



OPEN

Machine learning and bioinformatics to identify 8 autophagy-related biomarkers and construct gene regulatory networks in dilated cardiomyopathy

Fengjun Zhang^{1,10}, Mingyue Xia^{2,10}, Jiarong Jiang³, Shuai Wang⁴, Qiong Zhao⁵, Cheng Yu⁶, Jinzhen Yu⁷, Dexian Xian¹, Xiao Li⁸, Lin Zhang⁹, Yuan Liu^{5,10}✉ & Min Peng^{5,10}✉

Dilated cardiomyopathy (DCM) is a condition of impaired ventricular remodeling and systolic diastole that is often complicated by arrhythmias and heart failure with a poor prognosis. This study attempted to identify autophagy-related genes (ARGs) with diagnostic biomarkers of DCM using machine learning and bioinformatics approaches. Differential analysis of whole gene microarray data of DCM from the Gene Expression Omnibus (GEO) database was performed using the NetworkAnalyst 3.0 platform. Differentially expressed genes (DEGs) matching ($|\log_2\text{FoldChange}| \geq 0.8$, $p \text{ value} < 0.05$) were obtained in the GSE4172 dataset by merging ARGs from the autophagy gene libraries, HADb and HAMdb, to obtain autophagy-related differentially expressed genes (AR-DEGs) in DCM. The correlation analysis of AR-DEGs and their visualization were performed using R language. Gene Ontology (GO) enrichment analysis and combined multi-database pathway analysis were served by the Enrichr online enrichment analysis platform. We used machine learning to screen the diagnostic biomarkers of DCM. The transcription factors gene regulatory network was constructed by the JASPAR database of the NetworkAnalyst 3.0 platform. We also used the drug Signatures database (DSigDB) drug database of the Enrichr platform to screen the gene target drugs for DCM. Finally, we used the DisGeNET database to analyze the comorbidities associated with DCM. In the present study, we identified 23 AR-DEGs of DCM. Eight (*PLEKHF1*, *HSPG2*, *HSF1*, *TRIM65*, *DICER1*, *VDAC1*, *BAD*, *TFEB*) molecular markers of DCM were obtained by two machine learning algorithms. Transcription factors gene regulatory network was established. Finally, 10 gene-targeted drugs and complications for DCM were identified.

Abbreviations

DCM Dilated cardiomyopathy
ARGs Autophagy-related genes

¹College of Acupuncture and Massage, Shandong University of Traditional Chinese Medicine, Jinan, China. ²College of Traditional Chinese Medicine, Shandong University of Traditional Chinese Medicine, Jinan, China. ³Department of Cardiology, PLA Rocket Force Characteristic Medical Center, Beijing, China. ⁴Department of Pediatric Surgery, Shandong Provincial Hospital affiliated to Shandong First Medical University, Jinan, China. ⁵Department of Traditional Chinese Medicine, Shandong Provincial Hospital affiliated to Shandong First Medical University, Jinan, China. ⁶Department of Traditional Chinese Medicine Classics, Shandong University of Traditional Chinese Medicine Affiliated Hospital, Jinan, Shandong, China. ⁷First Clinical Medical College, Shandong University of Traditional Chinese Medicine, Jinan, China. ⁸Department of Cardiology, Shandong University of Traditional Chinese Medicine Affiliated Hospital, Jinan, Shandong, China. ⁹Department of Clinical Pharmacy, Shaoxing People's Hospital, Shaoxing Hospital, Zhejiang University School of Medicine, Shaoxing, China. ¹⁰These authors contributed equally: Fengjun Zhang, Mingyue Xia, Yuan Liu and Min Peng. ✉email: liuyuanly0429@163.com; pengmin186@126.com

GEO	Gene Expression Omnibus
DEGs	Differentially expressed genes
HADb	Human Autophagy Database
HAMdb	Human Autophagy Modulator Database
AR-DEGs	Autophagy-related differentially expressed genes
GO	Gene ontology
DSigDB	Drug Signatures database
KEGG	Kyoto Encyclopedia of Genes and Genomes
EMB	Endomyocardial biopsy
BP	Biological process
CC	Chromosomal location
MF	Molecular function
LASSO	Least absolute shrinkage and selection operator
SVM-RFE	Support vector machine-recursive feature elimination
TF	Transcription factor
ALP	Autophagic lysosomal pathway
CypD	Chaperone procyclin D
cytosol	Cytoplasmic lysosol
LSD	Lysosomal storage disorders
YAP	Lysosomal storage disorders

Dilated cardiomyopathy (DCM), which manifests clinically as ventricular dilatation and impaired progressive systolic diastole, is one of the most prevalent disease worldwide. It has a heterogeneous etiology, with viral infections, inflammatory reactions, genetic factors¹, etc. It can also cause arrhythmias and atrioventricular block, resulting in sudden cardiac death and heart failure. These circumstances often occur with a poor prognosis². It is reported that men with DCM have a higher mortality rate than women patients³. Endomyocardial biopsy (EMB) is the gold standard for the diagnosis of myocarditis and DCM. However, in clinical practice, DCM is not diagnosed and treated promptly, considering the high cardiac complications of performing EMB and treatment limitation⁴. Therefore, developing innovative, non-invasive biomarkers for DCM is essential to improve diagnostic accuracy.

Autophagy is a cellular self-degradation process that removes errant proteins and damaged organelles. It also eliminates intracellular pathogens and is often considered a survival mechanism⁵. Numerous studies have shown that autophagy genes are involved in various phenotypes and human diseases⁶, including neurodegenerative diseases⁷, liver diseases⁸, muscle diseases⁹, cancer¹⁰, and cardiac diseases¹¹. Evidence shows that autophagy is essential in maintaining cardiomyocyte homeostasis¹² and regulating the prognostic efficacy of cardiac diseases. In addition, an increasing number of animal models and clinical studies have reported the involvement of autophagy-related genes (ARGs) in the ventricular remodeling process, which is related to the mechanism of action of DCM^{13,14}. However, ARGs' diagnostic performance and prognostic efficacy in DCM have not been fully elucidated.

In this study, we downloaded gene expression profile data of DCM from the Gene Expression Omnibus (GEO) database, applied bioinformatics to search for AR-DEGs in DCM, and visualized the correlation between genes. Subsequently, gene enrichment analysis was performed on tagging gene functions and exploring pathogenesis. Machine learning algorithms were afterward executed to filter and identify diagnostic biomarkers of DCM. In addition, based on the diagnostic biomarkers of DCM, transcription factors gene regulatory network and gene-targeted drugs were predicted to provide ideas for clinical precision therapy and experimental studies. The DisGeNET database was used for association analysis of DCM with other related diseases to provide a transcriptomic basis for further investigation of the potential pathogenesis of the disease.

The flow chart of this study was shown in Fig. 1.

Results

Identification of autophagy-related differentially expressed genes (AR-DEGs) for dilated cardiomyopathy (DCM). The GSE4172 dataset was used to screen for DEGs in DCM. Based on the threshold set to $|\log_2\text{FoldChange}| \geq 0.8$, $p\text{-value} < 0.05$, 770 DEGs were acquired, containing 366 up-regulated genes and 404 down-regulated genes. In addition, the heatmap (Fig. 2a) showed the expression of the top 60 DEGs and the asymptotic volcano plot (Fig. 2b) showed the distribution of DEGs.

803 ARGs were obtained through two autophagy-related gene databases, HADb and HAMdb. The Venn diagram obtained by the Omicshare online tool demonstrated 23 AR-DEGs of DCM (*ADIPOQ*, *TRIM17*, *PPFIA4*, *CAPN12*, *PLEKHF1*, *RCAN1*, *RAB12*, *CXCR4*, *HSPG2*, *EIF4EBP1*, *HSF1*, *ZC3H12A*, *PRKAB1*, *TRIM65*, *ARSA*, *GABARAPL1*, *DICER1*, *VDAC1*, *CHMP4B*, *AGTR1*, *BAD*, *TFEB*, *AP2M1*) (Fig. 3). The relevant functions of 23 AR-DEGs were shown in Supplementary Table S1.

Correlation matrix analysis of 23 AR-DEGs and the expression of these genes in the disease and control groups were demonstrated in Fig. 4a. The absolute values of relative coefficients between genes exceeding 0.5 were considered to be of typical significance and were labeled in Fig. 4b. Moreover, some genes showed a strong association with others.

Gene Ontology (GO), Pathway Enrichment Analysis. GO analysis and multiple databases (KEGG, Wikipathway, Bioplanet, Reactome) pathway analysis were implemented through the Enrichr database. Three categories of GO analysis were obtained by clustering AR-DEGs of DCM, namely biological process (BP), chro-

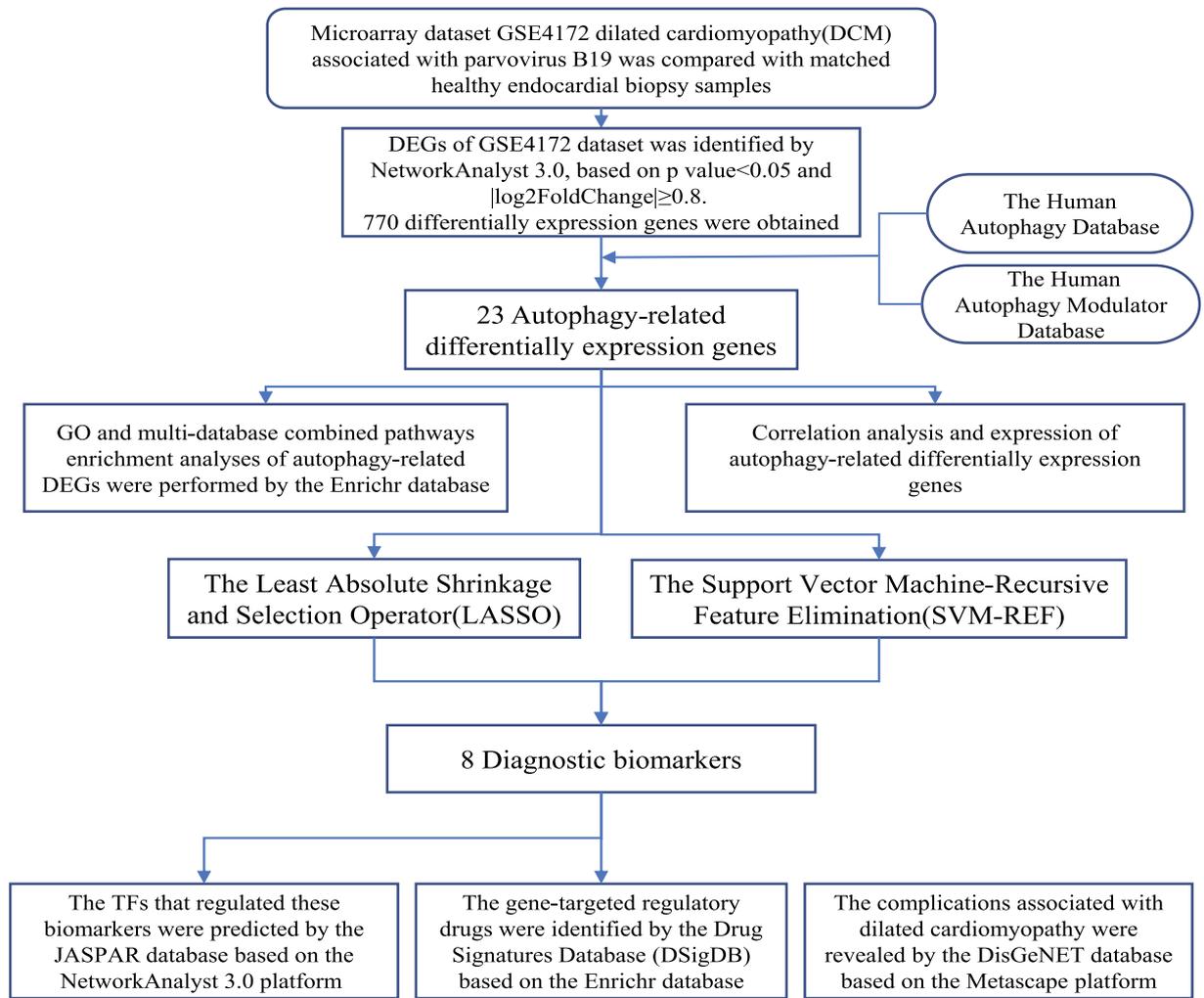


Figure 1. Workflow diagram of the current study. GO, go ontology; TFs, transcription factors.

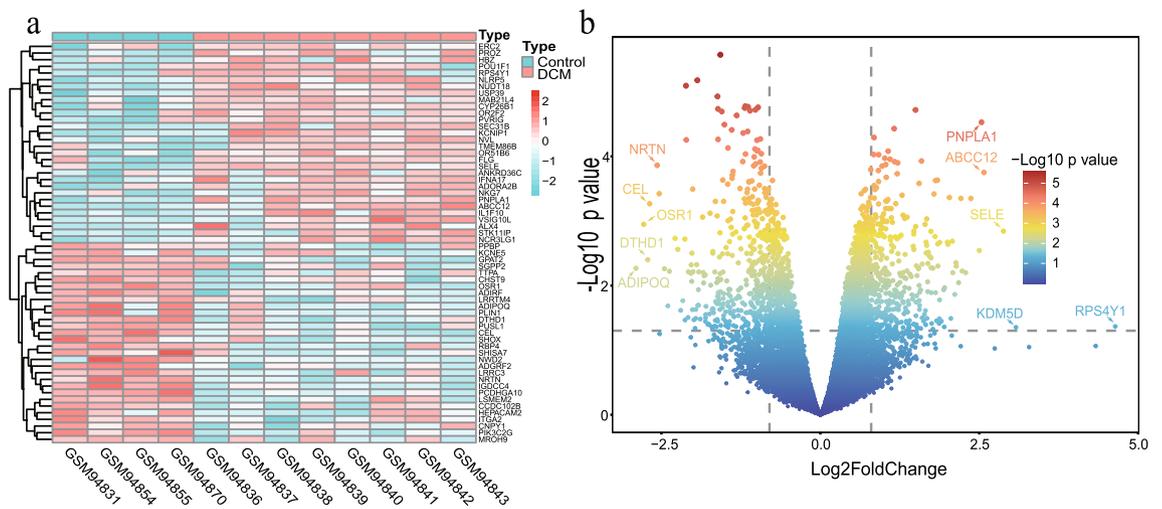


Figure 2. DEGs differential analysis of GSE4172 dataset. (a) Heatmap of DEGs in GSE4172 dataset (n = 60, p < 0.05, |log2 FoldChange| ≥ 0.8). (b) Asymptotic volcano map of gene expression in the GSE4172 dataset. The two vertical lines indicated gene expression ploidy changes > 0.8 and < -0.8, respectively, and the horizontal line indicated a p value of 0.05. The color of the dots represented the level of the p value. The top 10 significantly expressed genes among the DEGs were labeled on the graph.

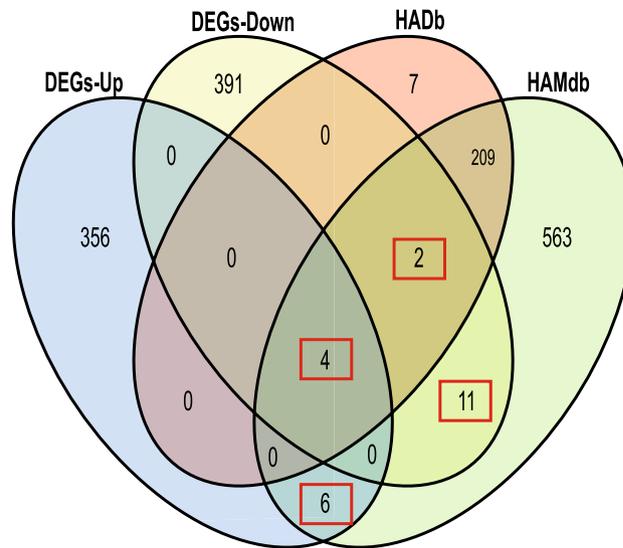


Figure 3. AR-DEGs were shown by Venn diagram. 366 DEGs-Up and 404 DEGs-Down were intersected with 232 and 796 autophagy-associated genes from the HADb and HAMDb autophagy gene pools, with 23 genes being identical. The number of intersecting genes was marked in the red box. DEGs-Up, differentially expressed up-regulated genes; DEGs-Down, differentially expressed down-regulated genes.

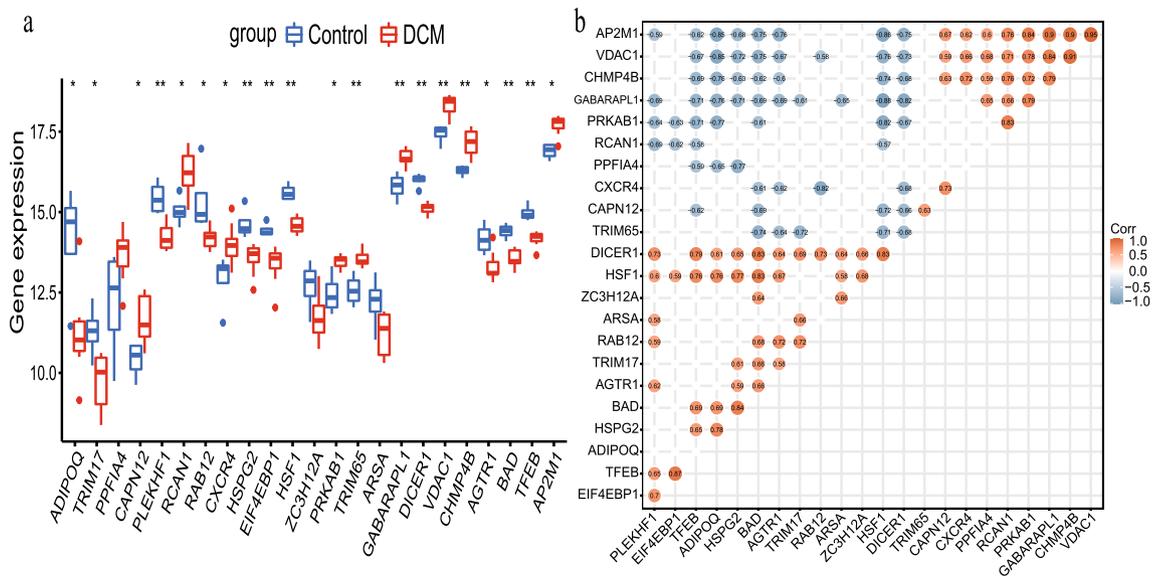


Figure 4. 23 AR-DEGs in dilated cardiomyopathy (DCM) group and control group and their correlation. (a) Box plot of the expression levels of 23 DEGs-Down in the DCM and control groups. The blue box plots above the corresponding gene names indicated expression in control groups, whereas the red box plots indicated expression in DCM groups. (b) Correlation heatmap of 23 AR-DEGs. The color within the circle shape and the magnitude of the correlation value represented the strength of the correlation; red represented positive correlation and blue represented negative correlation. The darker the color, the larger the absolute value of the correlation value represented a stronger correlation.

mosomal location (CC), and molecular function (MF). The top ten terms of each category were predicted in Table 1.

Based on the number of gene interactions, BP was mainly focused on the regulation of autophagy, positive regulation of autophagy, positive regulation of the cellular catabolic process, and macroautophagy. For cellular components, lysosome and lytic vacuole were significantly associated with autophagy-related differential genes, ultimately pointing to inflammatory cardiomyopathy in response to the human heart. Molecular functional studies revealed that AR-DEGs were most concentrated in low-density lipoprotein particle receptor binding. A similar concentration level could be found in lipoprotein particle receptor binding and endoribonuclease activity.

Go category	GO pathways	GO ID	p-value	AR-DEGs
Biological process	Regulation of autophagy	GO:0010506	5.24E-09	<i>PLEKHF1;BAD;ZC3H12A;TFEB;VDAC1;PRKAB1;TRIM65</i>
	Positive regulation of autophagy	GO:0010508	5.2E-08	<i>PLEKHF1;BAD;ZC3H12A;TFEB;TRIM65</i>
	Positive regulation of cellular catabolic process	GO:0031331	4.93E-07	<i>PLEKHF1;BAD;ZC3H12A;TFEB;TRIM65</i>
	Positive regulation of cold-induced thermogenesis	GO:0120162	4.29E-06	<i>ADIPOQ;HSF1;CXCR4;PRKAB1</i>
	Positive regulation of metabolic process	GO:0009893	7.87E-06	<i>ADIPOQ;HSF1;CXCR4;PRKAB1</i>
	Macroautophagy	GO:0016236	9.99E-06	<i>GABARAPL1;CHMP4B;VDAC1;PRKAB1</i>
	Response to sodium arsenite	GO:1903935	1.26E-05	<i>HSF1;ZC3H12A</i>
	Cellular response to sodium arsenite	GO:1903936	1.26E-05	<i>HSF1;ZC3H12A</i>
	Cellular response to salt	GO:1902075	1.89E-05	<i>HSF1;ZC3H12A</i>
Negative regulation of tumor necrosis factor production	GO:0032720	1.95E-05	<i>ADIPOQ;ZC3H12A;DICER1c</i>	
Molecular function	Low-density lipoprotein particle receptor binding	GO:0050750	0.000315	<i>HSPG2;AP2M1</i>
	Lipoprotein particle receptor binding	GO:0070325	0.00047	<i>HSPG2;AP2M1</i>
	Endoribonuclease activity	GO:0004521	0.000694	<i>ZC3H12A;DICER1</i>
	Regulatory RNA binding	GO:0061980	0.000961	<i>ZC3H12A;DICER1</i>
	Endonuclease activity	GO:0004519	0.001559	<i>ZC3H12A;DICER1</i>
	Kinase binding	GO:0019900	0.001742	<i>BAD;HSF1;VDAC1;PRKAB1</i>
	Protein kinase binding	GO:0019901	0.002446	<i>BAD;HSF1;VDAC1;PRKAB1</i>
	Ribonuclease activity	GO:0004540	0.002595	<i>ZC3H12A;DICER1</i>
	C-X-C chemokine receptor activity	GO:0016494	0.005737	<i>CXCR4</i>
Intronic transcription regulatory region sequence-specific DNA binding	GO:0001161	0.005737	<i>HSF1</i>	
Cellular component	lysosome	GO:0005764	7.4E-07	<i>ARSA;PLEKHF1;RAB12;TFEB;CXCR4;HSPG2;AP2M1</i>
	Lytic vacuole	GO:0000323	0.000105	<i>ARSA;PLEKHF1;RAB12;CXCR4</i>
	Lytic vacuole membrane	GO:0098852	0.003419	<i>PLEKHF1;TFEB;AP2M1</i>
	Lysosomal lumen	GO:0043202	0.00436	<i>ARSA;HSPG2</i>
	Endosome membrane	GO:0010008	0.005917	<i>PLEKHF1;RAB12;AP2M1</i>
	Lysosomal membrane	GO:0005765	0.006172	<i>PLEKHF1;TFEB;AP2M1</i>
	Nuclear stress granule	GO:0097165	0.006881	<i>HSF1</i>
	AP-2 adaptor complex	GO:0030122	0.008023	<i>AP2M1</i>
	Clathrin coat of endocytic vesicle	GO:0030128	0.008023	<i>AP2M1</i>
Mitochondrial outer membrane	GO:0005741	0.009136	<i>BAD;VDAC1</i>	

Table 1. GO category, GO pathways, corresponding p-values, and AR-DEGs.

Notably, the results of the pathway analysis in this study were joint (Table 2). Through the previously set database, the Longevity regulating pathway, Macroautophagy, PI3K-AKT-mTOR signaling pathway, therapeutic opportunities, and AMPK signaling were identified as the top pathways.

A comparison of GO terms was presented in Fig. 5a. Figure 5b provided pathway analysis from multiple databases.

Machine learning screened for autophagy-related biomarkers of DCM. The expression matrices of 23 AR-DEGs were used to construct the best diagnostic model using both least absolute shrinkage selection operator (LASSO) regression and support vector machine recursive feature elimination (SVM-RFE) algorithms to finally obtain potential diagnostic biomarkers of DCM. The LASSO regression algorithm narrowed down the range of AR-DEGs of DCM and obtained 9 variables as potential diagnostic biomarkers for DCM (Fig. 6a). The SVM-RFE algorithm was implemented to identify 13 signature genes (Fig. 6b).

Finally, 8 overlapping genes (*PLEKHF1*, *HSPG2*, *HSF1*, *TRIM65*, *DICER1*, *VDAC1*, *BAD*, *TFEB*) were obtained (Fig. 6c).

Construction of transcription factor (TF)-gene regulatory network. Based on the JASPAR TF binding site profile database, TF-gene regulatory network was constructed using the NetworkAnalyst 3.0 platform. The TF-gene regulatory network was constructed based on 8 diagnostic biomarkers of DCM (*PLEKHF1*, *HSPG2*, *HSF1*, *TRIM65*, *DICER1*, *VDAC1*, *BAD*, *TFEB*) (Figure). The network included 46 loci with 76 edges. In detail, these loci are combined by 8 seed genes and 38 transcription factors. *TFEB* was regulated by 19 transcription factors and *DICER1* was regulated by 15 transcription factors. Figure 7 showed the TF-gene regulatory network.

Databases	Pathways	p-value	Genes
KEGG	Longevity regulating pathway	2.12E-04	ADIPOQ;EIF4EBP1;PRKAB1
	AMPK signaling pathway	3.42E-04	ADIPOQ;EIF4EBP1;PRKAB1
	Apelin signaling pathway	5.04E-04	GABARAPL1;AGTR1;PRKAB1
	Insulin signaling pathway	5.04E-04	BAD;EIF4EBP1;PRKAB1
	cGMP-PKG signaling pathway	8.96E-04	BAD;AGTR1;VDAC1
	Calcium signaling pathway	0.002531	AGTR1;CXCR4;VDAC1
	Acute myeloid leukemia	0.002673	BAD;EIF4EBP1
	Mitophagy	0.002752	GABARAPL1;TFEB
	Adipocytokine signaling pathway	0.002832	ADIPOQ;PRKAB1
	Endocytosis	0.002905	CXCR4;CHMP4B;AP2M1
BioPlanet	AMPK signaling	6.34E-05	ADIPOQ;EIF4EBP1;PRKAB1
	Mitochondrial pathway of apoptosis: BH3-only Bcl-2 family	1.83E-04	CAPN12;BAD;VDAC1
	Phosphoinositides and their downstream targets	3.73E-04	BAD;AP2M1
	PKB-mediated events	4.70E-04	EIF4EBP1;PRKAB1
	TOR signaling	6.94E-04	EIF4EBP1;PRKAB1
	Endocytosis	0.001527	CXCR4;CHMP4B;AP2M1
	Calcineurin-dependent NFAT signaling role in lymphocytes	0.00181	RCAN1;BAD
	Acute myeloid leukemia	0.001943	BAD;EIF4EBP1
	ERK1/ERK2 MAPK pathway	0.002518	BAD;EIF4EBP1
	Adipocytokine signaling pathway	0.002673	ADIPOQ;PRKAB1
Reactome	Macroautophagy Homo sapiens R-HSA-1632852	6.07E-05	GABARAPL1;CHMP4B;PRKAB1
	Cellular responses to stress Homo sapiens R-HSA-2262752	7.49E-04	GABARAPL1;HSF1;CHMP4B;PRKAB1
	mTOR signalling Homo sapiens R-HSA-165159	9.13E-04	EIF4EBP1;PRKAB1
	PKB-mediated events Homo sapiens R-HSA-109703	9.61E-04	EIF4EBP1;PRKAB1
	Disease Homo sapiens R-HSA-1643685	0.001206	BAD;CXCR4;CHMP4B;HSPG2;AP2M1
	HIV Infection Homo sapiens R-HSA-162906	0.002028	CXCR4;CHMP4B;AP2M1
	PI3K Cascade Homo sapiens R-HSA-109704	0.003693	EIF4EBP1;PRKAB1
	Degradation of the extracellular matrix Homo sapiens R-HSA-1474228	0.006547	CAPN12;HSPG2
	Infectious disease Homo sapiens R-HSA-5663205	0.007144	CXCR4;CHMP4B;AP2M1
	Nef Mediated CD8 Down-regulation Homo sapiens R-HSA-182218	0.008023	AP2M1
WikiPathway	PI3K-AKT-mTOR signaling pathway and therapeutic opportunities WP3844	5.28E-06	BAD;TFEB;EIF4EBP1
	Leptin and adiponectin WP3934	5.66E-05	ADIPOQ;PRKAB1
	AMP-activated protein kinase (AMPK) signaling WP1403	6.62E-05	ADIPOQ;EIF4EBP1;PRKAB1
	The influence of laminopathies on Wnt signaling WP4844	7.36E-04	ADIPOQ;DICER1
	Target Of Rapamycin (TOR) Signaling WP1471	7.78E-04	EIF4EBP1;PRKAB1
	Synaptic signaling pathways associated with autism spectrum disorder WP4539	0.001499	EIF4EBP1;PRKAB1
	RAC1/PAK1/p38/MMP2 Pathway WP3303	0.002752	BAD;EIF4EBP1
	Peptide GPCRs WP24	0.003249	AGTR1;CXCR4
	Leptin signaling pathway WP2034	0.003424	BAD;EIF4EBP1
	IL-18 signaling pathway WP4754	0.003602	BAD;ADIPOQ;ZC3H12A

Table 2. Top 10 pathways from KEGG, BioPlanet, Reactome, WikiPathways databases and their corresponding p-values and genes for AR-DEGs.

Gene targeted drugs screening. Based on the DSigNET drug database, the Enrichr (<https://maayanlab.cloud/Enrichr/>) web platform was used to identify drug molecules associated with 8 diagnostic biomarkers for DCM. Gene-targeted drugs were collected based on P-values. The combined score is proportional to the gene-drug association when the p-value is satisfied. The analysis showed that Melatonin CTD 00006260 and metformin CTD 00006282 had high gene binding to DCM. Table 3 listed the top 10 drugs for DCM by the DSigDB database.

Genetic disease association analysis. Gene list enrichments were identified in the DisGeNET dataset. All genes in the genome had been used as the enrichment background. Terms with a p-value < 0.01, a minimum count of 3, and an enrichment factor > 1.5 (the enrichment factor is the ratio between the observed counts and the counts expected by chance) were collected and grouped into clusters based on their membership similarities. The top 10 enriched clusters were shown in the Fig. 8. The algorithm used here was the same for pathway and

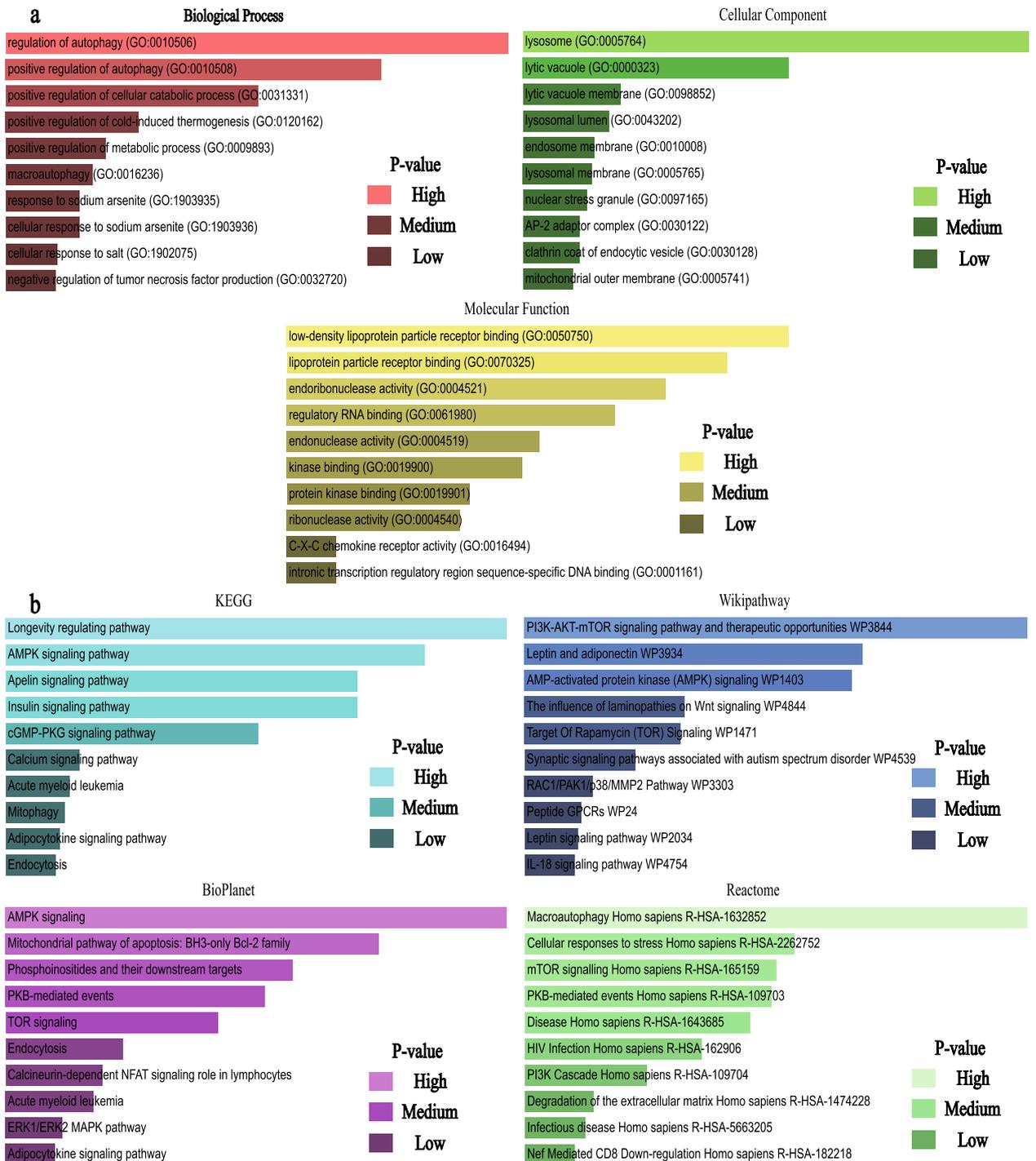


Figure 5. (a) Identification results of GO terms related to biological processes, cellular components and molecular functions based on gene enrichment analysis. Higher p value indicated a higher number of genes involved in this GO ontology. (b) Identification of results from combined multi-pathway analysis by KEGG, WikiPathways, BioPlanet and Reactome.

process enrichment analysis. Cyst, Uveal melanoma, Diabetes Mellitus, Experimental, Adult T-Cell Lymphoma/Leukemia, and Amyloidosis were identified as top 5 comorbidities of DCM.

Discussion

It is well known that DCM is impaired ventricular dilation and systolic diastole, leading to arrhythmias and heart failure in severe cases. Unfortunately, with the low prevalence of EMB, most patients with early-stage cardiomyopathy are not effectively treated. The gold standard for myocarditis and DCM is often poor prognosis in cases of concomitant arrhythmias and heart failure². Therefore, early diagnosis, precise evaluation, and therapeutic

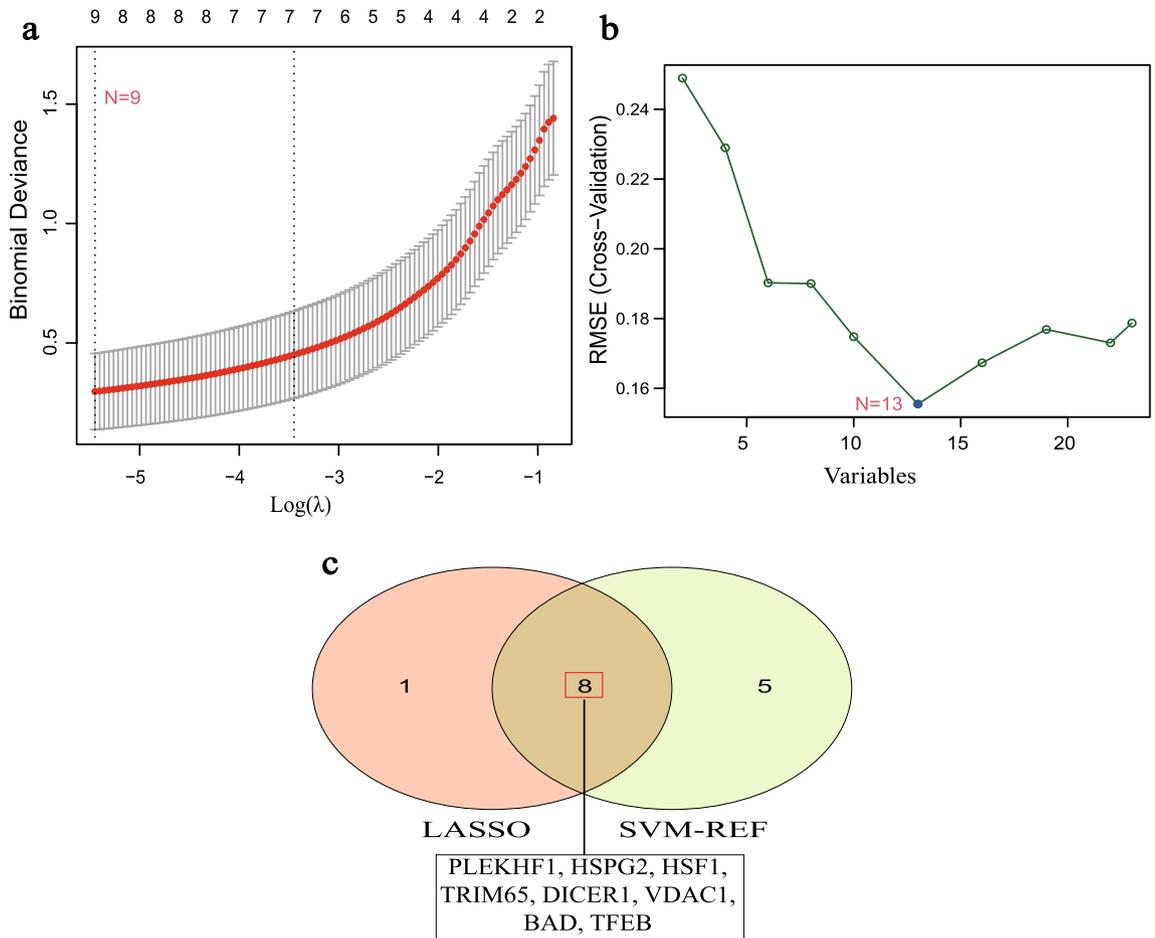


Figure 6. Screening of diagnostic biomarkers for DCM by machine learning algorithms. **(a)** Screening of optimal genes by LASSO regression model. **(b)** Plot of the best gene selected by SVM-RFE algorithm. **(c)** Venn diagram embodying the eight diagnostic biomarkers common to both machine learning algorithms. LASSO, least absolute shrinkage and selection operator; SVM-RFE, support vector machine-recursive feature elimination.

management of patients with DCM appear crucial. Hence, researchers are increasingly looking for diagnostic markers of DCM. Meanwhile, the molecular pathogenesis of DCM, viral infections, and other factors in disease progression and prognosis are still incompletely studied¹⁵.

It is well known that autophagy plays an important role in cancer, neurodegenerative diseases, inflammatory diseases, and cardiac diseases⁵. Among these, autophagy mechanisms are increasingly studied in cardiac diseases, and autophagy plays a crucial role in maintaining typical cardiac structure, function, and therapy^{16,17}. Two key autophagy-related molecules, mTOR and Beclin1, had been shown to play a regulatory role in myocardial ischemia-reperfusion injury¹⁷. Among them, mTOR is involved in the *PI3K* and *Akt* pathway to regulate myocardial ischemia/reperfusion-induced apoptosis and autophagy¹⁸. In addition, Beclin1 exerts a positive impact on myocardial ischemia and an adverse effect during myocardial ischemia/reperfusion¹⁹. Currently, studies on the role of autophagy in cardiomyopathy-related diseases are increasing¹³, and research has shown that damage to the autophagic lysosomal pathway (ALP) and activation of inflammatory vesicles were important factors contributing to DCM¹⁴. Improved left ventricular size and cardiac function in mice with DCM deficient in *NCOA4* (nuclear receptor coactivator 4, an autophagy-associated gene that mediates ferritin degradation) inhibit free ferrous iron overload and increased lipid peroxidation²⁰. Carolina et al.²¹ found that autophagy-related genes, such as *CALCOCO2* and *NRBP2*, the former of which regulates the expression of the latter, adversely affected left ventricular function parameters in patients with DCM.

In recent years, the exploration of the diagnostic and prognostic role of genetic biomarkers targeting DCM has been on the rise. For example, *CYR61* and *APN* were identified as two target genes for DCM by gene expression profiling studies in the GSE4172 dataset raw data²². It had been shown that *RBM20* induced aberrant TNN splicing as a determinant of DCM and increased the risk of arrhythmias²³. In previous bioinformatics studies, genes or transcription factors such as *CTGF*, *POSTN*, *CORIN*, and *FIGF* were closely associated with DCM²⁴. However, few studies have been conducted on the value of autophagy-related genes in diagnosing DCM.

To the best of our knowledge, this study is the first to investigate the diagnostic role of ARGs in DCM by mining the GEO database and integrating machine learning and bioinformatics approaches. We used the NetworkAnalyst 3.0 platform to deeply analyze the GSE4172 dataset, which compares gene expression in DCM with

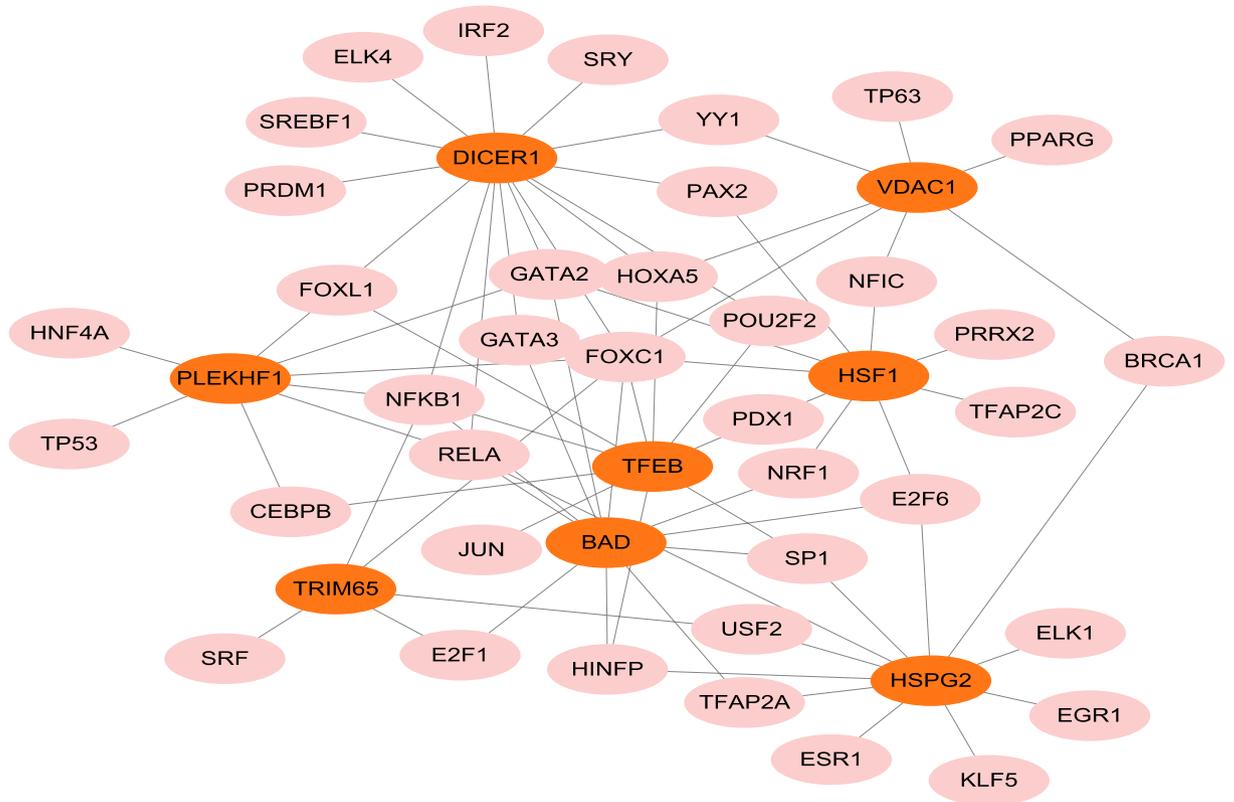


Figure 7. Network of transcription factors interacting with 8 potential diagnostic biomarkers. The highlighted orange nodes indicated the 8 potential diagnostic biomarkers and the other pink nodes indicated transcription factors. The network consisted of 8 core genes, 46 nodes and 76 edges.

Term	p-value	Combined score	Genes
Arsenous acid CTD 00000922	5.77E-05	76.35849	<i>PLEKHF1;BAD;ADIPOQ;HSF1;TFEB;EIF4EBP1;CXCR4;PPFIA4</i>
Melatonin CTD 00006260	0.000058	462.545	<i>BAD;ADIPOQ;EIF4EBP1</i>
Metformin CTD 00006282	0.000115	338.4126	<i>BAD;ADIPOQ;EIF4EBP1</i>
Tretinoin HL60 UP	0.000152	154.1536	<i>RCAN1;GABARAPL1;CXCR4;PPFIA4</i>
Imatinib CTD 00003267	0.000231	244.7684	<i>BAD;HSF1;EIF4EBP1</i>
Arsenous acid CTD 00000922	0.000374	61.03121	<i>GABARAPL1;HSF1;EIF4EBP1;CXCR4;VDAC1;TRIM65</i>
Wortmannin CTD 00000504	0.000606	154.125	<i>BAD;HSF1;EIF4EBP1</i>
Isoflupredone HL60 UP	0.000867	371.9284	<i>RCAN1;CXCR4</i>
Telmisartan CTD 00003021	0.001059	325.1773	<i>ADIPOQ;AGTR1</i>
Rosiglitazone CTD 00003139	0.001142	68.5849	<i>GABARAPL1;BAD;ADIPOQ;CXCR4</i>

Table 3. Drugs of choice for dilated cardiomyopathy.

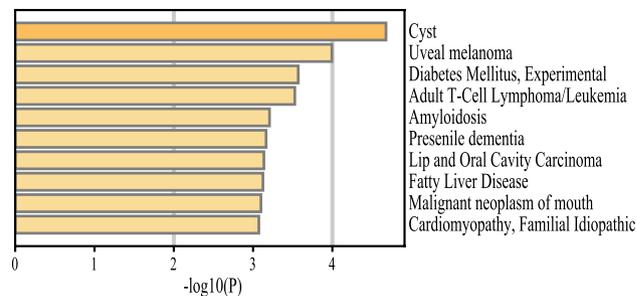


Figure 8. The process of identifying comorbidities in DCM.

healthy samples infected by the fine virus B19. Using differential analysis, we obtained 770 DEGs and combined them with the gene set from the autophagy databases to obtain 23 AR-DEGs of DCM. Finally, by machine learning methods such as LASSO regression and SVM-RFE, we obtained 8 (*PLEKHF1*, *HSPG2*, *HSF1*, *TRIM65*, *DICER1*, *VDAC1*, *BAD*, *TFEB*) diagnostic biomarkers of DCM. Previous studies showed significant relevance regarding DCM or cardiomyocyte remodeling in the above eight genes.

PLEKHF1 (Pleckstrin homology and FYVE domain containing 1) is located in the lysosome and plays a vital role in caspase-independent apoptosis, a process involved in autophagy²⁵. In previous studies, *PLEKHF1* is a susceptibility gene for several diseases. For example, Qi et al., identified *PLEKHF1* as a potential biomarker for diabetic atherosclerosis²⁶; also, *PLEKHF1* was shown to be a potential biomarker for chronic graft-versus-host disease, the accuracy of which was confirmed by several clinical independent validation studies²⁷. In addition, it had been shown that levosimendan ameliorated myocardial infarction and ventricular remodeling in diabetic rats, and the expression of the gene *Plekhhf1* received regulation by levosimendan, showing the potential of *Plekhhf1* as a target gene for myocardial infarction and diabetic cardiomyopathy²⁸.

HSPG2 (Heparan sulfate proteoglycan 2) plays an important role in cancer growth, development, and metastasis²⁹. Previous studies had shown that *HSPG2* was identified in key cardiac-related regions controlled by chromosome 1p36³⁰, and related studies had demonstrated that chromosome 1p36 deletion was responsible for cardiovascular malformations and cardiomyopathy³¹, suggesting an important role for *HSPG2* in the pathogenesis and prognostic impact of cardiomyopathy³⁰. In addition, *HSPG2* also plays an independent predictive role in a variety of diseases. For example, *HSPG2* was overexpressed in acute myeloid leukemia and can be used as a prognostic biomarker³². Recent studies had shown that *HSPG2* deficiency was a risk factor for aortic coarctation³³.

HSF1 (Heat shock transcription factor 1) is a significant heat stress response factor that plays an important role in inhibiting apoptosis and pathological remodeling of cardiomyocytes and is a protective factor for cardiomyocytes. In a previous quantitative transcriptomic analysis, *HSF1* was found to be significantly enriched in cardiomyocytes³⁴. It had been shown that *HSF1* could be isolated by the death trap method, preventing hydrogen peroxide-induced cardiomyocyte death. It was found that overexpression of *HSG1* in transgenic mice reduced ischemia–reperfusion-induced cardiomyocyte injury³⁵. In the present study, *HSF1* expression was lower in the DCM group compared with the healthy control group, which was also consistent with the findings of previous studies. In addition, it had been shown that overexpression of *HSF1* in *BAG* mutation-associated DCM helped to attenuate pathological remodeling of cardiomyocytes and alleviate proteostatic stress³⁶. In contrast, recent studies had shown that *HSF1* overexpression lead to reduced expression of myofilament localization-associated *BAG3*. Decreased expression of *BAG3* was strongly associated with non-inherited heart failure and was more susceptible in male patients with DCM³⁷. Therefore, the study of relevant molecules and pathways targeting *HSF1* contributes to our understanding of DCM.

TRIM65 is an E3 ubiquitin ligase involved in the positive regulation of autophagy and was expressed in vascular endothelial cells, located in the cytoplasmic lysate and nucleoplasm. Unfortunately, there are relatively few studies related to *TRIM65*. From the available literature, it appeared that *TRIM65* was mainly involved in proteopathy and ubiquitination regulation to regulate disease progression and as a target for a variety of diseases^{38,39}. Interestingly, although fewer studies are addressing the mechanisms associated with *TRIM65* and DCM, according to recent studies, *TRIM65* was closely linked to the inflammatory vesicle *NRLP3*⁴⁰, which is known to play a role in a variety of DCM¹⁴. *TRIM65* was associated with antiviral innate immune mechanisms⁴¹. In addition, it had been shown that *TRIM65* regulated *VCAM-1* to control inflammatory responses⁴². All these studies point the way to exploring the molecular mechanism of *TRIM65* and DCM.

DICER1 is a member of the ribonuclease III (RNaseIII) family and is involved in the production of microRNAs, which regulate gene expression at the post-transcriptional level and are more frequently studied in oncological diseases⁴³. Evidence suggested that *DICER* deletion resulted in a dramatic decrease in the level of miRNAs it regulates, which led to severe DCM and heart failure in mice, a trend that was also seen in the expression of *DICER* proteins in diseased populations, implying an important role of *DICER* family genes in the pathogenesis of DCM⁴⁴. Follow-up studies had shown that microRNAs act as negative regulators of genes and that specific regulation of microRNA expression could inhibit the loss of cardiac function due to *DICER* deficiency^{45,46}, leading to cardioprotection. These studies suggested that endogenous microRNA competitive regulation of *DICER* family genes will be an essential strategy for gene targeting therapy in DCM.

VDAC (voltage dependent anion channel), including *VDAC1* and *VDAC2*, is a mitochondrial outer membrane pore-forming protein present in all eukaryotes. As a mitochondrial transporter protein, *VDAG* is mostly expressed in cardiac tissue and has significant tissue specificity^{47,48}. It is well known that Ca^{2+} played a detrimental role in heart failure and myocardial ischemia/reperfusion, and Ca^{2+} overload activated the complex matrix chaperone procyclin D (CypD), which regulated the *VDAC1*, Grp75, and *IP3R1* complex and thus caused damage to cardiomyocytes, whereas inhibition of the CypD, *VDAC1*, Grp75, and *IP3R1* complex could protect cardiomyocytes⁴⁹. Numerous studies had shown^{50,51} that regulation of *VDAC1* expression through microRNA targeting could regulate mitochondrial function and promoted the release of mitochondrial calcium for cell protection. Furthermore, in DCM mice, the lncRNA H19/miR-675 axis competitively downregulated *VDAC1*, reducing apoptosis. The above report provides a new strategy to explore the role of *VDAC1* in DCM. It was shown that *VDAC1* expression was upregulated in the hearts of patients with hypertrophic cardiomyopathy⁵². In the present study, the expression of *VDAC1* was also upregulated in samples from patients with DCM. These findings could explain the unique role played by *VDAC1* as a target gene for DCM.

BAD (Bcl-2 associated agonist of cell death) often follows Bcl-2 and plays an anti-apoptotic role. In a TNF- α -mediated mouse model of DCM in which apoptosis occurs, the expression of *BAD* was reduced in association with Bcl-2d⁵³, which was consistent with the findings of the present study. According to previous studies, *BAD* played a key role in inducing β -cell apoptosis in Friedreich's ataxia, a neurodegenerative disease closely related to cardiomyopathy and diabetes⁵⁴. It is well known that microRNAs regulate protein expression of mRNAs through

negative regulation and play an important role in cardiovascular diseases, especially in heart failure and cardiac remodeling⁵⁵. Studies had shown that multiple microRNAs played a regulatory role on *BCL2*⁵⁶ and all of them were upregulated in heart failure⁵⁷. As an antagonist of apoptosis, the protective role of *BAD* and Bcl-2 in the pathogenesis of DCM depended on further studies.

TFEB (transcription factor EB), a transcription factor located within the cytoplasmic lysosol (cytosol), is the master gene of the autophagic machinery of lysosomal biogenesis and coordinates the autophagic process, including autophagosome formation, autophagosome-lysosome fusion, and substrate degradation by driving the expression of autophagy and lysosomal genes⁵⁸. According to reports, *TFEB* expression was highest in 18-week-old fetal heart tissue, with significant tissue specificity⁵⁹. There is growing evidence that *TFEB* plays an important role in various types of DCM. Lysosomal storage disorders (LSD) lead to cardiac involvement in hypertrophic cardiomyopathy and DCM⁶⁰. Further studies had shown that the Yes-associated protein (YAP) and *Feb* signaling pathway played a role in LSD disease by eliminating autophagic lysosomes, reducing cell death, and restoring cardiac function⁶¹. Also, it was found that *TFEB* deficiency led to cardiomyocyte hypertrophy and DCM causing heart failure⁶². Therefore, the role of *TFEB* in targeting DCM is extremely significant.

In addition, we performed a functional enrichment analysis of the pathogenesis of DCM and related molecular pathways and found that AR-DEGs of DCM were mainly enriched in autophagy regulatory pathways and cell growth signaling, such as regulation of autophagy, macroautophagy, AMPK signaling pathway, PKB-mediated events, etc. AMPK (Adenosine monophosphate-activated protein kinase) signaling pathway had been reported to be an important intracellular signaling pathway in the heart⁶³. As an emerging target recognized for the treatment of heart failure⁶⁴, AMPK plays an important role in regulating cardiomyocyte growth⁶⁵. Numerous studies had shown that the AMPK pathway and its binding autophagy-related pathways played a protective role in the pathological development of cardiomyopathy^{66–69}. These studies have provided ideas to explore the mechanistic studies of autophagy-related DCM. *PKB* (protein kinase B), also known as serine/threonine kinase Akt, serves as a central node for a variety of biological processes⁷⁰. It had been reported that *PKB* was involved in protective mechanisms against myocardial ischemia/reperfusion⁷¹. However, relatively few studies have been conducted on the association of *PKB*-mediated events with DCM. According to previous studies, Pleiotrophin, a pro-angiogenic factor, was significantly expressed in rat models of myocardial infarction and DCM patients. It is considered that Pleiotrophin protects the myocardium by inhibiting endogenous AKT/*PKB* activity⁷². In contrast, Alexander et al. found that *PKB* phosphorylation expression restored cardiac contractility in a zebrafish model of DCM⁷³.

In addition, we constructed TF-gene regulatory networks based on 8 autophagy-related genes in DCM and predicted them to target drugs, such as Melatonin and metformin. Studies showed that Melatonin had a better inhibitory effect on left heart dysfunction and ventricular remodeling in DCM rats with cardiorenal syndrome⁷⁴. Metformin was able to partially reverse ventricular remodeling in mice with DCM through an autophagic mechanism⁷⁵. These studies provided a basis and direction for clinical precision targeting therapy and novel drug development in DCM. In addition, we explored the comorbidities associated with DCM, such as fatty liver disease. Some scholars found that⁷⁶ NAFLD affected the cardiovascular system through metabolic and inflammatory responses, and also increased the abnormalities of cardiac anatomy including cardiomyopathy⁷⁷. Furthermore, the disease pathways between the two need further investigation.

However, there are certain shortcomings in our study. First, our data set of DCM was mined and analyzed secondarily by bioinformatics means, and the results of the study need to be validated with external evidence. In addition, the results of this study need to be combined with single-cell sequencing as the multi-omics study progresses. Finally, the mechanism of action and interrelationship between these 8 DCM genes and autophagy-related genes need further investigation.

Methods

Dilated cardiomyopathy dataset acquisition. The dataset of DCM was downloaded from the GSE4172 dataset of the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo>) database, which was contributed by Wittchen et al.²², piggybacked on the GPL570 [HG- U133_Plus_2] platform using Affymetrix Human Genome U133 Plus 2.0 Array, containing eight endomyocardial myocardial biopsy samples from patients with microvirus B19-associated cardiac inflammation as experimental group and four healthy human samples as a control group. Clinical information of patients from the GSE4172 dataset was presented in Table 4.

Autophagy genes acquisition. A total of 232 autophagy genes were downloaded from the Human Autophagy Database (HADb, <http://autophagy.lu/>). Similarly, 796 autophagy genes were obtained from the Human Autophagy Modulator Database (HAMdb, <http://hamdb.scbdd.com>)⁷⁸. A total of 803 autophagy-related genes were obtained as the autophagy gene set for this study by taking the intersection of the two.

Identification of differentially expressed genes (DEGs) in autophagy-related genes (ARGs). NetworkAnalyst 3.0 is a user-friendly bioinformatics visualization web platform for transcriptome analysis, gene network construction, and meta-analysis of gene expression data⁷⁹. The expression data and grouping information of the GSE4172 dataset were submitted to NetworkAnalyst 3.0 for identification of the DCM groups and the healthy control groups for DEGs. For mRNA in microarrays, the threshold was set to $|\log_2\text{FoldChange}| \geq 0.8$ with a p value < 0.05 , and genes meeting this criterion were considered as DEGs. We used the ggplot2 package (R package version 4.1.3) and pheatmap package (R package version 4.1.3) to draw the asymptotic volcano map and heatmap to show the DEGs. Autophagy-related genes (ARGs) and DEGs from the GSE4172 dataset were taken to intersect to obtain the set of autophagy-related differentially expression genes (AR-DEGs). Venn plots were created by using the Omicshare online tool (<https://www.omicshare.com/>). The expression of 23 AR-DEGs in GSE4172 was demonstrated using box plots through the ggpubr package as well as

Sample	Group	Age	Gender	Ejection fraction	Left ventricular end diastolic diameter	Inflammation/PVB19
GSM94836	DCM	45	Male	34	62	Positive
GSM94837	DCM	62	Male	51	73	Positive
GSM94838	DCM	31	Male	52	57	Positive
GSM94839	DCM	67	Male	43	59	Positive
GSM94840	DCM	60	Male	34	76	Positive
GSM94841	DCM	69	Male	35	60	Positive
GSM94842	DCM	55	Female	31	61	Positive
GSM94843	DCM	31	Female	56	71	Positive
GSM94831	Healthy control	36	Female	68	47	Negative
GSM94854	Healthy control	46	Female	61	49	Negative
GSM94855	Healthy control	26	Female	74	47	Negative
GSM94870	Healthy control	36	Male	64	50	Negative

Table 4. Clinical information for the GSE4172 dataset.

the associated helper R packages. The correlation analysis of AR-DEGs was visualized using the corrplot package (R package 4.1.3).

Functional enrichment analysis. Functional enrichment consists of performing biological processes, molecular functions, and chromosomal location analysis⁸⁰. Gene annotation uses gene ontology (GO) terminology and consists of biological processes, molecular functions, and cells. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway was used to understand metabolic pathways and plays an important role in the gene annotation process^{81,82}. In addition, the BioCarta, WikiPathways⁸³, and Reactome⁸⁴ databases were also used to analyze KEGG pathways. The Enrichr (<https://amp.pharm.mssm.edu/Enrichr/>) platform provides a comprehensive gene enrichment analysis applied databases containing rich gene set annotation, pathway information analysis, and screening of gene target drugs^{85,86}. The GO terms of the AR-DEGs of DCM and all pathway information for this study were obtained from the Enrichr platform.

Machine learning identifies molecular markers of AR-DEGs in DCM. In this study, the least absolute shrinkage and selection operator (LASSO) logistic regression was used for feature gene selection to reduce the number of genes in the disease prediction model, solve the multicollinearity problem in the regression analysis, and screen the molecular markers of DCM genes⁸⁷. The "glmnet" package was used to implement the LASSO regression algorithm with α set to 1 which was used to control the traits of the model when dealing with highly correlated data. In addition, the Support Vector Machine-Recursive Feature Elimination (SVM-RFE) algorithm model was also used in this study to characterize the AR-DEGs and remove irrelevant genes to make the diagnostic prediction model more robust⁸⁸. The SVM-RFE was implemented by the e1071 Package R software.

Transcription Factor (TF)-gene regulatory network construction. The JASPAR (<http://jaspar.genereg.net/>) database was used to generate a visual analysis of the TF-gene co-regulatory network⁸⁹. Based on 8 biomarkers of DCM, TFs that regulated the activity of functional pathways and gene expression levels in DCM were identified from the JASPAR database to form the TF-gene regulatory network. It is important to note that the JASPAR database is included in the NetworkAnalyst 3.0 platform.

Target drug screening. Gene target-based drug screening has become a new approach for drug molecular identification study, which helps to expand the scope of relevant drugs and reduce the process of drug development. In this study, molecular markers of DCM were screened for drug candidates through the drug Signatures database (DSigDB), which consists of 17,389 drugs and 19,531 genes associated with the drugs⁹⁰. The DSigDB database can be accessed by visiting Enrichr (<https://www.amp.pharm.mssm.edu/Enrichr/>) website to enter relevant gene targets and download target drug information. Drugs with p-values less than 0.05 and with larger combined scores were considered to be typically significant. The combined score represents the degree to which the small molecule drug is closely linked to the gene of interest.

Genetic disease association analysis. The DisGeNET (<http://www.disgenet.org>) database is an open and versatile platform for studying specific human diseases and their comorbidities through genetic and molecular pathways, probing the characteristics of disease genes and offering the possibility to elucidate the mechanisms of disease⁹¹. In the present study, molecular markers of DCM were uploaded to the Metascape (<https://metascape.org/gp/index.html#/main/step1>) platform⁹², which contains the DisGeNET database. We have revealed DCM-related comorbidities through the DisGeNET database, laying the foundation for the mechanistic study of DCM.

Copyright permission of KEGG. We have contacted Kanehisa Laboratories. We do not directly use these KEGG Pathway map "images" in the article, we need not obtain copyright permission of KEGG. However, they believe that we have written our article using their data, they kindly ask us to cite the following articles in it^{81,93,94}.

Data availability

The dataset GSE4172 for this study can be found in the GEO database (<https://www.ncbi.nlm.nih.gov/geo>). All data generated or analysed during this study are included in this published article.

Received: 2 April 2022; Accepted: 23 August 2022

Published online: 02 September 2022

References

- McNally, E. M. & Mestroni, L. Dilated cardiomyopathy: Genetic determinants and mechanisms. *Circ. Res.* **121**, 731–748. <https://doi.org/10.1161/CIRCRESAHA.116.309396> (2017).
- Ammirati, E. *et al.* Clinical presentation and outcome in a contemporary cohort of patients with acute myocarditis: Multicenter lombardy registry. *Circulation* **138**, 1088–1099. <https://doi.org/10.1161/CIRCULATIONAHA.118.035319> (2018).
- Long, C., Liu, X., Xiong, Q., Su, Y. & Hong, K. Sex differences in dilated cardiomyopathy prognosis. *Int. Heart J.* **63**, 36–42. <https://doi.org/10.1536/ihj.20-448> (2022).
- Yilmaz, A. *et al.* Comparative evaluation of left and right ventricular endomyocardial biopsy: Differences in complication rate and diagnostic performance. *Circulation* **122**, 900–909. <https://doi.org/10.1161/CIRCULATIONAHA.109.924167> (2010).
- Glick, D., Barth, S. & Macleod, K. F. Autophagy: Cellular and molecular mechanisms. *J. Pathol.* **221**, 3. <https://doi.org/10.1002/path.2697> (2010).
- Levine, B. & Kroemer, G. Autophagy in the pathogenesis of disease. *Cell* **132**, 27–42. <https://doi.org/10.1016/j.cell.2007.12.018> (2008).
- He, S. *et al.* Signatures of 4 autophagy-related genes as diagnostic markers of MDD and their correlation with immune infiltration. *J. Affect. Disord.* **295**, 11–20. <https://doi.org/10.1016/j.jad.2021.08.005> (2021).
- Ruart, M. *et al.* Impaired endothelial autophagy promotes liver fibrosis by aggravating the oxidative stress response during acute liver injury. *J. Hepatol.* **70**, 458–469. <https://doi.org/10.1016/j.jhep.2018.10.015> (2019).
- Fukuda, T. *et al.* Autophagy and mistargeting of therapeutic enzyme in skeletal muscle in Pompe disease. *Mol. Therapy* **14**, 831–839 (2006).
- Liang, X. H. *et al.* Induction of autophagy and inhibition of tumorigenesis by beclin 1. *Nature* **402**, 672–676 (1999).
- Terman, A. & Brunk, U. T. Autophagy in cardiac myocyte homeostasis, aging, and pathology. *Cardiovasc. Res.* **68**, 355–365 (2005).
- Nakai, A. *et al.* The role of autophagy in cardiomyocytes in the basal state and in response to hemodynamic stress. *Nat. Med.* **13**, 619–624. <https://doi.org/10.1038/nm1574> (2007).
- Du, J., Liu, Y. & Fu, J. Autophagy, myocarditis, and cardiomyopathy. *Adv. Exp. Med. Biol.* **1207**, 229–235. https://doi.org/10.1007/978-981-15-4272-5_17 (2020).
- Caragnano, A. *et al.* Autophagy and inflammasome activation in dilated cardiomyopathy. *J. Clin. Med.* **8**, 1519. <https://doi.org/10.3390/jcm8101519> (2019).
- Tschope, C. *et al.* Myocarditis and inflammatory cardiomyopathy: Current evidence and future directions. *Nat. Rev. Cardiol.* **18**, 169–193. <https://doi.org/10.1038/s41569-020-00435-x> (2021).
- Wu, X., Liu, Z., Yu, X. Y., Xu, S. & Luo, J. Autophagy and cardiac diseases: Therapeutic potential of natural products. *Med. Res. Rev.* **41**, 314–341. <https://doi.org/10.1002/med.21733> (2021).
- Shi, B., Ma, M., Zheng, Y., Pan, Y. & Lin, X. mTOR and Beclin1: Two key autophagy-related molecules and their roles in myocardial ischemia/reperfusion injury. *J. Cell Physiol.* **234**, 12562–12568. <https://doi.org/10.1002/jcp.28125> (2019).
- Ye, G., Fu, Q., Jiang, L. & Li, Z. Vascular smooth muscle cells activate PI3K/Akt pathway to attenuate myocardial ischemia/reperfusion-induced apoptosis and autophagy by secreting bFGF. *Biomed. Pharmacother.* **107**, 1779–1785. <https://doi.org/10.1016/j.biopha.2018.05.113> (2018).
- Matsui, Y. *et al.* Distinct roles of autophagy in the heart during ischemia and reperfusion: Roles of AMP-activated protein kinase and Beclin 1 in mediating autophagy. *Circ. Res.* **100**, 914–922. <https://doi.org/10.1161/01.RES.0000261924.76669.36> (2007).
- Ito, J. *et al.* Iron derived from autophagy-mediated ferritin degradation induces cardiomyocyte death and heart failure in mice. *Elife* **10**, 5. <https://doi.org/10.7554/eLife.62174> (2021).
- Gil-Cayuela, C. *et al.* The altered expression of autophagy-related genes participates in heart failure: NRBP2 and CALCOCO2 are associated with left ventricular dysfunction parameters in human dilated cardiomyopathy. *PLoS ONE* **14**, e0215818. <https://doi.org/10.1371/journal.pone.0215818> (2019).
- Wittchen, F. *et al.* Genomic expression profiling of human inflammatory cardiomyopathy (DCMi) suggests novel therapeutic targets. *J. Mol. Med. (Berl)* **85**, 257–271 (2007).
- van den Hoogenhof, M. M. G. *et al.* RBM20 mutations induce an arrhythmogenic dilated cardiomyopathy related to disturbed calcium handling. *Circulation* **138**, 1330–1342. <https://doi.org/10.1161/CIRCULATIONAHA.117.031947> (2018).
- Zhao, J. *et al.* Identification of target genes in cardiomyopathy with fibrosis and cardiac remodeling. *J. Biomed. Sci.* **25**, 63. <https://doi.org/10.1186/s12929-018-0459-8> (2018).
- Lin, W. J. *et al.* Lysosomal targeting of pphafin1 mediated by Rab7 induces autophagosome formation. *Biochem. Biophys. Res. Commun.* **417**, 35–42. <https://doi.org/10.1016/j.bbrc.2011.11.043> (2012).
- Huang, Q. *et al.* Comprehensive identification of key genes involved in development of diabetes mellitus-related atherosclerosis using weighted gene correlation network analysis. *Front. Cardiovasc. Med.* **7**, 580573. <https://doi.org/10.3389/fcvm.2020.580573> (2020).
- Pidala, J. *et al.* A combined biomarker and clinical panel for chronic graft versus host disease diagnosis. *J. Pathol. Clin. Res.* **3**, 3–16. <https://doi.org/10.1002/cjp2.58> (2017).
- Vahtola, E. *et al.* Effects of levosimendan on cardiac gene expression profile and post-infarct cardiac remodeling in diabetic Goto-Kakizaki rats. *Basic Clin. Pharmacol. Toxicol.* **109**, 387–397. <https://doi.org/10.1111/j.1742-7843.2011.00743.x> (2011).
- Elgundi, Z. *et al.* Cancer metastasis: The role of the extracellular matrix and the heparan sulfate proteoglycan perlecan. *Front. Oncol.* **9**, 1482. <https://doi.org/10.3389/fonc.2019.01482> (2019).
- Zaveri, H. P. *et al.* Identification of critical regions and candidate genes for cardiovascular malformations and cardiomyopathy associated with deletions of chromosome 1p36. *PLoS ONE* **9**, e85600. <https://doi.org/10.1371/journal.pone.0085600> (2014).
- Kang, S. H. *et al.* Identification of proximal 1p36 deletions using array-CGH: A possible new syndrome. *Clin. Genet.* **72**, 329–338. <https://doi.org/10.1111/j.1399-0004.2007.00876.x> (2007).
- Zhou, X. *et al.* HSPG2 overexpression independently predicts poor survival in patients with acute myeloid leukemia. *Cell Death Dis.* **11**, 492. <https://doi.org/10.1038/s41419-020-2694-7> (2020).
- Nonaka, R., Iesaki, T., Kerever, A. & Arikawa-Hirasawa, E. Increased risk of aortic dissection with Perlecan deficiency. *Int. J. Mol. Sci.* **23**, 315. <https://doi.org/10.3390/ijms23010315> (2021).
- Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell Proteom.* **13**, 397–406. <https://doi.org/10.1074/mcp.M113.035600> (2014).
- Zou, Y. *et al.* Heat shock transcription factor 1 protects cardiomyocytes from ischemia/reperfusion injury. *Circulation* **108**, 3024–3030. <https://doi.org/10.1161/01.CIR.0000101923.54751.77> (2003).

36. McDermott-Roe, C. *et al.* Investigation of a dilated cardiomyopathy-associated variant in BAG3 using genome-edited iPSC-derived cardiomyocytes. *JCI Insight* **4**, 22. <https://doi.org/10.1172/jci.insight.128799> (2019).
37. Martin, T. G., Tawfik, S., Moravec, C. S., Pak, T. R. & Kirk, J. A. BAG3 expression and sarcomere localization in the human heart are linked to HSF-1 and are differentially affected by sex and disease. *Am. J. Physiol. Heart Circul. Physiol.* **320**, H2339–H2350. <https://doi.org/10.1152/ajpheart.00419.2020> (2021).
38. Wang, X.-L. *et al.* Knockdown of TRIM65 inhibits lung cancer cell proliferation, migration and invasion: A therapeutic target in human lung cancer. *Oncotarget* **7**, 81527–81540. <https://doi.org/10.18632/oncotarget.13131> (2016).
39. Wei, W. S. *et al.* TRIM65 supports bladder urothelial carcinoma cell aggressiveness by promoting ANXA2 ubiquitination and degradation. *Cancer Lett.* **435**, 10–22. <https://doi.org/10.1016/j.canlet.2018.07.036> (2018).
40. Tang, T. *et al.* The E3 ubiquitin ligase TRIM65 negatively regulates inflammasome activation through promoting ubiquitination of NLRP3. *Front. Immunol.* **12**, 741839. <https://doi.org/10.3389/fimmu.2021.741839> (2021).
41. Lang, X. *et al.* TRIM65-catalyzed ubiquitination is essential for MDA5-mediated antiviral innate immunity. *J. Exp. Med.* **214**, 459–473. <https://doi.org/10.1084/jem.20160592> (2017).
42. Li, Y. *et al.* TRIM65 E3 ligase targets VCAM-1 degradation to limit LPS-induced lung inflammation. *J. Mol. Cell Biol.* **12**, 190–201. <https://doi.org/10.1093/jmcb/mjz077> (2020).
43. Rio-Frio, T. *et al.* DICER1 mutations in familial multinodular goiter with and without ovarian Sertoli-Leydig cell tumors. *JAMA* **305**, 68–77. <https://doi.org/10.1001/jama.2010.1910> (2011).
44. Chen, J.-F. *et al.* Targeted deletion of Dicer in the heart leads to dilated cardiomyopathy and heart failure. *Proc. Natl. Acad. Sci. USA* **105**, 2111–2116. <https://doi.org/10.1073/pnas.0710228105> (2008).
45. Roy, S. *et al.* Suppression of induced microRNA-15b Prevents rapid loss of cardiac function in a dicer depleted model of cardiac dysfunction. *PLoS ONE* **8**, e66789. <https://doi.org/10.1371/journal.pone.0066789> (2013).
46. Ali, R. *et al.* miR-1 mediated suppression of Sorcin regulates myocardial contractility through modulation of Ca²⁺ signaling. *J. Mol. Cell Cardiol.* **52**, 1027–1037. <https://doi.org/10.1016/j.yjmcc.2012.01.020> (2012).
47. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteom.* **13**, 397–406. <https://doi.org/10.1074/mcp.M113.035600> (2014).
48. Huizing, M. *et al.* Human mitochondrial transmembrane metabolite carriers: Tissue distribution and its implication for mitochondrial disorders. *J. Bioenerg. Biomembr.* **30**, 277–284 (1998).
49. Paillard, M. *et al.* Depressing mitochondria-reticulum interactions protects cardiomyocytes from lethal hypoxia-reoxygenation injury. *Circulation* **128**, 1555–1565. <https://doi.org/10.1161/CIRCULATIONAHA.113.001225> (2013).
50. Chaudhuri, A. D., Choi, D. C., Kabaria, S., Tran, A. & Junn, E. MicroRNA-7 regulates the function of mitochondrial permeability transition pore by targeting VDAC1 expression. *J. Biol. Chem.* **291**, 6483–6493. <https://doi.org/10.1074/jbc.M115.691352> (2016).
51. Lin, D., Cui, B., Ma, J. & Ren, J. MiR-183-5p protects rat hearts against myocardial ischemia/reperfusion injury through targeting VDAC1. *BioFactors* **46**, 83–93. <https://doi.org/10.1002/biof.1571> (2020).
52. Lim, D. S., Roberts, R. & Marian, A. J. Expression profiling of cardiac genes in human hypertrophic cardiomyopathy: Insight into the pathogenesis of phenotypes. *J. Am. Coll. Cardiol.* **38**, 1175–1180 (2001).
53. Das, S. *et al.* TNF-alpha-mediated signal transduction pathway is a major determinant of apoptosis in dilated cardiomyopathy. *J. Cell. Mol. Med.* **14**, 1988–1997. <https://doi.org/10.1111/j.1582-4934.2009.00904.x> (2010).
54. Igoillo-Esteve, M. *et al.* Unveiling a common mechanism of apoptosis in β -cells and neurons in Friedreich's ataxia. *Hum. Mol. Genet.* **24**, 2274–2286. <https://doi.org/10.1093/hmg/ddu745> (2015).
55. Divakaran, V. & Mann, D. L. The emerging role of microRNAs in cardiac remodeling and heart failure. *Circ. Res.* **103**, 1072–1083. <https://doi.org/10.1161/CIRCRESAHA.108.183087> (2008).
56. Katoh, M. Cardio-miRNAs and onco-miRNAs: Circulating miRNA-based diagnostics for non-cancerous and cancerous diseases. *Front. Cell Dev. Biol.* **2**, 61. <https://doi.org/10.3389/fcell.2014.00061> (2014).
57. Leptidis, S. *et al.* A deep sequencing approach to uncover the miRNOME in the human heart. *PLoS ONE* **8**, e57800. <https://doi.org/10.1371/journal.pone.0057800> (2013).
58. Settembre, C. *et al.* TFEB links autophagy to lysosomal biogenesis. *Science* **332**, 1429–1433. <https://doi.org/10.1126/science.1204592> (2011).
59. Szabo, L. *et al.* Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. *Genome Biol.* **16**, 126. <https://doi.org/10.1186/s13059-015-0690-5> (2015).
60. Nair, V., Belanger, E. C. & Veinot, J. P. Lysosomal storage disorders affecting the heart: A review. *Cardiovasc. Pathol.* **39**, 12–24. <https://doi.org/10.1016/j.carpath.2018.11.002> (2019).
61. Rabinovich-Nikitin, I. & Kirshenbaum, L. A. YAP/TFEB pathway promotes autophagic cell death and hypertrophic cardiomyopathy in lysosomal storage diseases. *J. Clin. Investig.* **131**, 5. <https://doi.org/10.1172/JCI146821> (2021).
62. Song, R. *et al.* TFEB insufficiency promotes cardiac hypertrophy by blocking autophagic degradation of GATA4. *J. Biol. Chem.* **297**, 101189. <https://doi.org/10.1016/j.jbc.2021.101189> (2021).
63. Corton, J. M., Gillespie, J. G. & Hardie, D. G. Role of the AMP-activated protein kinase in the cellular stress response. *Curr. Biol.* **4**, 315–324. [https://doi.org/10.1016/S0960-9822\(00\)00070-1](https://doi.org/10.1016/S0960-9822(00)00070-1) (1994).
64. Li, Y. *et al.* AMPK blunts chronic heart failure by inhibiting autophagy. *Biosci. Rep.* **38**, 4. <https://doi.org/10.1042/BSR20170982> (2018).
65. Takano, A. P., Diniz, G. P. & Barreto-Chaves, M. L. AMPK signaling pathway is rapidly activated by T3 and regulates the cardiomyocyte growth. *Mol. Cell. Endocrinol.* **376**, 43–50. <https://doi.org/10.1016/j.mce.2013.05.024> (2013).
66. Wang, L. *et al.* Chikusetsu saponin IVa attenuates isoprenaline-induced myocardial fibrosis in mice through activation autophagy mediated by AMPK/mTOR/ULK1 signaling. *Phytomedicine* **58**, 152764. <https://doi.org/10.1016/j.phymed.2018.11.024> (2019).
67. Liu, N. *et al.* LncRNA LncHrt preserves cardiac metabolic homeostasis and heart function by modulating the LKB1-AMPK signaling pathway. *Basic Res. Cardiol.* **116**, 48. <https://doi.org/10.1007/s00395-021-00887-3> (2021).
68. Mao, J.-Y. *et al.* The effects of UCP2 on autophagy through the AMPK signaling pathway in septic cardiomyopathy and the underlying mechanism. *Ann. Transl. Med.* **9**, 259. <https://doi.org/10.21037/atm-20-4819> (2021).
69. Zheng, H. *et al.* Nkx2-3 induces autophagy inhibiting proliferation and migration of vascular smooth muscle cells via AMPK/mTOR signaling pathway. *J. Cell. Physiol.* **236**, 7342–7355. <https://doi.org/10.1002/jcp.30400> (2021).
70. Manning, B. D. & Cantley, L. C. AKT/PKB signaling: Navigating downstream. *Cell* **129**, 1261–1274. <https://doi.org/10.1016/j.cell.2007.06.009> (2007).
71. Zeng, H. *et al.* Activated PKB/GSK-3 β synergizes with PKC- δ signaling in attenuating myocardial ischemia/reperfusion injury via potentiation of NRF2 activity: Therapeutic efficacy of dihydrotanshinone-I. *Acta Pharm. Sin. B* **11**, 71–88. <https://doi.org/10.1016/j.apsb.2020.09.006> (2021).
72. Li, J. *et al.* The pro-angiogenic cytokine pleiotrophin potentiates cardiomyocyte apoptosis through inhibition of endogenous AKT/PKB activity. *J. Biol. Chem.* **282**, 34984–34993. <https://doi.org/10.1074/jbc.M703513200> (2007).
73. Pott, A. *et al.* Therapeutic chemical screen identifies phosphatase inhibitors to reconstitute PKB phosphorylation and cardiac contractility in ILK-deficient Zebrafish. *Biomolecules* **8**, 153. <https://doi.org/10.3390/biom8040153> (2018).
74. Chua, S. *et al.* The cardioprotective effect of melatonin and exendin-4 treatment in a rat model of cardiorenal syndrome. *J. Pineal Res.* **61**, 438–456. <https://doi.org/10.1111/jpi.12357> (2016).

75. Kanamori, H. *et al.* Metformin enhances autophagy and provides cardioprotection in delta-sarcoglycan deficiency-induced dilated cardiomyopathy. *Circ. Heart Fail.* **12**, e005418. <https://doi.org/10.1161/CIRCHEARTFAILURE.118.005418> (2019).
76. Magida, J. A. & Leinwand, L. A. Metabolic crosstalk between the heart and liver impacts familial hypertrophic cardiomyopathy. *EMBO Mol. Med.* **6**, 482–495. <https://doi.org/10.1002/emmm.201302852> (2014).
77. Anstee, Q. M., Mantovani, A., Tilg, H. & Targher, G. Risk of cardiomyopathy and cardiac arrhythmias in patients with nonalcoholic fatty liver disease. *Nat. Rev. Gastroenterol. Hepatol.* **15**, 425–439. <https://doi.org/10.1038/s41575-018-0010-0> (2018).
78. Wang, N.-N. *et al.* HAMdb: A database of human autophagy modulators with specific pathway and disease information. *J. Cheminform.* **10**, 34. <https://doi.org/10.1186/s13321-018-0289-4> (2018).
79. Zhou, G. *et al.* NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* **47**, W234–W241. <https://doi.org/10.1093/nar/gkz240> (2019).
80. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
81. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
82. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361. <https://doi.org/10.1093/nar/gkw1092> (2017).
83. Slenter, D. N. *et al.* WikiPathways: A multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.* **46**, D661–D667. <https://doi.org/10.1093/nar/gkx1064> (2018).
84. Jassal, B. *et al.* The reactome pathway knowledgebase. *Nucleic Acids Res.* **48**, D498–D503. <https://doi.org/10.1093/nar/gkz1031> (2020).
85. Chen, E. Y. *et al.* Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **14**, 128. <https://doi.org/10.1186/1471-2105-14-128> (2013).
86. Kuleshov, M. V. *et al.* Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–97. <https://doi.org/10.1093/nar/gkw377> (2016).
87. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1 (2010).
88. Sanz, H., Valim, C., Vegas, E., Oller, J. M. & Reverter, F. SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinform.* **19**, 432. <https://doi.org/10.1186/s12859-018-2451-4> (2018).
89. Castro-Mondragon, J. A. *et al.* JASPAR 2022: The 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **50**, D165–D173. <https://doi.org/10.1093/nar/gkab1113> (2021).
90. Yoo, M. *et al.* DSigDB: Drug signatures database for gene set analysis. *Bioinform. (Oxf., Engl.)* **31**, 3069–3071. <https://doi.org/10.1093/bioinformatics/btv313> (2015).
91. Pintero, J. *et al.* DisGeNET: A comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.* **45**, D833–D839. <https://doi.org/10.1093/nar/gkw943> (2017).
92. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523. <https://doi.org/10.1038/s41467-019-09234-6> (2019).
93. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951. <https://doi.org/10.1002/pro.3715> (2019).
94. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551. <https://doi.org/10.1093/nar/gkaa970> (2021).

Acknowledgements

We thank Dr. Guangli Sun, Chief Physician for her critical comments on the manuscript. This study was supported by the National Natural Science Foundation of China (82074388); the project of shaoxing medical key discipline construction plan (2019SZD06); the Project of Health and Family Planning Commission of Zhejiang province (2021KY1139); Shaoxing Medical and Health Science and Technology Plan Project (2020A13026); Shandong Traditional Chinese Medicine Science and Technology Project (2020M044); and the Natural Science Foundation of Shandong Province (ZR2020QH306).

Author contributions

F.J.Z., M.Y.X. co-wrote the manuscript and revised it. J.R.J., S.W. and J.Z.Y. performed data analysis and data interpretation. X.L., C.Y. and Q.Z. conducted the bioinformatics and statistical analyses. D.X.X prepared Figs. 1, 2, 3, 4, 5, 6, 7 and 8. L.Z., Y.L. and M.P. supervised the project and revised the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-19027-5>.

Correspondence and requests for materials should be addressed to Y.L. or M.P.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2022