Research letter

Cardiology

# Machine-learning analysis and *in vivo* validation of shared immune signatures in viral myocarditis and dilated cardiomyopathy

Beibei Du[1,2], Rong Miao[1,2], Xiaojie Ren[3], Xinjing Chen[4]*, Lemin Zheng[1,2]*

[1]The Institute of Cardiovascular Sciences, the Institute of Systems Biomedicine, School of Basic Medical Sciences, Health Science Center, Peking University, Beijing, China
[2]State Key Laboratory of Vascular Homeostasis and Remodeling, NHC Key Laboratory of Cardiovascular Molecular Biology and Regulatory Peptides, Peking University, Beijing, China
[3]China National Clinical Research Center for Neurological Diseases, Beijing Tiantan Hospital, Advanced Innovation Center for Human Brain Protection, The Capital Medical University, Beijing, China
[4]Department of Cardiology, Provincial Clinical Medical College of Fujian Medical University, Fuzhou University Affiliated Provincial Hospital, Fuzhou, China

**Submitted:** 26 May 2026; **Accepted:** 26 September 2026
**Online publication:** 22 October 2026

Arch Med Sci 2026; 22 (5): 3058–3063
DOI: https://doi.org/10.5114/aoms/240346
Copyright © 2026 Termedia & Banach

***Corresponding authors:**
Lemin Zheng
The Institute of Cardiovascular Sciences
the Institute of Systems Biomedicine
School of Basic Medical Sciences
Health Science Center
State Key Laboratory of Vascular Homeostasis and Remodeling
NHC Key Laboratory of Cardiovascular Molecular Biology and Regulatory Peptides
Peking University
Beijing, China
E-mail: zhengl@bjmu.edu.cn

Xinjing Chen
Department of Cardiology
Provincial Clinical Medical College of Fujian Medical University
Fuzhou University Affiliated Provincial Hospital, Fuzhou, China
E-mail: 3319326@163.com

Viral myocarditis (VMC) can evolve into dilated cardiomyopathy (DCM), but the molecular features linking acute myocardial inflammation to chronic ventricular remodelling remain incompletely defined [1–3]. Persistent antiviral signalling, maladaptive immune activation, and fibrosis have been implicated [4, 5], yet biomarkers that are consistently identified across human disease and experimental models are scarce. We therefore integrated transcriptomic datasets from VMC and DCM, applied complementary machine-learning methods, and performed external and experimental validation to identify shared immune-related biomarkers and candidate therapeutic targets.

**Methods.** To identify molecular signals shared by VMC and DCM, we analysed the human myocardial datasets GSE57338 (136 controls and 82 DCM samples) and GSE1145 (11 controls and 7 VMC samples). Differentially expressed genes were identified using limma with a Benjamini-Hochberg-adjusted $p$-value < 0.05 and |log2 fold change| > 0.5, followed by functional and interaction analyses of genes dysregulated in both diseases [6]. We then focused on the immune-related subset and used GSE57338 as the training cohort to identify human DCM classifiers through 10-fold cross-validated LASSO regression, support-vector-machine recursive feature elimination, and random forest. To distinguish these classifiers from biologically conserved candidates, expression patterns were independently examined in GSE35182, a longitudinal CVB3 mouse dataset sampled during acute and chronic disease [7]. Immune-cell deconvolution and receiver operating characteristic analyses were used as complementary assessments of immune context and discriminatory performance [8]. The overall study design is summarized in Figure 1 A.

***Experimental validation.*** Seven genes supported by the cross-species analysis were evaluated in a separate CVB3-induced VMC mouse model. Six-week-old male BALB/c mice received an intraperitoneal injection of CVB3 or phosphate-buffered saline. At day 30, cardiac structure and systolic function were assessed by M-mode echocardiography, and



Creative Commons licenses: This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY -NC -SA 4.0). License (http://creativecommons.org/licenses/by-nc-sa/4.0/).

myocardial expression was measured by RT-qPCR. All procedures complied with the US National Institutes of Health Guide for the Care and Use of Laboratory Animals. Drug-gene interactions were explored with DGIdb; this analysis was not used to infer treatment efficacy. The mouse experiment tested expression direction together with a cardiac phenotype rather than reproducing the human classification analysis.

**Results.** We identified 465 differentially expressed genes in DCM and 2309 in VMC, including 134 shared genes. The shared genes were enriched for inflammatory, antiviral, and extracellular-matrix processes and for JAK/STAT, PI3K/Akt, and IL-17 signalling [9]. Gene-set enrichment showed activation of interferon-α and interferon-γ responses in both datasets. IL-6/JAK/STAT3 signalling, however, was suppressed in VMC and

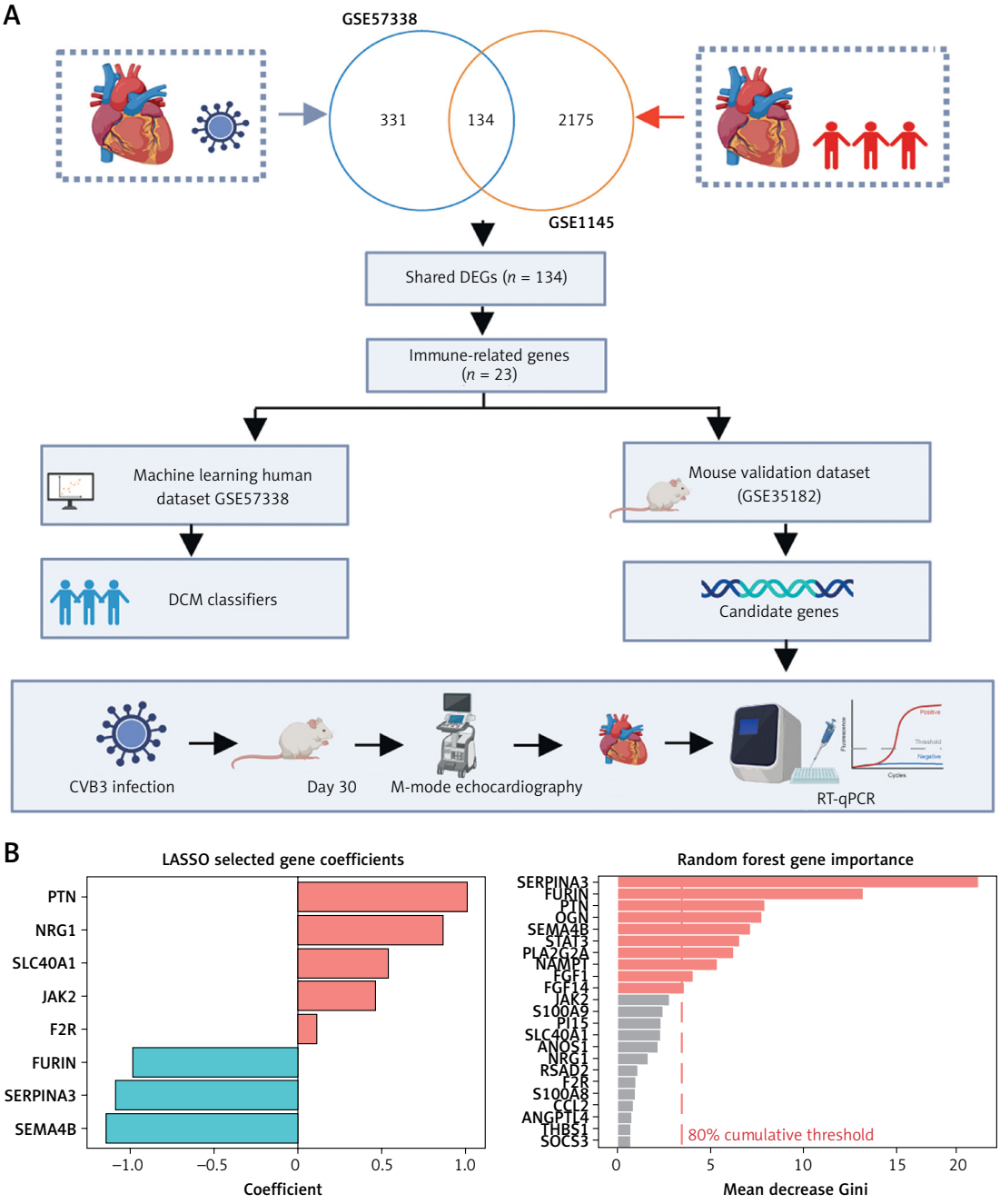


**Figure 1.** Machine-learning analysis and experimental validation of shared immune-related genes in viral myocarditis and dilated cardiomyopathy. **A** – Study design integrating the human DCM dataset GSE57338, human VMC dataset GSE1145, machine-learning analysis, validation in the longitudinal CVB3 mouse dataset GSE35182, and independent experimental validation. **B** – LASSO coefficient profiles and random-forest feature importance for immune-related genes in GSE57338. Data are presented as mean ± SD. Statistical significance is indicated as $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and $****p < 0.0001$; ns, not significant

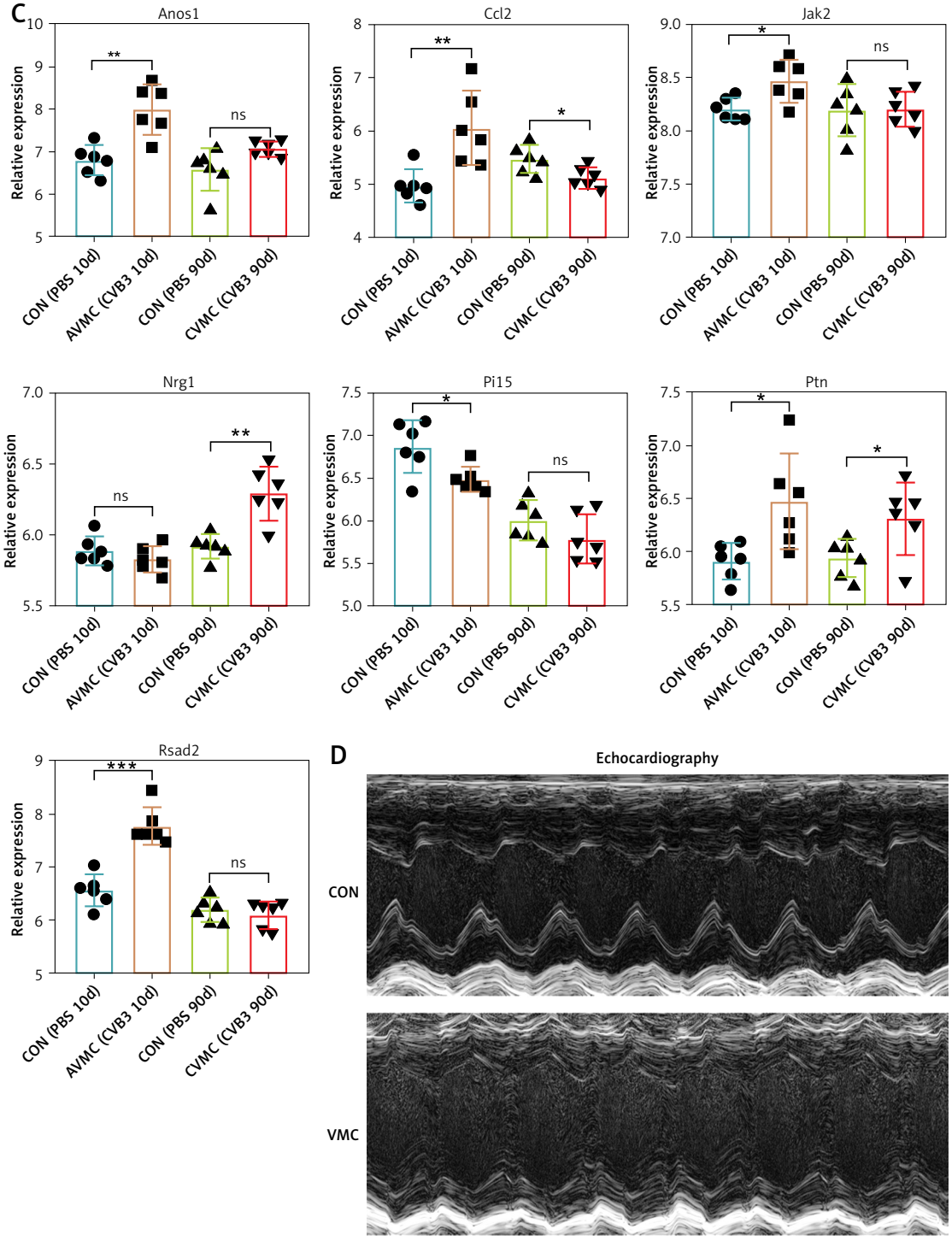


**Figure 1.** Cont. **C** – Expression of Anos1, Ccl2, Jak2, Nrg1, Pi15, Ptn, and Rsad2 in the longitudinal CVB3 mouse dataset GSE35182 during acute and chronic disease. **D** – Representative M-mode echocardiographic images of control and CVB3-infected mice at day 30. Data are presented as mean ± SD. Statistical significance is indicated as *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$; ns, not significant

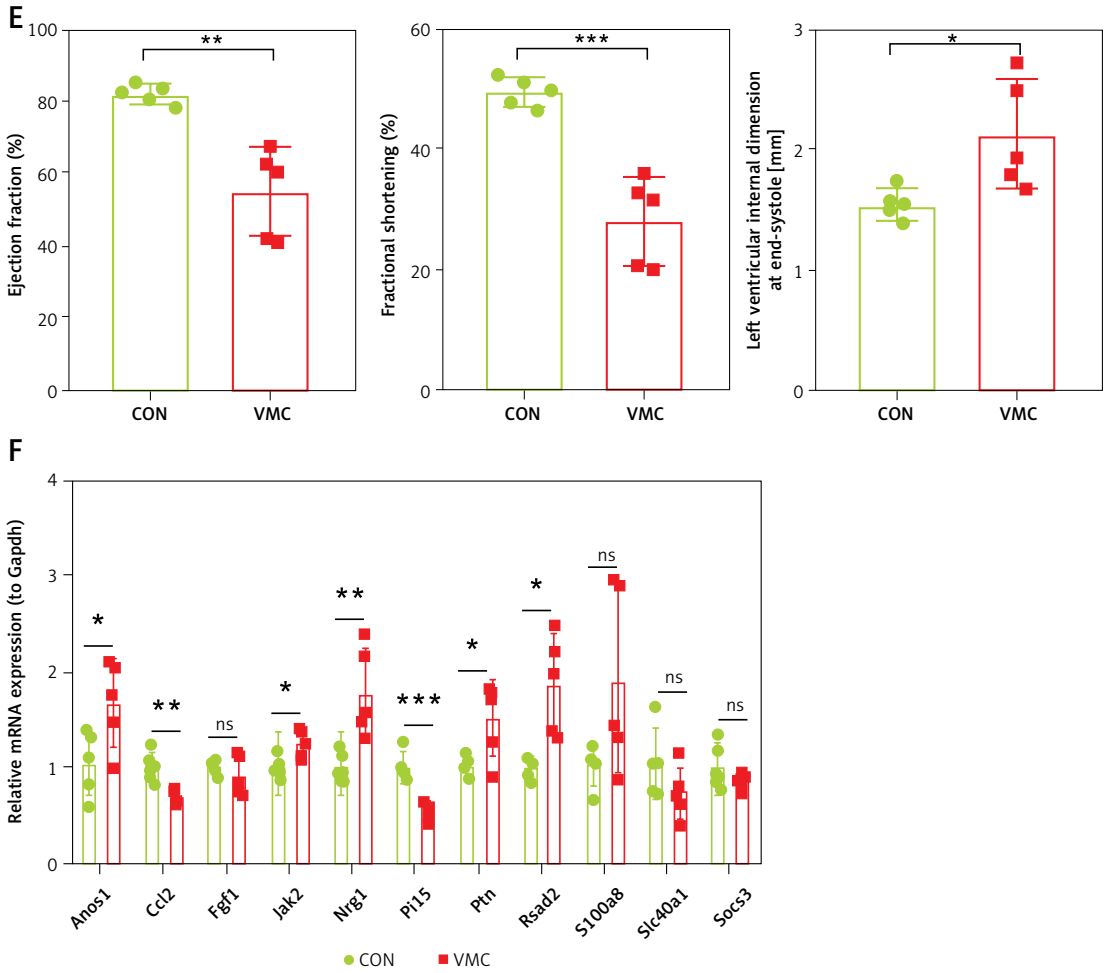


**Figure 1.** Cont. **E** – Echocardiographic assessment of cardiac function in control and CVB3-infected mice at day 30, including ejection fraction, fractional shortening, and left ventricular internal dimension at end-systole. **F** – Myocardial RT-qPCR validation of Anos1, Ccl2, Jak2, Nrg1, Pi15, Ptn, and Rsad2 in the independent CVB3-induced VMC mouse model. Data are presented as mean ± SD. Statistical significance is indicated as *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$; ns, not significant

activated in DCM. The protein-protein interaction network contained 85 nodes and 163 edges, with JAK2, STAT3, CCL2, and extracellular-matrix genes among the hubs. Of 23 shared immune-related genes, 10 were upregulated and 13 were downregulated in both diseases. ANOS1 correlated positively with OGN and negatively with FURIN in both datasets. Together, these findings placed interferon and JAK/STAT signals within a broader pattern of inflammation and matrix remodelling. In DCM, 242 genes were upregulated and 223 were downregulated; the corresponding counts in VMC were 1164 and 1145. Hierarchical clustering of the leading differentially expressed genes separated disease from control samples in both cohorts, indicating that the shared set was embedded in a wider disease-associated transcriptional pattern.

Intersection with the ImmPort database identified 23 immune-related genes that were concordantly dysregulated in VMC and DCM. LASSO retained eight genes. FURIN, SEMA4B, and SERPINA3 were selected by all three machine-learning methods (Figure 1 B). These three genes were regarded as classifiers of human DCM but were not significantly differentially expressed in the mouse dataset GSE35182. Separately, human-to-mouse expression validation supported seven genes – ANOS1, CCL2, JAK2, NRG1, PI15, PTN, and RSAD2 – whose expression changes in GSE35182 were concordant with those in the human datasets. The same directional changes were subsequently examined in the independent CVB3-induced mouse model. In GSE57338, the exploratory areas under the curve for these seven genes ranged from 0.741 to 0.909, with PTN showing the highest value. Because these estimates were obtained from the training dataset, they were not considered evidence of external diagnostic validation.

Immune deconvolution suggested changes in both innate and adaptive cell populations in DCM, whereas the main differences in VMC involved

macrophages, natural killer cells, and dendritic cells. These estimates represent relative cell proportions rather than direct measurements of cell abundance. In the GSE35182 time course, Jak2, Anos1, and Rsad2 increased mainly during acute VMC, whereas Pi15 decreased. Ptn remained increased during both the acute and chronic stages, whereas Nrg1 increased during the chronic stage. Ccl2 increased acutely but decreased at the chronic time point (Figure 1 C). Ccl2 expression correlated positively with Nppb, Il6, and Tnf, whereas Nrg1 correlated inversely with Col3a1. These exploratory findings indicated that the candidate genes were regulated differently during the acute-to-chronic transition.

In the independent mouse model, CVB3 infection produced ventricular dilation and systolic dysfunction, with reduced ejection fraction and fractional shortening and increased left ventricular dimensions (Figures 1 D, E). Myocardial expression of Nppb, Col3a1, Il6, Il1b, and Tnf was also increased, indicating inflammatory and fibrotic myocardial injury. Anos1, Jak2, Nrg1, Ptn, and Rsad2 were upregulated, whereas Ccl2 and Pi15 were downregulated (Figure 1 F). These changes reproduced the expression directions observed in the human datasets and provided experimental support for the human-to-mouse validation findings.

The drug-gene network contained 18 reported interactions involving the seven candidate genes. JAK2 had the highest degree and betweenness centrality and was linked to five inhibitors, including baricitinib, fedratinib, and ruxolitinib. CCL2 and NRG1 each had three reported interactions, whereas the remaining genes had one or two. These database-derived associations do not demonstrate treatment efficacy or support immediate drug repurposing. Nevertheless, JAK2 warrants further mechanistic investigation because JAK/STAT signalling contributes to antiviral defence but may also aggravate myocardial injury when it is excessive or prolonged [7, 8, 10]. CCL2 and PTN have previously been associated with inflammatory-cell recruitment and fibrotic remodelling, whereas NRG1 may have a compensatory cardioprotective role [11–13] (Supplementary Files S1, S2).

**Discussion.** The machine-learning analysis and the human-to-mouse validation addressed different questions. FURIN, SEMA4B, and SERPINA3 were selected as classifiers of human DCM but were not reproduced in GSE35182. The seven genes examined in the experimental model were instead selected because their expression directions were consistent in the human and mouse datasets. The two groups should therefore not be interpreted interchangeably. JAK/STAT signalling contributes to antiviral defence, but excessive or prolonged activation may also aggravate myocardial injury [9, 13]. The CCL2 and PTN findings are consistent with reports linking monocyte recruitment and fibroblast-derived PTN to myocarditis-related fibrosis and pathological remodelling [14, 15], whereas NRG1 may have a compensatory cardioprotective role [10]. JAK2 therefore warrants mechanistic investigation, but neither its network centrality nor its expression pattern establishes it as a therapeutic target.

Several limitations should be considered. GSE1145 was small and did not specify the causative virus. The human datasets were cross-sectional and therefore cannot demonstrate progression from VMC to DCM. Differences in microarray platform, species, sampling time, and clinical characteristics may also have affected the comparisons. Feature selection and diagnostic performance were assessed using retrospective public data, while the drug-gene network represents reported associations rather than treatment effects. The findings therefore extend the existing VMC-DCM transcriptomic framework through machine learning and independent *in vivo* validation, but they do not establish a diagnostic panel or support JAK2-directed treatment. Prospective longitudinal studies, protein-level confirmation, and cell-specific experiments are required to determine whether these genes contribute directly to disease progression.

## Acknowledgments

The authors thank the investigators who made the GEO datasets publicly available.

Beibei Du, Rong Miao and Xiaojie Ren contributed equally.

## Data availability

The datasets analysed in this study are available in the Gene Expression Omnibus under accession numbers GSE57338, GSE1145, and GSE35182.

## Funding

This work was supported by the Fujian Science and Technology Innovation Joint Funding Project Plan (2024Y9010), the Fujian Provincial Natural Science Funding Project (2025J01068), the Medical Innovation Project of Fujian Province Health Science and Technology Plan in 2025 (2025CXA003), and the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0527100).

## Ethical approval

Approval number: IACUC approval number: [BCJG0171].

## Conflict of interest

The authors declare no conflict of interest.

## References


1. Luo H, Wong J, Wong B. Protein degradation systems in viral myocarditis leading to dilated cardiomyopathy. Cardiovasc Res 2010; 85: 347-56.
2. Ammirati E, Moslehi JJ. Diagnosis and treatment of acute myocarditis: a review. JAMA 2023; 329: 1098-113.
3. Tschöpe C, Ammirati E, Bozkurt B, et al. Myocarditis and inflammatory cardiomyopathy: current evidence and future directions. Nat Rev Cardiol 2021; 18: 169-93.
4. Gigli M, Stolfo D, Merlo M, et al. Pathophysiology of dilated cardiomyopathy: from mechanisms to precision medicine. Nat Rev Cardiol 2025; 22: 183-98.
5. Harding D, Chong M, Lahoti N, et al. Dilated cardiomyopathy and chronic cardiac inflammation: pathogenesis, diagnosis and therapy. J Intern Med 2023; 293: 23-47.
6. Lei Z, Cao J, Wu J, et al. Identification of the communal pathogenesis and immune landscape between viral myocarditis and dilated cardiomyopathy. ESC Heart Fail 2024; 11: 282-92.
7. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.
8. Onyimba JA, Coronado MJ, Garton AE, et al. The innate immune response to coxsackievirus B3 predicts progression to cardiovascular disease and heart failure in male mice. Biol Sex Differ 2011; 2: 2.
9. Xiong C, Li B, Song R, et al. IFITM3 mediates inflammation-induced myocardial injury through JAK2/STAT3 signaling pathway. Mol Immunol 2024; 167: 1-15.
10. Lemmens K, Doggen K, De Keulenaer GW. Role of neuregulin-1/ErbB signaling in cardiovascular physiology and disease: implications for therapy of heart failure. Circulation 2007; 116: 954-60.
11. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015; 12: 453-7.
12. Hao J, Li W, Jiao W, et al. IL-17 affects the immune regulation of CD4+ T cells in dilated cardiomyopathy through JAK/STAT pathway. Cardiol Young 2025; 35: 16-23.
13. Yasukawa H, Yajima T, Duplain H, et al. The suppressor of cytokine signaling-1 is a novel therapeutic target for enterovirus-induced cardiac injury. J Clin Invest 2003; 111: 469-78.
14. Luo Y, Zhang H, Yu J, et al. Stem cell factor/mast cell/CCL2/monocyte/macrophage axis promotes coxsackievirus B3 myocarditis and cardiac fibrosis by increasing Ly6C-high monocyte influx and fibrogenic mediator production. Immunology 2022; 167: 590-605.
15. Sheng K, Ran Y, Feng X, et al. PTN secreted by cardiac fibroblasts promotes myocardial fibrosis and inflammation of pressure overload-induced hypertrophic cardiomyopathy through the PTN-SDC4 pathway. Life Sci 2025; 363: 123356.