



OPEN

## Development and verification of the nomogram for dilated cardiomyopathy gene diagnosis

Li-qiang Zhou, Chuan Liu, Yeqing Zou & Zhi-qing Chen✉

Dilated cardiomyopathy (DCM) is a primary myocardial disease of unclear mechanism and poor prevention. The purpose of this study is to explore the potential molecular mechanisms and targets of DCM via bioinformatics methods and try to diagnose and prevent disease progression early. We screened 333 genes differentially expressed between DCM and normal heart samples from GSE141910, and further used Weighted correlation network analysis to identify 197 DCM-related genes. By identifying the key modules in the protein–protein interaction network and Least Absolute Shrinkage and Selection Operator regression analysis, seven hub DCM genes (CX3CR1, AGTR2, ADORA3, CXCL10, CXCL11, CXCL9, SAA1) were identified. Calculating the area under the receiver's operating curve revealed that these 7 genes have an excellent ability to diagnose and predict DCM. Based on this, we built a logistic regression model and drew a nomogram. The calibration curve showed that the actual incidence is basically the same as the predicted incidence; while the C-index values of the nomogram and the four external validation data sets are 0.95, 0.90, 0.96, and 0.737, respectively, showing excellent diagnostic and predictive ability; while the decision curve indicated the wide applicability of the nomogram is helpful for clinicians to make accurate decisions.

Dilated cardiomyopathy (DCM) is a primary cardiomyopathy characterized by left ventricular dilation and systolic insufficiency, accompanied by heart remodeling and fibrosis<sup>1</sup>. DCM is usually progressive aggravation and gradually developed into severe congestive heart failure, which will seriously threaten the survival rate of patients<sup>2</sup>. Currently, heart transplantation is the most effective treatment for dilated cardiomyopathy with long-term heart failure that has failed medical treatment, and the prognosis is greatly improved<sup>3</sup>. Early diagnose and intervention of DCM are particularly important for improving the prognosis and quality life of DCM patients.

DCM is a genetic disease involving post-inflammatory or genetic causes. The genetic characteristics of DCM can be accurately diagnosed in advance, and preventive measures can be taken in time to avoid the further occurrence and development of the disease<sup>4</sup>. Elizabeth et al. re-evaluated a number of genes that are important in the progression of DCM, which has implications for the diagnosis and treatment of DCM<sup>5</sup>. And it has become feasible to use the second-generation sequencing to study the genetics of DCM, which deeply understand the molecular biological mechanism of DCM and find the biological targets for the diagnosis and treatment of DCM.

In this study, we conducted an in-depth analysis of DCM samples and normal samples in the GSE141910 data set through a variety of statistical methods, obtained differentially expressed genes that were significantly correlated with DCM, and studied the underlying mechanism of these genes' impression of DCM. Furthermore, we obtained the biomarkers of DCM and analyzed its diagnostic predictive value. Based on this, we constructed a genetic diagnosis model of DCM, and through multiple methods of testing, the results showed that it has excellent predictive sensitivity. This research can help clinicians diagnose DCM early and has far-reaching significance for the prevention and treatment of DCM.

### Materials and methods

**Statement.** All methods were carried out in accordance with relevant guidelines and regulations.

**Data collection and processing.** We obtained the dataset GSE141910 from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>), this dataset contained 166 samples are surgical myectomy tissues from patients with dilated cardiomyopathy (DCM), and 166 samples are patients without heart failure (NF) as control cardiac tissues. Subsequently, we implemented the eBayes method of the "limma" package

Jiangxi Key Laboratory of Molecular Medicine, The Second Affiliated Hospital of Nanchang University, 1 Minde Road, Donghu District, Nanchang 330006, Jiangxi, China. ✉email: czq033021@163.com

with R 3.6.1 software<sup>6</sup>. Simultaneously, we set  $|\log_2 \text{fold change (FC)}| > 1.5$  and false discovery rate (FDR)  $< 0.05$  as the filter condition to identify Differentially expressed genes (DEGs) between the DCM and NF samples.

**Weighted gene co-expression network analysis (WGCNA).** We used the “WGCNA” package to performed a network construction<sup>7</sup>. WGCNA is an analysis method that analyzes gene expression patterns in multiple samples. It can cluster genes with similar expression patterns and analyze the association between modules and specific traits or phenotypes. WGCNA started at the level of DEGs, dividing them into several modules based on their patterns of connection with other genes, thus identifying the module most relevant to DCM. An unsupervised co-expression relationship was initially built on the basis of the adjacency matrix of connection strengths by using Pearson’s correlation coefficients for gene pairs. After selecting the appropriate power value, the adjacency matrix of the gene expression data of GC patients is clustered using topological overlap matrix analysis. Finally, the dynamic tree cut algorithm was applied to the dendrogram for module identification with the mini-size of module gene numbers set as 50 and a cut height of 0.85. We select the genes in the modules with the highest correlation with DCM for further analysis.

**Gene function enrichment analysis.** In order to study the functions of genes in the key module, we use the “ClusterProfiler” package to perform gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis on these genes<sup>8</sup>. We used Database for Annotation, Visualization and Integrated discovery GO function is carefully evaluated from three aspects: biological process (BP), molecular function (MF) and cellular component (CC). KEGG is helpful for studying an integrated network of gene and expression information, systematically analyzing the metabolic pathways of gene products in cells and a database of the functions of these gene products. The enrichment results with *P*-value and FDR value below 0.05 are considered meaningful. The visual GO and KEGG enrichment results were performed by “GPlot” R packages<sup>9</sup>.

**Construction protein protein interaction (PPI) network.** STRING is a database for searching known and predicted protein interactions (Version 10.5; <http://stringdb.org/>)<sup>10</sup>. We used this database to build a PPI network to mine the core regulatory genes of DCM. We uploaded the gene of the key module to the STRING database to get a PPI play. Then, we used Cytoscape software (version: 3.7.2, <http://www.cytoscape.org/>) to visualize the PPI and perform further analysis<sup>11</sup>. Cytoscape’s network analysis analyzes the degree value of all nodes. All nodes are arranged in order from the inside to the outside according to the degree size, and the nodes with degree  $> 10$  are selected. Cytoscape’s pluggable unit molecular complex detection (MCODE) is a module for screening PPI networks<sup>12</sup>. Take the cut-off degree = 2, cut-off point = 0.2, *k*-core = 2, max depth = 100 as the network score and cluster search parameter settings, then analyze the hub modules. If there is a gene with degree  $> 10$  and exists in the hub module, it is regarded as a candidate gene, and further analysis is performed.

**LASSO regression to identify the hub gene.** The least absolute shrinkage and selection operator (LASSO) method, which is suitable for the reduction of high-dimensional data, was used to select the most useful predictive features from the primary data set in this study. Using the LASSO regression model, among the candidate genes, all candidate genes are reduced to a limited potential predictor. If the penalization coefficient lambda ( $\lambda$ ) is large, there is no effect on the estimated regression parameters, but as the  $\lambda$  gets smaller, some coefficients may be shrunk toward zero. We use “glmnet” package to screen the hub genes that predict DCM from these candidate genes<sup>13</sup>. And use the “ROC” package to analyze the ROC values of these hub genes and evaluate the sensitivity of their predictions.

**Construction and verification of DCM diagnostic signature and nomogram.** We used the “rms” package to perform multiple logistic regression on the obtained central genes and obtained the corresponding probability (OR) and regression coefficient ( $\beta$ ). In order to construct a diagnostic model for predicting DCM, according to the formula: Risk score =  $\beta_1 * \text{EXP1} + \beta_2 * \text{EXP2} + \dots + \beta_i * \text{EXP}_i$ , thereby predicting the incidence of DCM. We used a nomogram to visualize the DCM prediction model, assigned a corresponding score to each gene expression, and then added the scores of all hub genes to get the total score. By drawing a vertical line at the total score obtained, the risk value of DCM can be clarified. In order to quantify the predictive performance of the nomogram, a calibration curve was drawn to evaluate the consistency between the prediction and observation of the nomogram, and the calibration of the model was described according to the consistency between the predicted risk and the actual result. Then, the consistency index (C-index) was calculated in GSE116250<sup>14</sup>, GSE43435<sup>15</sup>, GSE21610<sup>16</sup> to verify the prediction performance. In addition, we used the “rmda” package to draw a decision curve (DCA) to calculate the net benefit of diagnostic signatures and evaluate diagnostic performance.

**Consent for publication.** All authors agree to the publication of this article.

## Results

**Identify DEGs of DCM.** This study analyzed the hub genes of DCM and its underlying mechanisms via various statistical methods and analyzed its ability to predict DCM. The entire workflow is shown in Fig. 1. We identified 333 differentially expressed genes in 166 NF and 166 DCM samples that met the screening criteria (Supplementary Table 1). A heat map was aimed for assessing the levels of the DEGs expression to distinguish the DCM and control samples (Fig. 2B). Among these DEGs, 228 were upregulated and 105 were downregulated (Fig. 2A).

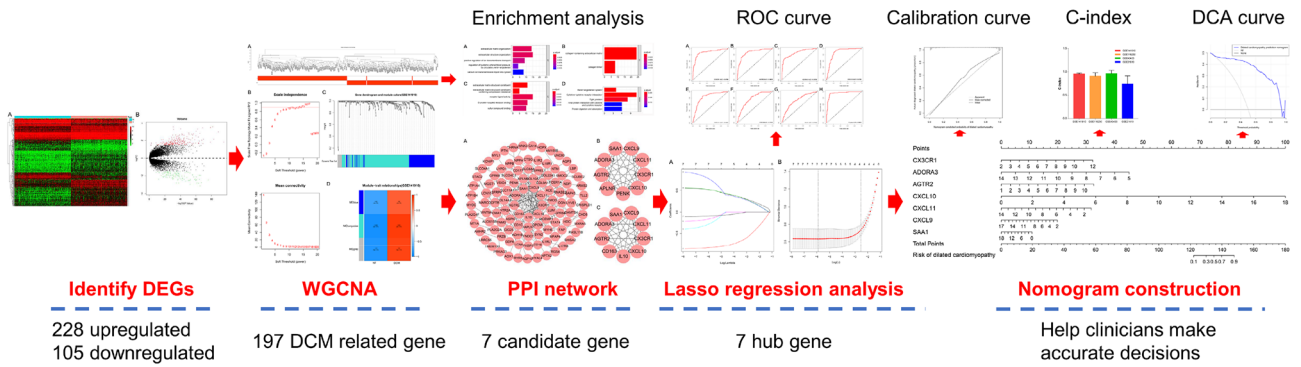


Figure 1. Flow chart of this research.

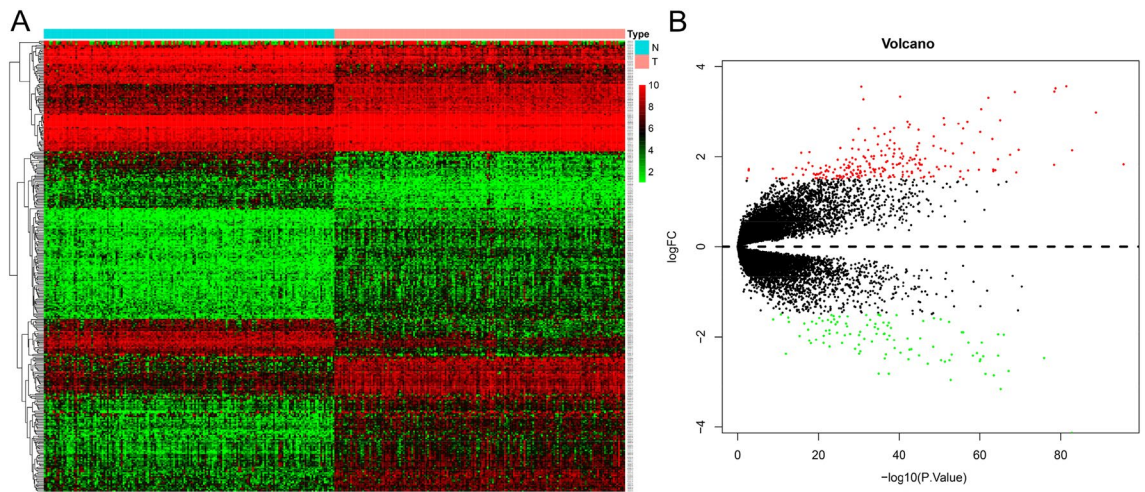
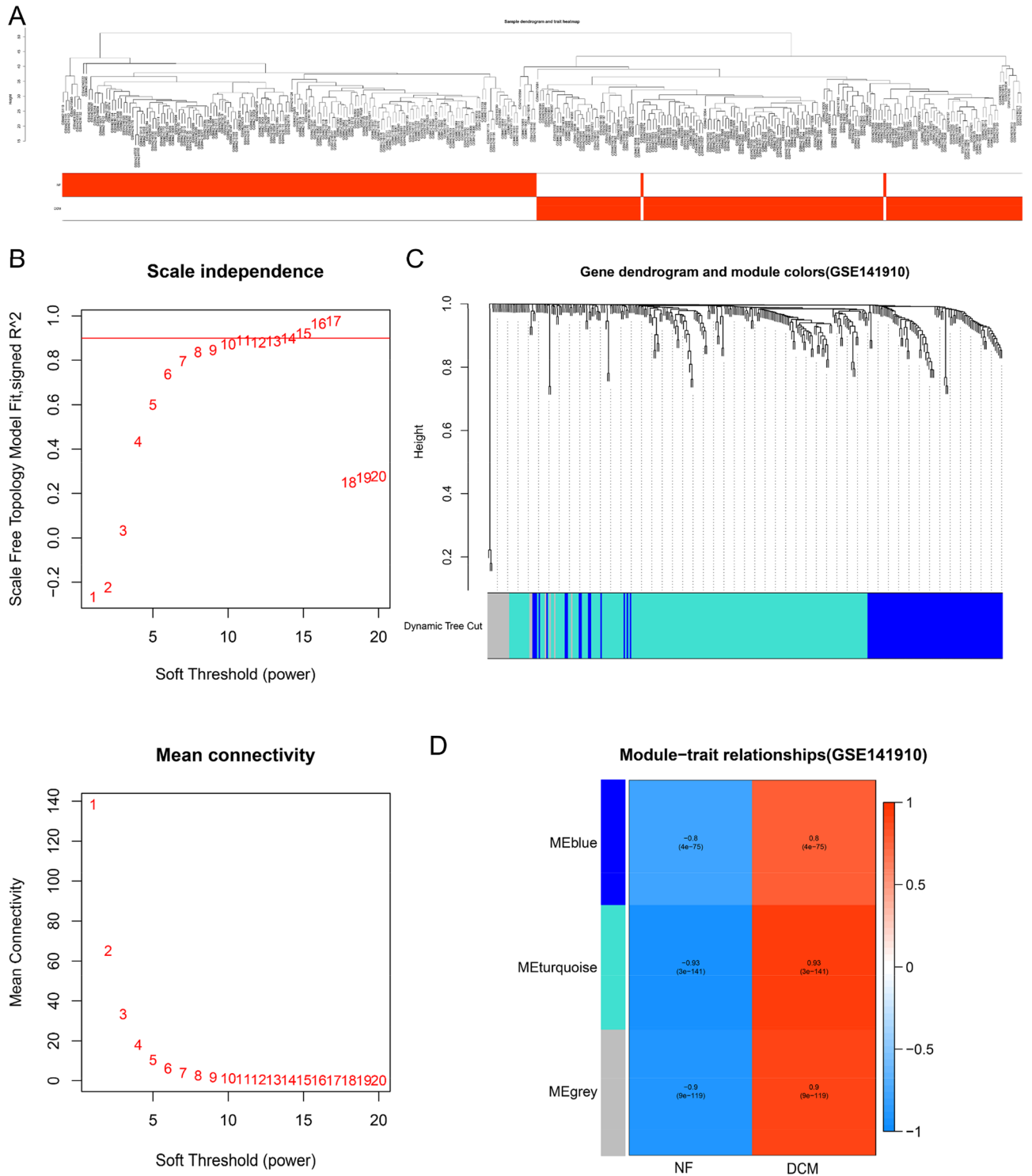


Figure 2. Identify the 333 differential genes of dilated cardiomyopathy. (A) Heat map. (B) Volcano plot.

**WGCNA identification of key modules.** We analyzed whether there are outliers in each sample, and then performed hierarchical clustering. The results suggested that there are no samples with outliers, and the dendrogram contains all 332 samples (Fig. 3A). In WGCNA analysis, we selected the soft-thresholding power to identify the relatively balanced scale independence and mean connectivity. As shown in Fig. 3B, under the condition that the cutting height is set to 0.85, power = 15 could be as the power value of soft-threshold. Then, a total of three modules were generated by average linkage hierarchical clustering based on the 333 input genes (Fig. 3C). Of which, after calculating the MS of each module-trait correlation by WGCNA, the three modules related to DCM are MEblue module ( $r=0.80$ ,  $P=4.00E-75$ ), MEturquoise module ( $r=0.93$ ,  $P=3.00E-141$ ), MEgrey module ( $r=0.90$ ,  $P=9.00E-119$ ). The MEturquoise module had the highest correlation, and it contained 197 genes that could potentially be associated with DCM (Fig. 3D).

**GO and KEGG function analysis.** GO and KEGG functional enrichment analysis to understand the potential mechanism of genes in the MEturquoise module (Fig. 3D) affecting DCM, we found the enrichment results of Top5. There are 93 items in GO enrichment analysis. The biological processes are enriched in 82 items such as extracellular matrix organization, extracellular structure organization, positive regulation of ion transmembrane transport, regulation of systemic arterial blood pressure by circulatory renin-angiotensin, calcium ion transmembrane import into cytosol, etc. Entry (Fig. 4A); Cellular components are enriched in collagen-containing extracellular matrix, collagen trimer (Fig. 4B); molecular functions are enriched in extracellular matrix structural constituent, extracellular matrix structural constituent conferring compression resistance, receptor ligand activity, G protein-coupled receptor binding, sulfur compound binding, lipopeptide 8 items such as binding (Fig. 4C). KEGG results suggested that these genes may affect DCM through renin-angiotensin system, cytokine-cytokine receptor interaction, Tight junction, Viral protein interaction with cytokine and cytokine receptor, Protein digestion and absorption (Fig. 4D).

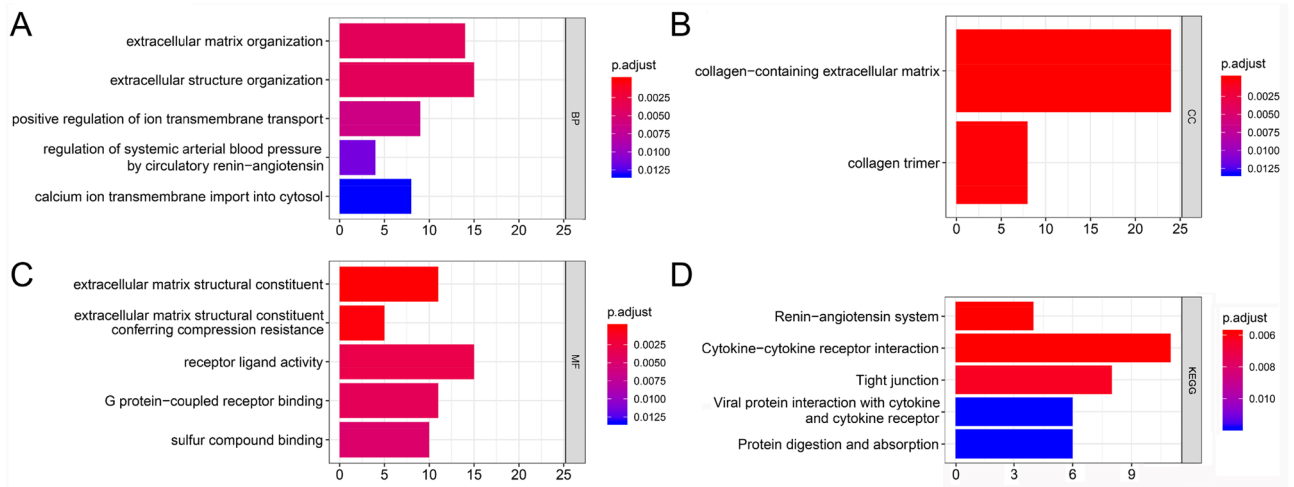
**PPI network screening candidate genes.** We used String to form a PPI network with 109 nodes and 206 edges from 197 genes in the turquoise module (Fig. 5A). And we analyzed the degree value of each node and selected 9 genes greater than 10 to form a sub-network with 9 nodes and 36 edges (Fig. 5B). Next, we analyzed the key modules of the entire PPI network using MCODE, and obtained a network composed of 9 genes



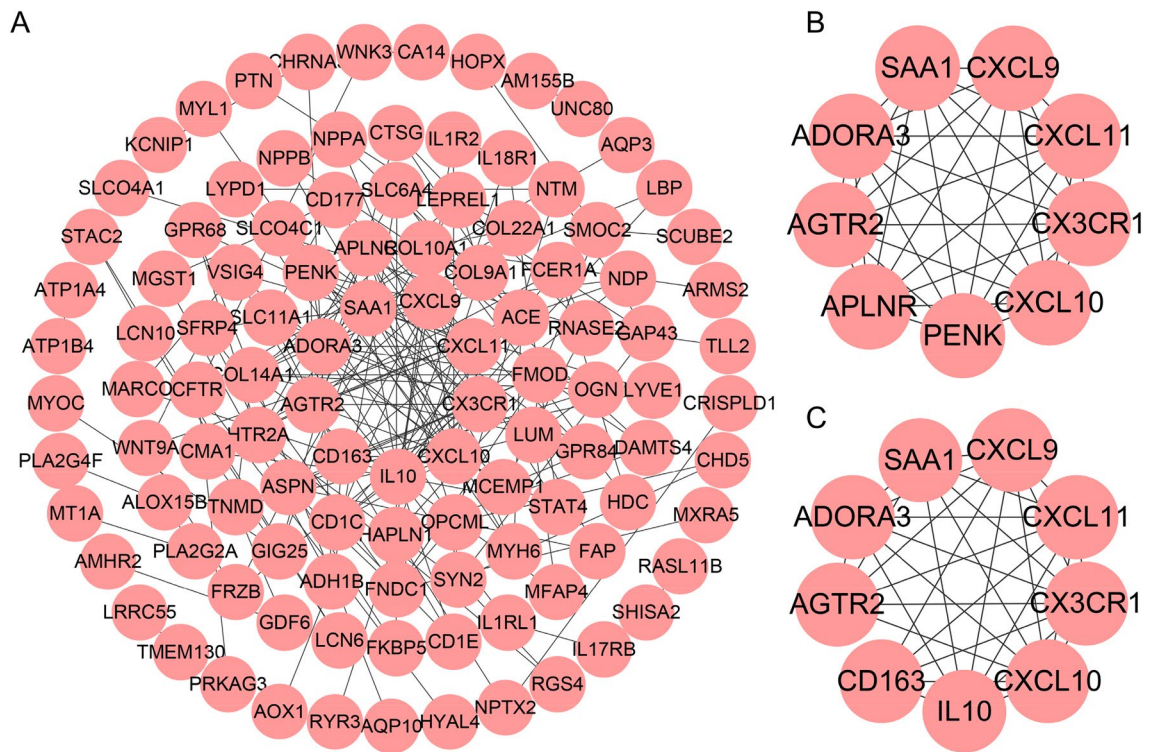
**Figure 3.** WGCNA identified 197 genes closely related to dilated cardiomyopathy. (A) The distribution of different samples. (B) Analyze the best Power value. (C) Dendrogram of all clusters expressing genes based on the difference metric (1-TOM). (D) The heatmap shows the correlation between each module and dilated cardiomyopathy.

(Fig. 5C). The genes with degree > 10 were consistent with the genes of the key modules, and they were used for further analysis. What is interesting is that the degree > 10 genes are the same as 7 genes in the key modules, including CXCR1, AGTR2, ADORA3, CXCL10, CXCL11, CXCL9, SAA1. We used these genes for further analysis to screen for genes that best predicted DCM.





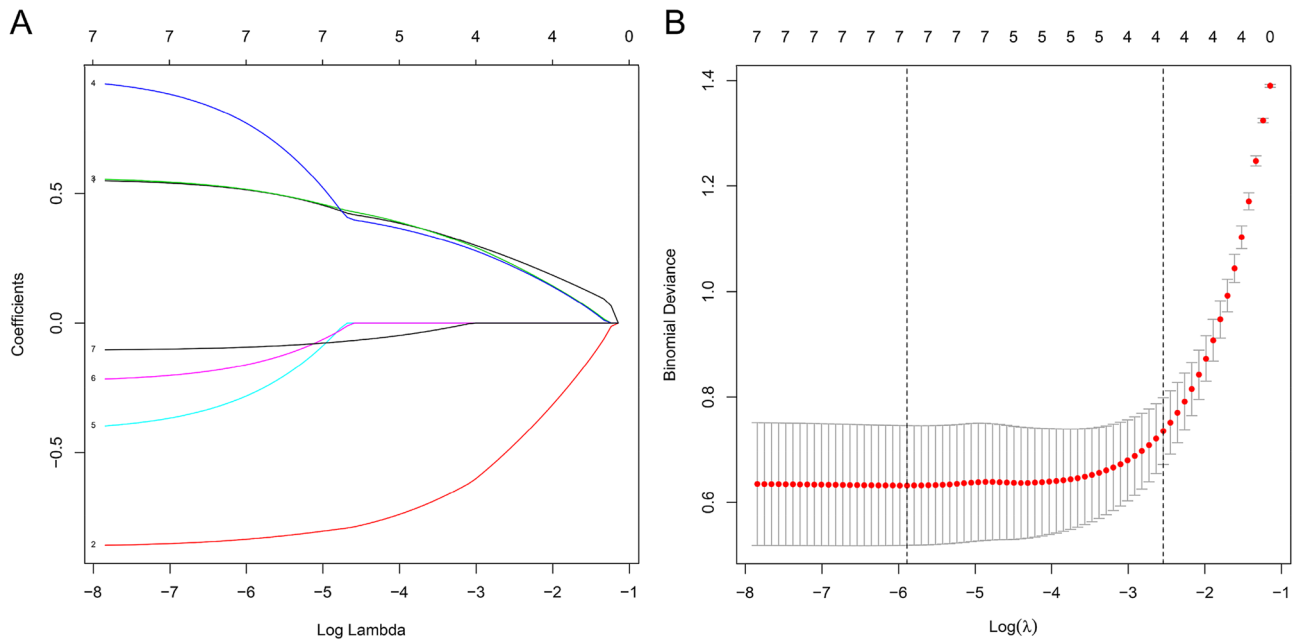
**Figure 4.** Gene function enrichment analysis. (A) Biological process. (B) Cellular components. (C) Molecular function. (D) KEGG.



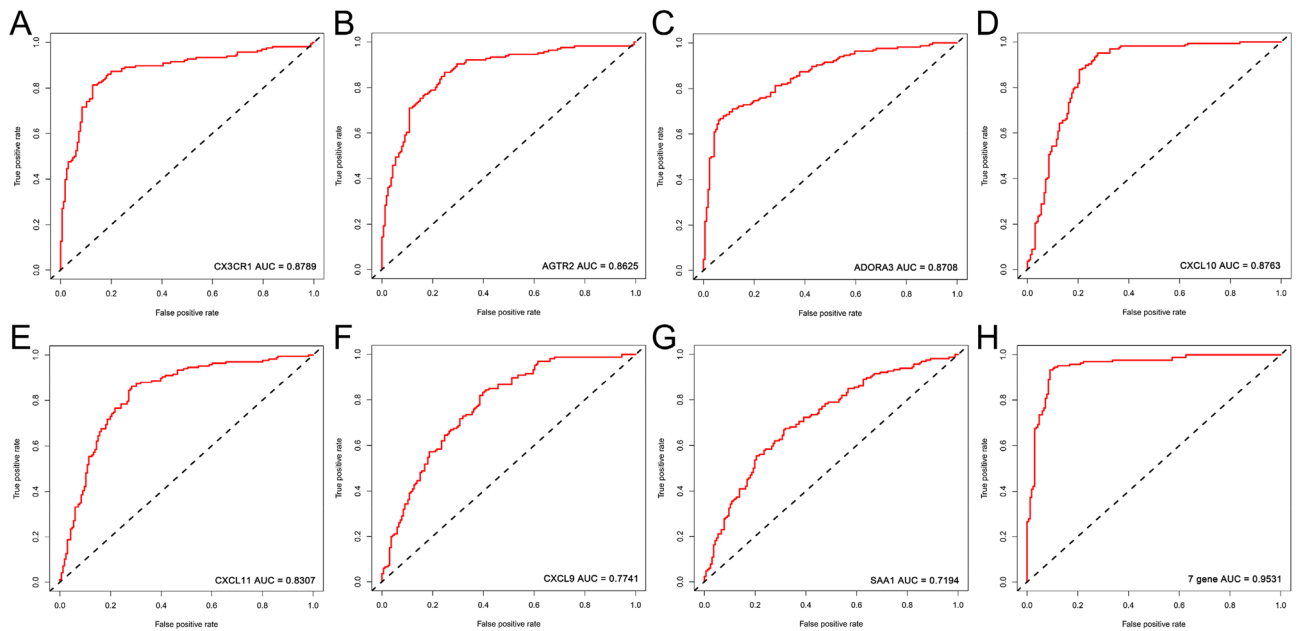
**Figure 5.** PPI network screening 7 candidate genes. (A) A sub-network composed of genes whose degree is greater than 10. (B) The key module identified by MCODE.

**Construction DCM diagnosis model.** Lasso regression is used to screen genes that predict DCM. We set the value of  $\lambda$  to  $2.09E-03$ , it indicated that the regression coefficients of most variables in the model tend to be stable, and the results showed that the above 7 genes are all predictors of DCM (Fig. 6). To this end, we drew the receiver operating characteristic curve (ROC) of these 7 genes, and calculated the area under the curve (AUC). The AUC value of each gene (Fig. 7A–G): CXCR1 (AUC=0.88), AGTR2 (AUC=0.86), ADORA3 (AUC=0.87), CXCL10 (AUC=0.88), CXCL11 (AUC=0.83), CXCL9 (AUC=0.77), SAA1 (AUC=0.72). We found that the AUC value after integration of 7 genes is 0.95 (Fig. 7H), which has the best sensitivity compared to individual genes. Based on these 7 hub genes, we used a multiple logistic regression to construct a predictive model to predict the risk of DCM (Table 1).

**DCM nomogram establishment and verification.** In order to understand the relationship between each gene and the onset of DCM more intuitively, we constructed a nomogram based on multiple logistic regres-



**Figure 6.** Lasso regression removes collinearity to identify 7 hub genes.

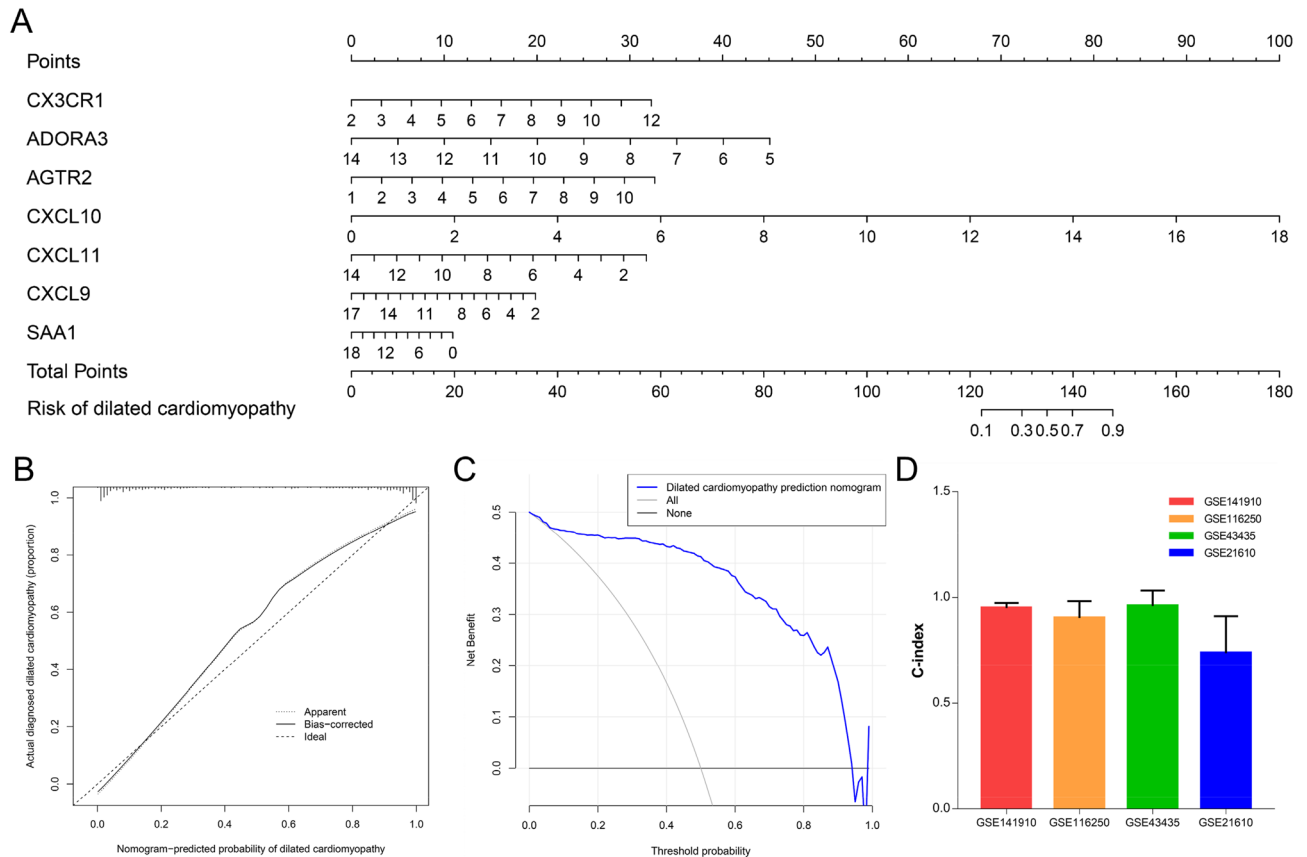


**Figure 7.** Subject's operable curve analysis 7 hub genes individually and jointly predict the ability of dilated cardiomyopathy. (A) CX3CR1. (B) AGTR2. (C) ADORA3. (D) CXCL10. (E) CXCL11. (F) CXCL9. (G) SAA1. (H) Combine 7 genes.

sion (Fig. 8A). By detecting the expression of each hub gene, the corresponding score was obtained, and the incidence of DCM can be judged according to the total score obtained. In order to verify the accuracy of the nomogram prediction, we draw a Calibration curve. According to Fig. 8B, the predicted curve is almost the ideal curve. In addition, we also calculated the C-index. The C-index in the GSE141910 data set used in this study is 0.95, and in the external data set GSE116250(C-index = 0.90), GSE43435(C-index = 0.96), GSE21610(C-index = 0.74), all indicate excellent predictive performance (Fig. 8D). A decision curve means that > 1% to < 98% of the population can follow the nomogram to predict the incidence of DCM (Fig. 8C).

Gene ID	Regression coefficients ( $\beta$ )	Odds ratio (OR)	95% Confidence interval (CI)	P-value
CX3CR1	0.56	1.74	1.25–2.47	1.35E-03
ADORA3	-0.86	0.42	0.28–0.61	1.78E-05
AGTR2	0.56	1.76	1.31–2.42	2.89E-04
CXCL10	0.96	2.60	1.66–4.31	8.35E-05
CXCL11	-0.42	0.65	0.53–0.98	4.43E-02
CXCL9	-0.23	0.80	0.58–1.09	1.59E-01
SAA1	-0.10	0.90	0.79–1.03	1.25E-01

**Table 1.** Multiple logistic regression of dilated cardiomyopathy.



**Figure 8.** The establishment and verification of the diagnosis and prediction model of dilated cardiomyopathy. (A) 7 Predictive nomograms of genetic diagnosis. (B) Calibration curve. (C) Decision curve. (D) 4 C-index of external validation data set.

## Discussion

DCM is most common in middle-aged people, and the onset more than 10 years. It is characterized by enlarged left or right ventricle or bilateral ventricles, accompanied by myocardial hypertrophy. Decreased ventricular systolic function, with or without congestive heart failure. Ventricular or atrial arrhythmia is more common. And death can occur at any stage of the disease. At present, dilated cardiomyopathy is diagnosed through exclusion method, that is, excluding other specific causes of cardiac enlargement and cardiac insufficiency, and the diagnosis based on clinical manifestations and auxiliary examinations<sup>1</sup>. It is difficult to prevent DCM due to the unclear mechanism. Therefore, the early diagnosis of DCM will improve the prognosis and quality of life of DCM patients. With the development of genomics, it has been widely used to predict disease susceptibility, drug response and early diagnosis of disease<sup>4</sup>. In this study, 333 differentially expressed genes were identified by analyzing the GSE141910 microarray, and then 197 genes significantly related to DCM were obtained using WGCNA. Among the genes obtained from WGCNA, RGS, FKBP5, STAT4, MYH6, COL9A1 and other genes were confirmed to be associated with DCM progression<sup>17–21</sup>. In addition, ACE-targeted inhibitors have been extensively studied in the treatment of DCM.

We have observed that the expression of NPPA and NPPB are elevated. The DCM patients' ventricular enlargement and reduced myocardial systolic function force the atria and cardiomyocytes to express and release

ANP and BNP, thereby exerting diuretic, vasodilatory and immunosuppressive effects. But interestingly, we also found an increase in the expression of ACE and ATGR2, which is opposed to the function of ANP to relax blood vessels. Among them, the overexpression of ATGR2 may be the leading cause that DCM forces the heart to remodel, fibroblasts are the main cell type, and ATGR2 is the main inducer of this cell<sup>22</sup>. In addition, Yan et al. showed that ATGR2 is up-regulated in ventricular myocytes, and its expression is closely related to the phosphorylated protein levels of PKC- $\alpha$ , PKC- $\beta$  and p70S6 kinase, which promotes the development of dilated cardiomyopathy and heart failure in vivo<sup>23</sup>. There is also evidence that DCM is closely related to immunity. The disorder of the cellular and humoral immune system should be gone hand in hand with the development of DCM. The infiltration of lymphocytes and monocytes and the increased expression of cell adhesion molecules are often found in DCM patients. The production of autoantibodies and the corresponding side effects have also been widely discussed<sup>24–26</sup>. Nakayama et al. confirmed that the increase in myocardial immune activation was positively correlated with poor prognosis of DCM, and CD163-specifically labeled M2 macrophages were closely related to the ventricular remodeling of DCM<sup>27</sup>. The important cause of DCM also includes genetic predisposition. About 20–35% of DCM cases are reported as familial inheritance<sup>28</sup>. Zhao et al. through sequencing revealed that 7 gene mutations, including MYH6, are closely related to the pathogenesis of DCM<sup>18</sup>.

Subsequently, we conducted enrichment analysis of these DCM related genes and explored their potential mechanisms in DCM. Among them, BP and CC are mainly enriched in the extracellular matrix or structure related to cardiac fibrosis, the transport or regulation of different ions, the regulation of renin–angiotensin, and the related immune response and regulation. In MF, in addition to the structural components of the extracellular matrix, there are also G protein-coupled receptors and CXCR chemokine receptors. G protein-coupled receptors are the largest cell membrane surface receptor family in the human body, which are closely related to the occurrence of diseases such as cancer, heart disease, diabetes, and Alzheimer's disease. Among them, autoantibodies to G-protein-coupled  $\beta$ 1-adrenergic receptors and M2-cholinergic receptors are specifically expressed in patients with DCM and are involved in the pathological and physiological functions of DCM<sup>29</sup>. Similarly, CXCR chemokine receptors are also a subfamily of G protein-coupled receptors, and this family of receptors is related to cardiac fibrosis. In patients with DCM and end-stage heart failure, there is a close relationship between persistent atrial fibrillation and extracellular remodeling. In atrial fibrosis and atrial fibrillation, the biochemical mediators that play a central role include angiotensin II, CTGF, PDGF and TGF- $\beta$ . Cytokines and Cytokine Receptors interaction is also enriched in KEGG, which mediates and initiates cell signal transduction. The cytokines exert their biological effects by binding to corresponding cytokine receptors on the cell surface. According to their structure and signal transduction pathway, cytokine receptors were divided into different families or superfamily, such as type I cytokine receptor, type II cytokine receptor, tumor necrosis factor receptor and chemotactic cytokine receptor. As confirmed by Ohtsuka, the increase in serum VEGF and IL-13 may be closely related to the changes in DCM myocardial tissue structure<sup>30</sup>. Chen et al. found that TNF- $\alpha$  gene polymorphism (G-308A) possibly be related to the susceptibility to DCM<sup>31</sup>. These evidence clarified the underlying mechanism of the physiological and pathological development of DCM.

In addition, we conducted a series of analyses to identify 7 hub genes in DCM. Among them, ATGR2 induces cardiac fibroblasts to remodel the DCM heart and promote the development of mental failure<sup>22,23</sup>. While CXCL9, CXCL10, and CXCL11 belong to the chemokine family, CX3CR1, ADORA3 and SAA1 have not been reported in dilated cardiomyopathy, and further research is needed. Finally, we constructed a nomogram for predicting the incidence of DCM based on the expression levels of these 7 genes, and conducted multiple verifications, which not only confirmed its good predictive performance, but also revealed its wide applicability. It can help clinicians and patients diagnose DCM as early as possible, and carry out timely interventions to obtain a better prognosis and improve the quality of life of patients.

## Data availability

All data can be obtained from the corresponding author's office and public databases.

Received: 26 November 2021; Accepted: 20 May 2022

Published online: 26 May 2022

## References

1. Jefferies, J. L. & Towbin, J. A. Dilated cardiomyopathy. *Lancet (London, England)*. **375**, 752–762 (2010).
2. Richardson, P. et al. Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. *Circulation* **93**, 841–842 (1996).
3. Hunt, S. A. et al. 2009 Focused update incorporated into the ACC/AHA 2005 Guidelines for the Diagnosis and Management of Heart Failure in Adults A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines Developed in Collaboration With the International Society for Heart and Lung Transplantation. *J. Am. Coll. Cardiol.* **53**, e1–e90 (2009).
4. Merlo, M. et al. Dilated cardiomyopathy: A dynamic disease - clinical course, reverse remodeling and prognostic stratification. *G. Ital. Cardiol.* **2016**(17), 15–23 (2006).
5. Jordan, E. et al. Evidence-based assessment of genes in dilated cardiomyopathy. *Circulation* **144**, 7–19 (2021).
6. Ritchie, M. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
7. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
8. Yu, G., Wang, L., Han, Y. & He, Q. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
9. Walter, W., Sánchez-Cabo, F. & Ricote, M. GPlot: An R package for visually combining expression data with functional analysis. *Bioinformatics (Oxford, England)* **31**, 2912–2914 (2015).
10. Szklarczyk, D. et al. The STRING database in 2021: Customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–D612 (2021).



11. Cline, M. *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* **2**, 2366–2382 (2007).
12. Bader, G. D. & Hogue, C. W. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinform.* **4**, 2 (2003).
13. Engebretsen, S. & Bohlin, J. Statistical predictions with glmnet. *Clin. Epigenet.* **11**, 123 (2019).
14. Sweet, M. E. *et al.* Transcriptome analysis of human heart failure reveals dysregulated cell adhesion in dilated cardiomyopathy and activated immune pathways in ischemic heart failure. *BMC Genomics* **19**, 812 (2018).
15. Koczor, C. A. *et al.* Detection of differentially methylated gene promoters in failing and nonfailing human left ventricle myocardium using computation analysis. *Physiol. Genomics* **45**, 597–605 (2013).
16. Schwientek, P. *et al.* Global gene expression analysis in nonfailing and failing myocardium pre- and postpulsatile and nonpulsatile ventricular assist device support. *Physiol. Genomics* **42**, 397–405 (2010).
17. Fu, K. *et al.* Signal transducer and activator of transcription 4 gene polymorphisms associated with dilated cardiomyopathy in Chinese Han population. *Sichuan da xue xue bao Yi xue ban (J. Sichuan Univ. Med. Sci. Ed.)* **44**, 558–562 (2013).
18. Zhao, T. *et al.* Young and early-onset dilated cardiomyopathy with malignant ventricular arrhythmia and sudden cardiac death induced by the heterozygous LDB3, MYH6, and SYNE1 missense mutations. *Ann. Noninvasive Electrocardiol.* **26**, e12840 (2021).
19. Wada, K. *et al.* Blood-based epigenetic markers of FKBP5 gene methylation in patients with dilated cardiomyopathy. *J. Am. Heart Assoc.* **10**, e021101 (2021).
20. Mittmann, C. *et al.* Expression of ten RGS proteins in human myocardium: Functional characterization of an upregulation of RGS4 in heart failure. *Cardiovasc. Res.* **55**, 778–786 (2002).
21. He, M., Li, M. & Guo, Z. STAT4 regulates cardiomyocyte apoptosis in rat models of diabetic cardiomyopathy. *Acta Histochem.* **124**, 151872 (2022).
22. Tsutsumi, Y. *et al.* Angiotensin II type 2 receptor is upregulated in human heart with interstitial fibrosis, and cardiac fibroblasts are the major cell type for its expression. *Circ. Res.* **83**, 1035–1046 (1998).
23. Yan, X. *et al.* Ventricular-specific expression of angiotensin II type 2 receptors causes dilated cardiomyopathy and heart failure in transgenic mice. *Am. J. Physiol. Heart Circ. Physiol.* **285**, H2179–H2187 (2003).
24. Heymans, S. *et al.* Inflammation as a therapeutic target in heart failure? A scientific statement from the Translational Research Committee of the Heart Failure Association of the European Society of Cardiology. *Eur. J. Heart Fail.* **11**, 119–129 (2009).
25. Kallwellis-Opara, A. *et al.* Autoimmunological features in inflammatory cardiomyopathy. *Clin. Res. Cardiol.* **96**, 469–480 (2007).
26. Jahns, R., Boivin, V., Schwarzbach, V., Ertl, G. & Lohse, M. J. Pathological autoantibodies in cardiomyopathy. *Autoimmunity* **41**, 454–461 (2008).
27. Nakayama, T. *et al.* Clinical impact of the presence of macrophages in endomyocardial biopsies of patients with dilated cardiomyopathy. *Eur. J. Heart Fail.* **19**, 490–498 (2017).
28. Maron, B. J. *et al.* Contemporary definitions and classification of the cardiomyopathies. *Circulation* **113**, 1807–1816 (2006).
29. Wallukat, G. & Schimke, I. Agonistic autoantibodies directed against G-protein-coupled receptors and their relationship to cardiovascular diseases. *Semin. Immunopathol.* **36**, 351–363 (2014).
30. Ohtsuka, T. *et al.* Serum markers of angiogenesis and myocardial ultrasonic tissue characterization in patients with dilated cardiomyopathy. *Eur. J. Heart Fail.* **7**, 689–695 (2005).
31. Chen, M. *et al.* Tumor necrosis factor- $\alpha$  gene polymorphism (G-308A) and dilated cardiomyopathy. *Int. Heart J.* **60**, 656–664 (2019).

## Author contributions

L.Z.: Conception, design, data analysis and manuscript writing, C.L. and Y.Z.: Data collection, Z.C.: Guide, supervise and review.

## Funding

The Second Affiliated Hospital of Nanchang University National Natural Science Foundation Incubation Project, Grant/Award Numbers: 2021YNYFY12005; Natural Science Foundation of Jiangxi Province, Grant/Award Numbers: 20192BAB205079; Natural Science Foundation of China, Grant/Award Numbers: 81860530.

## 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/s41598-022-13135-y>.

**Correspondence** and requests for materials should be addressed to Z.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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