, many areas regarding the causes and molecular mechanisms of heart failure remain unknown. The causes of heart failure are diverse, encompassing coronary artery disease, cardiomyopathy, cardiac surgery, substance abuse, ageing, and numerous other factors. Moreover, recent scholarship has indicated a possible link between the onset of heart failure and variations in genes that regulate circadian rhythms [4]. Recent evidence indicates that circadian rhythms play a role in the development of multiple cardiovascular diseases, such as myocardial infarction, ischemia-reperfusion injury post-myocardial infarction, and HF [5]. The cardiovascular system is particularly sensitive to circadian rhythm regulation, impacting various cardiovascular processes, including heart rate, signal transduction, contractility, metabolism and vascular tone [6,7]. Thus, comprehending circadian rhythm pattern alterations in patients with HF could better elucidate the relationship between circadian rhythm and HF.

Circadian rhythms are regulated by the molecular clock mechanism across various organs and tissues [8,9]. Mouse models have directly demonstrated significant cardiac phenotypes resulting from the key functions of peripheral core clocks in the heart. These core transcription factors, BMAL1 and CLOCK, regulate the circadian rhythms in the heart through transcription and translation [10]. Gene knockouts that disrupt the cardiac clock have been shown to result in various cardiomyopathies, including dilated cardiomyopathy and time-dependent cardiomyopathy [11,12]. Song et al. found that alterations in REVERB/BMAL1 rhythms are linked with the severity of dilated cardiomyopathy [13]. The deficiency of Bmal1 protein disrupts the rhythmic expression of genes encoding for Pik3r1, Akt, and Gsk3β proteins, ultimately leading to impaired cardiac contractility [14]. Moreover, the disturbance of circadian rhythms in the heart impacts mitochondrial function, leading to elevated levels of oxidative stress and the activation of pathways related to apoptosis, autophagy, inflammation, and remodeling [15]. Disruption of circadian rhythms also impacts the rhythm of cardiac metabolic activities, including fatty acid oxidation and glucose metabolism [16,17]. Moreover, the circadian clock can regulate immune and inflammatory responses, with the circadian/metabolic axis potentially being a key driver of immune function rhythms [18]. It is now known that circadian rhythm disturbances are risk factors for inflammatory bowel disease [19]. However, research on the interrelationships among circadian rhythms, immune responses, and metabolism in HF is still lacking.

The rapid advances in microarray and high-throughput sequencing technologies have expanded our comprehension of the genome, transcriptome, and epigenetics, resulting in a broader knowledge base of these fields [20,21]. A increasing number of researchers are willing to share their sequencing data through the Gene Expression Omnibus (GEO) database. This presents a prospect to unveil comprehensive insights into gene function, disease mechanisms and biological processes [22]. However, microarray or high-throughput sequencing data comprises a vast number of features at every level, which augments data dimensionality, redundancy, and noise. Subsequently, data analysis, visualization, and interpretation pose complex challenges. To tackle these issues, machine learning algorithms are significantly applied for feature selection and creating classification models, thereby enhancing the reliability of feature selection [23,24]. Iain S. Forrest and his co-authors used the Boruta machine learning algorithm to measure biomarkers associated with atherosclerosis and mortality risk using continuous spectra, without invasive procedures [25]. Additionally, LASSO models, SVM models, and Random Forest (RF) have also been employed for constructing risk models for ischemic stroke [26]. Based on this, the cross-combination of machine learning feature selection algorithms, when they converge on certain genes, indicates a stronger consensus regarding the importance of these genes. Therefore, we employ multiple machine learning algorithms to select more robust candidate genes.

This study systematically evaluates changes in circadian rhythm patterns in HF patients. Firstly, DEGs screened based on the publicly available microarray dataset GSE57338 and intersected with CRGs to obtain DE-CRGs. Subsequently, four machine learning algorithms were employed to screen for genes that could be used as features in disease onset prediction. Nomogram models were constructed, and the correlation between five risk genes and infiltration of immune cells was discussed. Moreover, samples from patients with HF were clustered based on the expression profiles of CRGs, and two different types of HF with distinct biological functions were identified. The variation in immune cell populations between these two distinct groups was further analyzed. Finally, using NMF clustering methods based on DEGs among subclusters, HF patients were divided into two subgroups with distinct molecular and clinical features, and hub genes related to the phenotype were screened through WGCNA. Through these studies, it was found that CRGs have a significant impact on the immune microenvironment and metabolic processes of HF patients.

2. Methods

2.1. Data collection and processing

Microarray data of heart failure patients analyzed in this study were retrieved from the GEO database (www.ncbi.nlm.nih.gov/geo/

). Four datasets were included: GSE57338 [27],GSE79962 [28],GSE26887 [29] and GSE5406 [30]. Table 1 shows the detailed information of these datasets. To obtain symbols for genes corresponding to each probe, download the corresponding platform files and construct a matrix of gene expression. All microarray datasets included in this study were normalized with the "Limma" R package [31] and transformed to a log2 scale for subsequent analysis. Furthermore, CRGs included in this study were derived from GeneCard (https://www.genecards.org/).

2.2. Identification and functional analysis of DE-CRGs

Dataset GSE57338 contains 177 HF patient samples and 136 control samples. Using the R package "Limma", differential gene analysis was performed and DEGs were identified by applying cutoff criteria of $|\log FC| > 0.5$ and adjusted P-value <0.05. By intersecting DEGs with CRGs, we obtained DE-CRGs. The R package "ClusterProfiler" was utilized to conduct enrichment analysis and provide insights into the biological functions of DEGs [32].

2.3. Establishment and validation of CRGs diagnostic signature

The prediction of both the onset and progression risk of HF relies on the identification of stable and sensitive diagnostic markers. Firstly, four machine learning algorithms (namely, Boruta [33], Random Forest(RF) [34], Support Vector Machine_Recursive Feature Elimination (SVM_RFE) [35] , and XGBoost [36]) were utilized to filter potential genes for HF diagnosis based on the expression profiles of differentially expressed genes related to circadian rhythms. We performed gene selection of features using the default parameters. Boruta, RF, and SVM_RFE machine learning algorithms demonstrated superior classification performance, whilst the XGBoost model displayed below-par performance and limited ability to generalize. Hence, employing disease classification as the dependent variable and circadian related DEGs as the explanatory variables, we randomly divided the entire GSE57338 dataset into a training set (70%) and an internal validation set (30%) to search for the best hyperparameter combination. In the training dataset, we employed grid search techniques to adjust the parameters and identify the optimal combination. The adjusted parameters include eta, max_depth, min_child_weight, gamma, subsample, colsample_bytree, alpha and lambda. Next, we evaluated the model's performance on the validation dataset. Finally, we utilized the optimized XGBoost predictive model to select feature genes and gauge variables' impact on the onset and progression of HF. We measured these impacts through Shapley Additive Explanation (SHAP) values. Ultimately, we analyzed the genes that the four algorithms identified at the intersection in greater depth.

We utilized the "rms" and "forestplot" R packages to construct nomograms and forest plots, respectively. Discriminative ability of the model was evaluated by analyzing receiver operating characteristic (ROC) curves, while calibration was assessed using calibration plots. The model's accuracy was further validated in both an internal training set and two external validation sets.

2.4. Evaluation of immune cell infiltration

The infiltration levels of 28 types of immune cells were calculated using the ssGSEA method. Furthermore, statistical analyses were conducted by performing T-tests to compare the levels of immune cell infiltration. The optimal immune cell variables were determined using the least absolute shrinkage and selection operator (LASSO) algorithm [37] and ten-fold cross-validation to screen for the best λ value. Additionally, the study utilized correlation analysis to determine how hub genes are related to immune cells and investigate their underlying connections.

To enhance comprehension of immune infiltration variations across different clustering patterns in heart failure patients, the ssGSEA method was employed to investigate the immune microenvironment in HF, and T-tests were executed to compare disparities in immune cell infiltration.

2.5. Unsupervised clustering analysis of DE-CRGs

Consensus clustering analysis was conducted using the R package "ConsensusClusterPlus" to recognize related patterns of different CRGs clustering in HF patients [38]. The optimum number of subtypes was based on evaluating the performance of cumulative distribution function (CDF) curves, relative changes in the area under the CDF curve, consensus matrix, and consistent cluster scores. Subsequently, the differences among the number of clustering samples were visualized using the "ggpubr" package.

Table 1
Details of the dataset used in this study, including number, sequencing platform, number of samples, etc.

ID	GSE number	Platform	Samples	Group
1	GSE57338	GPL11532	Control: 136 Patient: 177	Discovery cohort
2	GSE79962	GPL6244	Control: 11 Patient: 40	Validation cohort
3	GSE26887	GPL6244	Control: 5 Patient: 12	Validation cohort
4	GSE5406	GPL96	Control:16 Patient: 194	Validation cohort

2.6. GSEA and GSVA enrichment analysis

The "clusterprofiler" R package was used to perform GSEA enrichment analysis on the DEGs between the two clusters generated by the first clustering to explore biological processes associated with these genes.

To comprehend the biological functional differences within subgroups, we employed the "GSVA" [39] and "Limma" packages to investigate the underlying causes for these phenotypic differences. The background gene lists, Hall.v2023.1.Hs.symbols.gmt, C5.go.bp.v2023.1.Hs.symbols.gmt, and C2.cp.kegg.v2023.1.Hs.symbols.gmt, were downloaded from the Molecular Signatures Database (MSigDB) (https://www.gsea-msigdb.org/gsea/msigdb/).

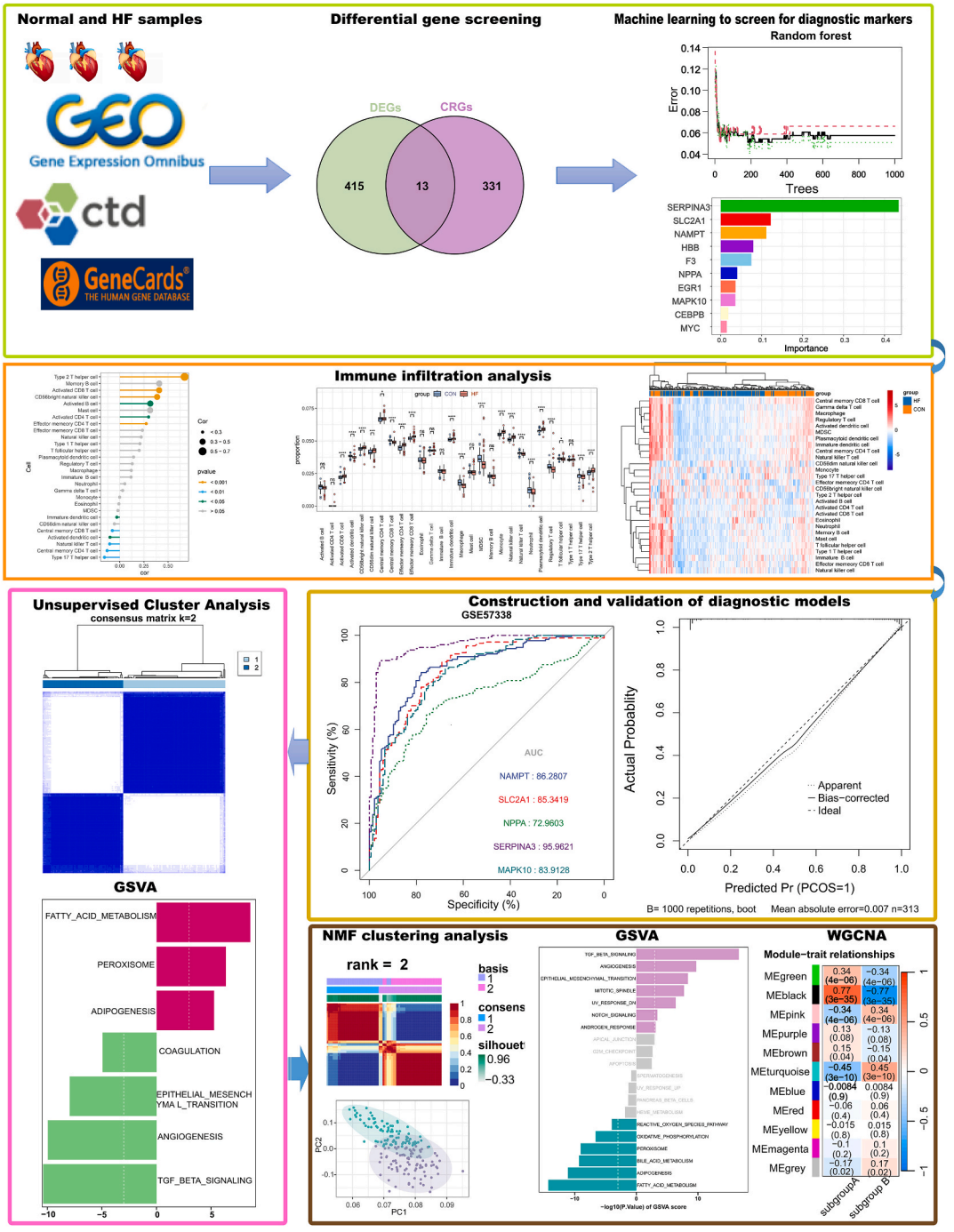

Fig. 1. Flowchart of this study.

2.7. Non-negative matrix factorization algorithm

The NMF algorithm extracts biologically relevant coefficients within the gene expression matrix and organizes genes and samples to capture the internal structural features of the data for sample grouping [40]. Using the empirical Bayesian method, DEGs between clusters A and B were acquired through the R package "Limma", with selection criteria of $|\log FC| > 0.5$ and adjusted p-value <0.05. The "NMF" R package was subsequently employed to conduct a clustering analysis of all HF samples using the DEGs identified between the subclusters, with the aim of uncovering possible molecular subtypes. The "brunet" algo rithm was chosen with 100 iterations for each value within the range. The number of clusters was set from 2 to 10. The optimal number of clusters is determined by the

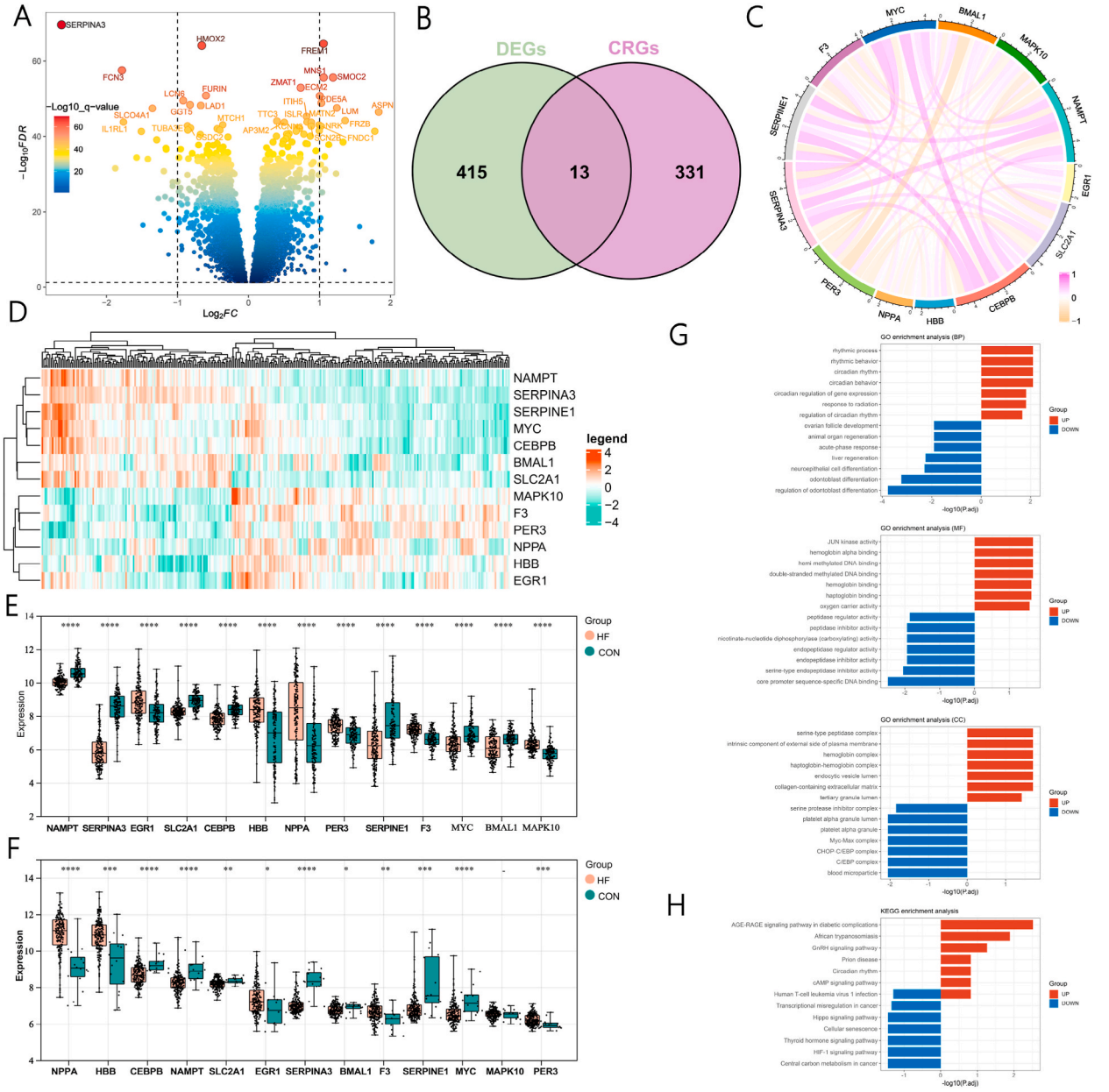

Fig. 2. Identification and functional enrichment analysis of DE-CRGs. (A) DEGs between HF and normal samples. (B) The common genes of DEGs and CRGs in HF. (C) Correlation chord diagram of DE-CRGs. Red represents positive correlations, and orange represents negative correlations. (D) Heatmap showing the expression levels of 13 CRGs in HF and control samples. (E, F) Boxplots and scatter plots displaying the expression differences of thirteen CRGs in GSE57338 and GSE5406. (G) GO enrichment analysis of DE-CRGs. (H) KEGG enrichment analysis of DE-CRGs. DE-CRGs, differentially expressed circadian rhythm-related genes; CRGs: circadian rhythm-related gene; DEGs: differentially expressed genes; HF: heart failure; GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes. *P < 0.05, **P < 0.01, ***P < 0.001. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

combination of cophenetic, dispersion and silhouette.

2.8. Weighted gene Co-expression network

We used the median absolute deviation method to select the top 5000 genes with the largest differences between samples. First, low-quality data were removed after assessing the data quality of both samples and genes. Second, samples with obvious outliers were excluded after using the "hclust" function to perform hierarchical clustering on the samples. Using the "pickSoftThreshold" function in the "WGCNA" R package, an appropriate soft-thresholding power β was chosen according to the standard of a scale-free network, with the range varying from 1 to 20. Selection criteria were as follows: $R^2 > 0.85$, slope approximately equal to −1, and choosing the value at the inflection point of the R^2 and power relationship plot. Next, a WGCNA was constructed to identify gene modules, with a

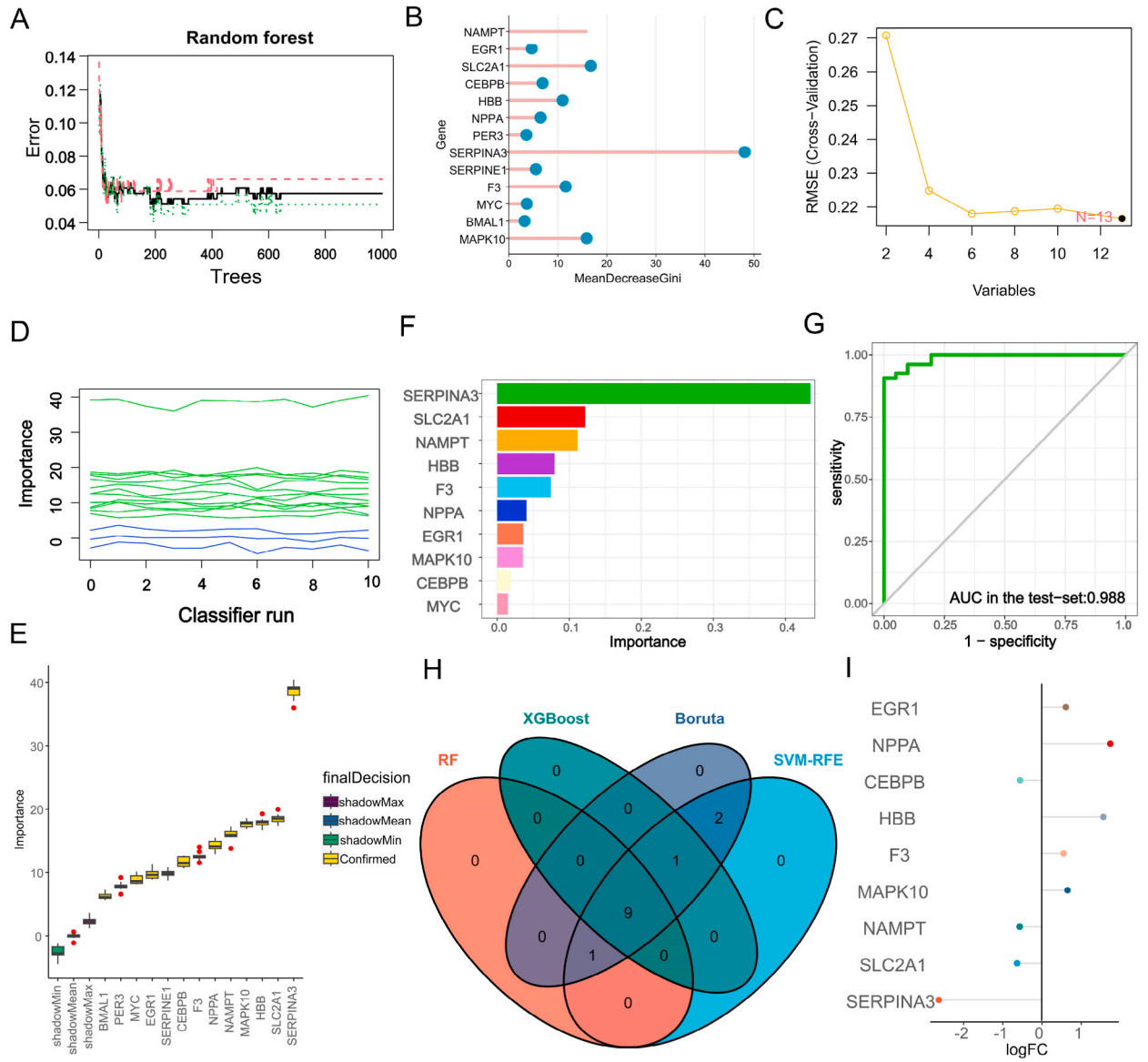

Fig. 3. Multiple machine learning methods for screening circadian rhythm-related diagnostic biomarkers. (A) Random forest tree. (B) Gini importance measure. (C) SVM-RFE algorithm selects feature genes. N = 13 represents the optimal number of selectable variables. (D) Z-score changes during Boruta operation. Green represents confirmed features, and blue represents the importance of minimum, average, and maximum shadow features, respectively. (E) MZSA distinguishes important and non-important features. Yellow is determined to be an important feature. (F) XGBoost ranks the importance of genes. (G) ROC curve evaluating the accuracy of the XGBoost model. (H) Intersection of genes selected by the four machine learning algorithms. (I) Fold change of genes with potential diagnostic value. MZSA, maximum Z score among shadow attributes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

combined cutting height of 0.25 and a minimum number of module genes set to 30.

Additionally, we computed the interaction network of genes within the black module using the STRING online database (https://string-db.org/) and ranked the importance of these genes using the MCC algorithm in the CytoHubba plugin. Furthermore, we confirmed the association of these genes with HF through the Comparative Toxicogenomics Database (CTD) (http://ctdbase.org/).

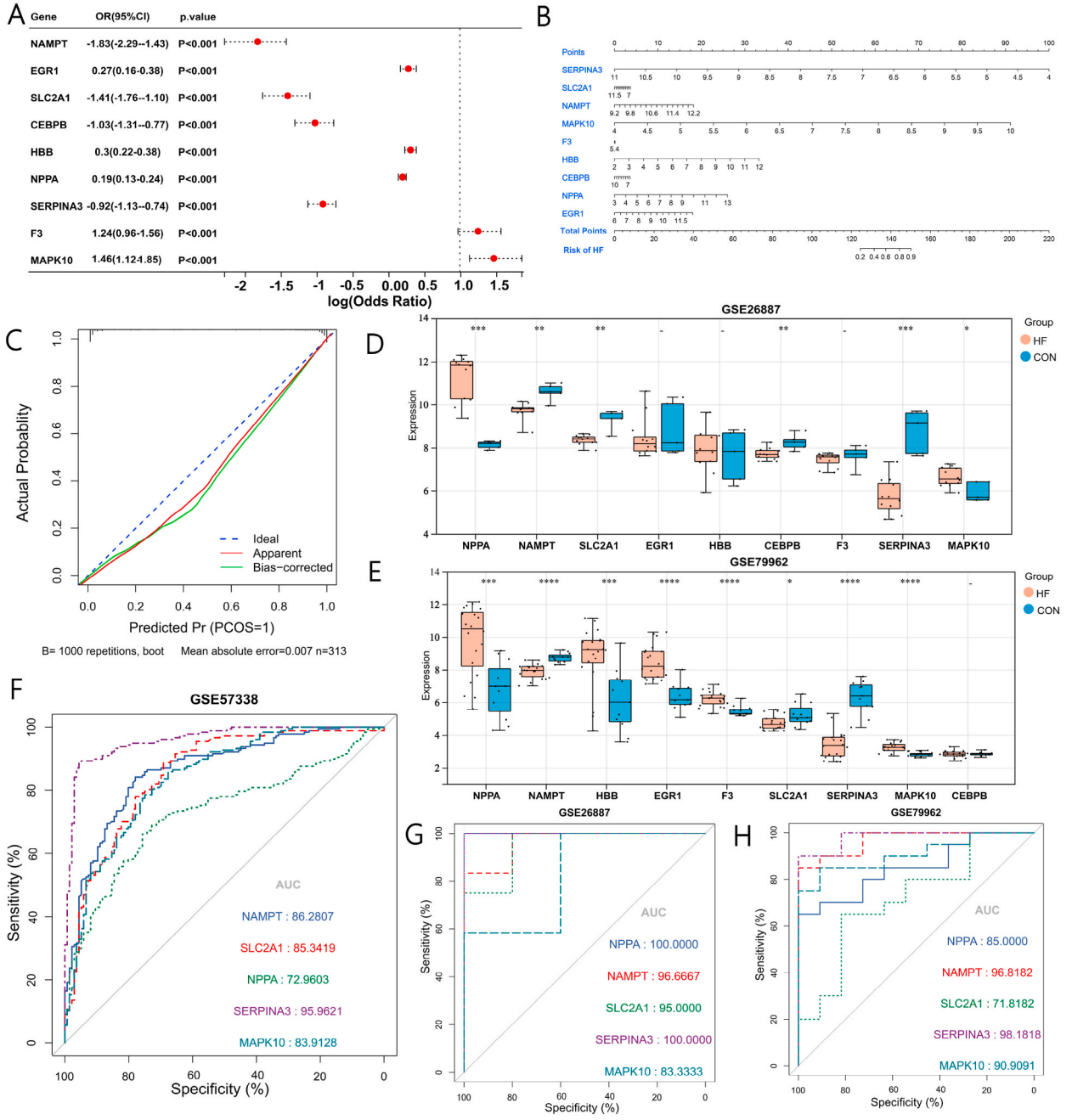

Fig. 4. Development and validation of CRGs as diagnostic features for HF. (A) Forest plot of multivariable logistic regression analysis. (B) Nomogram showing the probability of HF occurrence based on nine genes. (C) Calibration curve validation of the Nomogram. (D, E) Expression levels of nine diagnostic markers in two external datasets. *P < 0.05 **P < 0.01 ***P < 0.001. (F) Receiver operating characteristic (ROC) curves of diagnostic markers that showed expression differences in both external datasets. (G) ROC curves showing the diagnostic value of five feature genes in the GSE26887 dataset. (H) ROC curves of the five feature genes in the GSE79962 dataset.

3. Results

3.1. Expression landscape and functional enrichment analysis of CRGs

Fig. 1 shows the flowchart of this study. A differential analysis was performed on the expression profile data of 136 normal samples and 177 HF normalized samples, resulting in a total of 428 DEGs identified, including 234 significantly upregulated genes and 194 significantly downregulated genes. (Fig. 2A). Furthermore, by using correlation scores to examine the relationship between genes and circadian rhythms, we obtained a list of 344 circadian rhythm genes with correlation scores exceeding 3.5. To extract circadian rhythm genes that play a crucial role in regulating transcription and have potential research value in HF, we overlapped differentially expressed genes (DEGs) with circadian rhythm genes (CRGs). Eventually, this resulted in a set of circadian rhythm-related genes that are differentially expressed in HF patients. (Fig. 2B and Supplementary Table 1). Correlation analysis indicated that genes SERPINA3, CEBPB, MYC, and NAMPT were significantly positively correlated (Fig. 2C). Fig. 2D and E demonstrate the significant differences in the 13 CRGs between the control group and HF samples. To further verify the expression of the 13 CRGs, an independent HF cohort was used for validation. Fig. 2F revealed that a significant proportion of the CRGs displayed consistent expression levels, with notable differences.

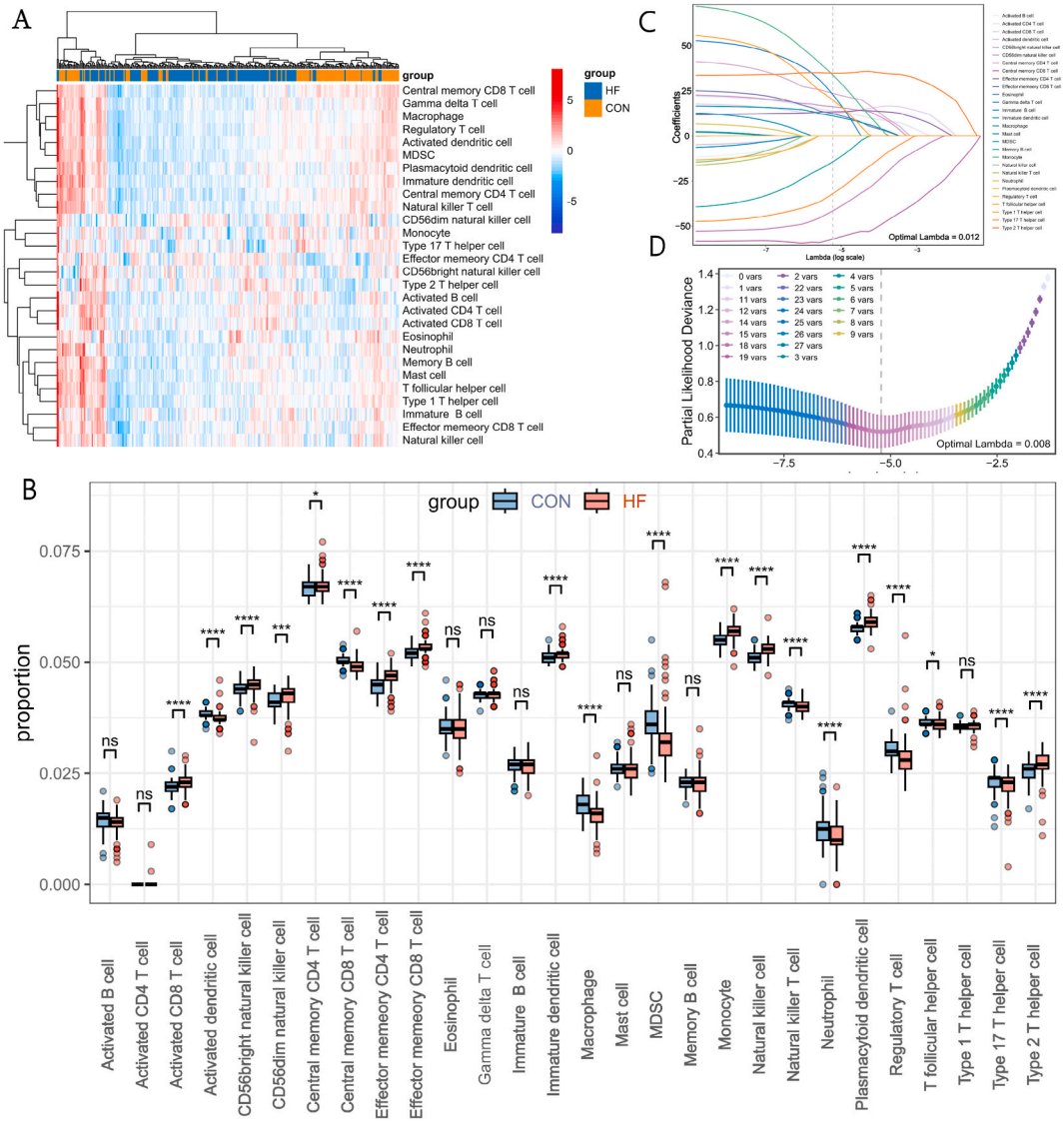

Fig. 5. SsGSEA analysis of HF and normal samples. (A) Heatmap demonstrating the difference in immune cell infiltration in HF and normal samples (B) Infiltration scores of 28 immune cells in cardiac tissue from HF patients and normal. ns = not significant, *p < 0.05, **p < 0.01, ***p < 0.001 (C, D) The best variable with a non-zero coefficient in the immune cell subpopulation was obtained by LASSO regression. ssGSEA , Single Sample Gene Set Enrichment Analysis.

To more finely explore the potential roles of DEGs, we conducted a functional enrichment analysis on the DE-CRGs. GO analysis indicated significant enrichments in rhythmic process, rhythmic behavior, circadian rhythm, and circadian regulation of gene expression (Fig. 2G). Additionally, KEGG results revealed that upregulated genes were significantly enriched in pathways, including the AGE-RAGE signaling pathway in diabetic complications, African trypanosomiasis, GnRH signaling pathway, and Circadian rhythm, while downregulated genes were mainly enriched in the HIF-1 signaling pathway and Human T-cell leukemia virus 1 infection signal pathways (Fig. 2H).

3.2. Development and selection of diagnostic markers

To screen for genes related to diagnosis from the 13 intersecting genes, three machine learning algorithms with feature selection were employed. The importance of the 13 features was ranked using the RF algorithm, and the top ten contributing genes were selected for subsequent analysis (Fig. 3A and B). According to the results of the SVM-REF algorithm, the classifier error was minimal when the number of features was 13 (Fig. 3C). Boruta algorithm results showed that the maximum Z-score of shadow features clearly distinguished important and unimportant features. The Z-Score values of all variables in a single random forest run were higher than the maximum Z-Score among shadow attributes (MZSA) in the random variable shadow features (Fig. 3D and E). Therefore, all 13 variables were confirmed as important features. To determine the machine learning model for predicting HF, the dataset consisting of 313 samples was randomly grouped into training and test cohorts, and an XGBoost machine learning model was established to predict the outcomes. Fig. 3G demonstrates that the XGBoost model exhibited perfect accuracy (AUC = 0.988). The impact of each feature variable on the predictive model based on the XGBoost model was elucidated using SHAP values in the training set. Based on the SHAP summary plot, the feature variable importance ranking indicated that SERPINA3 was the variable with the highest contribution to the XGBoost model (Fig. 3F). Ultimately, the intersection of the results of the four machine learning algorithms yielded the final nine diagnostic markers (Fig. 3H and Supplementary Table 2). The fold changes of the nine biomarkers are presented in Fig. 3I.

3.3. Assessment and validation of diagnostic features of CRGs for HF

While complex machine learning models can model the features of the data at a finer level, in practice, simple linear regression

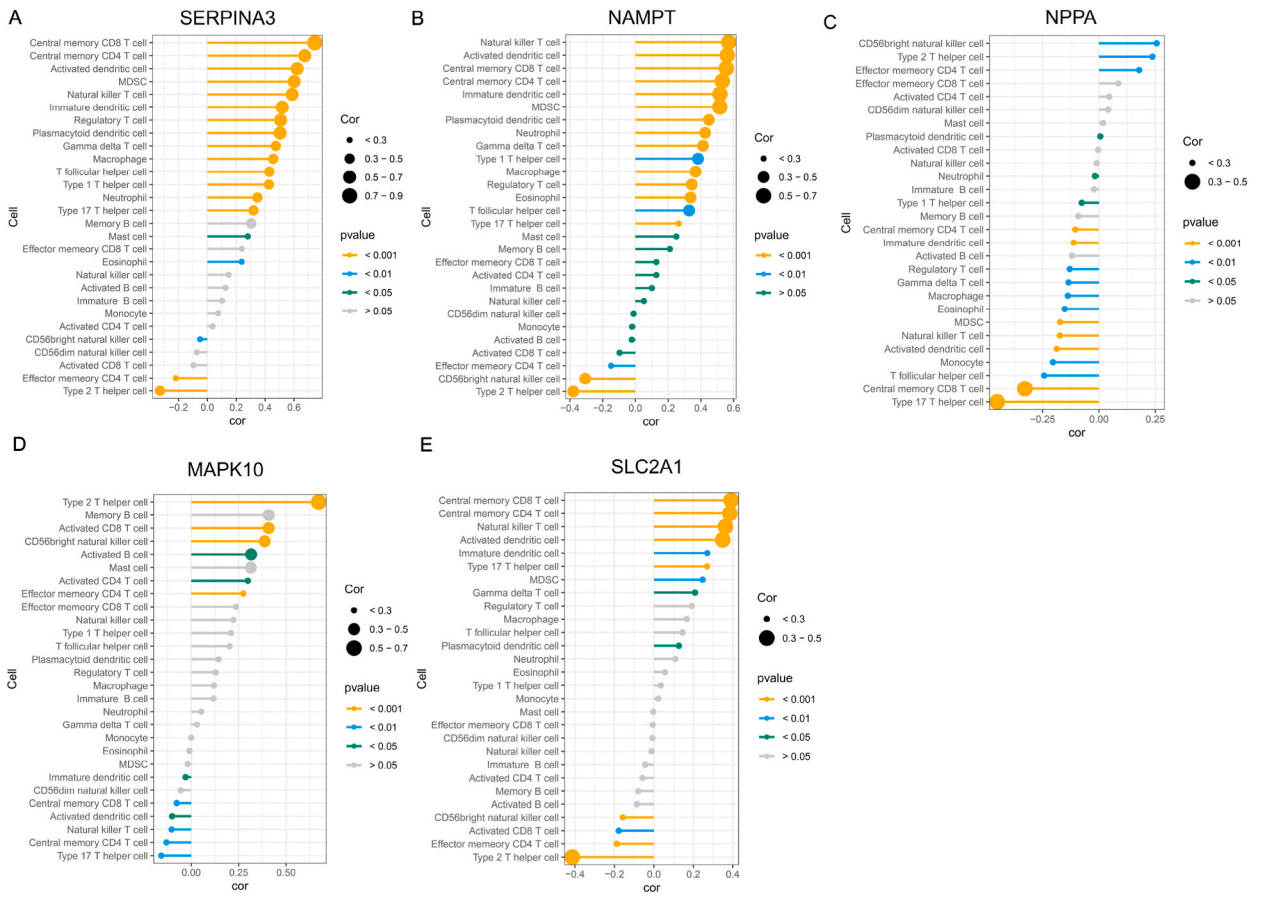

Fig. 6. (A–E) Correlation between immune cells and SERPINA3, NAMPT, NPPA, MAPKI0 and SLC2A1.

models are interpretable, highly predictive, and easier to translate into actionable models for clinical practice. Thus, a logistic regression model was constructed to explore the diagnostic value of characterized genes. The expression of nine markers was independently associated with HF, as revealed by multivariate logistic regression analysis shown in the forest plot (Fig. 4A). The nomogram assigns a score point to each feature variable value, and the total risk score for developing HF is obtained by adding up the scores of all

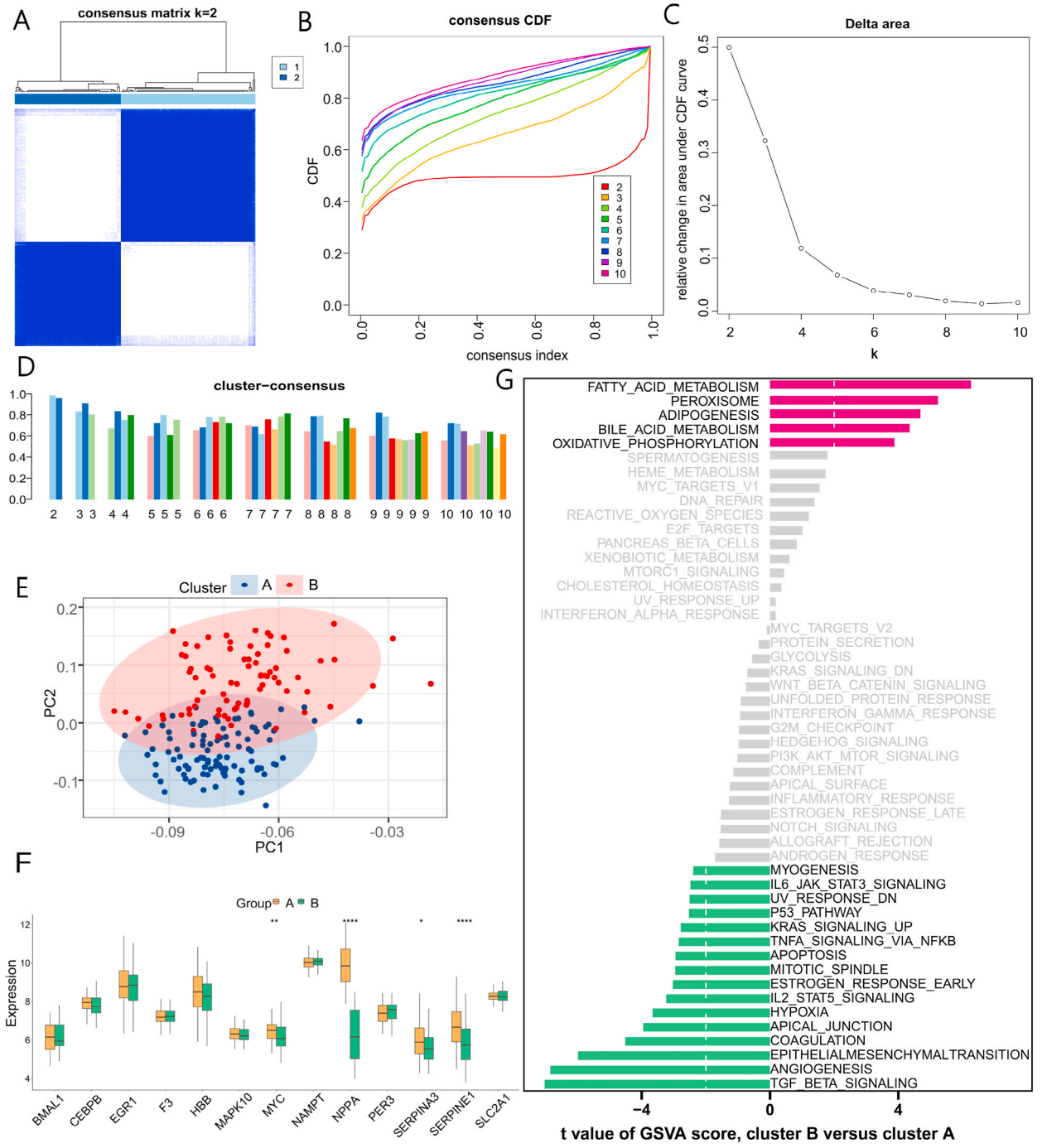

Fig. 7. Unsupervised clustering analysis based on DE-CRGs expression spectrum. (A) Consensus clustering matrix for k = 2, defining two different subtypes of circadian expression patterns. (B) Cumulative distribution function (CDF) curves for k = 1–9. (C) CDF delta area curves. (D) Consensus clustering scores when k is 2–9. (E) Principal component analysis (PCA) visualization of the sample distribution of the two clusters. (F) Expression of the 13 CRGs in the two clusters. (G) Gene set variance analysis (GSVA) based on the HALLMARK gene set. DE-CRGs, differentially expressed circadian rhythm-related genes.

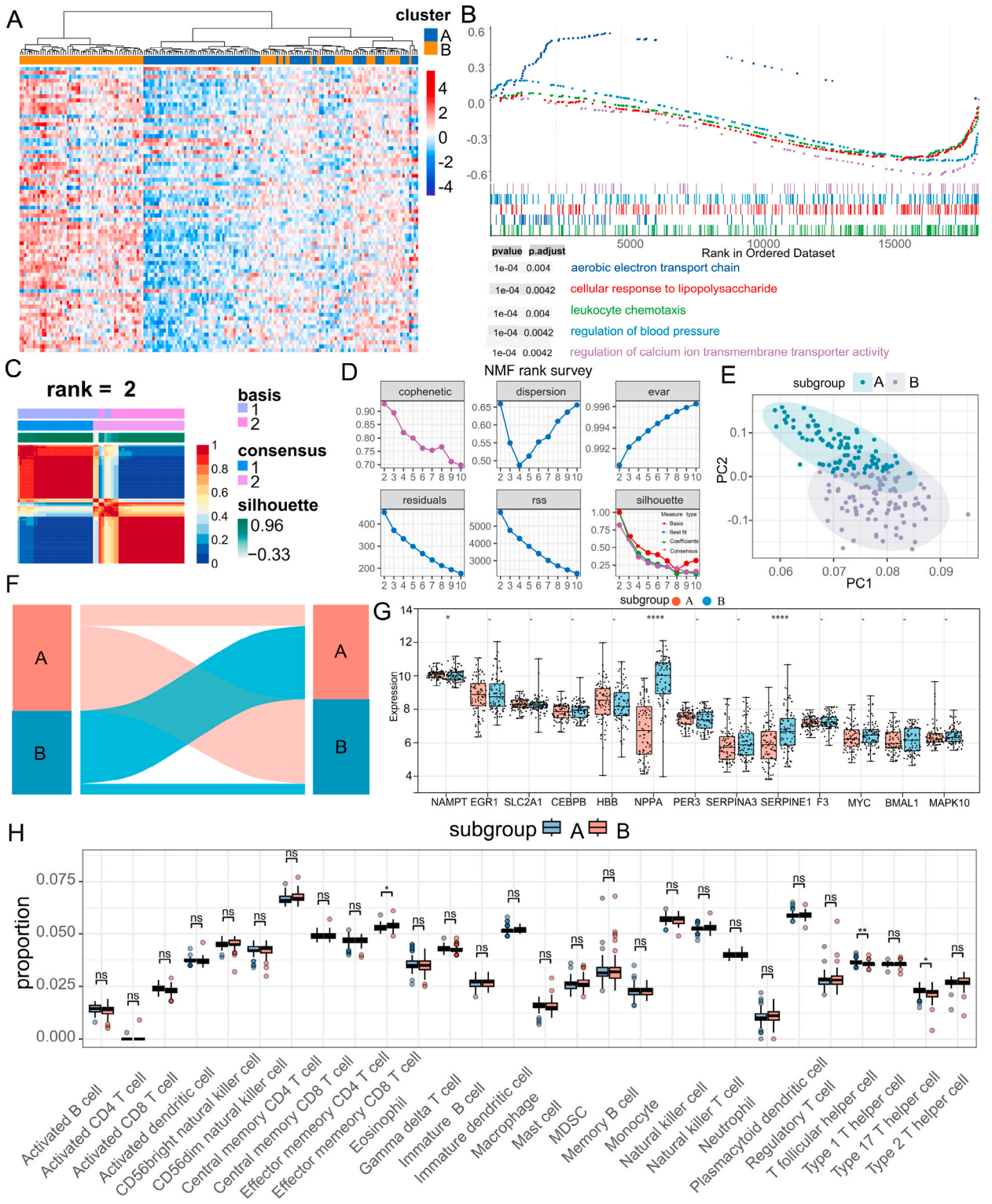

Fig. 8. Identification of HF gene subgroups based on DEGs associated with CRG subtypes. (A) Heatmap of differential expression of DEGs between the two CRG subtypes. (B) GSEA enrichment analysis of DEGs between two CRG subtypes. (C) Heatmap of NMF clustering at k = 2. (D) Distribution of cophenetic, residuals, RSS, silhouette, Var and dispersion with a rank of 2–10. (E) PCA plots showing differences in sample clustering between two HF gene subgroups. (F) Variation in gene subgroups for two different clustering methods. (G) Expression box plots and scatter plots of 13 CRGs. (H) Immune cell infiltration scores in 28 between the two subgroups. NMF, non-negative matrix decomposition; GSEA, gene set enrichment analysis.

feature variables (Fig. 4B). The accuracy of the nomogram in diagnosing HF was confirmed by the calibration curve (Fig. 4C). To better illustrate the diagnostic ability of the feature genes for HF, we further verified their diagnostic capacity in two external datasets. The results indicate that there were significant differences in the expression levels of NPPA, NAMPT, SLC2A1, MAKP10, and SERPINA3 in both datasets (Fig. 4D and E). Subsequently, ROC analysis was performed in the training cohort and the two external cohorts, and AUC values were calculated to assess the accuracy of each diagnostic gene. Our findings indicated that the five diagnostic genes had high predictive values, particularly SERPINA3 (Fig. 4F–H).

3.4. Immune infiltration analysis between normal and HF samples

As shown in Fig. 5A and B, significant differences in the abundance of 20 immune infiltrating cells were observed between HF and normal samples, including activated CD4 T cells, activated dendritic cells, CD56bright natural killer cells, CD56dim natural killer cells, central memory CD4 T cells, central memory CD8 T cells, effector memory CD8 T cells, effector memory CD4 T cells, immature dendritic cells, natural killer T cells, regulatory T cells, type 2 T helper cells, type 17 T helper cells, natural killer cells, macrophages, MDSCs, monocytes, natural killer cells, neutrophils, and plasmacytoid dendritic cells (Fig. 5A and B). Furthermore, we employed the LASSO algorithm to identify characteristic immune cells associated with HF progression. Ultimately, nine optimal variables with non-zero coefficients were identified from the immune cell subgroups mentioned above: activated CD8 T cells, CD56bright natural killer cells, CD56dim natural killer cells, central memory CD4 T cells, central memory CD8 T cells, effector memory CD4 T cells, effector memory CD8 T cells, type 17 T helper cells, and type 2 T helper cells (Fig. 5C and D).

The correlation analysis between feature genes and immune cells revealed that SERPINA3 and SLC2A1 were strongly positively correlated with central memory CD8 T cells and negatively correlated with type 2 T helper cells (Fig. 6A and E). NPPA was significantly negatively correlated with type 17 T helper cells (Fig. 6C). MAPK10 was significantly positively correlated with type 2 T helper cells (Fig. 6D). Interestingly, NAMPT exhibited significant correlations with all immune cells (Fig. 6B). Taken together, these results suggest the presence of immune dysregulation in HF, influenced by changes in the expression of circadian rhythm genes.

3.5. Consensus clustering analysis of circadian rhythm gene sets

Using the R package "ConsensusClusterPlus," we conducted consensus clustering analysis on 177 HF samples based on 13 DE-CRGs. The results of the clustering demonstrated that a perfect cluster number appeared to be $k = 2$, which accurately dispersed HF patients into two clusters: Cluster A ($n = 99$) and Cluster B ($n = 78$) (Fig. 7A–D). Clear differences between Cluster A and Cluster B were revealed through PCA analysis (Fig. 7E). Furthermore, several CRGs exhibited significant heterogeneity between the two clusters (Fig. 7F), specifically MYC, NPPA, SERPINA3, and SERPINE1, which displayed higher expression in Cluster A than in Cluster B.

3.6. Identification of distinct circadian rhythm patterns of immune microenvironment and biological function features

The HALLMARK gene set was used to perform GSVA analysis, which revealed significant enrichment of TGF-β signaling, angiogenesis, NFκB-mediated signaling, TNF-α signaling, coagulation function, epithelial-mesenchymal transition, and IL6-JAK-STAT3 signaling pathways in Cluster A. In contrast, Cluster B showed primary enrichment in metabolic-related pathways such as fatty acid metabolism and oxidative phosphorylation (Fig. 7G). Moreover, we conducted further analysis of the biological function and pathway enrichment differences between Cluster A and Cluster B. Regulation of insulin biosynthesis and protein acylation were significantly enriched in Cluster B, while positive regulation of cardiac muscle contraction, negative regulation of platelet activation, and phagosome maturation mediated by regulatory particles were significantly enriched in Cluster A (Supplementary Fig. 1A). Based on KEGG, GSVA results indicated significant enrichment of TGF-β signaling, ECM receptor interactions, and regulation of actin cytoskeleton-related pathways in Cluster A. Conversely, metabolic pathways such as lipid metabolism, valine and isoleucine degradation, and butanoate metabolism were enriched in Cluster B (Supplementary Fig. 1B). Furthermore, we compared changes in infiltrating immune cell abundance between the two clusters, observing increased T follicular helper cell infiltration in Cluster B, as shown in Supplementary Figs. 2A and B. Overall, the study revealed that there are two distinct circadian rhythm patterns which correlate with different immune features and pathways. This finding suggests a critical role of circadian rhythm in the regulation of immune microenvironment and metabolic processes in HF.

3.7. Determination of HF gene subgroups based on DEGs related to circadian rhythm subtypes

To further investigate potential biological behaviors of different circadian rhythm-related phenotypes, we utilized the R package "Limma" to identify 72 DEGs associated with circadian rhythm phenotypes (Supplementary Fig. 2C). A heatmap displayed expression differences of the 72 DEGs between the two clusters (Fig. 8A). Subsequently, we conducted GO and KEGG analyses to determine biological functions (Supplementary Fig. 2E and Supplementary Table 3). Moreover, significant enrichment for leukocyte chemotaxis, regulation of blood pressure, aerobic electron transport chain, cellular response to lipopolysaccharide, and regulation of calcium ion transmembrane transporter activity was revealed by GSEA enrichment analysis (Fig. 8B). Thus, CRGs may affect HF by modulating cardiac material and energy metabolism, as well as mediating immune-related processes.

In addition, we performed cluster analysis based on the expression patterns of 72 DEGs in Clusters A and B using the NMF algorithm. The optimal cluster number, $k = 2$, was determined based on cophenetic, dispersion, and silhouette metrics (Fig. 8C and D). The HF samples were divided into two subgroups, Subgroup A ($n = 88$) and Subgroup B ($n = 89$), according to the distinct expression

patterns of the 72 DEGs. PCA demonstrated significant expression differences between these subgroups (Fig. 8E). A dendrogram showed a similar grouping tendency for the two clusters and gene subgroups (Fig. 8F). The expression of 13 CRGs among the two subgroups is depicted in Fig. 8G and Supplementary Fig. 2D, with genes NPPA, NAMPT, and SERPINE1 exhibiting notable differences between the subgroups.

3.8. Identification of immune microenvironment and biological function features in different subgroups

Firstly, we conducted an analysis of the differences in immune microenvironments between Subgroup A and Subgroup B. In comparison to Subgroup A, Subgroup B exhibited higher levels of infiltration by monocyte immune cells as demonstrated by the

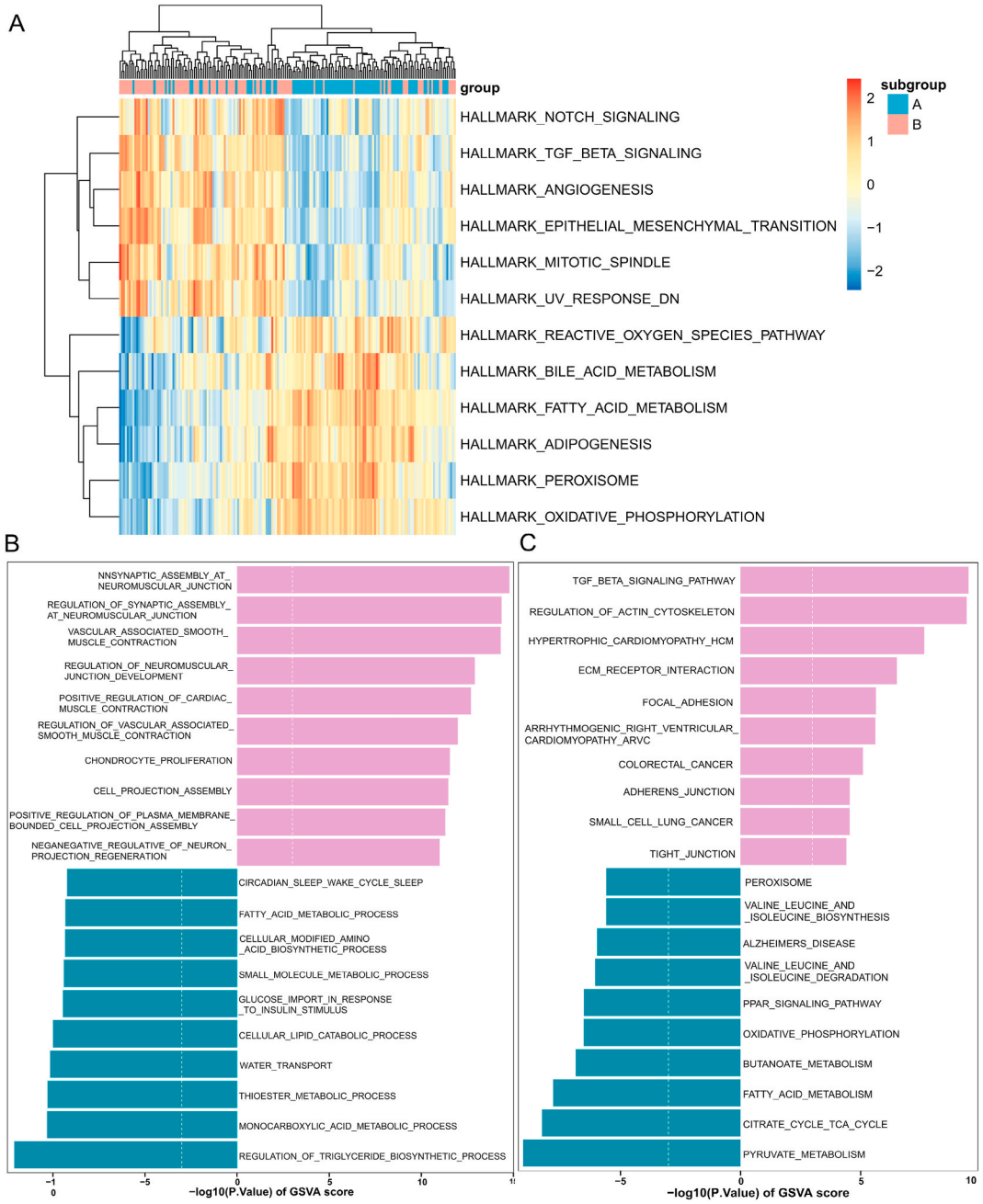

Fig. 9. GSVA analysis to assess biological function and pathway differences between the two subgroups. (A) GSVA analysis based on the HALLMARK gene set. (C) Biological process analysis based on the GO gene set. (C) Enrichment pathway based on the KEGG pathway. GSVA, genomic variation analysis.

results. Conversely, In Subgroup A, higher infiltration levels were observed in activated dendritic cells and type 17 T helper cells (Fig. 8H). Next, we performed GSVA between the different subgroups. Enrichment analysis of the HALLMARK gene set showed that Subgroup B was enriched in Notch signaling pathways, TNF-β signaling pathways, angiogenesis, and epithelial-mesenchymal transition. In contrast, cardiac metabolic processes such as fatty acid metabolism, oxidative phosphorylation, peroxidation products, and lipid metabolism were significantly enriched in Subgroup A (Fig. 9A). Biological processes upregulated in Subgroup B included Tnfsf11-mediated signaling pathways, positive regulation of the epithelial-to-mesenchymal transition during endocardial formation, and regulatory mechanisms for vascular-associated smooth muscle contraction (Fig. 9B). In Subgroup B, the expression levels of cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy, and regulation of the actin cytoskeleton were significantly upregulated, whereas in Subgroup A, there was a significant enrichment of metabolic-related pathways such as fatty acid metabolism

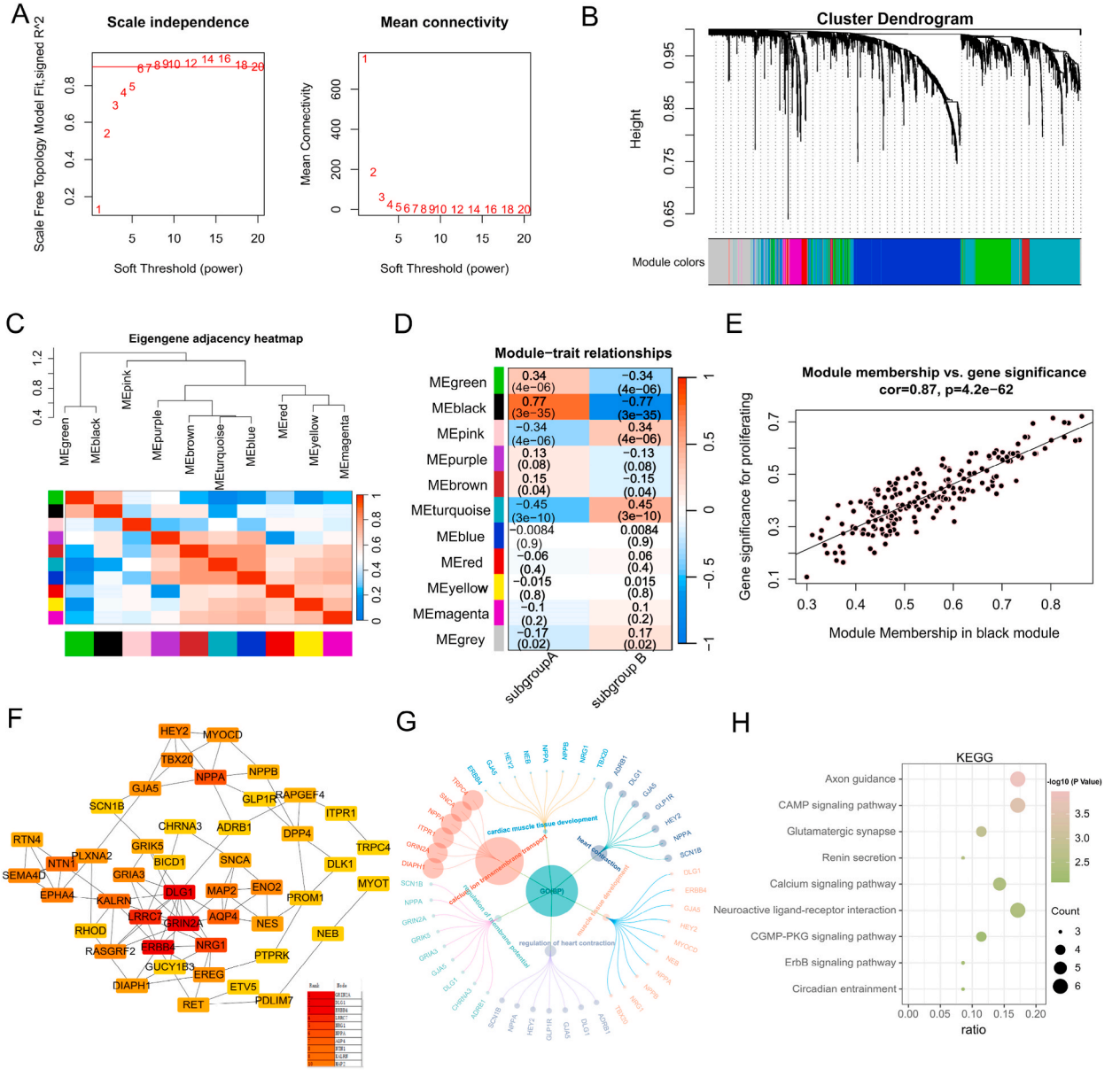

Fig. 10. Identification of hub genes between the two subgroups by the WGCNA method. (A) The network constructed when power index 6 was chosen as the appropriate soft threshold was more consistent with a scale-free topology. (B) Co-expression networks were constructed based on the optimal soft threshold to divide the genes into 11 different modules. (C) Cluster tree and correlation heat map between modules. (D) Correlation and significance between modules and features. the highest correlation was found between MEblack and features. (E) Scatter plot of module feature genes in the black module. mm and gs were positively correlated. (F) Protein-protein interaction network (PPI) ranking of gene importance. (G, H) GO and KEGG enrichment analysis of the top 50 significant genes in the PPI network. mm, module membership; GS, gene significance; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

and butyrate metabolism (Fig. 9C). These findings suggest that patients with HF in Subgroup A have higher metabolic levels, whereas those in Subgroup B have higher levels of immune infiltration.

3.9. Identification and functional enrichment analysis of hub genes between two subgroups

The clustering results of the 177 samples are shown in Supplementary Fig. 3. Two outlier samples (GSM379815 and GSM380018)

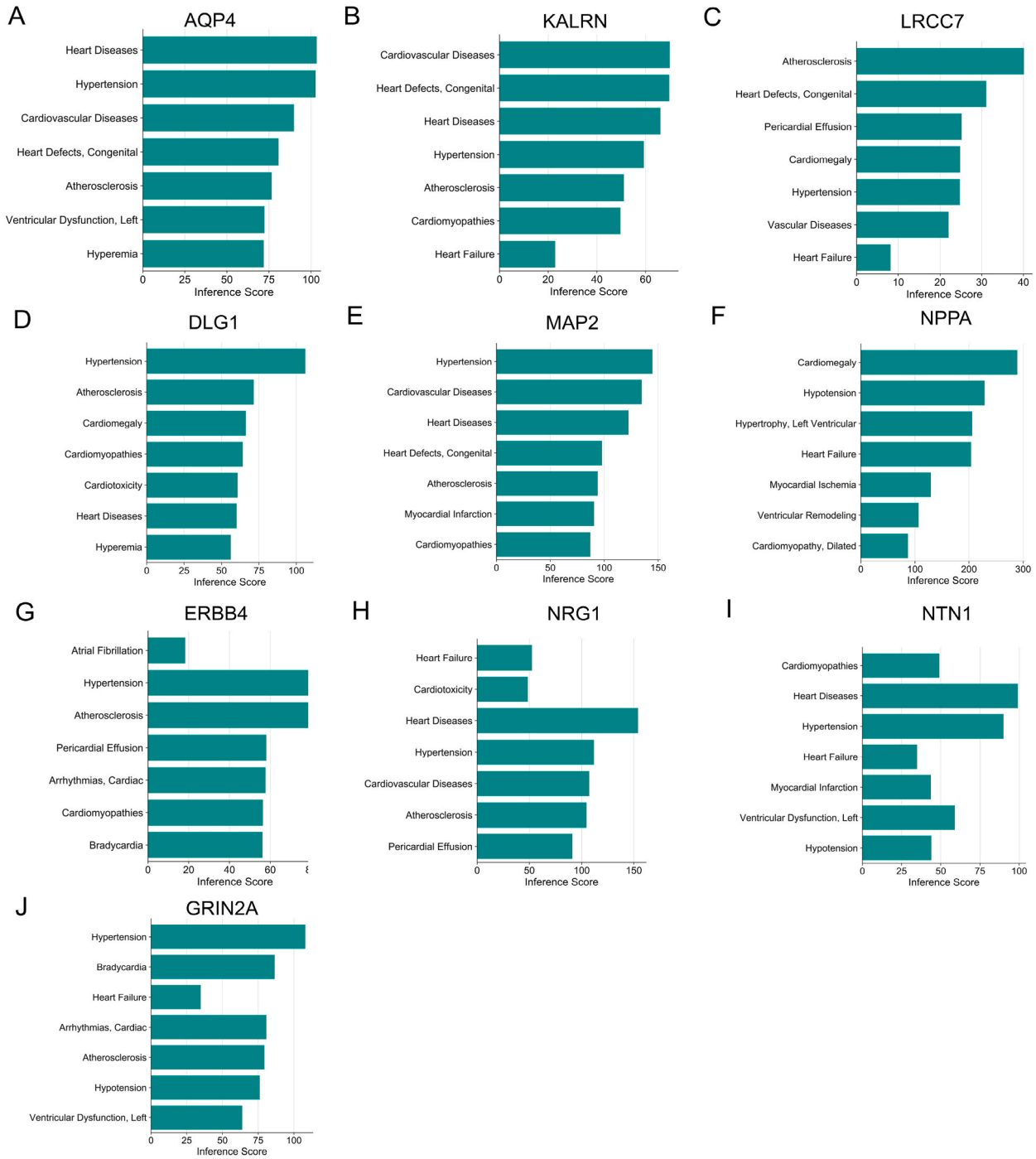

Fig. 11. Association of hub genes with cardiovascular disease, evidence from the CTD database. (A–J) Relationship between expression levels of AQP4, KALRN, LRCC7, DLG1, MAP2, NPPA, ERBB4, NRG1, NTN1, and GRIN2A and inferred scores of cardiovascular disease. CTD, Comparative Toxicogenomics Database.

were excluded, and 175 samples were included. At a soft threshold of $\beta = 6$, the scale-free R^2 value reached 0.9, resulting in a greater average connectivity (Fig. 10A). Hierarchical clustering identified a total of eleven modules (Fig. 10B). Among them, the black and green modules showed positive correlations (Fig. 10C). From the module-trait correlation heatmap, the black module was positively correlated with Subgroup A ($R^2 = 0.77$) (Fig. 10D and E). Therefore, we constructed a PPI network by utilizing genes extracted from the black module and screened for hub genes using the MCC algorithm in cytoHubba plugin (Fig. 10F). Additionally, we performed biological function annotation for the top 50 ranked genes. The hub genes were found to be primarily associated with cardiac muscle tissue development, heart contraction, calcium ion transmembrane transport, and regulation of heart contraction according to GO analysis (Fig. 10G). KEGG analysis revealed significant enrichment of axon guidance, calcium signaling pathway, cAMP signaling pathway, and circadian entrainment pathways (Fig. 10H). We confirmed through the CTD that most hub genes participate in the occurrence and development of HF or other cardiovascular diseases (Fig. 11A–J).

4. Discussion

Circadian rhythms are a manifestation of the body's biological clock, regulating many physiological processes, including metabolism, immunity, and cell death [15,41,42]. Existing research suggests that there is a close relationship between circadian rhythms and HF, but the exact mechanisms of molecular clocks in these phenomena are yet to be fully understood. Therefore, we conducted bioinformatics approaches to investigate the role of CRGs in HF and reveal their association with immune and metabolic processes. In this study, we employed various machine learning algorithms to screen five diagnostic biomarkers that demonstrated excellent diagnostic ability for HF, yielding good predictive results in the validation set. Secondly, we performed unsupervised clustering of HF samples based on CRGs expression profiles for the first time, identifying two subclusters with different expression patterns. GSVA revealed that Cluster A was mainly associated with immune processes while Cluster B was mainly associated with metabolic processes. Finally, based on the DEGs between the two clusters from the first clustering, we further identified two gene subgroups with unique immune and metabolic features and revealed hub genes related to circadian rhythm changes through WGCNA.

In recent years, the impact of CRG rhythmic expression changes on cardiac phenotypes has been demonstrated in knockout mouse models. Mice without BMAL1 exhibit disrupted circadian rhythms in blood pressure and heart rate, accompanied by cardiac electrophysiological and metabolic abnormalities [14,43]. Moreover, REV-ERBα agonists and antagonists have been shown to provide cardiac protection in mouse models, preventing heart failure [44,45]. Therefore, the exploration of novel biomarkers may bring unexpected benefits for HF treatment. Machine learning can perfectly solve data feature dimensionality and redundancy issues, enabling more accurate identification of diagnostically valuable genes [46]. Based on this, we combined four machine learning algorithms (including RF, Boruta, SVM-REF, and XGBoost) to obtain nine diagnostic biomarkers. In addition, we further validated the nine diagnostic biomarkers using multivariate logistic regression analysis and two external validation datasets, finding that only NAMPT, SERPINA3, MAPK10, NPPA, and SLC2A1 could accurately predict HF.

Current evidence indicates that immune and inflammatory activation have a significant impact on the development of HF, and immune responses are strictly controlled by circadian rhythm mechanisms [47,48]. Therefore, we conducted an analysis to investigate the correlation between gene expression and immune cell infiltration. The correlation analysis revealed that SERPINA3 and SLC2A1 were strongly positively correlated with Central memory CD8 T cells and negatively correlated with Type 2 T helper cells. SERPINA3 mainly activates the immune process by regulating neutrophil cathepsin G and leukocyte elastase [49]. Proteomic results also highlighted the role of SERPINA3 in the epicardial adipose tissue of HF patients, where regulation of SERPINA3 metabolism promotes epicardial adipose inflammation associated with HF [50]. Furthermore, SERPINA3 can improve prognostic stratification for patients with newly diagnosed or worsening heart failure [51]. Consistent with previous research, dysregulated or imbalanced SERPINA3 gene expression is associated with the progression of HF patients. SLC2A1 is an insulin-independent glucose transporter that predominantly facilitates basal cardiac glucose uptake in quiescent muscle cells [52]. Reduced SLC2A1 expression in the human heart indicates that reduced SLC2A1-mediated glucose uptake and utilization could contribute to the development of HF [53]. Notably, glucose availability is a key factor in regulating innate and adaptive immune responses as well as cancer. It has been shown that SLC2A1 is the rate-limiting glucose transporter on pro-inflammatory polarized macrophages. SLC2A1 overexpression results in increased glucose uptake and metabolism, exacerbating the inflammatory response [54]. A significant negative correlation was observed between NPPA and Type 17 T helper cells. Atrial natriuretic peptide (ANP), which is encoded by NPPA, is a member of the natriuretic peptide family and has been utilized as a diagnostic biomarker for HF [55], further confirming the accuracy of this study. MAPK10 was significantly positively correlated with Type 2 T helper cells. The gene is primarily expressed in the heart, brain, and testes and is a member of the JNK family [56]. MAPK10 knockdown alleviates diabetes-induced atrial fibrillation by inhibiting inflammation, fibrosis, electrical dysfunction, and apoptosis [57]. Specific knockdown in mice can lead to myocardial dysfunction, including diastolic dysfunction and cell apoptosis [58,59]. Furthermore, inhibiting MAPK10 phosphorylation can effectively alleviate synaptic dysfunction, inflammatory responses, and various neurodegenerative pathologies associated with Alzheimer's disease [60]. However, few reports exist on its role in HF patients. Based on the above analysis, MAPK10 may regulate the occurrence of heart failure by modulating inflammatory responses, myocardial cell apoptosis, and electrical signal conduction. Interestingly, NAMPT is significantly correlated with all immune cells. Various types of immune cells regulate NAMPT through diverse inflammatory signals. Inflammatory cytokines released by inflammatory cells in response to chronic inflammation can suppress the expression of NAMPT [61]. TNF-α inhibits the activation of clock genes induced by the CLOCK-BMAL1 complex, thereby disrupting the function of clock genes [62]. In addition, knocking down or overexpressing NAMPT both result in reduced protein acetylation, suppressed metabolic genes, and mitochondrial dysfunction [63]. Moreover, the expression of NAMPT is controlled by the transcription factor Kruppel-like factor 15 (KLF15) [64]. Research suggests that KLF15 regulates catabolic metabolism and oxidative stress responses in the heart through transcriptional activation of

NAD-producing enzyme NAMPT [65]. In our study, we also observed an increase in NAMPT expression levels in metabolic subtype A. This may indicate a disruption of NAD homeostasis in HF patients. Consequently, NAMPT emerges as a crucial hub gene enabling regulation of cardiac metabolism and active oxygen scavenging. Integrating this information, we believe that disrupted normal CRGs expression may disrupt the cardiac energy balance and immune function imbalance, thereby exacerbating HF development.

To further explore the distinctions in immune features between normal and HF samples, as well as the relationship between gene features connected to circadian rhythm and immune cells, we performed quantification analysis of 28 immune cell infiltration scores utilizing the single-sample gene set enrichment analysis (ssGSEA) algorithm. Our findings indicate that the quantity of a majority of immune cells varies in HF and healthy control samples, comprising neutrophils, macrophages, and most lymphocytes. Research indicates that circadian rhythm fluctuations can convey information to the immune system, mediating both innate and adaptive immune responses [66,67]. The fluctuations of the circadian rhythm in neutrophils may aid their migration to various tissues during their active phase, whilst safeguarding blood vessels against the effect of these extremely active cells during their resting phase [68]. Macrophages release the inflammatory cytokines IL-1β, IL-6 and TNF-α, dependent on the molecular clock phase. Most genes expressed in a circadian rhythm in macrophages play a role in immune response at varying levels, including Toll-like receptor (TLR) expression [69]. In our investigation, we discovered that there were varying degrees of macrophage infiltration between two groups, and that the Toll-like receptor signaling pathway was significantly enriched. As a result, we hypothesize that circadian rhythms impact the regulation of macrophage expression in the context of HF. Research has also shown that CD8 T cells lacking Bmal1 produce reduced levels of IL-2, IFN-γ and TNF-α when co-cultured with wild-type cells [70]. Additionally, circadian rhythm genes regulate the periodic migration of T cells to lymph nodes and the timing of lymph node transitions [71]. Chronic disruption of circadian rhythms during HF could significantly disrupt various immune responses, perhaps by regulating immune cell cytokine release rhythmically. Although our study offers fresh insight into the link between circadian rhythms and immunity in HF, further research is necessary to determine the specific mechanisms.

To better understand the heterogeneity among HF patients, we carried out unsupervised clustering analysis based on CRGs. Cluster A was significantly enriched in TGF-β signaling pathway, angiogenesis, NFκB-mediated signal transduction, TNF-α signaling pathway, blood coagulation function, epithelial and stromal cell transition, and IL6-JAK-STAT3 signaling pathway. Cluster B mainly focused on metabolism-related pathways, such as fatty acid metabolism and oxidative phosphorylation. Additionally, NMF clustering based on differentially expressed genes also effectively classified HF patients into two subgroups. The Notch signaling pathway, TNF-β signaling pathway, angiogenesis, and epithelial and stromal cell transition were significantly enriched in Subgroup B, while cardiac metabolic processes such as fatty acid metabolism, lipid metabolism, reactive oxygen species, and oxidative phosphorylation were enriched in Subgroup A. Prior research has established that numerous genes related to metabolism, particularly those involved in fatty acid metabolism, display circadian oscillations [13]. Research has also shown that intramyocardial lipid content is higher in the early morning [72]. The immune system also participates in the process of heart failure. Hemodynamic changes in HF patients trigger cardiac immune responses [73]. Immune cell differentiation and cytokine release also exhibit time dependence [74]. Existing research confirms that targeting circadian rhythm mechanisms inhibits the formation of cardiac NLRP3 inflammasomes and subsequent cascade reactions [75]. Therefore, CRG-based subtyping can better help us understand the influence of different biological processes on heart failure, thereby adopting appropriate intervention measures to slow down heart failure development.

However, we must acknowledge the limitations of our study. Firstly, our results have not been validated in vivo or in vitro due to the absence of experimental conditions for follow-up trials. Second, the dataset was obtained from a public database, and we did not obtain clinical information related to patients and therefore could not perform a prognostic analysis. Lastly, transcriptomic features may vary at different stages of heart failure, and individual differences must be considered when including samples.

5. Conclusion

The machine learning based diagnostic model has higher accuracy and can perfectly discriminate HF patients from normal patients. In addition, we investigated the relationship between CRGs and immune infiltration, as well as the differences in biological function and immune cell infiltration between different clusters and subgroups. The subclusters of HF patients showed different immune and metabolic patterns among themselves. In summary, our study provides comprehensive analysis for the first time on the role of CRGs in HF, which offers valuable references for targeted inhibition of immune response and energy metabolism through regulating circadian rhythm mechanisms in the treatment of HF.

Source of funding

This work was supported by the Shanghai Shuguang Project [No. 21SG37] and the National Natural Science Foundation of China [No. 81770244].

Data availability

The data relevant to this study (GSE57338, GSE79962, GSE26887, and GSE5406) have all been deposited in the publicly available repository, the GEO database (www.ncbi.nlm.nih.gov/geo/), as they originate from public databases. The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Summary

This study has shed light on the connection between circadian rhythms and heart failure (HF) pathogenesis. The study used machine learning algorithms to identify differentially expressed circadian rhythm-related genes (DE-CRGs) in HF patients, and diagnostic models were constructed based on these feature genes. Nine feature genes were obtained through cross-referencing results from four distinct machine learning algorithms, and multivariate LASSO regression and external dataset validation were performed to select five key genes with diagnostic value, including NAMPT, SERPINA3, MAPK10, NPPA, and SLC2A1. Consensus clustering analysis could divide HF patients into two distinct clusters, which exhibited different biological functions and immune characteristics. Additionally, two subgroups were distinguished using the non-negative matrix factorization (NMF) algorithm based on circadian rhythm associated differentially expressed genes. Studies on immune infiltration showed marked variances in levels of immune infiltration between these subgroups. Finally, the Weighted Gene Co-expression Network Analysis (WGCNA) method was utilized to discern the modules that had the closest association with the two observed subgroups, and hub genes were pinpointed via protein-protein interaction (PPI) networks. GRIN2A, DLG1, ERBB4, LRRC7, and NRG1 were circadian rhythm-related hub genes closely associated with HF.

The study provides valuable references for further elucidating the pathogenesis of HF, offering beneficial insights for targeting circadian rhythm mechanisms to regulate immune responses and energy metabolism in HF treatment. Five genes identified by the study as diagnostic features could be potential targets for therapy for HF. However, further research is needed to fully understand the relationship between circadian rhythms and heart failure.

CRediT authorship contribution statement

Xuefu Wang: Writing – original draft, Visualization, Software, Formal analysis, Data curation. **Jin Rao:** Methodology, Formal analysis, Data curation. **Li Zhang:** Validation, Supervision, Project administration, Methodology, Investigation. **Xuwen Liu:** Validation, Supervision, Software, Investigation. **Yufeng Zhang:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Many thanks to the uploaders of GEO data for providing us with transcriptome data that we can continue to explore.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e27049.

References

[1] T.A. McDonagh, M. Metra, M. Adamo, R.S. Gardner, A. Baumbach, M. Böhm, H. Burri, J. Butler, J. Čelutkienė, O. Chioncel, J.G.F. Cleland, A.J.S. Coats, M. G. Crespo-Leiro, D. Farmakis, M. Gilard, S. Heymans, A.W. Hoes, T. Jaarsma, E.A. Jankowska, M. Lainscak, C.S.P. Lam, A.R. Lyon, J.J.V. McMurray, A. Mebazaa, R. Mindham, C. Muneretto, M. Francesco Piepoli, S. Price, G.M.C. Rosano, F. Ruschitzka, A. Kathrine Skibelund, 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure, Eur. Heart J. 42 (2021) 3599–3726.

[2] F. Orso, G. Fabbri, A.P. Maggioni, Epidemiology of heart failure, Handb. Exp. Pharmacol. 243 (2017) 15–33.

[3] P. Rossignol, A.F. Hernandez, S.D. Solomon, F. Zannad, Heart failure drug treatment, Lancet 393 (2019) 1034–1044.

[4] T.A. McDonagh, M. Metra, M. Adamo, R.S. Gardner, A. Baumbach, M. Böhm, H. Burri, J. Butler, J. Čelutkienė, O. Chioncel, J.G.F. Cleland, M.G. Crespo-Leiro, D. Farmakis, M. Gilard, S. Heymans, A.W. Hoes, T. Jaarsma, E.A. Jankowska, M. Lainscak, C.S.P. Lam, A.R. Lyon, J.J.V. McMurray, A. Mebazaa, R. Mindham, C. Muneretto, M. Francesco Piepoli, S. Price, G.M.C. Rosano, F. Ruschitzka, A.K. Skibelund, Focused Update of the 2021 ESC Guidelines for the Diagnosis and Treatment of Acute and Chronic Heart Failure, Eur Heart J, 2023, 2023.

[5] C.W. Tsao, A.W. Aday, Z.I. Almarzooq, C.A.M. Anderson, P. Arora, C.L. Avery, C.M. Baker-Smith, A.Z. Beaton, A.K. Boehme, A.E. Buxton, Y. Commodore-Mensah, M.S.V. Elkind, K.R. Evenson, C. Eze-Nliam, S. Fugar, G. Generoso, D.G. Heard, S. Hiremath, J.E. Ho, R. Kalani, D.S. Kazi, D. Ko, D.A. Levine, J. Liu, J. Ma, J.W. Magnani, E.D. Michos, M.E. Mussolino, S.D. Navaneethan, N.I. Parikh, R. Poudel, M. Rezk-Hanna, G.A. Roth, N.S. Shah, M.P. St-Onge, E.L. Thacker, S.S. Virani, J.H. Voeks, N.Y. Wang, N.D. Wong, S.S. Wong, K. Yaffe, S.S. Martin, Heart disease and stroke statistics-2023 update: a report from the American heart association, Circulation 147 (2023) e93–e621.

[6] S. Crnko, B.C. Du Pré, J.P.G. Sluijter, L.W. Van Laake, Circadian rhythms and the molecular clock in cardiovascular biology and disease, Nat. Rev. Cardiol. 16 (2019) 437–447.

[7] N. El Jamal, R. Lordan, S.L. Teegarden, T. Grosser, G. FitzGerald, The circadian biology of heart failure, Circ. Res. 132 (2023) 223–237.

[8] J.S. Takahashi, Transcriptional architecture of the mammalian circadian clock, Nat. Rev. Genet. 18 (2017) 164–179.

[9] J. Bass, M.A. Lazar, Circadian time signatures of fitness and disease, Science 354 (2016) 994–999.

[10] I. Rabinovich-Nikitin, B. Lieberman, T.A. Martino, L.A. Kirshenbaum, Circadian-regulated cell death in cardiovascular diseases, Circulation 139 (2019) 965–980.

[11] F.J. Alibhai, C.J. Reitz, W.T. Peppler, P. Basu, P. Sheppard, E. Choleris, M. Bakovic, T.A. Martino, Female ClockΔ19/Δ19 mice are protected from the development of age-dependent cardiomyopathy, Cardiovasc. Res. 114 (2018) 259–271.

[12] T.A. Martino, G.Y. Oudit, A.M. Herzenberg, N. Tata, M.M. Koletar, G.M. Kabir, D.D. Belsham, P.H. Backx, M.R. Ralph, M.J. Sole, Circadian rhythm disorganization produces profound cardiovascular and renal disease in hamsters, Am. J. Physiol. Regul. Integr. Comp. Physiol. 294 (2008) R1675–R1683.
[13] S. Song, C.L. Tien, H. Cui, P. Basil, N. Zhu, Y. Gong, W. Li, H. Li, Q. Fan, J. Min Choi, W. Luo, Y. Xue, R. Cao, W. Zhou, A.R. Ortiz, B. Stork, V. Mundra, N. Putluri, B. York, M. Chu, J. Chang, S. Yun Jung, L. Xie, J. Song, L. Zhang, Z. Sun, Myocardial rev-erb-mediated diurnal metabolic rhythm and obesity paradox, Circulation 145 (2022) 448–464.
[14] M.E. Young, R.A. Brewer, R.A. Peliciari-Garcia, H.E. Collins, L. He, T.L. Birky, B.W. Peden, E.G. Thompson, B.J. Ammons, M.S. Bray, J.C. Chatham, A.R. Wende, Q. Yang, C.W. Chow, T.A. Martino, K.L. Gamble, Cardiomyocyte-specific BMAL1 plays critical roles in metabolism, signaling, and maintenance of contractile function of the heart, J. Biol. Rhythm. 29 (2014) 257–276.
[15] L. Zhang, M.K. Jain, Circadian regulation of cardiac metabolism, J. Clin. Invest. 131 (2021).
[16] M.E. Young, P. Razeghi, A.M. Cedars, P.H. Guthrie, H. Taegtmeyer, Intrinsic diurnal variations in cardiac metabolism and contractile function, Circ. Res. 89 (2001) 1199–1208.
[17] W. Xu, M.K. Jain, L. Zhang, Molecular link between circadian clocks and cardiac function: a network of core clock, slave clock, and effectors, Curr. Opin. Pharmacol. 57 (2021) 28–40.
[18] S.L. Cox, J.R. O'Siorain, L.E. Fagan, A.M. Curtis, R.G. Carroll, Intertwining roles of circadian and metabolic regulation of the innate immune response, Semin. Immunopathol. 44 (2022) 225–237.
[19] L. Galluzzi, I. Vitale, S.A. Aaronson, J.M. Abrams, D. Adam, P. Agostinis, E.S. Alnemri, L. Altucci, I. Amelio, D.W. Andrews, M. Annicchiarico-Petruzzelli, A. V. Antonov, E. Arama, E.H. Baehrecke, N.A. Barlev, N.G. Bazan, F. Bernassola, M.J.M. Bertrand, K. Bianchi, M.V. Blagosklonny, K. Blomgren, C. Borner, P. Boya, C. Brenner, M. Campanella, E. Candi, D. Carmona-Gutierrez, F. Cecconi, F.K. Chan, N.S. Chandel, E.H. Cheng, J.E. Chipuk, J.A. Cidlowski, A. Ciechanover, G. M. Cohen, M. Conrad, J.R. Cubillos-Ruiz, P.E. Czabotar, V. D'Angiolella, T.M. Dawson, V.L. Dawson, V. De Laurenzi, R. De Maria, K.M. Debatin, R. J. DeBerardinis, M. Deshmukh, N. Di Daniele, F. Di Virgilio, V.M. Dixit, S.J. Dixon, C.S. Duckett, B.D. Dynlacht, W.S. El-Deiry, J.W. Elrod, G.M. Fimia, S. Fulda, A.J. García-Sáez, A.D. Garg, C. Garrido, E. Gavathiotis, P. Golstein, E. Gottlieb, D.R. Green, L.A. Greene, H. Gronemeyer, A. Gross, G. Hajnoczky, J.M. Hardwick, I.S. Harris, M.O. Hengartner, C. Hetz, H. Ichijo, M. Jäättelä, B. Joseph, P.J. Jost, P.P. Juin, W.J. Kaiser, M. Karin, T. Kaufmann, O. Kepp, A. Kimchi, R.N. Kitsis, D. J. Klionsky, R.A. Knight, S. Kumar, S.W. Lee, J.J. Lemasters, B. Levine, A. Linkermann, S.A. Lipton, R.A. Lockshin, C. López-Otín, S.W. Lowe, T. Luedde, E. Lugli, M. MacFarlane, F. Madeo, M. Malewicz, W. Malorni, G. Manic, et al., Molecular mechanisms of cell death: recommendations of the nomenclature committee on cell death 2018, Cell Death Differ. 25 (2018) 486–541.
[20] A.H. Saeidian, L. Youssefian, H. Vahidnezhad, J. Uitto, Research techniques made simple: whole-transcriptome sequencing by RNA-seq for diagnosis of monogenic disorders, J. Invest. Dermatol. 140 (2020) 1117–1126.e1.
[21] D.A. Hanauer, D.R. Rhodes, C. Sinha-Kumar, A.M. Chinnaiyan, Bioinformatics approaches in the study of cancer, Curr. Mol. Med. 7 (2007) 133–141.
[22] A. Joshi, M. Rienks, K. Theofilatos, M. Mayr, Systems biology in cardiovascular disease: a multiomics approach, Nat. Rev. Cardiol. 18 (2021) 313–330.
[23] V. Vijayasarveswari, A.M. Andrew, M. Jusoh, T. Sabapathy, R.A.A. Raof, M.N.M. Yasin, R.B. Ahmad, S. Khatun, H.A. Rahim, Multi-stage feature selection (MSFS) algorithm for UWB-based early breast cancer size prediction, PLoS One 15 (2020) e0229367.
[24] J. Li, Q. Xu, M. Wu, T. Huang, Y. Wang, Pan-Cancer classification based on self-normalizing neural networks and feature selection, Front. Bioeng. Biotechnol. 8 (2020) 766.
[25] I.S. Forrest, B.O. Petrazzini, Á. Duffy, J.K. Park, C. Marquez-Luna, D.M. Jordan, G. Rocheleau, J.H. Cho, R.S. Rosenson, J. Narula, G.N. Nadkarni, R. Do, Machine learning-based marker for coronary artery disease: derivation and validation in two longitudinal cohorts, Lancet 401 (2023) 215–225.
[26] X. Qin, S. Yi, J. Rong, H. Lu, B. Ji, W. Zhang, R. Ding, L. Wu, Z. Chen, Identification of anoikis-related genes classification patterns and immune infiltration characterization in ischemic stroke based on machine learning, Front. Aging Neurosci. 15 (2023) 1142163.
[27] Y. Liu, M. Morley, J. Brandimarto, S. Hannenhalli, Y. Hu, E.A. Ashley, W.H. Tang, C.S. Moravec, K.B. Margulies, T.P. Cappola, M. Li, RNA-Seq identifies novel myocardial gene expression signatures of heart failure, Genomics 105 (2015) 83–89.
[28] S.J. Matkovich, B. Al Khiami, I.R. Efimov, S. Evans, J. Vader, A. Jain, B.H. Brownstein, R.S. Hotchkiss, D.L. Mann, Widespread down-regulation of cardiac mitochondrial and sarcomeric genes in patients with sepsis, Crit. Care Med. 45 (2017) 407–414.
[29] S. Greco, P. Fasanaro, S. Castelvecchio, Y. D'Alessandra, D. Arcelli, M. Di Donato, A. Malavazos, M.C. Capogrossi, L. Menicanti, F. Martelli, MicroRNA dysregulation in diabetic ischemic heart failure patients, Diabetes 61 (2012) 1633–1641.
[30] S. Hannenhalli, M.E. Putt, J.M. Gilmore, J. Wang, M.S. Parmacek, J.A. Epstein, E.E. Morrisey, K.B. Margulies, T.P. Cappola, Transcriptional genomics associates FOX transcription factors with human heart failure, Circulation 114 (2006) 1269–1276.
[31] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015) e47.
[32] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (2012) 284–287.
[33] M.B. Kursa, W.R. Rudnicki, Feature selection with the Boruta package, J. Stat. Software 36 (2010) 1–13.
[34] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[35] W.S. Noble, What is a support vector machine? Nat. Biotechnol. 24 (2006) 1565–1567.
[36] T. Chen, C. Guestrin, Xgboost: a scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
[37] S. Engebretsen, J. Bohlin, Statistical predictions with glmnet, Clin. Epigenet. 11 (2019) 123.
[38] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (2010) 1572–1573.
[39] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7.
[40] R. Gaujoux, C. Seoighe, A flexible R package for nonnegative matrix factorization, BMC Bioinf. 11 (2010) 367.
[41] M. Wang, W. Pan, Y. Xu, J. Zhang, J. Wan, H. Jiang, Microglia-mediated neuroinflammation: a potential target for the treatment of cardiovascular diseases, J. Inflamm. Res. 15 (2022) 3083–3094.
[42] A. Kohsaka, P. Das, I. Hashimoto, T. Nakao, Y. Deguchi, S.S. Gouraud, H. Waki, Y. Muragaki, M. Maeda, The circadian clock maintains cardiac function by regulating mitochondrial metabolism in mice, PLoS One 9 (2014) e112811.
[43] A.M. Curtis, Y. Cheng, S. Kapoor, D. Reilly, T.S. Price, G.A. Fitzgerald, Circadian variation of blood pressure and the vascular response to asynchronous stress, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 3450–3455.
[44] L.A. Solt, Y. Wang, S. Banerjee, T. Hughes, D.J. Kojetin, T. Lundasen, Y. Shin, J. Liu, M.D. Cameron, R. Noel, S.H. Yoo, J.S. Takahashi, A.A. Butler, T. M. Kamenecka, T.P. Burris, Regulation of circadian behaviour and metabolism by synthetic REV-ERB agonists, Nature 485 (2012) 62–68.
[45] L. Zhang, R. Zhang, C.L. Tien, R.E. Chan, K. Sugi, C. Fu, A.C. Griffin, Y. Shen, T.P. Burris, X. Liao, M.K. Jain, REV-ERBα ameliorates heart failure through transcription repression, JCI Insight 2 (2017).
[46] S. Liang, A. Ma, S. Yang, Y. Wang, Q. Ma, A review of matched-pairs feature selection methods for gene expression data analysis, Comput. Struct. Biotechnol. J. 16 (2018) 88–97.
[47] M. Keller, J. Mazuch, U. Abraham, G.D. Eom, E.D. Herzog, H.D. Volk, A. Kramer, B. Maier, A circadian clock in macrophages controls inflammatory immune responses, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 21407–21412.
[48] R. Orozco-Solis, L. Aguilar-Arnal, Circadian regulation of immunity through epigenetic mechanisms, Front. Cell. Infect. Microbiol. 10 (2020) 96.
[49] V. Sorokin, C.C. Woo, Role of Serpina3 in vascular biology, Int. J. Cardiol. 304 (2020) 154–155.
[50] L. Zhao, Z. Guo, P. Wang, M. Zheng, X. Yang, Y. Liu, Z. Ma, M. Chen, X. Yang, Proteomics of epicardial adipose tissue in patients with heart failure, J. Cell Mol. Med. 24 (2020) 511–520.
[51] L. Delrue, M. Vanderheyden, M. Beles, P. Paolisso, G. Di Gioia, R. Dierckx, S. Verstreken, M. Goethals, W. Heggermont, J. Bartunek, Circulating SERPINA3 improves prognostic stratification in patients with a de novo or worsened heart failure, ESC Heart Fail 8 (2021) 4780–4790.
[52] M. Mueckler, C. Makepeace, Transmembrane segment 6 of the Glut1 glucose transporter is an outer helix and contains amino acid side chains essential for transport activity, J. Biol. Chem. 283 (2008) 11550–11555.

[53] I. Luptak, J. Yan, L. Cui, M. Jain, R. Liao, R. Tian, Long-term effects of increased glucose entry on mouse hearts during normal aging and ischemic stress, Circulation 116 (2007) 901–909.
[54] A.J. Freemerman, A.R. Johnson, G.N. Sacks, J.J. Milner, E.L. Kirk, M.A. Troester, A.N. Macintyre, P. Goraksha-Hicks, J.C. Rathmell, L. Makowski, Metabolic reprogramming of macrophages: glucose transporter 1 (GLUT1)-mediated glucose metabolism drives a proinflammatory phenotype, J. Biol. Chem. 289 (2014) 7884–7896.
[55] J.P. Goetze, B.G. Bruneau, H.R. Ramos, T. Ogawa, M.K. de Bold, A.J. de Bold, Cardiac natriuretic peptides, Nat. Rev. Cardiol. 17 (2020) 698–717.
[56] M. Shvedova, Y. Anfinogenova, E.N. Atochina-Vasserman, I.A. Schepetkin, D.N. Atochin, c-Jun N-terminal kinases (JNKs) in myocardial and cerebral ischemia/reperfusion injury, Front. Pharmacol. 9 (2018) 715.
[57] F. Liu, Y. Deng, Y. Zhao, Z. Li, J. Gao, Y. Zhang, X. Yang, Y. Liu, Y. Xia, Time series RNA-seq analysis identifies MAPK10 as a critical gene in diabetes mellitus-induced atrial fibrillation in mice, J. Mol. Cell. Cardiol. 168 (2022) 70–82.
[58] F. Urano, X. Wang, A. Bertolotti, Y. Zhang, P. Chung, H.P. Harding, D. Ron, Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1, Science 287 (2000) 664–666.
[59] Y.W. Deng, F. Liu, Z.T. Li, J.H. Gao, Y. Zhao, X.L. Yang, Y.L. Xia, Hyperglycemia promotes myocardial dysfunction via the ERS-MAPK10 signaling pathway in db/db mice, Lab. Invest. 102 (2022) 1192–1202.
[60] Q. Shen, L. Liu, X. Gu, D. Xing, Photobiomodulation suppresses JNK3 by activation of ERK/MKP7 to attenuate AMPA receptor endocytosis in Alzheimer's disease, Aging Cell 20 (2021) e13289.
[61] S. Kralisch, J. Klein, U. Lossner, M. Bluher, R. Paschke, M. Stumvoll, M. Fasshauer, Interleukin-6 is a negative regulator of visfatin gene expression in 3T3-L1 adipocytes, Am. J. Physiol. Endocrinol. Metab. 289 (2005) E586–E590.
[62] G. Cavadini, S. Petrzilka, P. Kohler, C. Jud, I. Tobler, T. Birchler, A. Fontana, TNF-alpha suppresses the expression of clock genes by interfering with E-box-mediated transcription, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 12843–12848.
[63] J. Byun, S.I. Oka, N. Imai, C.Y. Huang, G. Ralda, P. Zhai, Y. Ikeda, S. Ikeda, J. Sadoshima, Both gain and loss of Nampt function promote pressure overload-induced heart failure, Am. J. Physiol. Heart Circ. Physiol. 317 (2019) H711–h725.
[64] L. Li, H. Li, C.L. Tien, M.K. Jain, L. Zhang, Kruppel-like factor 15 regulates the circadian susceptibility to ischemia reperfusion injury in the heart, Circulation 141 (2020) 1427–1429.
[65] L. Li, W. Xu, L. Zhang, KLF15 regulates oxidative stress response in cardiomyocytes through NAD( ), Metabolites 11 (2021).
[66] G. Cano, A.F. Sved, L. Rinaman, B.S. Rabin, J.P. Card, Characterization of the central nervous system innervation of the rat spleen using viral transneuronal tracing, J. Comp. Neurol. 439 (2001) 1–18.
[67] I.J. Elenkov, R.L. Wilder, G.P. Chrousos, E.S. Vizi, The sympathetic nerve–an integrative interface between two supersystems: the brain and the immune system, Pharmacol. Rev. 52 (2000) 595–638.
[68] J.M. Adrover, C. Del Fresno, G. Crainiciuc, M.I. Cuartero, M. Casanova-Acebes, L.A. Weiss, H. Huerga-Encabo, C. Silvestre-Roig, J. Rossaint, I. Cossío, A. V. Lechuga-Vieco, J. García-Prieto, M. Gómez-Parrizas, J.A. Quintana, I. Ballesteros, S. Martin-Salamanca, A. Aroca-Crevillen, S.Z. Chong, M. Evrard, K. Balabanian, J. López, K. Bidzhekov, F. Bachelerie, F. Abad-Santos, C. Muñoz-Calleja, A. Zarbock, O. Soehnlein, C. Weber, L.G. Ng, C. Lopez-Rodriguez, D. Sancho, M.A. Moro, B. Ibáñez, A. Hidalgo, A neutrophil timer coordinates immune defense and vascular protection, Immunity 50 (2019) 390–402.e10.
[69] R.W. Logan, D.K. Sarkar, Circadian nature of immune function, Mol. Cell. Endocrinol. 349 (2012) 82–90.
[70] C.E. Sutton, C.M. Finlay, M. Raverdeau, J.O. Early, J. DeCourcey, Z. Zaslona, L.A.J. O'Neill, K.H.G. Mills, A.M. Curtis, Loss of the molecular clock in myeloid cells exacerbates T cell-mediated CNS autoimmune disease, Nat. Commun. 8 (2017) 1923.
[71] D. Druzd, O. Matveeva, L. Ince, U. Harrison, W. He, C. Schmal, H. Herzel, A.H. Tsang, N. Kawakami, A. Leliavski, O. Uhl, L. Yao, L.E. Sander, C.S. Chen, K. Kraus, A. de Juan, S.M. Hergenhan, M. Ehlers, B. Koletzko, R. Haas, W. Solbach, H. Oster, C. Scheiermann, Lymphocyte circadian clocks control lymph node trafficking and adaptive immune responses, Immunity 46 (2017) 120–132.
[72] M. Ith, C. Stettler, J. Xu, C. Boesch, R. Kreis, Cardiac lipid levels show diurnal changes and long-term variations in healthy human subjects, NMR Biomed. 27 (2014) 1285–1292.
[73] L. Adamo, C. Rocha-Resende, S.D. Prabhu, D.L. Mann, Reappraising the role of inflammation in heart failure, Nat. Rev. Cardiol. 17 (2020) 269–285.
[74] J.A. Haspel, R. Anafi, M.K. Brown, N. Cermakian, C. Depner, P. Desplats, A.E. Gelman, M. Haack, S. Jelic, B.S. Kim, A.D. Laposky, Y.C. Lee, E. Mongodin, A. A. Prather, B.J. Prendergast, C. Reardon, A.C. Shaw, S. Sengupta, É. Szentirmai, M. Thakkar, W.E. Walker, L.A. Solt, Perfect timing: circadian rhythms, sleep, and immunity - an NIH workshop summary, JCI Insight 5 (2020).
[75] C.J. Reitz, F.J. Alibhai, T.N. Khatua, M. Rasouli, B.W. Bridle, T.P. Burris, T.A. Martino, SR9009 administered for one day after myocardial ischemia-reperfusion prevents heart failure in mice by targeting the cardiac inflammasome, Commun. Biol. 2 (2019) 353.