

Identification MNS1, FRZB, OGN, LUM, SERP1NA3 and FCN3 as the potential immune-related key genes involved in ischaemic cardiomyopathy by random forest and nomogram

Peng-Fei Zheng^{1,2,3,*}, Fen Liu^{3,4,*}, Zhao-Fen Zheng^{1,2,3}, Hong-Wei Pan^{1,2,3}, Zheng-Yu Liu^{1,2,3}

¹Cardiology Department, Hunan Provincial People's Hospital, Furong, Changsha 410000, Hunan, China

²Clinical Research Center for Heart Failure in Hunan Province, Furong, Changsha 410000, Hunan, China

³Institute of Cardiovascular Epidemiology, Hunan Provincial People's Hospital, Furong, Changsha 410000, Hunan, China

⁴The First Affiliated Hospital of Hunan Normal University (Hunan Provincial People's Hospital), Furong, Changsha 410000, Hunan, China

*Equal contribution

Correspondence to: Hong-Wei Pan, Zheng-Yu Liu; **email:** panhongwei@hunnu.edu.cn, liuzhengyu@hunnu.edu.cn

Keywords: ischaemic cardiomyopathy, random forest, immune cell infiltrates, bioinformatics analysis, immune-related genes

Received: November 27, 2022

Accepted: February 11, 2023

Published: February 27, 2023

Copyright: © 2023 Zheng et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

The immune molecular mechanisms involved in ischaemic cardiomyopathy (ICM) have not been fully elucidated. The current study aimed to elucidate the immune cell infiltration pattern of the ICM and identify key immune-related genes that participate in the pathologic process of the ICM. The differentially expressed genes (DEGs) were identified from two datasets (GSE42955 combined with GSE57338) and the top 8 key DEGs related to ICM were screened using random forest and used to construct the nomogram model. Moreover, the "CIBERSORT" software package was used to determine the proportion of infiltrating immune cells in the ICM. A total of 39 DEGs (18 upregulated and 21 downregulated) were identified in the current study. Four upregulated DEGs, including *MNS1*, *FRZB*, *OGN*, and *LUM*, and four downregulated DEGs, *SERP1NA3*, *RNASE2*, *FCN3* and *SLCO4A1*, were identified by the random forest model. The nomogram constructed based on the above 8 key genes suggested a diagnostic value of up to 99% to distinguish the ICM from healthy participants. Meanwhile, most of the key DEGs presented prominent interactions with immune cell infiltrates. The RT-qPCR results suggested that the expression levels of *MNS1*, *FRZB*, *OGN*, *LUM*, *SERP1NA3* and *FCN3* between the ICM and control groups were consistent with the bioinformatic analysis results. These results suggested that immune cell infiltration plays a critical role in the occurrence and progression of ICM. Several key immune-related genes, including the *MNS1*, *FRZB*, *OGN*, *LUM*, *SERP1NA3* and *FCN3* genes, are expected to be reliable serum markers for the diagnosis of ICM and potential molecular targets for ICM immunotherapy.

INTRODUCTION

Ischaemic cardiomyopathy (ICM) refers to a disease in which the heart cannot pump blood normally due to long-term ischaemia and hypoxia damage to myocardial cells caused by severe coronary artery stenosis [1]. According to the World Health Organization, ICM is

one of the leading causes of death worldwide. Moreover, heart failure (HF) secondary to ICM, in contrast to nonischaemic aetiologies, has been demonstrated to be independently associated with mortality [2, 3]. Due to the high hospitalization and mortality rates caused by coronary artery disease (CAD) and HF, prevention and timely therapy for ICM are especially crucial [4].

Presently, standardized drug therapy aimed at relieving symptoms and improving the prognosis is still the cornerstone of ICM treatment [5, 6]. However, even after complete revascularization for severe coronary artery stenosis and long-term standardized use of secondary prevention drugs for CAD, there are still some patients with a poor prognosis and deteriorating cardiac function, and immunologic and inflammatory responses may be partly responsible for the poor prognosis of these patients. It was reported that the pathophysiology of ICM involves a spectrum of metabolic, neurohumoral and inflammatory changes [6]. When damaged and necrotic cardiomyocytes die due to ischaemic shock to the myocardium, immune and nonimmune cells are activated [7]. Recently, Bansal et al. suggested that proinflammatory and antiangiogenic regulatory T-lymphocytes (Tregs) promote immune activation and pathological left ventricular remodelling in ischaemic heart failure [8]. Moreover, a previous study demonstrated that monocytes play a pathological role in mediating left ventricular remodelling, interstitial fibrosis, and progressive cardiac insufficiency [9]. Nehra et al. showed that CD8+ T cells were increased at eight weeks after myocardial infarction during ICM [10]. These evidences indicate that immune cells play a crucial role in the pathological process of ICM. Nevertheless, the pattern of immune cell infiltration in the ICM has not been fully elucidated. Therefore, further elucidating the pattern of immune cell infiltration in the ICM and identifying the key immune-related DEGs involved in the ICM will hopefully provide a new molecular target for ICM immunotherapy.

In recent years, with the rapid development of gene chip sequencing technology, microarray analysis has become a novel and practical method to screen susceptible genes involved in cardiovascular diseases [11, 12], which lays a solid foundation for establishing a new gene-based diagnostic model related to cardiovascular diseases. Random forest is a classification algorithm that uses a set of selected classification trees to perform gene selection and microarray data classification [13], which shows excellent performance even when most predictive variables are disordered [14, 15]. Importantly, the prediction error could be reduced by using the random forest algorithm on a selected subset of genes compared with the existing methods and other proposed methods [13, 16]. Moreover, compared with other methods, the random forest method is more effective in identifying key genes significantly associated with disease. [17]. Recently, as a widely applied analytical method, CIBERSORT has often been used to study the immune cell infiltration pattern in diseases based on RNA sequencing or microarray data and to evaluate the infiltrated proportion of 22 immune cells in each sample [18]. However, studies using a combination of random

forest and CIBERSORT to verify immune-related genes associated with ICM are very limited. Hence, in the current study, CIBERSORT was used to evaluate the immune cell infiltration patterns in ICM, and the key DEGs were identified using random forest and then we further analysed the correlation between the key DEGs and immune cells. Finally, we verified the expression levels of the screened key DEGs and their diagnostic efficiency in the testing sets and the collected clinical samples.

RESULTS

Identification of DEGs

The specific workflow is shown in Figure 1. After data normalization and removal of the batch differences, a total of 39 DEGs, including 18 upregulated and 21 downregulated DEGs, were identified (Supplementary Table 1) and could be visualized in the volcano plot and heatmap (Figure 2A, 2B). In addition, the specific expressions of 39 DEGs in the training set are also shown in the Supplementary Table 2.

Functional analysis

As shown in Figure 3A, the results of the GO analysis include biological processes, cellular components and molecular function. The top three enriched GO terms were muscle hypertrophy in response to stress, cardiac muscle adaptation, and cardiac muscle hypertrophy in response to stress (biological processes); collagen-containing extracellular matrix, vacuolar lumen, and I band (cellular components); extracellular matrix structural constituent, extracellular matrix structural constituent conferring, and compression resistance (molecular function). The KEGG enrichment analyses suggested that DEGs were involved in the following pathways, including the complement and coagulation cascades, the HIF-1 signalling pathway, the cGMP-PKG signalling pathway, and the Wnt signalling pathway (Figure 3B). The DO analysis also suggested that the DEGs were mainly concentrated in some cardiovascular system diseases such as cardiomyopathy, arteriosclerosis, arteriosclerosis cardiovascular disease, intrinsic cardiomyopathy, heart septal defect, and dilated cardiomyopathy (Figure 3C). In addition, the detailed results of the KEGG, DO and GO enrichment analyses are presented in the Supplementary Tables 4, 5.

Screening the key DEGs with the random forest classification model

The out-of-bag classification error was relatively stable when the decision tree was set at 450 in the random forest classification model (Figure 4A). A total of 15

DEGs were identified as potential diagnostic markers of ICM using the random forest algorithm (Mean Decrease Gini >2) and the top 8 DEGs (*MNS1*, *FRZB*, *OGN*, *LUM*, *SERPINA3*, *RNASE2*, *FCN3*, and *SLCO4A1*) were selected as key genes for further analysis (Figure 4B). In addition, the specific importance of 39 DEGs in the training set are also shown in the Supplementary Table 3. As shown in Figure 4C, the 8 key DEGs might be useful to distinguish between the ICM and control samples. Among them, *MNS1*, *FRZB*, *OGN*, and *LUM* were upregulated while *SERPINA3*, *RNASE2*, *FCN3*, and *SLCO4A1* were downregulated in the control samples compared with the ICM samples.

ICM nomogram construction and verification

The nomogram (Figure 5A) was constructed from the results of the multiple logistic regression to intuitively

understand the relationship between 8 key genes and the prognosis of ICM. The calibration curves were plotted based on the expression levels of the 8 key genes in the training (Figure 5B) and validation (Figure 5C) sets, respectively, to verify the accuracy of the nomogram prediction. The predicted curves were almost the ideal curve, which suggested good performance. The DCA confirmed the improved clinical utility of the nomogram in predicting the morbidity of ICM patients in the training (Figure 5D) and validation (Figure 5E) sets. Moreover, the clinical impact curves also demonstrated that the nomogram model had significant predictive ability and good clinical utility both in the training (Figure 5F) and validation (Figure 5G) sets. In addition, the area under the curve (AUC) was 0.992 (95% confidence interval (CI) 0.984-1.000) in the training (Figure 5H) and 0.995 (95% CI 0.985-1.000) in the validation (Figure 5I) sets.

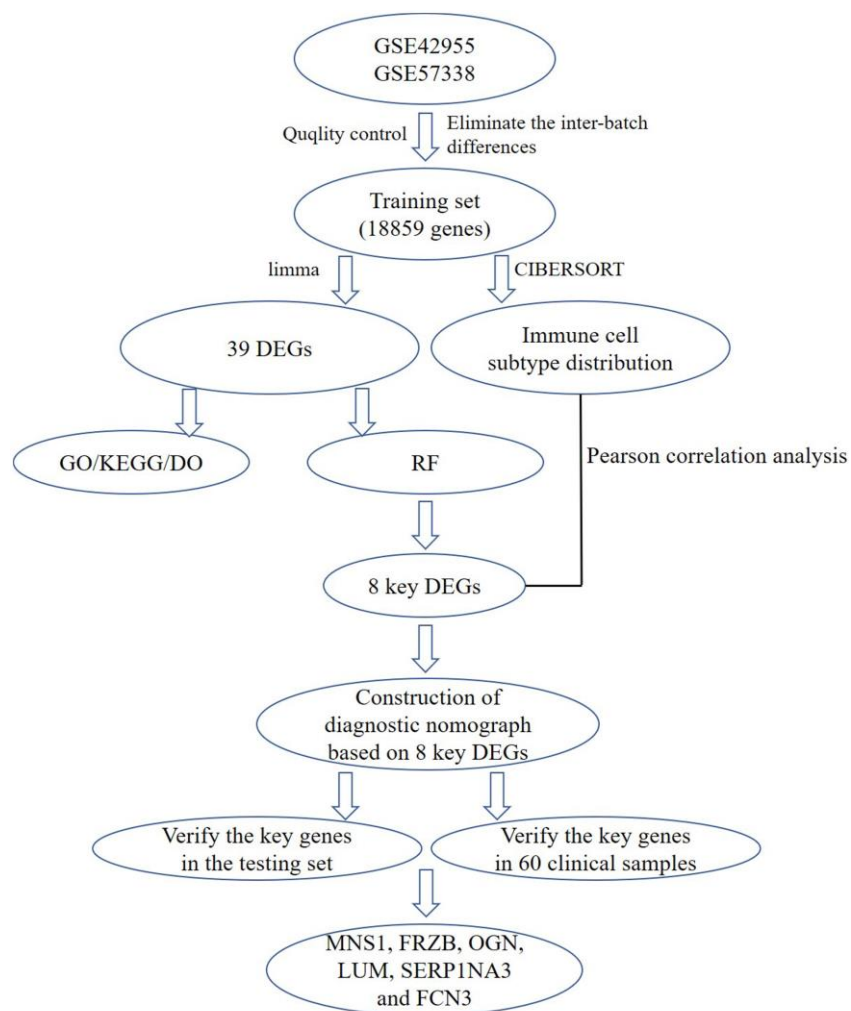


Figure 1. A flow chart for analysis. DEGs, differentially expressed genes; GO, gene ontology annotation; KEGG, kyoto encyclopedia of genes and genomes pathway enrichment analyses; DO, disease ontology analysis; RF, random forest; MNS1, meiosis-specific nuclear structural 1; FRZB, frizzled-related protein; OGN, osteoglycin; LUM: lumican; SERPINA3: serpin family A member 3; FCN3: ficolin-3.

Immune cell distribution pattern

The immune fractions expressed differentially between the control and ICM samples were evaluated using the CIBERSORT algorithm. The histogram intuitively shows that the total proportion of 22 different immune cell subtypes in 246 ICM samples was 1 (Supplementary Figure 3). The heatmap (Figure 6A) suggested that the proportions of several immune cells in both the ICM and control samples were obviously different. Based on the correlation matrix, we found that

CD4 memory resting T cells were negatively correlated with Tregs and CD8 T cells but positively correlated with plasma cells; Tregs were positively related to CD8 T cells and naive B cells but negatively related to plasma cells (Figure 6B). As shown in Figure 6C, naive CD4 T cells, M0 macrophages and resting mast cells were increased, while Tregs and monocytes were decreased in ICM samples compared with control samples ($p < 0.05$). Furthermore, Supplementary Table 6 shows the details of the immune cell infiltration pattern between the ICM and control samples.

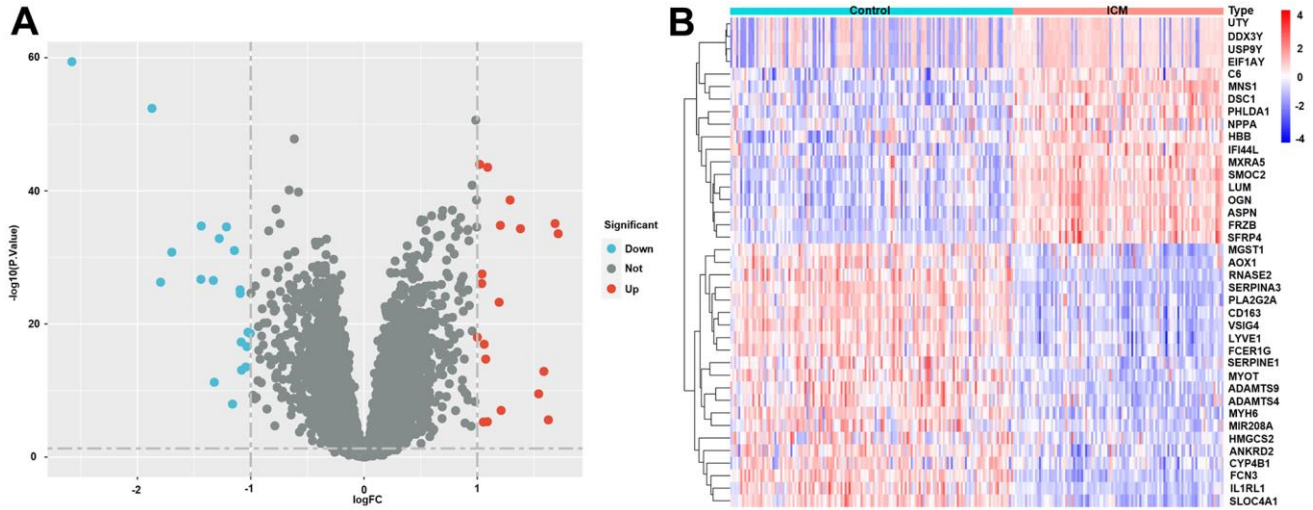


Figure 2. Differentially expressed genes (DEGs) in ICM and healthy samples. (A) Volcano plot of the 39 DEGs. Red dots represent significantly upregulated genes, and blue dots represent significantly downregulated genes. (B) Heatmap of the 39 lipid-related DEGs in ICM and control samples. Red blocks indicate high-expression genes, and blue blocks indicate low-expression genes.

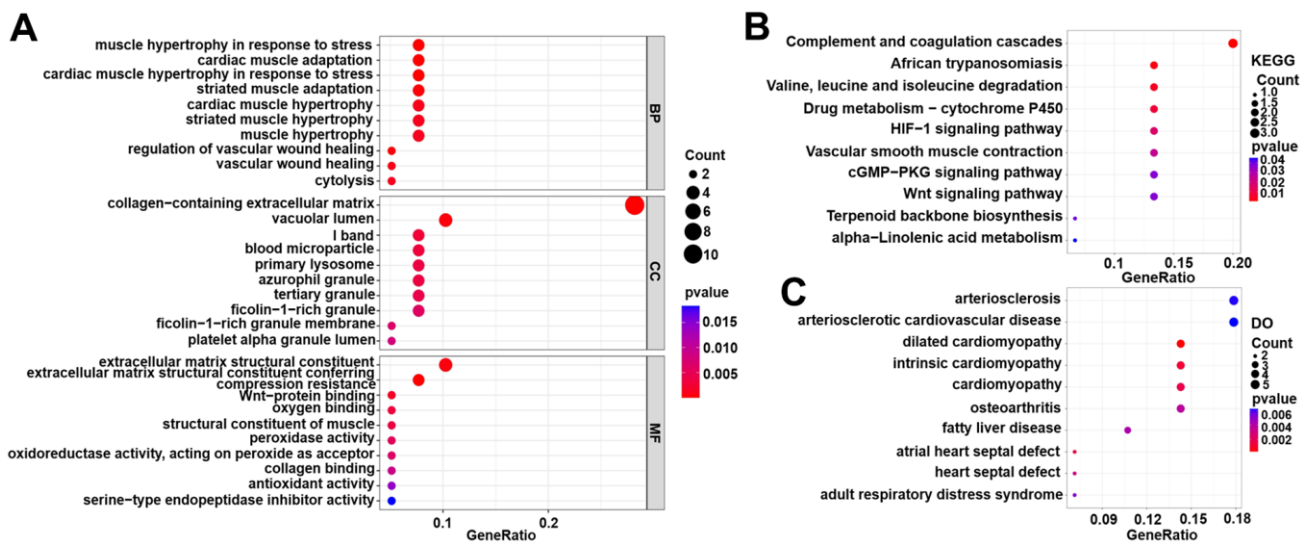


Figure 3. Functional enrichment analyses for DEG genes in the training set. (A) Gene Ontology (GO) enrichment analysis. (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. (C) Disease Ontology (DO) analysis.

The correlations between the 8 key DEGs and infiltrated immune cells are shown in Figure 7. The *SERPINA3* gene was negatively correlated with resting mast cells, activated dendritic cells, and activated NK cells but positively correlated with monocytes, M2 macrophages, eosinophils, and neutrophils; the *SLCO4A1* gene was negatively correlated with resting mast cells and positively correlated with neutrophils, monocytes, and eosinophils; the *FCN3* gene was positively correlated with monocytes and resting memory CD4 T cells but negatively correlated with M0 macrophages, CD8 T cells and naive CD4 T cells; the *FRZB* was positively correlated with resting mast cells and activated dendritic cells but negatively correlated with monocytes; the *LUM* gene was positively correlated with M1 macrophages and resting mast cells but negatively correlated with monocytes and resting NK cells; the *MNS1* gene was positively correlated with resting mast cells but negatively correlated with neutrophils and monocytes;

the *OGN* gene was positively correlated with M1 macrophages and resting mast cells but negatively correlated with monocytes; and the *RNASE2* gene was positively correlated with monocytes, neutrophils, eosinophils and M2 macrophages but negatively correlated with activated NK cells, M0 macrophages, and resting mast cells ($p < 0.05$ – 0.01 , respectively).

External validation of the top 8 key DEGs in the testing set

The expression levels of these 8 key DEGs are shown in Supplementary Figure 4. The expression levels of *MNS1* ($p = 0.00033$), *FRZB* ($p = 6.3e-09$), *OGN* ($p = 2.9e-08$), and *LUM* ($p = 2.3e-10$) were increased in ICM compared to the control samples, while *SERPINA3* ($p = 7.2e-09$), *RNASE2* ($p = 0.0043$), *FCN3* ($p = 2.9e-12$), and *SLCO4A1* ($p = 7.2e-09$) were decreased in the ICM samples.

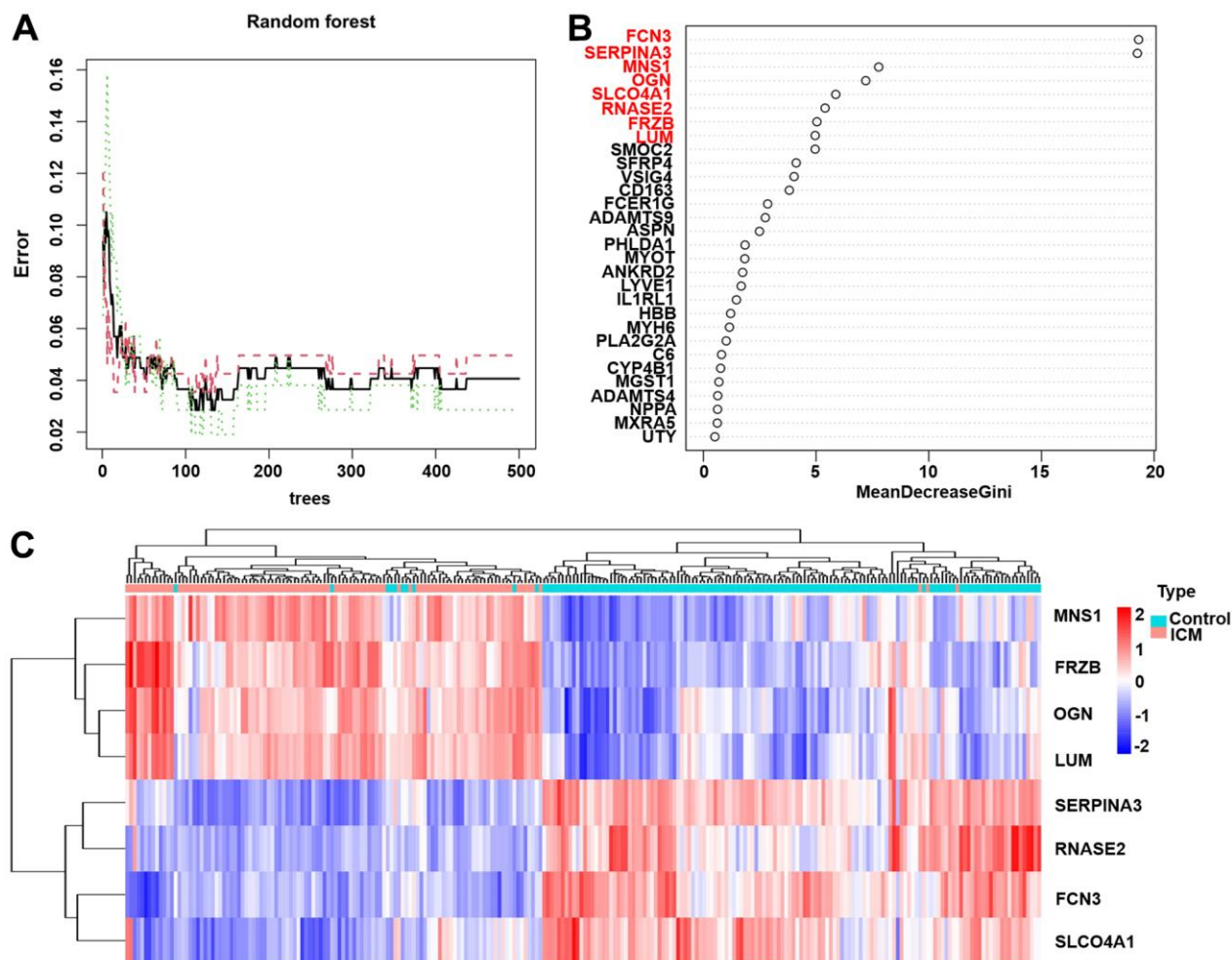


Figure 4. The results of the top 8 DEGs screened by random forest. (A) The plot of performance in log scale against epoch number. (B) The importance of the top 8 DEGs ranked by the mean decrease in accuracy. (C) Heatmap of the top 8 genes. Red blocks indicate high-expression genes, and blue blocks indicate low-expression genes.

Validation by RT-qPCR

The expression levels of *MNS1* ($p = 2e-05$), *FRZB* ($p = 2.3e-05$), *OGN* ($p = 0.0001$), and *LUM* ($p = 9.5e-06$) were increased in ICM compared to the control group, and *SERPNA3* ($p = 4.3e-05$) and *FCN3* ($p = 3.7e-07$) were higher in the control group (Figure 8). Nevertheless, the expression levels of *RNASE2* ($p = 0.49$) and *SLCO4A1* ($p = 0.23$) were similar between the two groups. As shown in Figure 9, the AUC values of *MNS1*, *FRZB*, *OGN*, *LUM*, *SERPINA3*, and *FCN3* were 0.821 (95% CI: 0.718–0.923), 0.818 (95% CI: 0.711–0.925), 0.792 (95% CI: 0.669–0.915), 0.833 (95% CI: 0.733–0.934), 0.808 (95% CI: 0.701–0.914), and 0.882 (95% CI: 0.790–0.974), respectively.

DISCUSSION

Ischaemic cardiomyopathy (ICM) is a chronic immune system activation state, and immune or inflammatory

mechanisms play crucial roles in its occurrence and development [8]. After ischaemic injury of myocardial tissue, the inflammatory immune response will activate the repair process, which is manifested as the removal of necrotic tissue debris, the formation of stable scars, and the initiation of neovascularization in the myocardium to achieve effective wound healing. However, incomplete immune resolution or a persistent low-grade inflammatory response can lead to poor remodelling of myocardial tissue and left ventricular dilatation, ultimately leading to ICM [10]. Scientists have observed that the innate inflammatory cascade is primarily mediated by neutrophils, monocytes, and macrophages, and it contributes to chronic inflammatory processes [19]. Notably, monocytes and macrophages have been found to play complex and diverse roles in myocardial injury [10]. He et al. showed that monocytes and reparative macrophages contribute to the removal of necrotic debris, granulation tissue formation, and angiogenesis in the early stage after myocardial infarction [20]. However,

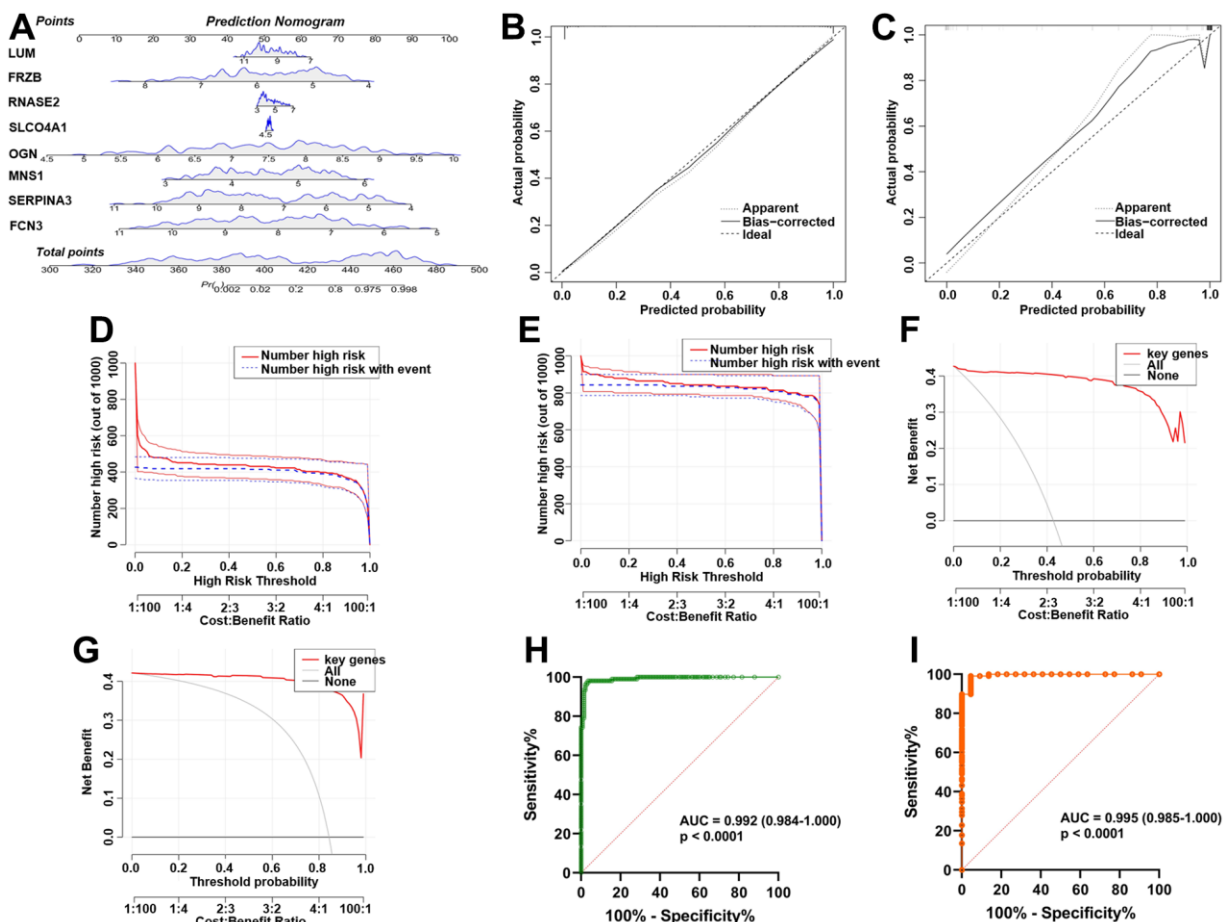


Figure 5. Establishment and verification of the ICM diagnostic signature and nomogram. (A) Eight predictive nomograms of genetic diagnosis. (B) Calibration curve in the training set. (C) Calibration curve in the testing set. (D) Decision curve analysis (DCA) in the training. (E) DCA in the testing set. (F) Clinical impact curve in the training set. (G) Clinical impact curve in the testing set. (H) ROC curve in the training set. (I) ROC curve in the testing set.

Ismahil et al. demonstrated that monocytes play a pathological role in mediating left ventricle remodelling, interstitial fibrosis and progressive cardiac dysfunction [9]. Depending on the stimulation, M0 macrophages polarize into M1 and M2 macrophages, which perform proinflammatory and anti-inflammatory functions, respectively [21]. In addition, other types of immune cells have also been found to play key roles in ventricular remodelling or heart failure. Sun et al. demonstrated that TNF- α secreted by mast cells could induce matrix metalloproteinase activation, leading to ventricular remodelling [22]. In heart failure,

chymotrypsin released by mast cells causes left ventricle dysfunction by promoting cardiomyocyte apoptosis and fibroblast proliferation [23]. CD4+ T cells impart proinflammatory and detrimental effects during chronic ischaemic heart failure and promote adverse ventricular remodelling [24, 25]. Tregs are the most critical immune cells found *in vivo* with powerful immune-suppressive potential [10]. A persuasive study indicated that Tregs are essential for modulating immune responses, promoting cardiac repair and initiating immune resolution [26]. It has been reported that adoptive transfer of Tregs after myocardial infarction can alleviate

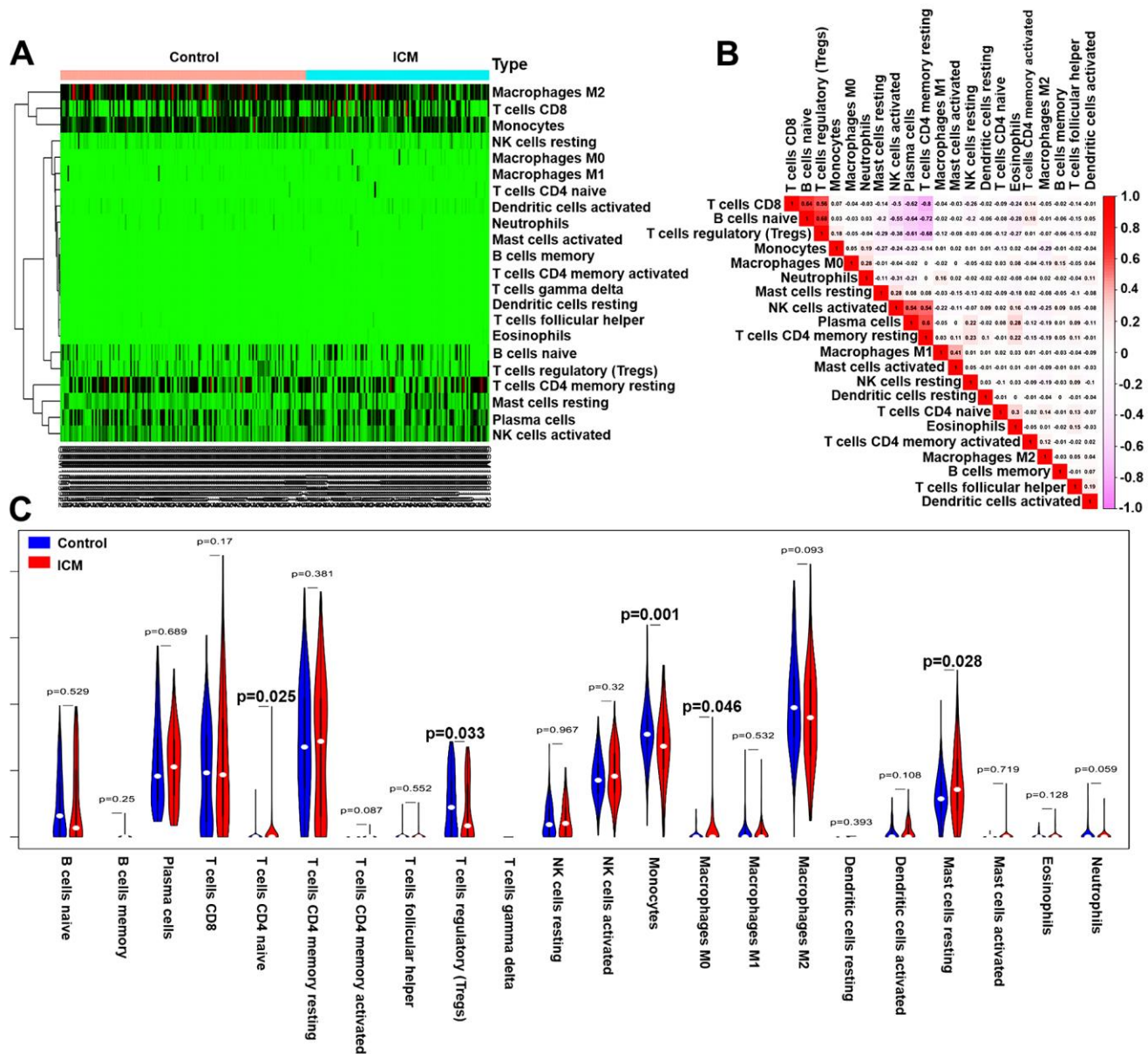


Figure 6. Pattern of immune cell subtype infiltration in the training set. (A) Heatmap of the 22 immune cell proportions in each sample. **(B)** Correlation heatmap of the 22 immune cells. The deeper the red, the stronger the positive correlation. The darker the pink, the stronger the negative correlation. **(C)** Representative violin plot showing the different fractions of infiltrated immune cells.

inflammation and promote cardiac repair. Importantly, partial depletion and successful reconstruction of Tregs resets their phenotypic and immunosuppressive potential, improving cardiac function, reducing left ventricle remodelling and partially reversing the

ICM [27]. These studies indicate that the immune microenvironment plays a complex and critical role in the occurrence and development of ICM. Nevertheless, the immune microenvironment of ICM has not been fully elucidated.

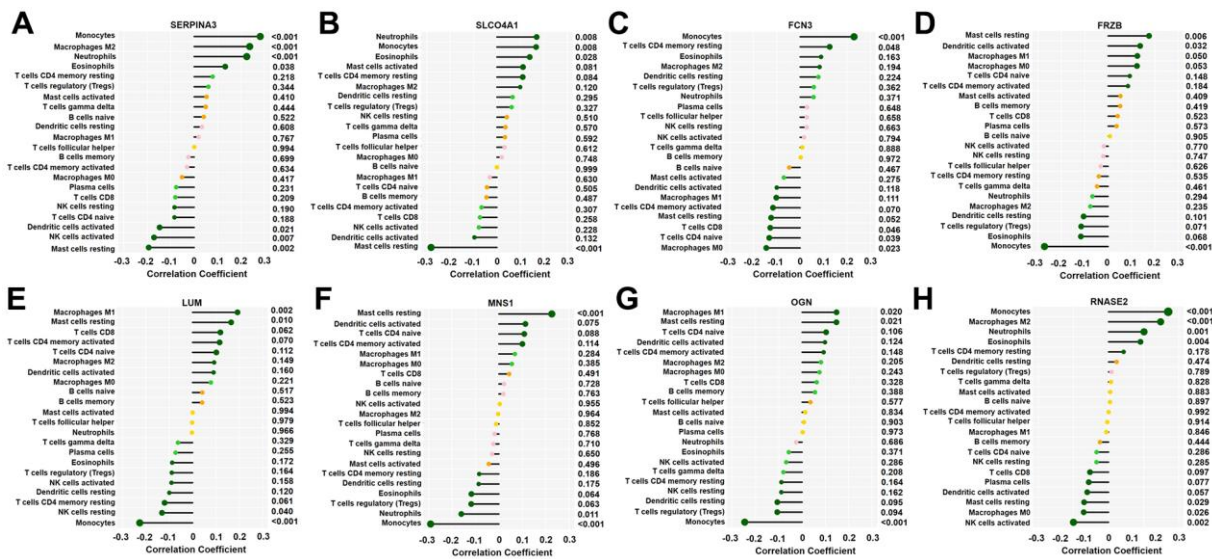


Figure 7. Correlation between 8 key genes and infiltrated immune cells. Representative graphs showing the correlation between the infiltrated immune cells and *SERPINA3* (A), *SLCO4A1* (B), *FCN3* (C), *FRZB* (D), *LUM* (E), *MNS1* (F), *OGN* (G) and *RNASE2* (H). The correlation strength is shown by the dot size; the *p* values are expressed as the changes in dot colour, and *p* < 0.05 was considered statistically significant. *SERPINA3*: serpin family A member 3; *SLCO4A1*, solute carrier organic anion transporter family member 4A1; *RNASE2*, ribonuclease A family member 2; *FCN3*, ficolin-3; *FRZB*, frizzled-related protein; *LUM*, lumican; *MNS1*, meiosis-specific nuclear structural 1; *OGN*, osteoglycin.

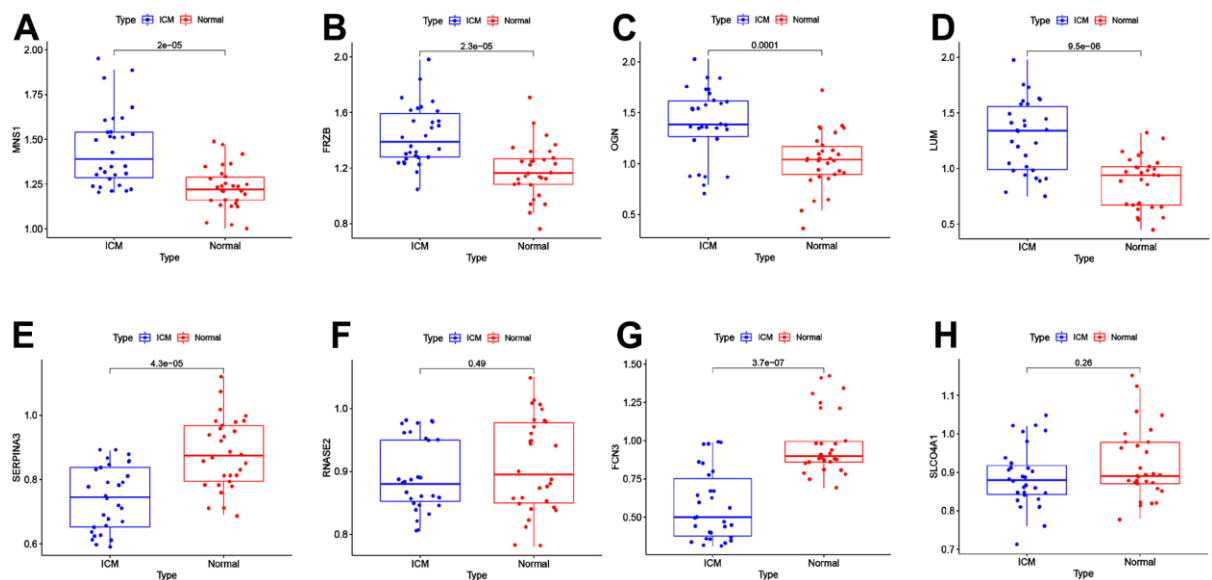


Figure 8. External validation of the key genes in the clinical samples. The expression levels of *MNS1* (A), *FRZB* (B), *OGN* (C), *LUM* (D), *SERPINA3* (E), *RNASE2* (F), *FCN3* (G) and *SLCO4A1* (H) in the clinical samples. *MNS1*, meiosis-specific nuclear structural 1; *FRZB*, frizzled-related protein; *OGN*, osteoglycin; *LUM*: lumican; *SERPINA3*: serpin family A member 3; *FCN3*: ficolin-3; *SLCO4A1*, solute carrier organic anion transporter family member 4A1; *RNASE2*, ribonuclease A family member 2.

In the current research, we found that Treg and monocyte infiltration were decreased and that naive CD4 T cells, resting mast cells and M0 macrophages were increased in the ICM compared with normal samples, which showed that Tregs and monocytes might play protective roles and that naive CD4 T cells, resting mast cells and M0 macrophages might play promoting roles in the pathological process of ICM. Moreover, several interactions among 22 types of infiltrating immune cells in the ICM were also noted. We found that CD4 memory resting T cells were negatively correlated with Tregs and CD8 T cells but positively correlated with plasma cells; Tregs were positively correlated with CD8 T cells and naive B cells but negatively correlated with plasma cells. However, our results have only partially elucidated the characteristics of the immune microenvironment in ICM, and the immune-related key molecular targets involved in ICM have not been fully identified. Therefore, identifying these molecular markers will provide a new perspective for the diagnosis and immunotherapy of ICM. In this study, a total of 8 key immune-related DEGs (*MNS1*, *FRZB*, *OGN*, *LUM*, *SERPINA3*, *RNASE2*, *FCN3*, and *SLCO4A1*) involved in ICM were screened by a random forest algorithm combined with a nomogram model. Meanwhile, a nomogram based on the above 8 key genes suggested a diagnostic value of up to 99% to distinguish ICM from healthy participants. In addition, the results of RT-qPCR and ROC curve

analysis showed that the expression levels of *MNS1*, *FRZB*, *OGN*, and *LUM* were increased in ICM patients compared with the control group, while the expression levels of *SERPINA3* and *FCN3* were decreased. Moreover, gene enrichment analysis showed that the potential molecular mechanism of these key genes involved in ICM was mainly related to inflammation or immunity, but the underlying molecular mechanisms of these genes might be slightly different.

Among the four upregulated genes in ICM, frizzled-related protein (*FRZB*) is a secreted protein that serves as a modulator of the Wnt signalling pathway by directly interacting with Wnt and it plays a key role in dorsoventral patterning of the mesoderm during vertebrate development [28]. It has been demonstrated that *FRZB* is a key molecule in the progression of abdominal aortic aneurysm [29], and it can reduce the growth and aggressiveness of fibrosarcoma cells [30]. Previous research has reported that *FRZB* is involved in congenital heart defects [31]. Yang et al. [32] suggested that *FRZB* could regulate dilated cardiomyopathy through the extracellular matrix signalling pathway. Meanwhile, Ma et al. [33] showed that *FRZB* was significantly upregulated in hypertrophic cardiomyopathy (HCM) and can serve as a biomarker and is a potential therapeutic target for HCM. This evidence strongly suggests that *FRZB* is significantly associated with

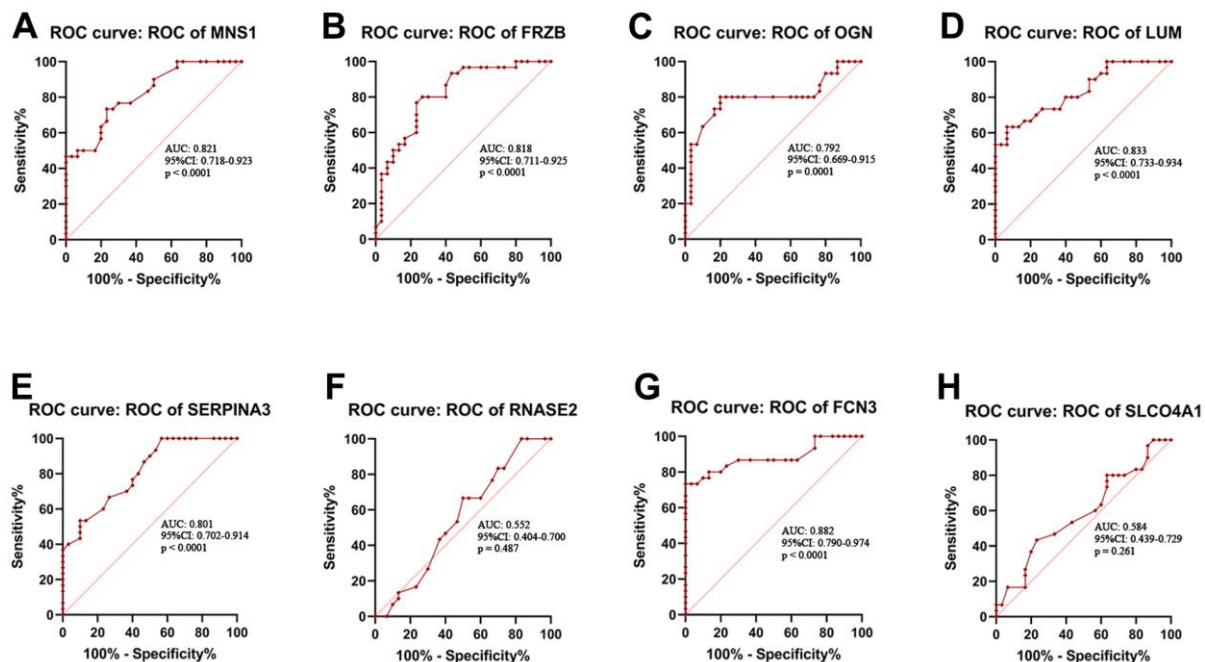


Figure 9. Receiver operating characteristic (ROC) curve analysis. ROC curve analysis of *MNS1* (A), *FRZB* (B), *OGN* (C), *LUM* (D), *SERPINA3* (E), *RNASE2* (F), *FCN3* (G) and *SLCO4A1* (H) in the clinical samples. *MNS1*, meiosis-specific nuclear structural 1; *FRZB*, frizzled-related protein; *OGN*, osteoglycin; *LUM*, lumican; *SERPINA3*: serpin family A member 3; *FCN3*: ficolin-3; *SLCO4A1*, solute carrier organic anion transporter family member 4A1; *RNASE2*, ribonuclease A family member 2.

cardiomyopathy development, but the relationship between *FRZB* and immune cells and ICM remains unclear. Our current study suggested that *FRZB* was upregulated in ICM samples; meanwhile, we also observed that was negatively related to monocytes but positively related to resting mast cells and activated dendritic cells. Nevertheless, further research is needed to confirm these discoveries.

Osteoglycin (*OGN*) is a protein that belongs to the small leucine-rich proteoglycan (SLRP) family, and these SLRPs play critical roles in shaping the organization and structure of the extracellular matrix in the heart and other organs [34]. In 2008, *OGN* was first discovered to be related to the ventricular mass [35], which laid a foundation for further research on the role of *OGN* in cardiovascular disease. A study showed that *OGN* expression was increased in patients with coronary heart disease compared with normal subjects and was related to the severity of the coronary lesions [36]. Rienks et al. demonstrated that 72 kDa chondroitin sulfate-*OGN* could aggravate cardiac inflammation by enhancing Toll-like receptor 4 activation [34]. It has been reported that the downregulation of *OGN* is beneficial for the formation of endothelial cells and it might become a novel treatment target for ischaemic vascular diseases [37]. In addition, *OGN* expression changes during the course of cytotoxicity events, and chemotactic stimulation was found in natural killer cells and neutrophils [38], but the relationship between *OGN* and immune cells and ICM susceptibility remains poorly understood. Herein, we noticed that the expression of *OGN* was significantly increased in ICM patients compared with healthy subjects and it was negatively related to monocytes but positively related to M1 macrophages and resting mast cells. Thus, we speculate that the upregulation of *OGN* expression might cause ICM by enhancing immune inflammation.

Lumican (*LUM*) belongs to the small leucine-rich repeat proteoglycan family [39, 40]. Previous studies reported that *LUM* could regulate the growth of cardiomyocytes by adjusting the pericellular extracellular matrix [41], and moderate *LUM* deficiency can weaken cardiac fibrosis and ameliorate diastolic dysfunction after excessive pressure load [42]. When heart failure occurs, fibroblasts produce large amounts of *LUM*, which is involved in cardiac remodelling processes induced by mechanical and proinflammatory stimulation [43]. Moreover, Zhang et al. [44] suggested that *LUM* may have a potentially good predictive ability for the diagnosis of dilated cardiomyopathy. The meiosis-specific nuclear structural 1 (*MNS1*) gene is involved in bile acid, fatty acid, and haem metabolism [45] and it affects the course of heart failure. A recent study showed that *MNS1* was screened through three machine

learning methods to be regarded as a possible bioinformatics marker for heart failure [46]. These results indicate that *LUM* and *MNS1* play vital roles in heart failure, but their association with the immune cells involved in ICM is still unclear. In the current research, we found that *LUM* and *MNS1* were highly expressed in ICM patients compared with healthy subjects. Meanwhile, we also noticed that *LUM* and *MNS1* were positively correlated with resting mast cells but negatively correlated with monocytes. These findings suggested that *LUM* and *MNS1* may participate in ICM by affecting resting mast cells and monocyte cells, but more studies are needed to verify our current findings.

Serpin family A member 3 (*SERPINA3*), an acute phase response gene, is upregulated during the process of inflammation [47] and has been demonstrated to promote myocardial ischaemia reperfusion injury [48]. Wågsäter et al. suggested that *SERPINA3* expression is significantly increased in human atherosclerotic lesions [49]. However, a recent study found that *SERPINA3* expression was decreased in heart failure patients compared with control subjects, indicating that *SERPINA3* played a protective role during the process of heart failure [7]. Meanwhile, Masanori et al. showed that *SERPINA3* might serve as a potential prognostic biomarker for heart failure [50]. Ficolin-3 (*FCN3*) is a recognition molecule in the lectin pathway of the complement system that is expressed in the liver and the lung [51]. Trine et al. demonstrated that *FCN3* is a serum protein that may be involved in the progression of systemic lupus erythematosus [52]. Shang et al. found that *FCN3* may activate the complement system and be overexpressed in type 2 diabetes plasma [53]. However, Song et al. showed that *FCN3* expression was significantly decreased in ICM patients compared with healthy subjects [54]. These studies indicate that *SERPINA3* and *FCN3* maintain different expression patterns in different stages of atherosclerosis and related diseases. Similarly, in the current study, the expression levels of *SERPINA3* and *FCN3* were significantly decreased in ICM patients compared with healthy subjects. Moreover, *SERPINA3* is positively related to neutrophils and monocytes but negatively related to activated NK cells; *FCN3* is positively related to monocytes but negatively related to M0 macrophages, CD8 T cells and naive CD4 T cells. Our findings indicated that *SERPINA3* and *FCN3* might play a protective role in the pathological process of ischaemic cardiomyopathy by regulating immune cell infiltration.

However, our current study also has some limitations. First, the included clinical samples were relatively small; therefore, our conclusions must be verified by a larger ICM cohort. Second, we did not detect the level of immune-related cytokines or chemokines in clinical

samples, so we could not further analyze the correlation between these key genes and cytokines and chemokines. Third, the key DEGs were only confirmed in clinical samples, and their potential roles were not demonstrated in ICM cells or animal models. Hence, more *in vivo* and *in vitro* studies are needed to clarify the association between the 6 key genes and the infiltrated immune cells and to understand the mechanism of these key genes during the pathological process of ICM.

In conclusion, we confirmed that *MNS1*, *FRZB*, *OGN*, *LUM*, *SERPINA3*, and *FCN3* are novel credible serum markers for the diagnosis of ICM using random forest combined with nomogram. We noticed that naive CD4 T cells, M0 macrophages and resting mast cells may be correlated with the occurrence and progression of ICM, while Tregs and monocytes may play a protective role in ICM. The interaction mechanisms between the above key DEGs and immune cells may be of great significance for the pathogenesis and progression of ICM.

MATERIALS AND METHODS

Data download and preprocessing

The Gene Expression Omnibus (GEO) website was used to download 4 ICM-related datasets (<https://www.ncbi.nlm.nih.gov/geo/>). The GSE42955 dataset, including 12 ICM and 5 control samples, was based on the GPL6244 platform (Affymetrix Human Gene 1.0 ST Array). GSE57338 included 95 ICM and 136 control samples and was retrieved from GPL11532 (Affymetrix Human Gene 1.1 ST Array). Based on the GPL96 platform (Affymetrix Human Genome U 133A Array), we also downloaded the GSE1869 dataset (including 10 ICM and 5 control samples) and the GSE5406 dataset (including 108 ICM and 16 controls) for subsequent validation. As shown in Supplementary Figure 1, the *normalize Between Arrays* function in the “limma” package was used to normalize samples, and the interbatch differences were eliminated using the “sva” package. Moreover, after removing two outlier samples (GSM1380018 and GSM1379815) (Supplementary Figure 2), the key genes involved in ICM were identified using a total of 246 samples in the training sets (GSE42955 combined with GSE57338). Meanwhile, after removing the interbatch differences between GSE1869 and GSE5406, the new integrated gene expression profiles were used as validation sets to verify the expression levels of the key genes.

Identification of DEGs

The DEGs between the ICM and control samples were identified using the “limma” package of R software [55].

The threshold values were $p < 0.05$ and $|\log \text{fold change (FC)}| > 1$. The “heatmap” and “ggplot2” R packages were used to draw heatmaps and volcano plots.

Functional enrichment analysis of DEGs

The “enrichplot” R package was used to conduct GO and KEGG pathway analyses [56]. Cell composition (CC), biological processes (BP), and molecular functions (MF) were included in the GO analysis. Then, DO analysis was performed using the “clusterProfiler” R package [56]. A p value < 0.05 was considered statistically significant.

Screening key DEGs

The “randomForest” R package was used to construct a random forest classification model to screen the key DEGs associated with ICM [57]. The Gini index was used as an importance measure [58], and the best number of trees in the random forest algorithm was set to 450. Since the expression of the ninth gene was not detected in the subsequent validation set, the ninth gene could not be further validated in the validation set. So that, the genes with importance values greater than 2 [57, 59] and ranking in the top eight were considered the key DEGs of ICM for subsequent model construction. The key genes screened in the training sets were reclassified into unsupervised hierarchical clusters using the “pheatmap” software package, and a heatmap was plotted.

Construction and verification of the ICM diagnostic signature and nomogram

A nomogram [60] was used to demonstrate the diagnostic efficacy of a diagnostic model based on eight key genes to distinguish ICM from healthy subjects. We visualized the ICM prediction model using the nomogram method and scored the expression values of the top 8 key genes. The scores of the 8 key genes were then summed to obtain the total score. The value of the risk of ICM can be determined by drawing a vertical line at the total score obtained. To quantify the nomogram prediction function, a calibration curve was plotted to evaluate the consistency between the predictions of the nomograms and the actual observations and to describe the calibration of the model according to the consistency between the predicted risk and the actual outcome in the training and testing sets. Decision curve analysis (DCA) was conducted to evaluate the net clinical benefit of the model for predicting ICM in the training and testing sets, and the model had diagnostic value if the drawn DCA curve was higher than the horizontal and dotted lines. Subsequently, a clinical impact curve was drawn to assess the clinical practicability and applicability net

benefits of the nomogram model in the training and testing sets. The “ROC” package was used to draw receiver operator characteristic (ROC) curves of the nomogram in the training and validation sets to evaluate the diagnostic efficacy of the nomogram.

Correlation between key genes and immune cells

The immune infiltration pattern in the ICM was determined using the “CIBERSORT” software package [18]. The “ggplot2” package in R software was used to draw the boxplot, heatmaps, and histograms after obtaining the immune cell expression matrix. The histogram suggests the percentage of 22 immune cells infiltrated in ICM samples, while the difference in immune cell infiltration between the ICM and control samples was displayed through boxplot and heatmap diagrams. The “corrplot” software package was used to calculate the correlation coefficient between the 8 key genes and each immune cell and visualize the results by the relevant heatmap.

Study population

A total of 60 participants were recruited from Hunan Provincial People’s Hospital, including 30 ICM patients and 30 healthy participants. The diagnostic criteria for ICM are as follows: a history of myocardial infarction, coronary revascularization (coronary artery bypass grafting or percutaneous coronary intervention), $\geq 75\%$ stenosis of the left main or proximal left anterior descending coronary artery, or $\geq 75\%$ stenosis of ≥ 2 epicardial vessels [61]. All enrolled ICM patients had a left ventricular ejection fraction of $\leq 40\%$ and a history of symptomatic heart failure (New York Heart Association [NYHA] functional class II or greater) [62]. Patients with a history of haematological disease, idiopathic cardiomyopathy, autoimmune disease, neoplasia, and severe renal or liver insufficiency were excluded. All participants provided written informed consent before the beginning of the study.

Validation of the top 8 key DEGs in the testing set and the collected clinical samples

First, boxplots were drawn to compare the relative expression levels of the top 8 key genes between ICM and the control samples in the testing set (GSE1869 combined with GSE5406). Then, RT-qPCR analysis was used to compare the relative expression levels of the top 8 key genes between ICM and the control samples. Total RNA was obtained from peripheral blood samples of the patients using a RNeasy™ Mini Kit (QIAGEN, Frankfurt, Germany). The total RNA was reverse transcribed into cDNA using the PrimeScript RT reagent kit (Takara Bio, Japan). RT-

qPCR was performed with a Taq PCR Master Mix Kit (Takara, Otsu, Japan) using an ABI 7500 instrument (Applied Biosystems, USA).

Statistical analysis

R software (version 4.6.0) was used to perform the bioinformatics analyses. ROC curve analyses were performed using GraphPad Prism software (version 9.0.0). A $p < 0.05$ was considered statistically significant.

Abbreviations

ICM: Ischaemic cardiomyopathy; DEGs: Differentially expressed genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; DO: Disease Ontology; HF: Heart failure; CAD: Coronary artery disease; Tregs: T-lymphocytes; GEO: Gene Expression Omnibus; FC: Fold change; CC: Cell composition; BP: Biological processes; MF: Molecular functions; DCA: Decision curve analysis; ROC: Receiver operator characteristic; NYHA: New York Heart Association; CI: Confidence interval; AUC: Area under the curve; FRZB: Frizzled-related protein; HCM: Hypertrophic cardiomyopathy; OGN: Osteoglycin; SLRP: Small leucine-rich proteoglycan; LUM: Lumican; MNS1: Meiosis-specific nuclear structural 1; SERPINA3: Serpin family A member 3; FCN3: Ficolin-3.

AUTHOR CONTRIBUTIONS

Z.P.-F and L.F. conceived the study, participated in the design, performed the statistical analyses, and drafted the manuscript. L.-Z.Y. and P.H.-W. conceived the study, participated in the design and helped to draft the manuscript. carried out the epidemiological survey, collected the samples and helped revise the manuscript. Z.Z.-F. performed the statistical analyses. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank all the participants of this study.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ETHICAL STATEMENT AND CONSENT

This research was approved by the Ethics Committee of the Hunan Provincial People’s Hospital (approval number: LL-20210615-144). All participants provided written informed consent before the beginning of the study.

FUNDING

This study was supported by grants from Key Research and Development Program of Hunan Province (No: 2019SK2021), General program of National Natural Science Foundation of China (No: 82070520), the Natural Science Foundation of Hunan Province (NO: 2020JJ8014 and NO: 2020JJ4406), and Young Doctor Fund Project of Hunan Provincial People's Hospital (No. BSJJ202212). There was no role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

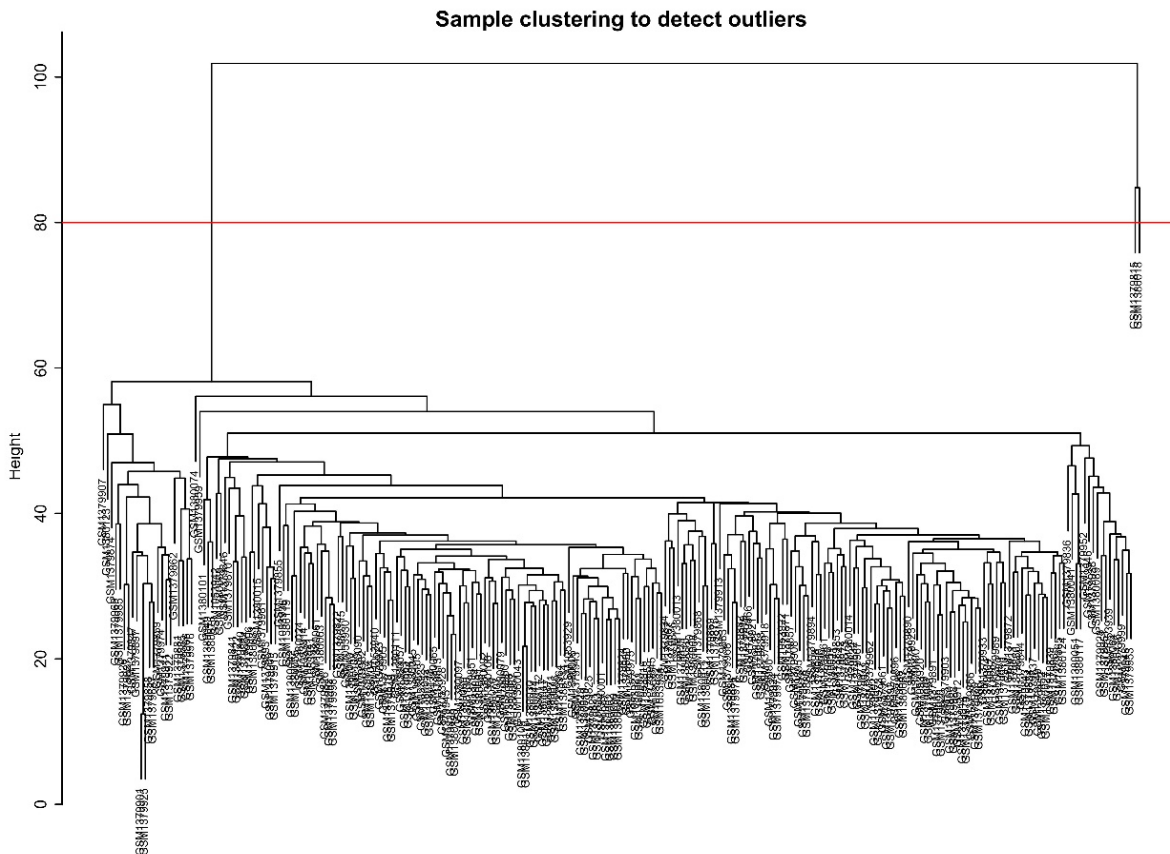
REFERENCES

1. Moroni F, Gertz Z, Azzalini L. Relief of Ischemia in Ischemic Cardiomyopathy. *Curr Cardiol Rep.* 2021; 23:80.
<https://doi.org/10.1007/s11886-021-01520-4>
PMID:[34081220](https://pubmed.ncbi.nlm.nih.gov/34081220/)
2. Doesch AO, Mueller S, Nelles M, Konstandin M, Celik S, Frankenstein L, Goeser S, Kaya Z, Koch A, Zugck C, Katus HA. Impact of troponin I-autoantibodies in chronic dilated and ischemic cardiomyopathy. *Basic Res Cardiol.* 2011; 106:25–35.
<https://doi.org/10.1007/s00395-010-0126-z>
PMID:[20957484](https://pubmed.ncbi.nlm.nih.gov/20957484/)
3. Luczak ED, Wu Y, Granger JM, Joiner MA, Wilson NR, Gupta A, Umapathi P, Murphy KR, Reyes Gaido OE, Sabet A, Corradini E, Tseng WW, Wang Y, et al. Mitochondrial CaMKII causes adverse metabolic reprogramming and dilated cardiomyopathy. *Nat Commun.* 2020; 11:4416.
<https://doi.org/10.1038/s41467-020-18165-6>
PMID:[32887881](https://pubmed.ncbi.nlm.nih.gov/32887881/)
4. Briceno N, Schuster A, Lumley M, Perera D. Ischaemic cardiomyopathy: pathophysiology, assessment and the role of revascularisation. *Heart.* 2016; 102:397–406.
<https://doi.org/10.1136/heartjnl-2015-308037>
PMID:[26740480](https://pubmed.ncbi.nlm.nih.gov/26740480/)
5. Razeghian-Jahromi I, Matta AG, Canitrot R, Zibaenezhad MJ, Razmkhah M, Safari A, Nader V, Roncalli J. Surfing the clinical trials of mesenchymal stem cell therapy in ischemic cardiomyopathy. *Stem Cell Res Ther.* 2021; 12:361.
<https://doi.org/10.1186/s13287-021-02443-1>
PMID:[34162424](https://pubmed.ncbi.nlm.nih.gov/34162424/)
6. Cabac-Pogorevici I, Muk B, Rustamova Y, Kalogeropoulos A, Tzeis S, Vardas P. Ischaemic cardiomyopathy. Pathophysiological insights, diagnostic management and the roles of revascularisation and device treatment. *Gaps and dilemmas in the era of advanced technology.* *Eur J Heart Fail.* 2020; 22:789–99.
<https://doi.org/10.1002/ejhf.1747> PMID:[32020756](https://pubmed.ncbi.nlm.nih.gov/32020756/)
7. Jiang Y, Zhang Y, Zhao C. Integrated gene expression profiling analysis reveals *SERPINA3*, *FCN3*, *FREM1*, *MNS1* as candidate biomarkers in heart failure and their correlation with immune infiltration. *J Thorac Dis.* 2022; 14:1106–19.
<https://doi.org/10.21037/jtd-22-22>
PMID:[35572891](https://pubmed.ncbi.nlm.nih.gov/35572891/)
8. Bansal SS, Ismahil MA, Goel M, Zhou G, Rokosh G, Hamid T, Prabhu SD. Dysfunctional and Proinflammatory Regulatory T-Lymphocytes Are Essential for Adverse Cardiac Remodeling in Ischemic Cardiomyopathy. *Circulation.* 2019; 139:206–21.
<https://doi.org/10.1161/CIRCULATIONAHA.118.036065>
PMID:[30586716](https://pubmed.ncbi.nlm.nih.gov/30586716/)
9. Ismahil MA, Hamid T, Bansal SS, Patel B, Kingery JR, Prabhu SD. Remodeling of the mononuclear phagocyte network underlies chronic inflammation and disease progression in heart failure: critical importance of the cardiosplenic axis. *Circ Res.* 2014; 114:266–82.
<https://doi.org/10.1161/CIRCRESAHA.113.301720>
PMID:[24186967](https://pubmed.ncbi.nlm.nih.gov/24186967/)
10. Nehra S, Gumina RJ, Bansal SS. Immune cell Dilemma in Ischemic Cardiomyopathy: To Heal or Not to Heal. *Curr Opin Physiol.* 2021; 19:39–46.
<https://doi.org/10.1016/j.cophys.2020.09.002>
PMID:[33103020](https://pubmed.ncbi.nlm.nih.gov/33103020/)
11. Osmak G, Baulina N, Koshkin P, Favorova O. Collapsing the list of myocardial infarction-related differentially expressed genes into a diagnostic signature. *J Transl Med.* 2020; 18:231.
<https://doi.org/10.1186/s12967-020-02400-1>
PMID:[32517814](https://pubmed.ncbi.nlm.nih.gov/32517814/)
12. Gronich N, Kumar A, Zhang Y, Efimov IR, Soldatov NM. Molecular remodeling of ion channels, exchangers and pumps in atrial and ventricular myocytes in ischemic cardiomyopathy. *Channels (Austin).* 2010.
<https://doi.org/10.4161/chan.4.2.10975>
PMID:[20090424](https://pubmed.ncbi.nlm.nih.gov/20090424/)
13. Díaz-Uriarte R, Alvarez de Andrés S. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics.* 2006; 7:3.
<https://doi.org/10.1186/1471-2105-7-3>
PMID:[16398926](https://pubmed.ncbi.nlm.nih.gov/16398926/)
14. Schwender H, Zucknick M, Ickstadt K, Bolt HM, and GENICA network. A pilot study on the application of statistical classification procedures to molecular epidemiological data. *Toxicol Lett.* 2004; 151:291–9.
<https://doi.org/10.1016/j.toxlet.2004.02.021>
PMID:[15177665](https://pubmed.ncbi.nlm.nih.gov/15177665/)

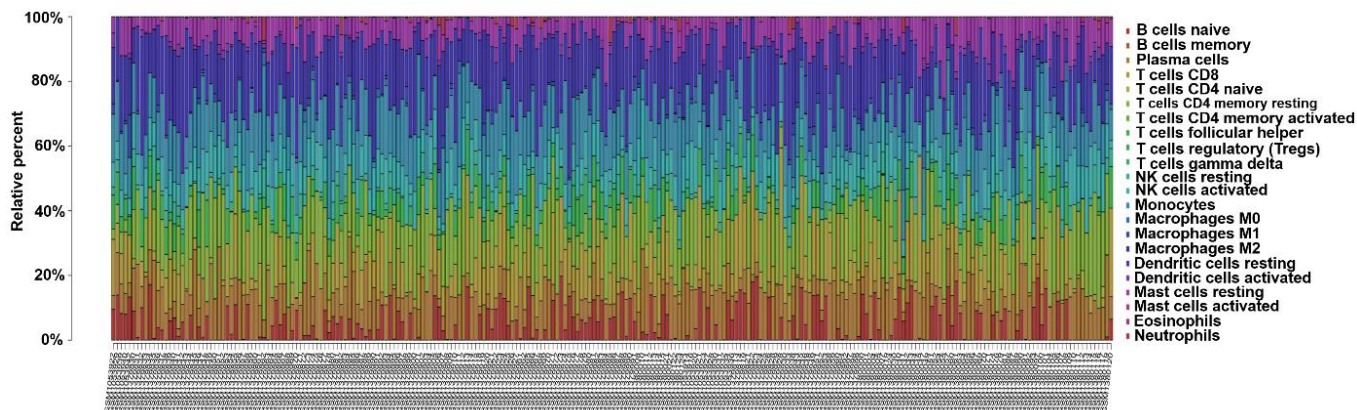
15. Jiang T, Gradus JL, Lash TL, Fox MP. Addressing Measurement Error in Random Forests Using Quantitative Bias Analysis. *Am J Epidemiol*. 2021; 190:1830–40.
<https://doi.org/10.1093/aje/kwab010>
PMID:[33517416](https://pubmed.ncbi.nlm.nih.gov/33517416/)
16. Moorthy K, Mohamad MS. Random forest for gene selection and microarray data classification. *Bioinformatics*. 2011; 7:142–6.
<https://doi.org/10.6026/97320630007142>
PMID:[22125385](https://pubmed.ncbi.nlm.nih.gov/22125385/)
17. Liang C, Wang S, Zhang M, Li T. Diagnosis, clustering, and immune cell infiltration analysis of m6A-related genes in patients with acute myocardial infarction—a bioinformatics analysis. *J Thorac Dis*. 2022; 14:1607–19.
<https://doi.org/10.21037/jtd-22-569>
PMID:[35693610](https://pubmed.ncbi.nlm.nih.gov/35693610/)
18. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015; 12:453–7.
<https://doi.org/10.1038/nmeth.3337>
PMID:[25822800](https://pubmed.ncbi.nlm.nih.gov/25822800/)
19. Flores-Arredondo JH, García-Rivas G, Torre-Amione G. Immune modulation in heart failure: past challenges and future hopes. *Curr Heart Fail Rep*. 2011; 8:28–37.
<https://doi.org/10.1007/s11897-010-0044-2>
PMID:[21221862](https://pubmed.ncbi.nlm.nih.gov/21221862/)
20. He S, Chousterman BG, Fenn A, Anzai A, Nairz M, Brandt M, Hilgendorf I, Sun Y, Ye YX, Iwamoto Y, Tricot B, Weissleder R, Macphee C, et al. Lp-PLA2 Antagonizes Left Ventricular Healing After Myocardial Infarction by Impairing the Appearance of Reparative Macrophages. *Circ Heart Fail*. 2015; 8:980–7.
<https://doi.org/10.1161/CIRCHEARTFAILURE.115.002334> PMID:[26232205](https://pubmed.ncbi.nlm.nih.gov/26232205/)
21. Tabas I, Bornfeldt KE. Macrophage Phenotype and Function in Different Stages of Atherosclerosis. *Circ Res*. 2016; 118:653–67.
<https://doi.org/10.1161/CIRCRESAHA.115.306256>
PMID:[26892964](https://pubmed.ncbi.nlm.nih.gov/26892964/)
22. Sun K, Li YY, Jin J. A double-edged sword of immuno-microenvironment in cardiac homeostasis and injury repair. *Signal Transduct Target Ther*. 2021; 6:79.
<https://doi.org/10.1038/s41392-020-00455-6>
PMID:[33612829](https://pubmed.ncbi.nlm.nih.gov/33612829/)
23. Janicki JS, Brower GL, Gardner JD, Forman MF, Stewart JA Jr, Murray DB, Chancey AL. Cardiac mast cell regulation of matrix metalloproteinase-related ventricular remodeling in chronic pressure or volume overload. *Cardiovasc Res*. 2006; 69:657–65.
<https://doi.org/10.1016/j.cardiores.2005.10.020>
PMID:[16376324](https://pubmed.ncbi.nlm.nih.gov/16376324/)
24. Kumar V, Prabhu SD, Bansal SS. CD4⁺ T-lymphocytes exhibit biphasic kinetics post-myocardial infarction. *Front Cardiovasc Med*. 2022; 9:992653.
<https://doi.org/10.3389/fcvm.2022.992653>
PMID:[36093172](https://pubmed.ncbi.nlm.nih.gov/36093172/)
25. Bansal SS, Ismahil MA, Goel M, Patel B, Hamid T, Rokosh G, Prabhu SD. Activated T Lymphocytes are Essential Drivers of Pathological Remodeling in Ischemic Heart Failure. *Circ Heart Fail*. 2017; 10:e003688.
<https://doi.org/10.1161/CIRCHEARTFAILURE.116.003688> PMID:[28242779](https://pubmed.ncbi.nlm.nih.gov/28242779/)
26. Matsumoto K, Ogawa M, Suzuki J, Hirata Y, Nagai R, Isobe M. Regulatory T lymphocytes attenuate myocardial infarction-induced ventricular remodeling in mice. *Int Heart J*. 2011; 52:382–7.
<https://doi.org/10.1536/ihj.52.382> PMID:[22188713](https://pubmed.ncbi.nlm.nih.gov/22188713/)
27. Saxena A, Dobaczewski M, Rai V, Haque Z, Chen W, Li N, Frangogiannis NG. Regulatory T cells are recruited in the infarcted mouse myocardium and may modulate fibroblast phenotype and function. *Am J Physiol Heart Circ Physiol*. 2014; 307:H1233–42.
<https://doi.org/10.1152/ajpheart.00328.2014>
PMID:[25128167](https://pubmed.ncbi.nlm.nih.gov/25128167/)
28. Wang S, Krinks M, Lin K, Luyten FP, Moos M Jr. Frzb, a secreted protein expressed in the Spemann organizer, binds and inhibits Wnt-8. *Cell*. 1997; 88:757–66.
[https://doi.org/10.1016/s0092-8674\(00\)81922-4](https://doi.org/10.1016/s0092-8674(00)81922-4)
PMID:[9118219](https://pubmed.ncbi.nlm.nih.gov/9118219/)
29. Oh CK, Ko Y, Park JJ, Heo HJ, Kang J, Kwon EJ, Kang JW, Lee Y, Myung K, Kang JM, Ko DS, Kim YH. FRZB as a key molecule in abdominal aortic aneurysm progression affecting vascular integrity. *Biosci Rep*. 2021; 41:BSR20203204.
<https://doi.org/10.1042/BSR20203204>
PMID:[33245093](https://pubmed.ncbi.nlm.nih.gov/33245093/)
30. Guo Y, Xie J, Rubin E, Tang YX, Lin F, Zi X, Hoang BH. Frzb, a secreted Wnt antagonist, decreases growth and invasiveness of fibrosarcoma cells associated with inhibition of Met signaling. *Cancer Res*. 2008; 68:3350–60.
<https://doi.org/10.1158/0008-5472.CAN-07-3220>
PMID:[18451162](https://pubmed.ncbi.nlm.nih.gov/18451162/)
31. Person AD, Garriock RJ, Krieg PA, Runyan RB, Klewer SE. Frzb modulates Wnt-9a-mediated beta-catenin signaling during avian atrioventricular cardiac cushion development. *Dev Biol*. 2005; 278:35–48.
<https://doi.org/10.1016/j.ydbio.2004.10.013>
PMID:[15649459](https://pubmed.ncbi.nlm.nih.gov/15649459/)
32. Yang Y, Liu P, Teng R, Liu F, Zhang C, Lu X, Ding Y. Integrative bioinformatics analysis of potential

- therapeutic targets and immune infiltration characteristics in dilated cardiomyopathy. *Ann Transl Med.* 2022; 10:348.
<https://doi.org/10.21037/atm-22-732> PMID:35433958
33. Ma Z, Wang X, Lv Q, Gong Y, Xia M, Zhuang L, Lu X, Yang Y, Zhang W, Fu G, Ye Y, Lai D. Identification of Underlying Hub Genes Associated with Hypertrophic Cardiomyopathy by Integrated Bioinformatics Analysis. *Pharmgenomics Pers Med.* 2021; 14:823–37.
<https://doi.org/10.2147/PGPM.S314880>
PMID:34285551
34. Rienks M, Papageorgiou A, Wouters K, Verhesen W, Leeuwen RV, Carai P, Summer G, Westermann D, Heymans S. A novel 72-kDa leukocyte-derived osteoglycin enhances the activation of toll-like receptor 4 and exacerbates cardiac inflammation during viral myocarditis. *Cell Mol Life Sci.* 2017; 74:1511–25.
<https://doi.org/10.1007/s00018-016-2423-7>
PMID:27878326
35. Petretto E, Sarwar R, Grieve I, Lu H, Kumaran MK, Muckett PJ, Mangion J, Schroen B, Benson M, Punjabi PP, Prasad SK, Pennell DJ, Kiesewetter C, et al. Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass. *Nat Genet.* 2008; 40:546–52.
<https://doi.org/10.1038/ng.134> PMID:18443592
36. Hu Y, Liu J, Zhao Q, Xu P, Chen Y, Zhou H, Li X. Correlation between mimecan expression and coronary artery stenosis in patients with coronary heart disease. *Int J Clin Exp Med.* 2015; 8:21641–6.
PMID:26885118
37. Wu QH, Ma Y, Ruan CC, Yang Y, Liu XH, Ge Q, Kong LR, Zhang JW, Yan C, Gao PJ. Loss of osteoglycin promotes angiogenesis in limb ischaemia mouse models via modulation of vascular endothelial growth factor and vascular endothelial growth factor receptor 2 signalling pathway. *Cardiovasc Res.* 2017; 113:70–80.
<https://doi.org/10.1093/cvr/cvw220> PMID:28069703
38. Mannino MP, Hart GW. The Beginner's Guide to O-GlcNAc: From Nutrient Sensitive Pathway Regulation to Its Impact on the Immune System. *Front Immunol.* 2022; 13:828648.
<https://doi.org/10.3389/fimmu.2022.828648>
PMID:35173739
39. Kafienah W, Cheung FL, Sims T, Martin I, Miot S, Von Ruhland C, Roughley PJ, Hollander AP. Lumican inhibits collagen deposition in tissue engineered cartilage. *Matrix Biol.* 2008; 27:526–34.
<https://doi.org/10.1016/j.matbio.2008.04.002>
PMID:18534835
40. Yeh JT, Yeh LK, Jung SM, Chang TJ, Wu HH, Shiu TF, Liu CY, Kao WW, Chu PH. Impaired skin wound healing in lumican-null mice. *Br J Dermatol.* 2010; 163:1174–80.
<https://doi.org/10.1111/j.1365-2133.2010.10008.x>
PMID:20738297
41. Dupuis LE, Berger MG, Feldman S, Doucette L, Fowlkes V, Chakravarti S, Thibaudeau S, Alcalá NE, Bradshaw AD, Kern CB. Lumican deficiency results in cardiomyocyte hypertrophy with altered collagen assembly. *J Mol Cell Cardiol.* 2015; 84:70–80.
<https://doi.org/10.1016/j.yjmcc.2015.04.007>
PMID:25886697
42. Mohammadzadeh N, Melleby AO, Palmero S, Sjaastad I, Chakravarti S, Engebretsen KVT, Christensen G, Lunde IG, Tønnessen T. Moderate Loss of the Extracellular Matrix Proteoglycan Lumican Attenuates Cardiac Fibrosis in Mice Subjected to Pressure Overload. *Cardiology.* 2020; 145:187–98.
<https://doi.org/10.1159/000505318> PMID:31968347
43. Engebretsen KV, Lunde IG, Strand ME, Waehre A, Sjaastad I, Marstein HS, Skrbic B, Dahl CP, Askevold ET, Christensen G, Bjørnstad JL, Tønnessen T. Lumican is increased in experimental and clinical heart failure, and its production by cardiac fibroblasts is induced by mechanical and proinflammatory stimuli. *FEBS J.* 2013; 280:2382–98.
<https://doi.org/10.1111/febs.12235> PMID:23480731
44. Zhang Q, Fan M, Cao X, Geng H, Su Y, Wu C, Pan H, Pan M. Integrated Bioinformatics Algorithms and Experimental Validation to Explore Robust Biomarkers and Landscape of Immune Cell Infiltration in Dilated Cardiomyopathy. *Front Cardiovasc Med.* 2022; 9:809470.
<https://doi.org/10.3389/fcvm.2022.809470>
PMID:35433865
45. Gallagher GL, Jackson CJ, Hunyor SN. Myocardial extracellular matrix remodeling in ischemic heart failure. *Front Biosci.* 2007; 12:1410–9.
<https://doi.org/10.2741/2157> PMID:17127391
46. Jiang C, Jiang W. Integrated Bioinformatics Identifies *FREM1* as a Diagnostic Gene Signature for Heart Failure. *Appl Bionics Biomech.* 2022; 2022:1425032.
<https://doi.org/10.1155/2022/1425032>
PMID:35726312
47. Kloth JN, Gorter A, Fleuren GJ, Oosting J, Uljee S, ter Haar N, Dreef EJ, Kenter GG, Jordanova ES. Elevated expression of SerpinA1 and SerpinA3 in HLA-positive cervical carcinoma. *J Pathol.* 2008; 215:222–30.
<https://doi.org/10.1002/path.2347>
PMID:18438953
48. Liu Z, Hou L, Liu Y, Gong J. LncRNA GAS5 exacerbates myocardial ischemia-reperfusion injury through

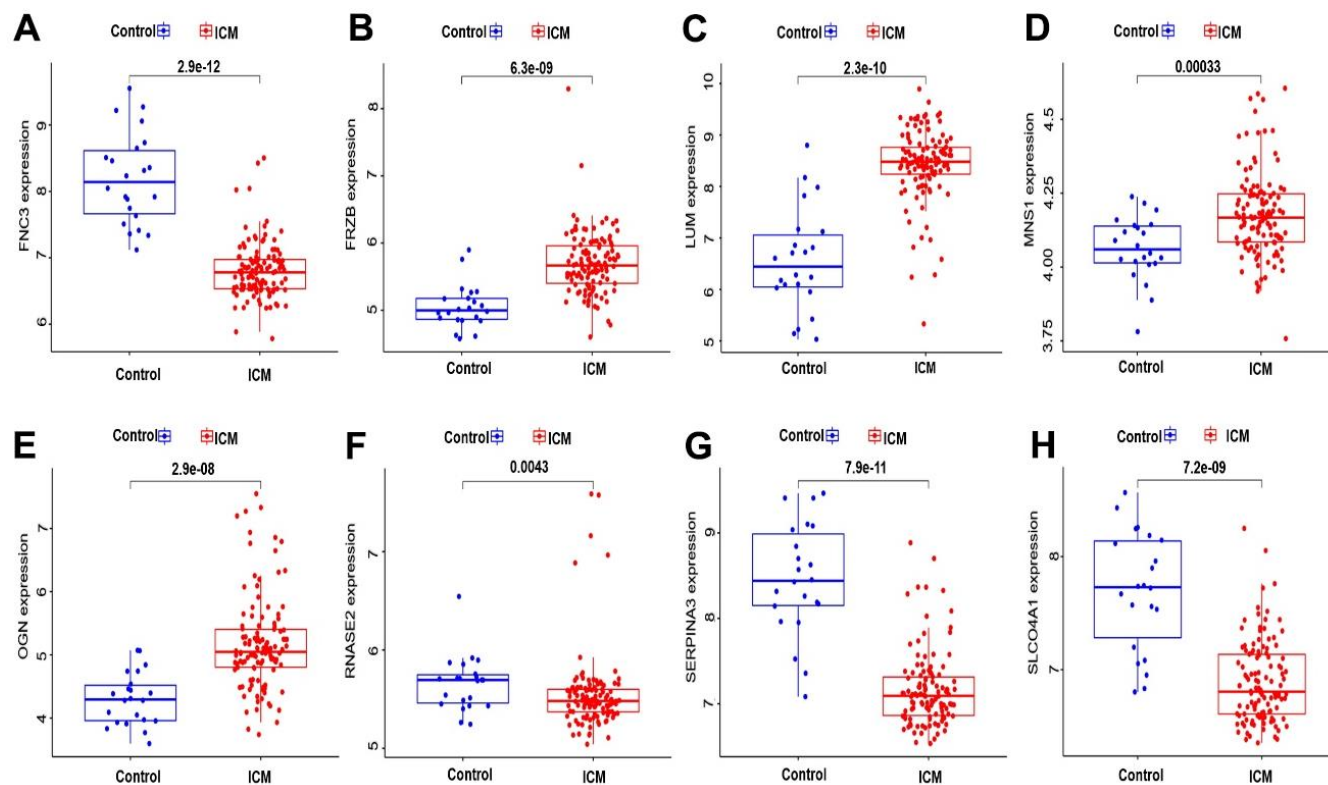
- regulating serpin3 by targeting miR-137. *Int J Cardiol.* 2020; 306:9.
<https://doi.org/10.1016/j.ijcard.2020.01.067>
PMID:[32276715](https://pubmed.ncbi.nlm.nih.gov/32276715/)
49. Wågsäter D, Johansson D, Fontaine V, Vorkapic E, Bäcklund A, Razuvaev A, Mäyränpää MI, Hjerpe C, Caidahl K, Hamsten A, Franco-Cereceda A, Wilbertz J, Swedenborg J, et al. Serine protease inhibitor A3 in atherosclerosis and aneurysm disease. *Int J Mol Med.* 2012; 30:288–94.
<https://doi.org/10.3892/ijmm.2012.994>
PMID:[22580763](https://pubmed.ncbi.nlm.nih.gov/22580763/)
50. Asakura M, Kitakaze M. Global gene expression profiling in the failing myocardium. *Circ J.* 2009; 73:1568–76.
<https://doi.org/10.1253/circj.cj-09-0465>
PMID:[19638707](https://pubmed.ncbi.nlm.nih.gov/19638707/)
51. Munthe-Fog L, Hummelshøj T, Honoré C, Madsen HO, Permin H, Garred P. Immunodeficiency associated with FCN3 mutation and ficolin-3 deficiency. *N Engl J Med.* 2009; 360:2637–44.
<https://doi.org/10.1056/NEJMoa0900381>
PMID:[19535802](https://pubmed.ncbi.nlm.nih.gov/19535802/)
52. Andersen T, Munthe-Fog L, Garred P, Jacobsen S. Serum levels of ficolin-3 (Hakata antigen) in patients with systemic lupus erythematosus. *J Rheumatol.* 2009; 36:757–9.
<https://doi.org/10.3899/jrheum.080361>
PMID:[19208603](https://pubmed.ncbi.nlm.nih.gov/19208603/)
53. Shang Q, Feng L, Yu W, Xu J, Liu X, Wang J. [Proteomics study on ficolin 3 in the human plasma of type 2 diabetics]. *Wei Sheng Yan Jiu.* 2016; 45:8–13.
PMID:[26987188](https://pubmed.ncbi.nlm.nih.gov/26987188/)
54. Song H, Chen S, Zhang T, Huang X, Zhang Q, Li C, Chen C, Chen S, Liu D, Wang J, Tu Y, Wu Y, Liu Y. Integrated Strategies of Diverse Feature Selection Methods Identify Aging-Based Reliable Gene Signatures for Ischemic Cardiomyopathy. *Front Mol Biosci.* 2022; 9:805235.
<https://doi.org/10.3389/fmolb.2022.805235>
PMID:[35300115](https://pubmed.ncbi.nlm.nih.gov/35300115/)
55. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015; 43:e47.
<https://doi.org/10.1093/nar/gkv007> PMID:[25605792](https://pubmed.ncbi.nlm.nih.gov/25605792/)
56. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012; 16:284–7.
<https://doi.org/10.1089/omi.2011.0118>
PMID:[22455463](https://pubmed.ncbi.nlm.nih.gov/22455463/)
57. Tian Y, Yang J, Lan M, Zou T. Construction and analysis of a joint diagnosis model of random forest and artificial neural network for heart failure. *Aging (Albany NY).* 2020; 12:26221–35.
<https://doi.org/10.18632/aging.202405>
PMID:[33401250](https://pubmed.ncbi.nlm.nih.gov/33401250/)
58. Yan J, Liu L, Wang W, Zhao Y, Li KK, Li K, Wang L, Yuan B, Geng H, Zhang S, Liu Z, Duan W, Zhan Y, et al. Radiomic Features From Multi-Parameter MRI Combined With Clinical Parameters Predict Molecular Subgroups in Patients With Medulloblastoma. *Front Oncol.* 2020; 10:558162.
<https://doi.org/10.3389/fonc.2020.558162>
PMID:[33117690](https://pubmed.ncbi.nlm.nih.gov/33117690/)
59. Wu Y, Chen H, Li L, Zhang L, Dai K, Wen T, Peng J, Peng X, Zheng Z, Jiang T, Xiong W. Construction of Novel Gene Signature-Based Predictive Model for the Diagnosis of Acute Myocardial Infarction by Combining Random Forest With Artificial Neural Network. *Front Cardiovasc Med.* 2022; 9:876543.
<https://doi.org/10.3389/fcvm.2022.876543>
PMID:[35694667](https://pubmed.ncbi.nlm.nih.gov/35694667/)
60. Park SY. Nomogram: An analogue tool to deliver digital knowledge. *J Thorac Cardiovasc Surg.* 2018; 155:1793.
<https://doi.org/10.1016/j.jtcvs.2017.12.107>
PMID:[29370910](https://pubmed.ncbi.nlm.nih.gov/29370910/)
61. Chou BP, Critsinelis A, Lamba HK, Long G, Civitello AB, Delgado RM, Chatterjee S. Continuous-Flow Left Ventricular Assist Device Support in Patients with Ischemic Versus Nonischemic Cardiomyopathy. *Tex Heart Inst J.* 2021; 48:e207241.
<https://doi.org/10.14503/THIJ-20-7241>
PMID:[34468765](https://pubmed.ncbi.nlm.nih.gov/34468765/)
62. Felker GM, Shaw LK, O'Connor CM. A standardized definition of ischemic cardiomyopathy for use in clinical research. *J Am Coll Cardiol.* 2002; 39:210–8.
[https://doi.org/10.1016/s0735-1097\(01\)01738-7](https://doi.org/10.1016/s0735-1097(01)01738-7)
PMID:[11788209](https://pubmed.ncbi.nlm.nih.gov/11788209/)



Supplementary Figure 2. Clustering dendrogram of samples. Cut height = 80 was used to divide 248 samples in the training sets (GSE42955 combined with GSE57338) into two different cluster types. Cluster 0 contains the following two samples (GSM1379815 and GSE1380018), cluster 1 contains the following 246 samples (GSM1053922, GSM1053929, GSM1053939, GSM1053940, GSM1053942, GSM1379830, GSM1379831, GSM1379832, GSM1379833, GSM1379834, GSM1379835, GSM1379836, GSM1379837, GSM1379838, GSM1379839, GSM1379840, GSM1379841, GSM1379842, GSM1379843, GSM1379844, GSM1379845, GSM1379846, GSM1379847, GSM1379848, GSM1379849, GSM1379850, GSM1379851, GSM1379852, GSM1379853, GSM1379854, GSM1379855, GSM1379856, GSM1379857, GSM1379858, GSM1379859, GSM1379860, GSM1379861, GSM1379862, GSM1379863, GSM1379864, GSM1379865, GSM1379866, GSM1379867, GSM1379868, GSM1379869, GSM1379870, GSM1379871, GSM1379872, GSM1379873, GSM1379874, GSM1379875, GSM1379876, GSM1379877, GSM1379878, GSM1379880, GSM1379881, GSM1379882, GSM1379883, GSM1379884, GSM1379885, GSM1379886, GSM1379887, GSM1379888, GSM1379889, GSM1379890, GSM1379891, GSM1379892, GSM1379893, GSM1379894, GSM1379895, GSM1379896, GSM1379897, GSM1379898, GSM1379899, GSM1379900, GSM1379901, GSM1379902, GSM1379903, GSM1379904, GSM1379905, GSM1379906, GSM1379907, GSM1379908, GSM1379909, GSM1379910, GSM1379911, GSM1379912, GSM1379913, GSM1379914, GSM1379915, GSM1379918, GSM1379919, GSM1379920, GSM1379921, GSM1379922, GSM1379923, GSM1379924, GSM1379925, GSM1379926, GSM1379927, GSM1379961, GSM1379962, GSM1379965, GSM1379966, GSM1379967, GSM1379968, GSM1379969, GSM1379970, GSM1379971, GSM1379972, GSM1379973, GSM1379974, GSM1379975, GSM1379976, GSM1379977, GSM1379978, GSM1379979, GSM1379980, GSM1379981, GSM1379982, GSM1379983, GSM1379984, GSM1379985, GSM1379986, GSM1379987, GSM1379988, GSM1379989, GSM1379990, GSM1379991, GSM1380008, GSM1380010, GSM1380011, GSM1380012, GSM1380013, GSM1380014, GSM1380015, GSM1380021, GSM1380027, GSM1380122, GSM1380123, GSM1380124, GSM1053914, GSM1053916, GSM1053920, GSM1053921, GSM1053923, GSM1053927, GSM1053928, GSM1053930, GSM1053931, GSM1053932, GSM1053934, GSM1053936, GSM1379813, GSM1379814, GSM1379819, GSM1379821, GSM1379822, GSM1379823, GSM1379824, GSM1379825, GSM1379826, GSM1379828, GSM1379829, GSM1379928, GSM1379930, GSM1379933, GSM1379938, GSM1379943, GSM1379945, GSM1379946, GSM1379948, GSM1379949, GSM1379951, GSM1379952, GSM1379953, GSM1379958, GSM1379959, GSM1379960, GSM1379964, GSM1379992, GSM1379997, GSM1379998, GSM1379999, GSM1380000, GSM1380001, GSM1380002, GSM1380004, GSM1380005, GSM1380007, GSM1380019, GSM1380024, GSM1380030, GSM1380040, GSM1380041, GSM1380042, GSM1380043, GSM1380044, GSM1380045, GSM1380047, GSM1380049, GSM1380051, GSM1380052, GSM1380054, GSM1380055, GSM1380057, GSM1380060, GSM1380061, GSM1380063, GSM1380064, GSM1380065, GSM1380066, GSM1380067, GSM1380069, GSM1380070, GSM1380071, GSM1380074, GSM1380076, GSM1380079, GSM1380081, GSM1380086, GSM1380087, GSM1380089, GSM1380090, GSM1380091, GSM1380092, GSM1380093, GSM1380097, GSM1380100, GSM1380101, GSM1380104, GSM1380105, GSM1380106, GSM1380107, GSM1380108, GSM1380109, GSM1380110, GSM1380111, GSM1380112, GSM1380113, GSM1380114, GSM1380115, GSM1380116, GSM1380117, GSM1380119 and GSM1380120).



Supplementary Figure 3. Infiltration pattern of immune cell subtypes in validation set. The bar plot visualizing the relative percent of 22 immune cell in each sample.



Supplementary Figure 4. External validation of the key genes in testing set. The expression levels of FCN3 (A), FRZB (B), LUM (C), MNS1 (D), OGN (E), RNASE2 (F), SERPINA3 (G), SLCO4A1 (H) in testing set.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 2, 4–6.

Supplementary Table 1. A total of 39 differentially expressed genes were identified.

Supplementary Table 2. The gene expression matrix in the training set (GSE42955 combined with GSE57338).

Supplementary Table 3. The importance of 39 DEGs.

Gene	Importance
FCN3	19.31250056
SERPINA3	19.25873157
MNS1	7.778561915
OGN	7.202295921
SLCO4A1	5.872656897
RNASE2	5.399599675
FRZB	5.037672655
LUM	4.95926197
SMOC2	4.958591477
SFRP4	4.116572428
VSIG4	4.024444085
CD163	3.813078851
FCER1G	2.852484032
ADAMTS9	2.749976824
ASPN	2.486912883
PHLDA1	1.847027445
MYOT	1.836173723
ANKRD2	1.743026675
LYVE1	1.677558479
IL1RL1	1.466302701
HBB	1.209413475
MYH6	1.156130419
PLA2G2A	1.012830974
C6	0.805501497
CYP4B1	0.76590613
MGST1	0.689610082
ADAMTS4	0.636695692
NPPA	0.627402182
MXRA5	0.616162724
UTY	0.508651928
DDX3Y	0.49215963
USP9Y	0.462653587
SERPINE1	0.436366429
HMGCS2	0.422313019
EIF1AY	0.38407382
MIR208A	0.371855071
AOX1	0.333867857
IFI44L	0.305582353
DSC1	0.269703898

Supplementary Table 4. KEGG and DO analysis of the 39 DEGs.

Supplementary Table 5. GO analysis of the 39 DEGs.

Supplementary Table 6. Immune cell distribution pattern between the ICM and control samples.