

RESEARCH

Open Access



Identification through machine learning of potential immune-related gene biomarkers associated with immune cell infiltration in myocardial infarction

Hao Dong¹, Shi-Bai Yan², Guo-Sheng Li³, Zhi-Guang Huang², Dong-Ming Li², Yu-lu Tang², Jia-Qian Le², Yan-Fang Pan⁴, Zhen Yang⁵, Hong-Bo Pan², Gang Chen² and Ming-Jie Li^{2*}

Abstract

Background To investigate the potential role of immune-related genes (IRGs) and immune cells in myocardial infarction (MI) and establish a nomogram model for diagnosing myocardial infarction.

Methods Raw and processed gene expression profiling datasets were archived from the Gene Expression Omnibus (GEO) database. Differentially expressed immune-related genes (DIRGs), which were screened out by four machine learning algorithms-partial least squares (PLS), random forest model (RF), k-nearest neighbor (KNN), and support vector machine model (SVM) were used in the diagnosis of MI.

Results The six key DIRGs (PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM) were identified by the intersection of the minimal root mean square error (RMSE) of four machine learning algorithms, which were screened out to establish the nomogram model to predict the incidence of MI by using the rms package. The nomogram model exhibited the highest predictive accuracy and better potential clinical utility. The relative distribution of 22 types of immune cells was evaluated using cell type identification, which was done by estimating relative subsets of RNA transcripts (CIBERSORT) algorithm. The distribution of four types of immune cells, such as plasma cells, T cells follicular helper, Mast cells resting, and neutrophils, was significantly upregulated in MI, while five types of immune cell dispersion, T cells CD4 naive, macrophages M1, macrophages M2, dendritic cells resting, and mast cells activated in MI patients, were significantly downregulated in MI.

Conclusion This study demonstrated that IRGs were correlated with MI, suggesting that immune cells may be potential therapeutic targets of immunotherapy in MI.

Keywords Immune-related gene, Immune cell infiltration, CIBERSORT, Nomogram

*Correspondence:

Ming-Jie Li

limingjie@stu.gxmu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2023. **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 Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

The mortality of coronary artery disease (CAD) has decreased in recent decades, but it remains the main cause of mortality worldwide [1]. For example, over the past 35 years, 1.7 million deaths annually have been attributed to CAD in the USA and Europe [2]. One of the biggest causes of mortality in CAD is MI [1]. With the accessibility of coronary interventions, coronary bypass surgery, and drugs, early diagnosis and risk stratification can significantly reduce mortality in MI. Traditional biomarkers in the early diagnosis of MI, such as high-sensitivity cardiac troponin T (hs-cTnT), high-sensitivity cardiac troponin I (hs-cTnI), and creatine kinase-MB, have been demonstrated with high sensitivity but without specificity [3]. As a result, there is a need to screen out the novel diagnostic biomarkers of MI.

Patients with CAD can be classified as either chronic coronary syndromes (CCS) or acute coronary syndromes (ACS), depending on the clinical symptoms the patient is presenting [4]. The literature has demonstrated that the conversion from CCS to ACS is typically initiated by an acute atherothrombotic event, which results in the rupture or erosion of atherosclerotic plaques [5]. Many conventional risk factors, such as smoking, high cholesterol, obesity, diabetes mellitus, and hypertension, are responsible for the incidence of CAD by participating in the immune microenvironment [6]. However, there is less evidence suggesting a relationship between the pathogenesis of MI and immune genes or immune cells or inflammatory mediators.

With the advancement of microarray analysis, an increasing number of studies have demonstrated that genes can be targeted for early diagnosis, classification, prognosis, prediction of disease severity, and new drugs. For example, several genes in peripheral blood mononuclear cells, such as ADAP2, KLRC1, MIR21, PDGFD, and CD14, were demonstrated as having a significant signature for categorizing MI patients and normal controls [7]. ASCC2, LRRC18, and SLC25A37 have not only been demonstrated as the diagnostic biomarkers of CAD, but also have closely participated in the pathogenesis and advancement of CAD [8]. There is little research that has analyzed the immune genes of CAD and MI.

Many pieces of research have indicated that immune cell infiltration is closely related to the onset of MI. For instance, increased apoptosis of lymphocytes in peripheral blood and infiltrated proinflammatory Th1 lymphocytes was observed in pig hearts after reperfusion within 48 h, and circulating T lymphocytes were significantly decreased in post-PCI MI patients within the first 24 h [9]. Cell type identification by estimating the relative subsets of RNA transcripts (CIBERSORT) has been widely implemented to portray immune cell

ratios from RNA-seq data of samples from various diseases [10]. However, there is little research on immune cell infiltration analysis in MI patients conducted using CIBERSORT.

In our study, the raw gene expression profiling datasets of MI patients and normal controls were archived from the GEO database, and the intersections between IRGs and differentially expressed genes (DEGs) were identified for further analysis. Diagnostic biomarkers were identified using four machine learning algorithms. Subsequently, the relative proportion of 22 different types of immune cells in patients with MI and in normal controls was calculated using CIBERSORT. Finally, the potential role of diagnostic IRGs associated with immune cell infiltration was verified in patients with MI using machine learning.

Material and methods

Microarray data source

The GEO database (<http://www.ncbi.nlm.nih.gov/geo>) was explored to download the datasets according to keywords such as “coronary artery disease” or “Myocardial infarction.” The gene dataset was considered eligible according to the following inclusion criteria: (1) datasets belonging to humans and (2) the source of tissue was blood. The exclusion criteria were as follows: (1) duplicated datasets, (2) without case control data, and (3) non-human data. The GSE113079 contains 97 CAD patients and healthy controls. A new data cohort merged by four GEO datasets (GSE29111, GSE48060, GSE66360, and GSE97320) contains 101 MI patients and 74 healthy controls. The information of five GEO datasets (GSE29111, GSE48060, GSE66360, GSE97320, and GSE113079) are included in Table 1. Here, 2,484 IRGs were archived from ImmPort (<https://www.immport.org>) [11]. The schematic diagram of our study is shown in Fig. 1.

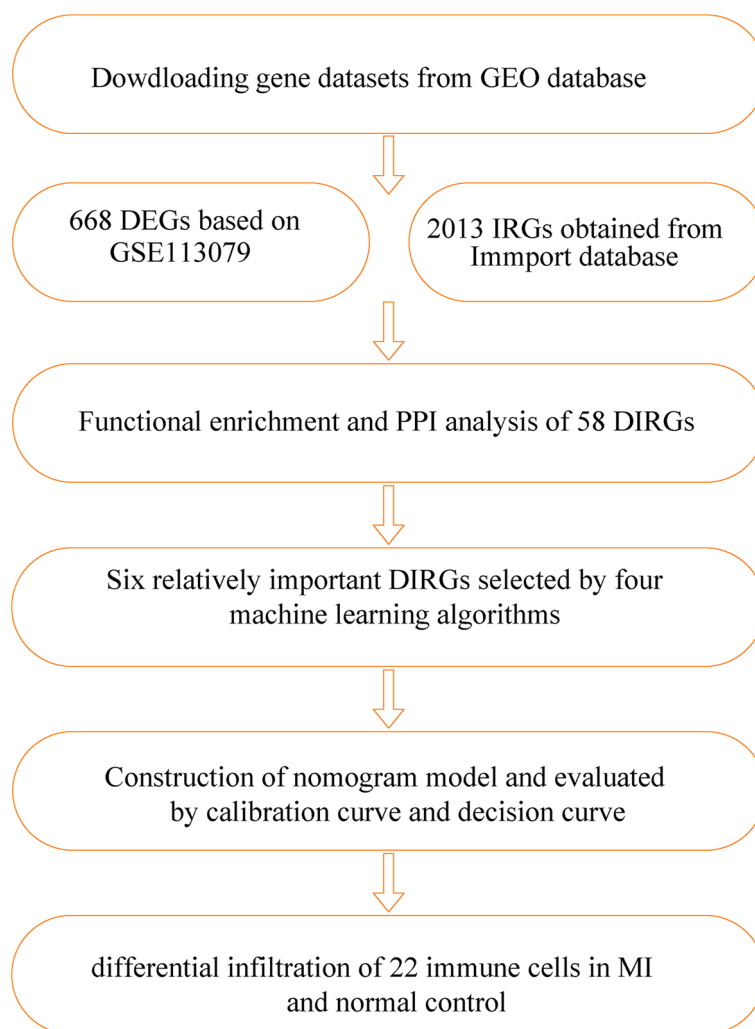
Data processing and differentially expressed gene analysis

The DEGs with a threshold of a P value of <0.05 and $\log(\text{fold change}) > 1$ or $\log(\text{fold change}) < -1$ between patients with CAD and normal controls in GSE113079 were obtained by the limma package [12]. DIRGs were obtained by overlapping DEGs and IRGs, as shown in Fig. 2. The relationship between DIRGs and the incidence of MI was verified using a new data cohort. First, the “RMA” function of the Affy package was applied to raw data in the new data cohort for background correction and quantile normalization [13], and the batch effect was removed by the “removeBatchEffect” function in the sva package [14]. The DEGs between MI patients and normal controls were screened by conducting the limma package. The volcano plot of DEGs was visualized by the ggplot2 package.

Table 1 Information of the involved gene datasets

ID	Year	Platform	Region	Tissue	Sample
GSE29111	2011	GPL570	United Kingdom	Peripheral blood	N:MI = 0:18
GSE48060	2014	GPL570	USA	Peripheral blood	N:MI = 21:31
GSE66360	2015	GPL570	USA	Peripheral blood	N:MI = 50:49
GSE97320	2017	GPL570	China	Peripheral blood	N:MI = 3:3
GSE113079	2018	GPL20115	China	Peripheral blood	N:CAD = 48:93

Abbreviations: N Normal control, MI Myocardial infarction, CAD coronary artery disease

**Fig. 1** Schematic diagram of study

Functional enrichment analysis of DIRGs

To explore the mechanism of the incidence of MI, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment [15–17] were performed to annotate the genes and gene products. GO analysis was conducted to identify the

biological process (BP), cellular components (CC), and molecular function (MF) of key DIRGs. KEGG, consisting of chemical and systemic functional information, was implemented to identify functional and metabolic pathways. DO (disease ontology) analysis was conducted to combine biomedical data with

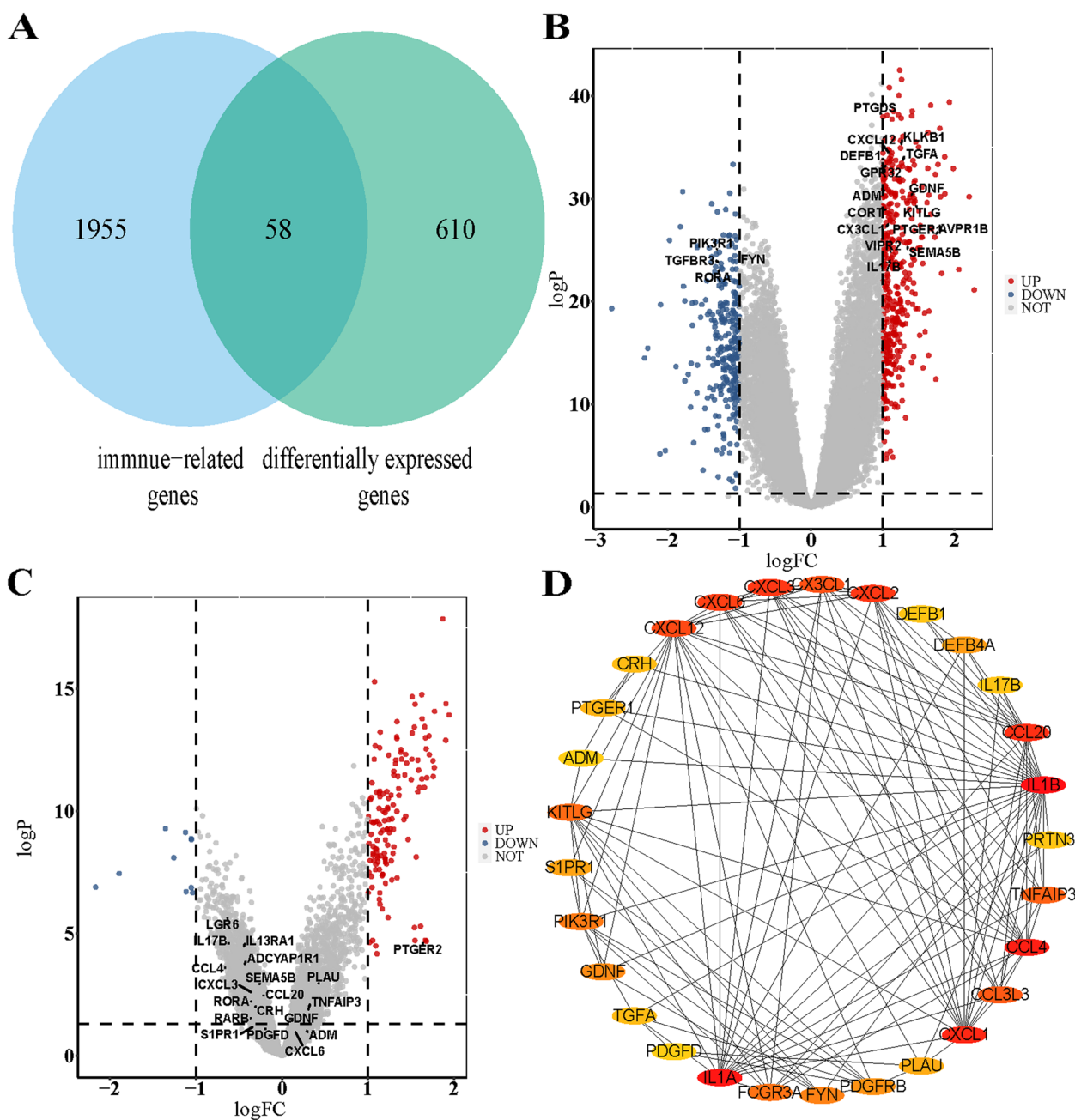


Fig. 2 Differentially expressed gene analysis and protein–protein interaction networks. **A** Intersection of 668 DEGs and 1213 IRGs. **B** volcano plot in GSE113079. **C** volcano plot in merged GP570 datasets. **D** The PPI network of the 58 DIRGs

human disease. Metascape (www.metascape.org) is an online website that provides a comprehensive gene annotation and analysis resource [18]. DIRGs were imported to Metascape to analyze the GO and KEGG pathway analysis, and the DOSE package was conducted to investigate the DO analysis of DIRGs [19].

Integration of PPI network to select hub genes

The STRING database (<http://string-db.org/>) aims to investigate protein interactions, both known and predicted. The protein–protein interaction network (PPI) of DIRGs was constructed by the STRING database. Cytoscape is one open-source software tool for

bioinformatics analysis, which is conducted for PPI network visualization and analysis. The key modules of the PPI network, here identified by the molecular complex detection (MCODE) plug-in of Cytoscape software, were widely implemented to screen out significant parts according to the default parameter settings (Degree Cutoff=2, Node Score Cutoff=0.2, K-Core=2, and Max Depth=100).

Key DIRGs selected by the PLS, GLM, and SVM algorithms

To develop the diagnosis model for patients with MI, we conducted four machine learning algorithms, such as partial least squares (PLS), random forest model (RF), k-nearest neighbor (KNN), and support vector machine model (SVM). PLS is a multivariate statistical data analysis, considered as the combination of Principal Component Analysis and multiple linear regression analysis, which could develop accurate prediction model when variables significantly correlated [20]. RF is a learning method for classification, regression and other tasks, which builds decision trees on different samples and scores the classification results, and RF model will use statistical analysis on the classification results to screen out high accuracy classification results of all single trees [21]. SVM was a supervised machine-learning technique for classification and regression. It can filter out the feature subset of the highest accuracy results in a large amount of data [22]. KNN is a non-parametric, supervised learning classifier which uses proximity to make classifications about the grouping of an individual data point to find a predetermined number of training samples closest in the distance to a new point and provide a value for the data [23]. PLS, KNN, RF, and SVM were constructed in a new data cohort merged by four GEO datasets (GSE29111, GSE48060, GSE66360, and GSE97320) by the DALEX package in R. Significant DIRGs were used as explanatory features to distinguish MI patients and normal controls in a new data cohort. The residual distribution was shown to have the best model with minimal residuals, and the importance of the model in DIRGs was selected by the root mean square error (RMSE). Finally, the six most important DIRGs were selected by the above four models for further study.

A nomogram model constructed and assessed for a diagnosis of MI

A nomogram model consisting of six significant DIRGs was constructed using the rms package for predicting the incidence of MI. "Points" were demonstrated separately as the score of the six most important DIRGs, and "Total Points" was the summation of the above DIRGs. The Area Under Curve (AUC) was implemented to assess the discrimination ability of nomogram model and each DIRGs.

A calibration curve was constructed to plot the predictive and actual probability of the nomogram model. Finally, the clinical usefulness and effectiveness of the nomogram model were demonstrated by decision curve analysis.

Distribution of immune cells

As a computational deconvolution algorithm method, CIBERSORT can characterize the cell composition of complex tissue from gene expression profiles. Twenty-two immune cell types in a new data cohort were calculated by CIBERSORT. We then compared the distribution of 22 types of immune cells between patients with MI and normal controls.

The association between key DIRGs and the immune cell infiltration of MI

Pearson correlation analysis was implemented to estimate the association between key DIRGs and immune cell infiltration, here by using the psych package and as visualized by the ggplot2 package.

Results

Data processing and differentially expressed gene analysis

According to the following criterion of a P value < 0.05 and $\log(\text{fold change}) > 1$ or $\log(\text{fold change}) < -1.668$, DEGs between patients with CAD and normal controls were archived in the GSE113079 dataset by the limma package. To determine the significant IRGs from the DEGs, 58 DIRGs were obtained after the intersection between 668 DEGs and 2013 IRG (Fig. 2A). Here, 58 DIRGs were shown by volcano plot in GSE113079 and the new dataset (Fig. 2B, C). The PPI network of the 58 DIRGs was investigated using STRING 11.0 and visualizing in Cytoscape 3.8.0 (Fig. 2D).

GO, KEGG, and DO pathway enrichment analysis

Fifty-eight DIRGs were imported to the Metascape website to explore the enrichment analysis visualized by the ggplot2 package. GO analysis, as explored by the Metascape online database, was conducted to cluster the functions of the BP, CC, and MF. The results demonstrated that 58 DIRGs were enriched in the inflammatory response of BP, side of the membrane of CC, and signaling receptor regulator activity of MF (Fig. 3A, B, C). The KEGG pathway showed that 58 DIRGs were mostly enriched in cytokine-cytokine receptor interactions (Fig. 3D). The DO pathway analysis was mainly enriched in atherosclerosis (Fig. 3E). These results indicated that inflammation may be significantly associated with the incidence of CAD.

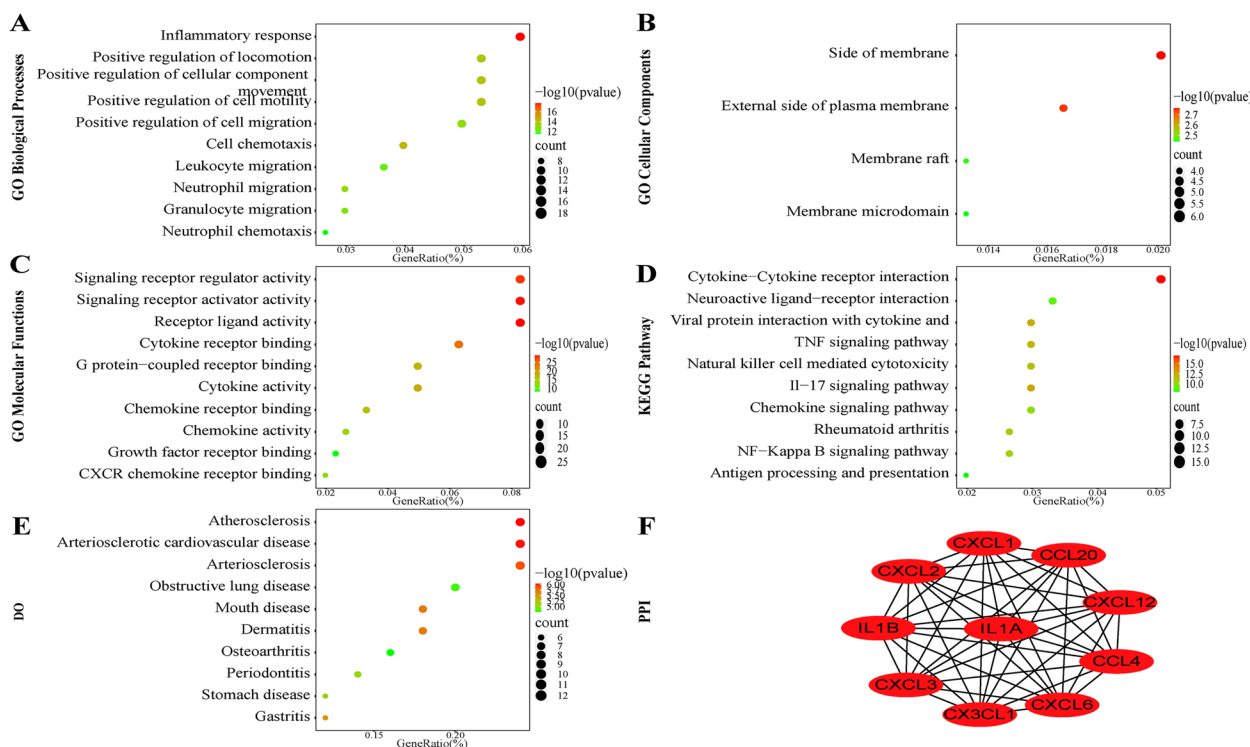


Fig. 3 GO, KEGG, and DO pathway enrichment analysis. **A** BP of 58 DIRGs. **B** CC of 58 DIRGs. **C**: MF of 58 DIRGs. **D** KEGG pathway of 58 DIRGs. **E** DO pathway of 58 DIRGs. **F** PPI of hub genes

Analysis of the PPI network

A PPI network of 58 DIRGs was utilized on the STRING website and visualized by Cytoscape. MCODE plug-ins were performed to screen out the most vital function modules in the total protein-protein networks. Significant modules consisted of key genes with 10 nodes and 90 edges, including IL1B, IL1A, CXCL1, CXCL2, CXCL6, CXCL3, CXCL12, CX3CL1, CCL4, and CCL20 (Fig. 3F).

Significant DIRGs selected by PLS, GLM, and SVM algorithms

To investigate the connection between key IRGs and the prevalence of MI, PLS, RF, KNN, and SVM were independently constructed to narrow down key IRGs using the new data cohort. The importance of explanatory features and residual distribution of four models were analyzed and visualized by the DALEX package in the new data cohort. The PLS model was demonstrated to be the best suitable model with the lowest residual (Fig. 4A and B). The importance of 58 DIRGs based on the above four models is shown in Fig. 4C. The intersection of six genes in the four models of minimal RMSE was PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM. These six genes were applied for further analysis.

Further analysis of six DIRGs

The six most vital explanatory variables (PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM) were selected for further analysis. The expression of PTGER2 and ADM in MI patients was more highly expressed than in healthy patients, while the other four genes (LGR6, IL17B, IL13RA1, and CCL4) were low expressed in MI patients (Fig. 5A). The correlations of those genes are analyzed in Fig. 5B, in which IL13RA1 was found to be positively related to LGR6, IL17B, CCL4, and ADM, and IL17B was positively related to ADM and CCL4.

A nomogram model constructed and assessed for diagnosing MI

The nomogram model was implemented for diagnosing MI based on six DIRGs (PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM) using the rms package in R (Fig. 6A). The nomogram constructed by multivariable logistic regression was shown in Table 2. Independent risk factors for incidence of MI included PTGER2 gene (OR:1.495,95%:1.239–1.866, $p < 0.001$), ADM gene (OR:5.817,95%:2.955–12.649, $p < 0.001$), IL17B gene (OR:0.322,95%:0.158–0.615, $p = 0.001$), IL13RA1 gene (OR:0.245,95%:0.098–0.562, $p = 0.001$), and CCL4 gene

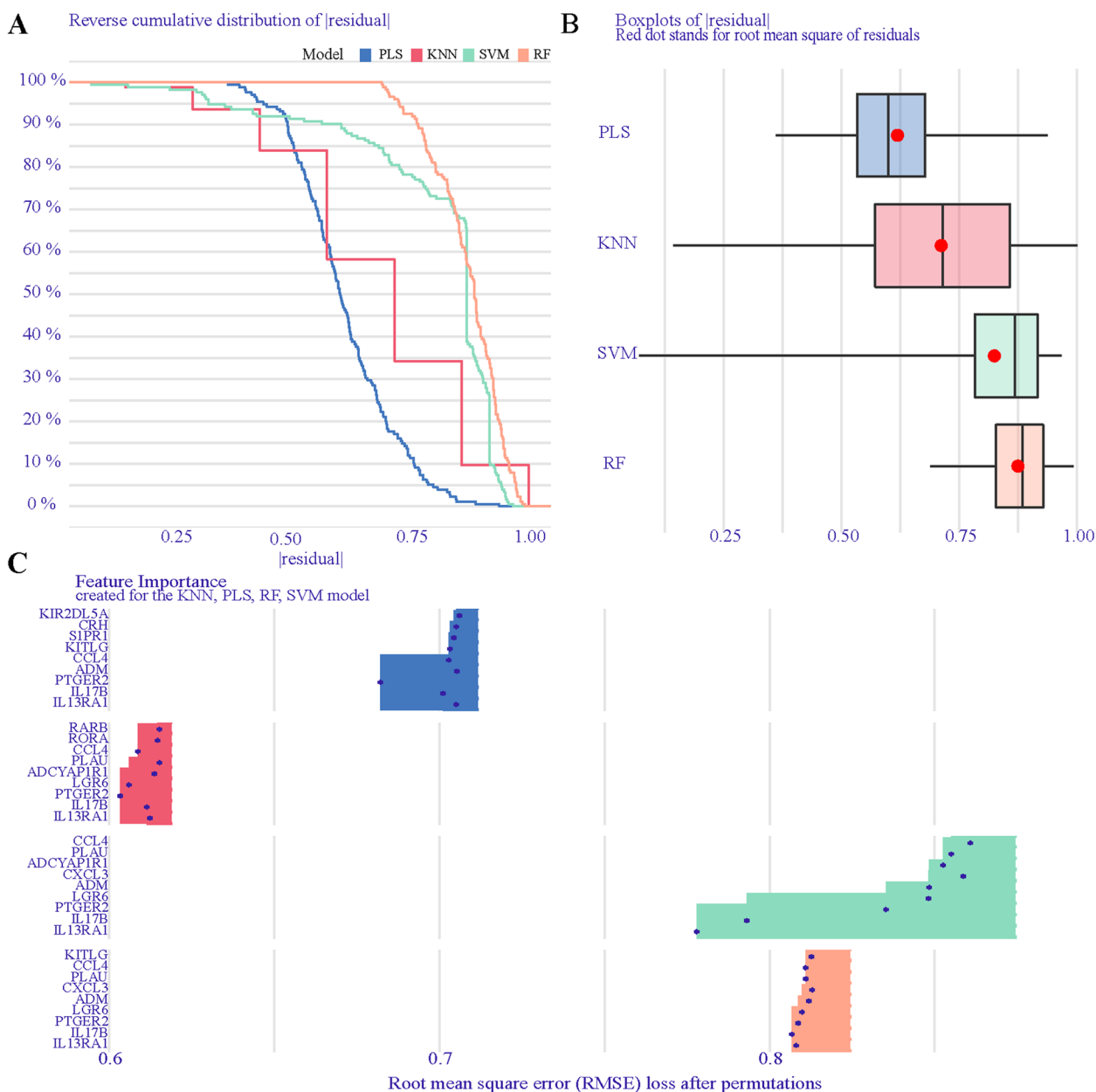


Fig. 4 Significant DIRGs of the RF, PLS, SVM, and KNN models. **A** Cumulative residual distribution of the RF, PLS, SVM, and KNN models. **B** Boxplots of the residuals in the RF, PLS, SVM, and KNN models. **C** The importance of significant DIRGs in the RF, PLS, SVM, and KNN models

(OR:0.399,95%:0.226–0.673, $p=0.001$). The Area Under Curve(AUC) was implemented to assess the discrimination ability of the nomogram model, which exhibited the highest predictive accuracy compared with the six above DIRGs (Fig. 6B), while the calibration curve demonstrated a small error between the actual MI incidence and predicted incidence (Fig. 6C). The decision curve (DCA) demonstrated that the nomogram model exhibited better potential clinical utility than the other curves, indicating that patients with MI can benefit from the nomogram

model at high-risk threshold probabilities ranging from 0 to 0.8 (Fig. 6D).

Distribution of immune cells in MI

To explore the association between immune cell infiltration and the incidence of MI, the 22 types of immune cell infiltration in each sample were calculated by CIBERSORT and visualized by histogram (Fig. 7A). Twenty-two immune cell infiltration between MI patients and normal

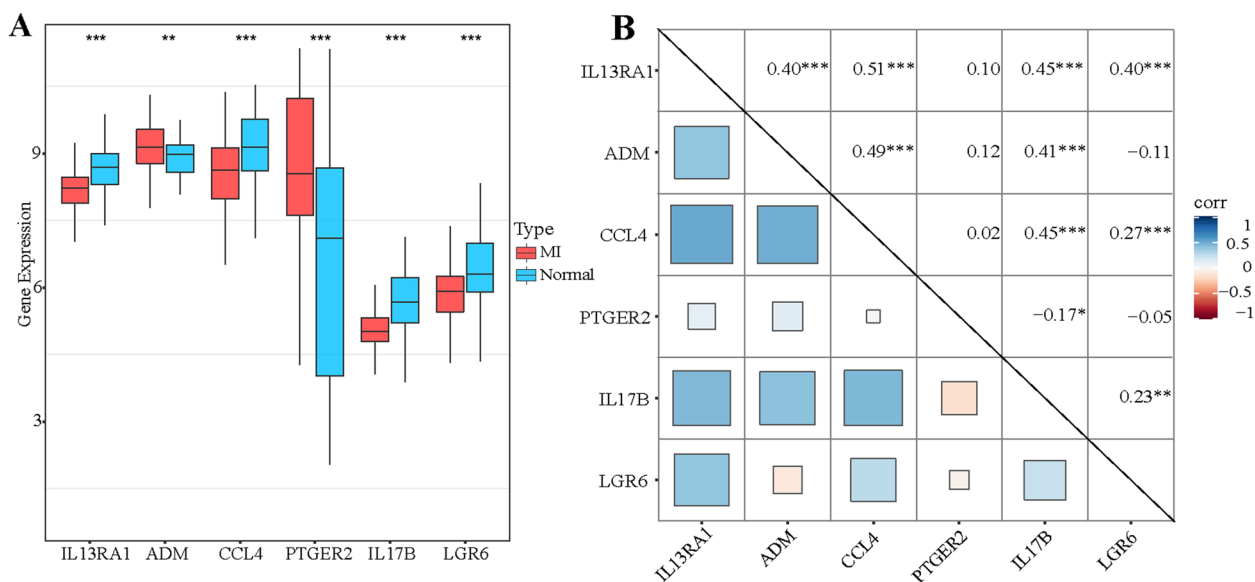


Fig. 5 **A** Relative expression level of PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM. **B** Correlation among PTGER2, LGR6, IL17B, IL13RA1, CCL4 and ADM

controls were visualized by boxplot (Fig. 7B). We found that plasma cells, T cells follicular helper, mast cells resting, and neutrophils were higher in MI patients than in healthy patients, while the T cells CD4 naive, macrophages M1, macrophages M2, dendritic cells resting, and mast cells activated in MI patients were lower.

The association between diagnostic DIRGs and infiltrating immune cells of MI

The IL13RA1 gene was positively related to T cell CD4 memory resting. The ADM gene was positively related to T cell CD4 memory resting. The CCL4 gene was negatively related to Natural Killer (NK) cell resting. The PTGER2 gene was positively correlated with T cell CD4 memory resting. The IL17B gene was negatively correlated with neutrophils. The LGR6 gene was negatively correlated with T cells and the follicular helper (Fig. 8).

Discussion

MI, which is considered one of the most severe complications of CAD, is the main cause of mortality globally [1]. The short-term mortality of patients with MI can be improved by early diagnosis through appropriate medical therapy and revascularization blockage of the coronary artery [2]. Infarction of cardiomyocytes can lead to left ventricular dilatation dysfunction, which eventually contributes to heart failure with high long-term mortality after MI [24]. MI is strongly related to the immune system, but the underlying mechanism of inflammatory mediators in the progress of MI is poorly understood.

Thus, the role of inflammatory mediators has been investigated using machine learning. A nomogram model was constructed for the diagnosis of MI based on the DIRGs.

Injured cardiomyocytes can secrete damage-associated molecular patterns (DAMPs) containing heat-shock proteins and mitochondrial DNA [25], which can initiate the immune response by recognizing and linking to extracellular or intracellular pattern recognition receptors (PRRs) on immune cells [24]. However, the role of other inflammatory mediators remains unknown in MI. In our study, 58 DIRGs were enriched in the GO, KEGG, and DO pathway analysis. The six key IRGs (PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM) were selected by four machine learning methods, such as PLS, RE, KNN, and SVM, and were constructed for further analysis.

PTGER2, which encodes a receptor for prostaglandin E2, can increase intracellular cAMP concentration and initiate smooth muscle cell relaxation [26]. Almudena reported that PTGER2 was significantly expressed in atherosclerotic plaques in humans. The chemokine production of human macrophages was suppressed by PGE2 through PTGER2 in the progression of atherosclerosis plaque [27]. Leitinger found that, when highly expressed in endothelial cells and monocytes, PTGER2 can induce vascular inflammation and atherogenesis [28]. PTGER2 may be a drug target in atherosclerosis. PTGER2 was upregulated in MI patients compared with normal controls in our study, which was positively related to T cell CD4 memory resting.

LGR6 is a subgroup of the leucine-rich-repeat-containing G protein-coupled receptor (LGR) superfamilies.

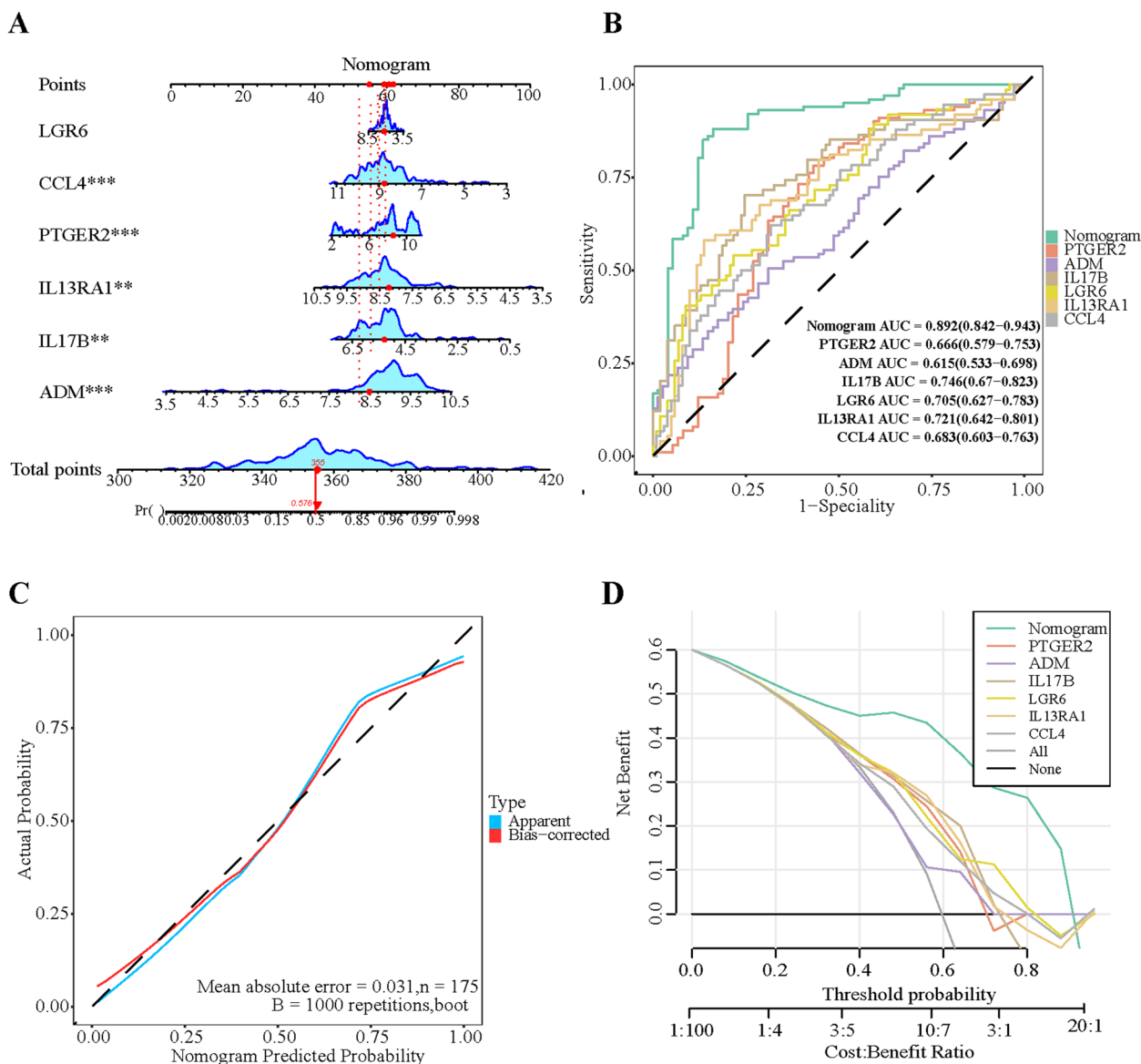


Fig. 6 Validation and assessment of a nomogram model for MI diagnosis. **A** Nomogram model. **B** The AUC of the nomogram model in predicting the incidence of MI. **C** Calibration curve to assess the predictive value. **D** DCA curve to evaluate the clinical value

Table 2 Results of multivariable logistic regression

	OR	2.50%	97.5%	95%CI	P value
PTGER2	1.495	1.239	1.866	1.495(1.239–1.866)	$p < 0.001$
ADM	5.817	2.955	12.649	5.817(2.955–12.649)	$p < 0.001$
IL17B	0.322	0.158	0.615	0.322(0.158–0.615)	0.001
LGR6	0.739	0.418	1.276	0.739(0.418–1.276)	0.282
IL13RA1	0.245	0.098	0.562	0.245(0.098–0.562)	0.001
CCL4	0.399	0.226	0.673	0.399(0.226–0.673)	0.001

Ruan et al. reported that LGR6 plays a significant role in the chemoresistance of ovarian cancer by potentiating the Wnt/ β -catenin signaling [29]. Chiang et al. reported that LGR6 can enhance phagocytosis and efferocytosis of M Φ and initiate intracellular phosphorylation signaling in neutrophils by binding to maresin 1 [30]. Our study showed that LGR6 was downregulated in MI and that LGR6 was positively related to NK cell activation.

IL-17B is a proinflammatory cytokine. Irez-Carrozzi et al. demonstrated that IL-17B is associated with inflammatory disease, which can trigger type 2 immune responses from NKT, CD4+CRTH2+ Th2, and innate

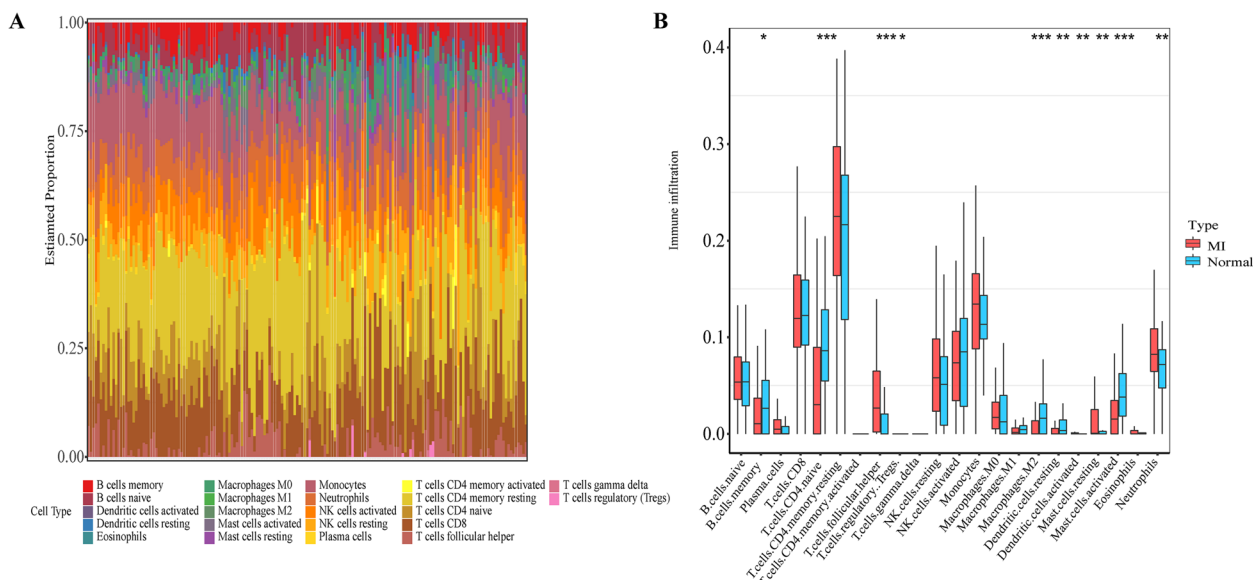


Fig. 7 A Distribution of the immune cells of all samples. B Different distribution of 22 immune cells between patients with MI and healthy controls

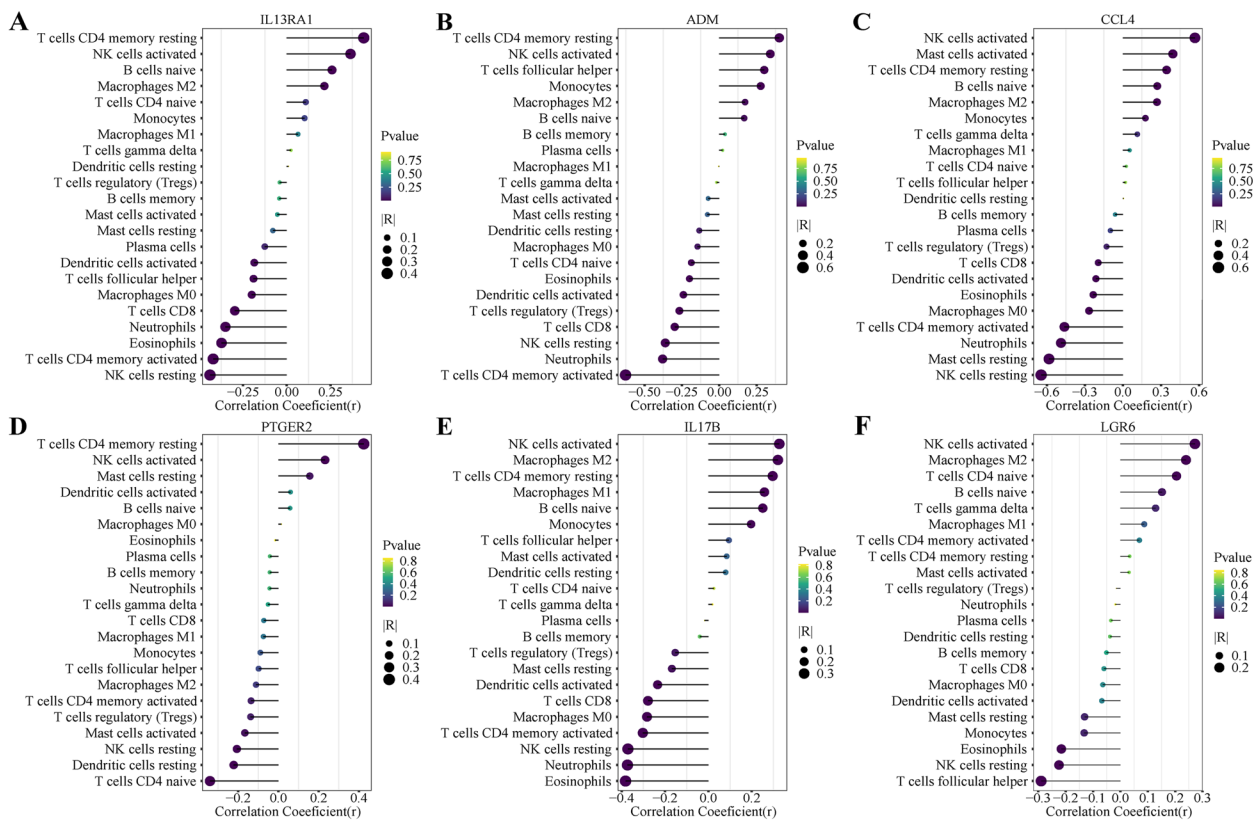


Fig. 8 Correlation between IL13RA1 (A), ADM (B), CCL4 (C), PTGER2 (D), IL17B (E), LGR6 (F) and infiltrating immune cells in myocardial infarction

type 2 lymphocytes (ILC2s) [31]. Zhou et al. found that the expression of IL-17B was significantly highly expressed in patients with community-acquired pneumonia, and IL-17B could upregulate the expression of IL-8 by initiating p38 mitogen-activated protein kinase (MAPK) and extracellular signal-regulated kinase (ERK) in human bronchial epithelial cells [32]. IL-17B was found to be downregulated in MI patients in the current study, and this was found to be positively related to NK cell activation.

IL-13 is a cytokine participating in normal immune function. Gwiggner et al. suggested that IL-13 was secreted by activated type 2 T helper (Th2) cells. The mechanism of IL-13 signaling, here via binding to IL-13 receptor α -1 and IL-13 receptor α -2, started by initiating phosphorylation of the signal transducer and activating transcription 6 (STAT6) via Janus kinases (JAK) [33]. Amit et al. reported that IL-13RA1 participated in the homeostasis and repair of the myocardium. IL-13RA1 signaling was significant for cardiac, including extracellular matrix integrity. Stimulation of IL-13RA1/STAT3 signaling can induce the excessive accumulation of included extracellular matrix, cardiac fibrosis, chronic cardiac stress, and heart failure [34]. The IL-13RA1 was found to be downregulated in MI patients in the current study, and IL-13RA1 was positively associated with T cell CD4 memory resting.

CCL4 is a significant chemotactic mediator for recruiting macrophages [35]. CCL4 is related to the pathogenesis of several diseases, including sarcoidosis [36], cystic fibrosis [37], and multiple sclerosis [38]. Kalinskaya et al. reported that non-ST-elevation myocardial infarction (NSTEMI) patients demonstrated significantly higher expression of CCL4 compared with ST-elevation myocardial infarction (STEMI) patients. The synergy of TNF- α and CCL4 in STEMI patients can be lowly expressed in monocytes that are mediated and increased through the adhesion of leukocytes by TNF- α [39]. CCL4 was downregulated in MI patients in the current study and was positively associated with NK cell activation.

Various tissues, such as the myocardium, adrenal medulla, and central nervous system, can secrete ADM [40]. ADM plays a prominent role in vasodilation, stimulation of angiogenesis, and the production of NO [41]. Ali reported that the ADM expression levels were highly elevated in the plasma of hypertension patients [42]. Previous studies have demonstrated that the vasodilation effects of ADM were mediated by cyclic adenosine 3,5-monophosphate and nitric-oxide-dependent mechanisms. ADM can regulate myocardial protection by disrupting mitochondrial metabolism and lowering the renin-aldosterone system levels in cardiovascular diseases [43]. The ADM was highly regulated in MI patients

in the current study and positively associated with T cell CD4 memory resting.

Acute myocardial infarction (AMI) is generally diagnosed by typical symptoms, electrocardiographic changes, and traditional biomarkers, such as hs-cTnT, hs-cTnI, and creatine kinase-MB, but traditional biomarkers have been demonstrated with high sensitivity but without specificity. Previous studies have shown that ncRNAs appears to provide better sensitivity and specificity in diagnosis of MI. MiR-1 participates in regulation of cardiac development and differentiation of other cell types to cardiomyocytes [44]. miR-1 was highly correlated with cTnT in MI patients and demonstrated with a high specificity value (0.82), and sensitivity value (0.73) [45]. miR-499 plays a significant role in cardiac cell recovery and stem cell differentiation [46]. miR-499 had the most accurate predictive value (AUC of 0.91, sensitivity of 0.83, and specificity of 0.90) in distinguishing MI and control [47]. LncRNA H19 had the relatively high predictive value (AUC of 0.753, sensitivity of 0.609, and specificity of 0.817) in MI patients and healthy control, and positively correlated with troponinT ($r=0.344$, $p<0.001$), CK ($r=0.261$, $p=0.001$) and CKMB ($r=0.24$, $p=0.002$) [47, 48]. Compared to ncRNAs, the six DIRGs had shown the relatively high predictive value, and the nomogram model exhibited the highest predictive accuracy (AUC of 0.892, sensitivity of 0.881, and specificity of 0.838) in my study. The nomogram model was constructed for MI diagnosis using PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM. The model had the highest predictive accuracy to distinguish MI and normal patients, which could be applied to identify high-risk groups from 0 to 1 (0 for normal, 1 for MI).

The distribution of 22 immune cells in patients with MI and normal controls was calculated using CIBERSORT. The ratio of T cells follicular helper, mast cells resting, and neutrophils was higher in MI patients, while the ratio of T cells CD4 naive, macrophages M1, macrophages M2, dendritic cells resting, and mast cells activated was lower. The correlation analysis between six key IRGs (PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM) and immune cells all were shown to be associated with T cell CD4 memory resting, NK cells activated, and macrophages M2. Admittedly, inflammatory mediators, such as IRGs and immune cells, have been shown to play an important role in the progression of MI. Intercellular adhesion molecule 1 was involved in lymphocyte migration into the intima. T lymphocytes were found to activate to secrete various cytokines that interact with macrophages. T lymphocytes and macrophages can be activated by the engagement of CD40/CD40L to initiate the production of tissue factor and cytokines that enhance the body's inflammatory responses [49].

Ortega-Rodríguez reported that the total level of NK cells was elevated in MI patients compared with normal controls. The various phenotypes of NK cells play a different role in the progress of MI. The level of natural killer group 2, member D(NKG2D)+NK cells in peripheral blood tends to decrease in MI patients, which may demonstrate that NKG2D cells migrate to the injured myocardium [50]. The other phenotypes of NK cells mobilize from the myocardium to the peripheral blood to mediate the inflammatory response. The evidence demonstrated that several types of infiltrating immune cells play significant roles in the progress of MI and can be the focus of future therapeutic targets [51].

There are also some limitations to our study. First, the sample size of the merged datasets of MI patients and normal controls was relatively small. Second, the expression of PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM should be validated in datasets with a larger sample size by quantitative polymerase chain reaction or Western blot. Third, the characteristics of clinical information did not exist in the GEO datasets, and the predictive value of the nomogram model and traditional biomarkers in the diagnosis of MI could not be compared. Finally, immune cells inferred from the DIRGs may have a prognostic utility in the clinic to identify high risk MI patients, although this needs to be validated in larger cohorts.

Conclusion

The six IRGs, such as PTGER2, LGR6, IL17B, IL13RA1, CCL4, and ADM, were selected by machine learning associated with the occurrence of MI. The nomogram constructed by the six IRGs above was demonstrated as having higher predictive accuracy in the diagnosis of MI. T cell CD4 memory resting, NK cells activated, and macrophages M2 may participate in the advancement of MI.

Authors' contributions

(I) Conception and design: H Dong; (II) Administrative support: G Chen, MJ Li; (III) Provision of study materials or patients: SB Yan; GS Li; (IV) Collection and assembly of data: ZG Huang, DM Li, YL Tang, JQ Le, YF Pan, Z Yang, HB Pan; (V) Data analysis and interpretation: H Dong, ZG Huang, DM Li, YL Tang, JQ Le, YF Pan, Z Yang, HB Pan; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Funding

This work was supported by Guangxi Higher Education Undergraduate Teaching Reform Project (2022JGA146, 2021JGA142) and Guangxi Medical University Education and Teaching Reform Project (2022XJGY31, 2022XJGY78).

Availability of data and materials

The gene datasets in this study including GSE29111, GSE48060, GSE66360, GSE97320, and GSE113079 can be found in Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>).

Declarations

Ethics approval and consent to participate

All procedures conducted in studies consisting of human participants were in accordance with the ethical standards of the institutional and/or national

research committee and with the 1964 Helsinki declaration as well as its later amendments or comparable ethical standards. All gene datasets of this study were based on data mining of public high-throughput sources that did not require approval from our institutional review board.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Cardiovascular Medicine, The First Affiliated Hospital of Guangxi Medical University, No.6 Shuangyong Road, Nanning, Guangxi Zhuang Autonomous Region 530021, People's Republic of China. ²Department of Pathology/ Forensic Medicine, The First Affiliated Hospital of Guangxi Medical University, No.6 Shuangyong Road, Nanning, Guangxi Zhuang Autonomous Region 530021, People's Republic of China. ³Department of Cardiothoracic Surgery, The First Affiliated Hospital of Guangxi Medical University, No.6 Shuangyong Road, Nanning, Guangxi Zhuang Autonomous Region 530021, People's Republic of China. ⁴Department of Pathology, Hospital of Guangxi Liugang Medical Co.LTD./Guangxi Liuzhou Dingshun Forensic Expert Institute, No.9, Queershan Rd, Liuzhou, Guangxi Zhuang Autonomous Region 545002, People's Republic of China. ⁵Department of Gerontology, NO.923 Hospital of Chinese People's Liberation Army, No. 1 Tangcheng Rd, Nanning, Guangxi Zhuang Autonomous Region 530021, People's Republic of China.

Received: 4 October 2022 Accepted: 22 March 2023

Published online: 28 March 2023

References

- Collaborators GBDCoD. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1736–88.
- Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, Chiuve SE, Cushman M, Dellings FN, Deo R, et al. Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. *Circulation*. 2018;137(12):e67–492.
- Collet JP, Thiele H, Barbato E, Barthelémy O, Bauersachs J, Bhatt DL, Dendale P, Dorobantu M, Edvardsen T, Folliguet T, et al. 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *Eur Heart J*. 2021;42(14):1289–367.
- Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, Prescott E, Storey RF, Deaton C, Cuisset T, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J*. 2020;41(3):407–77.
- Jensen RV, Hjortbak MV, Botker HE. Ischemic Heart Disease: An Update. *Semin Nucl Med*. 2020;50(3):195–207.
- Han H, Du R, Cheng P, Zhang J, Chen Y, Li G. Comprehensive Analysis of the Immune Infiltrates and Aberrant Pathways Activation in Atherosclerotic Plaque. *Front Cardiovasc Med*. 2020;7:602345.
- 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(1):231.
- Yang Y, Xu X. Identification of key genes in coronary artery disease: an integrative approach based on weighted gene co-expression network analysis and their correlation with immune infiltration. *Aging (Albany NY)*. 2021;13(6):8306–19.
- Forteza MJ, Trapero I, Hervas A, de Dios E, Ruiz-Sauri A, Minana G, Bonanad C, Gomez C, Oltra R, Rios-Navarro C, et al. Apoptosis and Mobilization of Lymphocytes to Cardiac Tissue Is Associated with Myocardial Infarction in a Reperfused Porcine Model and Infarct Size in Post-PCI Patients. *Oxid Med Cell Longev*. 2018;2018:1975167.
- Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243–59.

11. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, Hu Z, Zalocusky KA, Shankar RD, Shen-Orr SS, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci Data*. 2018;5:180015.
12. 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(7):e47.
13. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 2004;20(3):307–15.
14. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28(6):882–3.
15. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
16. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. 2019;28(11):1947–51.
17. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51(D1):D587–92.
18. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
19. Yu G, Wang LG, Yan GR, He QY. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*. 2015;31(4):608–9.
20. Deidda M, Piras C, Binaghi G, Congia D, Pani A, Boi A, Sanna F, Rossi A, Loi B, Cadeddu Dessalvi C, et al. Metabolomic fingerprint of coronary blood in STEMI patients depends on the ischemic time and inflammatory state. *Sci Rep*. 2019;9(1):312.
21. Cutler A, Stevens JR. Random forests for microarrays. *Methods Enzymol*. 2006;411:422–32.
22. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. *Cancer Genomics Proteomics*. 2018;15(1):41–51.
23. Wang C, Long Y, Li W, Dai W, Xie S, Liu Y, Zhang Y, Liu M, Tian Y, Li Q, et al. Exploratory study on classification of lung cancer subtypes through a combined K-nearest neighbor classifier in breathomics. *Sci Rep*. 2020;10(1):5880.
24. Anzai A, Ko S, Fukuda K. Immune and Inflammatory Networks in Myocardial Infarction: Current Research and Its Potential Implications for the Clinic. *Int J Mol Sci*. 2022;23(9):5214.
25. Epelman S, Liu PP, Mann DL. Role of innate and adaptive immune mechanisms in cardiac injury and repair. *Nat Rev Immunol*. 2015;15(2):117–29.
26. Bryson TD, Harding P. Prostaglandin E2 EP receptors in cardiovascular disease: An update. *Biochem Pharmacol*. 2022;195:114858.
27. Gomez-Hernandez A, Martin-Ventura JL, Sanchez-Galan E, Vidal C, Ortego M, Blanco-Colio LM, Ortega L, Tunon J, Egido J. Overexpression of COX-2, Prostaglandin E synthase-1 and prostaglandin E receptors in blood mononuclear cells and plaque of patients with carotid atherosclerosis: regulation by nuclear factor-kappaB. *Atherosclerosis*. 2006;187(1):139–49.
28. Leitinger N. A rancid culprit in vascular inflammation acts on the prostaglandin receptor EP2. *Circ Res*. 2006;98(5):587–9.
29. Ruan X, Liu A, Zhong M, Wei J, Zhang W, Rong Y, Liu W, Li M, Qing X, Chen G, et al. Silencing LGR6 Attenuates Stemness and Chemoresistance via Inhibiting Wnt/beta-Catenin Signaling in Ovarian Cancer. *Mol Ther Oncolytics*. 2019;14:94–106.
30. Chiang N, Libreros S, Norris PC, de la Rosa X, Serhan CN. Maresin 1 activates LGR6 receptor promoting phagocyte immunoresolvent functions. *J Clin Invest*. 2019;129(12):5294–311.
31. Ramirez-Carrozzi V, Ota N, Sambandam A, Wong K, Hackney J, Martinez-Martin N, Ouyang W, Pappu R. Cutting Edge: IL-17B Uses IL-17RA and IL-17RB to Induce Type 2 Inflammation from Human Lymphocytes. *J Immunol*. 2019;202(7):1935–41.
32. Zhou J, Ren L, Chen D, Lin X, Huang S, Yin Y, Cao J. IL-17B is elevated in patients with pneumonia and mediates IL-8 production in bronchial epithelial cells. *Clin Immunol*. 2017;175:91–8.
33. Gwiggner M, Martinez-Nunez RT, Whiteoak SR, Bondanese VP, Claridge A, Collins JE, Cummings JRF, Sanchez-Elsner T. MicroRNA-31 and MicroRNA-155 Are Overexpressed in Ulcerative Colitis and Regulate IL-13 Signaling by Targeting Interleukin 13 Receptor alpha-1. *Genes (Basel)*. 2018;9(2):85.
34. Amit U, Kain D, Wagner A, Sahu A, Nevo-Caspi Y, Gonen N, Molotski N, Konfino T, Landa N, Naftali-Shani N, et al. New Role for Interleukin-13 Receptor alpha1 in Myocardial Homeostasis and Heart Failure. *J Am Heart Assoc*. 2017;6(5):e005108.
35. Kochumon S, Wilson A, Chandy B, Shenouda S, Tuomilehto J, Sindhu S, Ahmad R. Palmitate Activates CCL4 Expression in Human Monocytic Cells via TLR4/MyD88 Dependent Activation of NF-kappaB/MAPK/ PI3K Signaling Systems. *Cell Physiol Biochem*. 2018;46(3):953–64.
36. Barczyk A, Pierzchala E, Caramori G, Sozanska E. Increased expression of CCL4/MIP-1beta in CD8+ cells and CD4+ cells in sarcoidosis. *Int J Immunopathol Pharmacol*. 2014;27(2):185–93.
37. Mrugacz M. CCL4/MIP-1beta levels in tear fluid and serum of patients with cystic fibrosis. *J Interferon Cytokine Res*. 2010;30(7):509–12.
38. Boven LA, Montagne L, Nottet HS, De Groot CJ. Macrophage inflammatory protein-1alpha (MIP-1alpha), MIP-1beta, and RANTES mRNA semiquantification and protein expression in active demyelinating multiple sclerosis (MS) lesions. *Clin Exp Immunol*. 2000;122(2):257–63.
39. Kalinskaya A, Dukhin O, Lebedeva A, Maryukhnich E, Rusakovich G, Vorobyeva D, Shpektor A, Margolis L, Vasilieva E. Circulating Cytokines in Myocardial Infarction Are Associated With Coronary Blood Flow. *Front Immunol*. 2022;13:837642.
40. Schonauer R, Els-Heindl S, Beck-Sickinger AG. Adrenomedullin - new perspectives of a potent peptide hormone. *J Pept Sci*. 2017;23(7–8):472–85.
41. Voors AA, Kremer D, Geven C, Ter Maaten JM, Struck J, Bergmann A, Pickkers P, Metra M, Mebazaa A, Dungen HD, et al. Adrenomedullin in heart failure: pathophysiology and therapeutic application. *Eur J Heart Fail*. 2019;21(2):163–71.
42. Ali F, Khan A, Muhammad SA, Hassan SSU. Quantitative Real-Time Analysis of Differentially Expressed Genes in Peripheral Blood Samples of Hypertension Patients. *Genes (Basel)*. 2022;13(2):187.
43. Murakami S, Kimura H, Kangawa K, Nagaya N. Physiological significance and therapeutic potential of adrenomedullin in pulmonary hypertension. *Cardiovasc Hematol Disord Drug Targets*. 2006;6(2):125–32.
44. Chistiakov DA, Orekhov AN, Bobryshev YV. Cardiac-specific miRNA in cardiogenesis, heart function, and cardiac pathology (with focus on myocardial infarction). *J Mol Cell Cardiol*. 2016;94:107–21.
45. Liu X, Fan Z, Zhao T, Cao W, Zhang L, Li H, Xie Q, Tian Y, Wang B. Plasma miR-1, miR-208, miR-499 as potential predictive biomarkers for acute myocardial infarction: An independent study of Han population. *Exp Gerontol*. 2015;72:230–8.
46. Zhang L, Chen X, Su T, Li H, Huang Q, Wu D, Yang C, Han Z. Circulating miR-499 are novel and sensitive biomarker of acute myocardial infarction. *J Thorac Dis*. 2015;7(3):303–8.
47. Zhao J, Yu H, Yan P, Zhou X, Wang Y, Yao Y. Circulating MicroRNA-499 as a Diagnostic Biomarker for Acute Myocardial Infarction: A Meta-analysis. *Dis Markers*. 2019;2019:6121696.
48. Wang XM, Li XM, Song N, Zhai H, Gao XM, Yang YN. Long non-coding RNAs H19, MALAT1 and MIAT as potential novel biomarkers for diagnosis of acute myocardial infarction. *Biomed Pharmacother*. 2019;118:109208.
49. Jebari-Benslaiman S, Galicia-Garcia U, Larrea-Sebal A, Olaetxea JR, Alloza I, Vandenbroeck K, Benito-Vicente A, Martin C. Pathophysiology of Atherosclerosis. *Int J Mol Sci*. 2022;23(6):3346.
50. Ortega-Rodriguez AC, Marin-Jauregui LS, Martinez-Shio E, Hernandez Castro B, Gonzalez-Amaro R, Escobedo-Urbe CD, Monsivais-Urenda AE. Altered NK cell receptor repertoire and function of natural killer cells in patients with acute myocardial infarction: A three-month follow-up study. *Immunobiology*. 2020;225(3):151909.
51. Montecucco F, Brauner-Reuther V, Lenglet S, Delattre BM, Pelli G, Bua-tois V, Guilhot F, Galan K, Vuilleumier N, Ferlin W, et al. CC chemokine CCL5 plays a central role impacting infarct size and post-infarction heart failure in mice. *Eur Heart J*. 2012;33(15):1964–74.

Publisher's Note

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