



OPEN ACCESS

EDITED BY

Margreet R. De Vries,
Leiden University Medical Center
(LUMC), Netherlands

REVIEWED BY

Daniel Petrovič,
University of Ljubljana, Slovenia
Teresa Padro,
Institut de Recerca de l'Hospital de la
Santa Creu i Sant Pau, Spain

*CORRESPONDENCE

Yue Guan
✉ guanyuedmu@163.com

†These authors have contributed
equally to this work and share first
authorship

SPECIALTY SECTION

This article was submitted to
Atherosclerosis and Vascular Medicine,
a section of the journal
Frontiers in Cardiovascular Medicine

RECEIVED 27 September 2022

ACCEPTED 15 December 2022

PUBLISHED 06 January 2023

CITATION

Liu C, Liu J, Zhang Y, Wang X and
Guan Y (2023) Immune-related
potential biomarkers and therapeutic
targets in coronary artery disease.
Front. Cardiovasc. Med. 9:1055422.
doi: 10.3389/fcvm.2022.1055422

COPYRIGHT

© 2023 Liu, Liu, Zhang, Wang and
Guan. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Immune-related potential biomarkers and therapeutic targets in coronary artery disease

Chaosheng Liu^{1†}, Jifeng Liu^{2†}, Yunshu Zhang³, Xi Wang⁴ and Yue Guan^{5*}

¹Department of Cardiology, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China, ²Department of General Surgery, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China, ³Department of Traditional Chinese Medicine, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China, ⁴Department of Critical Medicine, The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China, ⁵Department of Cardiology, The Third Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China

Background: Coronary artery disease (CAD) is a complex illness with unknown pathophysiology. Peripheral biomarkers are a non-invasive method required to track the onset and progression of CAD and have unbeatable benefits in terms of early identification, prognostic assessment, and categorization of the diagnosis. This study aimed to identify and validate the diagnostic and therapeutic potential of differentially expressed immune-related genes (DE-IRGs) in CAD, which will aid in improving our knowledge on the etiology of CAD and in forming genetic predictions.

Methods: First, we searched coronary heart disease in the Gene Expression Omnibus (GEO) database and identified GSE20680 (CAD = 87, Normal = 52) as the trial set and GSE20681 (CAD = 99, Normal = 99) as the validation set. Functional enrichment analysis using protein-protein interactions (PPIs), Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) was carried out on the identified differentially expressed genes. Optimal feature genes (OFGs) were generated using the support vector machine recursive feature elimination algorithm and the least absolute shrinkage and selection operator (LASSO) algorithm. Furthermore, immune infiltration in CAD patients and healthy controls was compared using CIBERSORT, and the relationship between immune cells and OFGs was examined. In addition, we constructed potential targeted drugs for this model through the Drug-Gene Interaction database (DGldb) database. Finally, we verify the expression of S100A8-dominated OFGs in the GSE20681 dataset to confirm the universality of our study.

Results: We identified the ten best OFGs for CAD from the DE-IRGs. Functional enrichment analysis showed that these marker genes are crucial for receptor-ligand activity, signaling receptor activator activity, and positive

control of the response to stimuli from the outside world. Additionally, CIBERSORT revealed that S100A8 could be connected to alterations in the immune microenvironment in CAD patients. Furthermore, with the help of DGIdb and Cytoscape, a total of 64 medicines that target five marker genes were subsequently discovered. Finally, we verified the expression of the OFGs genes in the GSE20681 dataset between CAD patients and normal patients and found that there was also a significant difference in the expression of S100A8.

Conclusion: We created a 10-gene immune-related prognostic model for CAD and confirmed its validity. The model can identify potential biomarkers for CAD prediction and more accurately gauge the progression of the disease.

KEYWORDS

coronary artery disease (CAD), immune-related genes (IRGs), optimal feature genes (OFGs), bioinformatics analysis, Gene Expression Omnibus (GEO)

1. Introduction

In coronary artery disease (CAD), myocardial ischemia, hypoxia, and necrosis are the result of narrowing or obstruction of the lumen caused by atherosclerotic lesions of the coronary arteries (1, 2). CAD is one of the deadliest diseases worldwide, with 12 million people per year expected to die from CAD by 2030 (3). The morbidity and death rates associated with CAD are rising in low- and middle-income countries and are now on par with those in developed nations, making it an issue worldwide (3, 4). Atherosclerosis progresses slowly over decades and due to sub-typical symptoms, the start of CAD is frequently overlooked (5). The pathogenesis of CAD is complicated and remains to be fully understood. Although coronary angiography is the most effective method for diagnosing CAD, it is intrusive and expensive. To develop predictive, diagnostic, or prognostic tools for CAD, more combinations of biomarkers must be included utilizing various techniques. Over time, atherosclerosis is suggested to be a lipid storage disease; however, it is becoming increasingly clear that inflammation links dyslipidemia and other risk factors to atherosclerosis formation, which ultimately leads to the formation of coronary heart disease (6). Therefore, it is highly feasible to analyze differences in the expression of immune-related genes (IRGs) between CAD patients and healthy controls to identify and validate the diagnostic and therapeutic in CAD.

CAD patients frequently experience changes in the immunological components of their peripheral blood throughout their illness, which have been linked to acute exacerbations, remission, and stability of the disease (7, 8). Thus, these immunological components can also be used as therapeutic targets or biomarkers for disease monitoring. The secondary prevention of CAD depends on accurate

detection and subsequent mechanistic investigation into these immunological components. The pathophysiology and development of CAD are thought to be significantly influenced by immunological diseases (9). Before the onset of the pathogenic phenomena, specific immunological changes occur (10). However, as the result of the current sample size and project scope remain constrained, these novel targets are rarely verified or mechanistically investigated further, which restricts their translation to the clinic (11, 12). Therefore, despite compelling data, further research is needed to fully understand the peripheral immunological characteristics and processes of CAD.

From the perspective of the primary prevention of CAD, our concern is whether new risk markers for atherosclerosis can improve CAD risk prediction. However, existing biomarkers, including heat shock C-reactive protein (hsCRP), are intertwined with inflammation, oxidation, hemostasis, and other processes involving atherosclerosis, and the results are not ideal (13). In addition, many histopathologies have also shown that inflammation is closely related to CAD, and monocytes and mast cells play a crucial role in the chronic inflammatory response (14). Therefore, it is of clinical significance to establish an immune-related prognostic model to identify and validate the diagnostic and therapeutic in CAD. In this study, we developed a new diagnostic model of optimal feature genes (OFGs) using differentially expressed immune-related genes (DE-IRGs) in CAD patients. We then performed correlation analysis and testing, which demonstrated that the model, particularly the S100A8 gene, has strong potential as a target for diagnostic and therapeutic interventions. This study aimed to identify and validate the diagnostic and therapeutic potential of DE-IRGs in CAD, which will aid in improving our knowledge on the etiology of CAD and in forming genetic predictions.

2. Materials and methods

2.1. Data preparation

Gene expression information for CAD and standard samples was retrieved from the Gene Expression Omnibus (GEO) database. The GSE20680 dataset, which has 139 models overall and includes 52 standard samples and 87 CAD samples, was used as the training set for the study's principal analysis (15). We then verified the expression of marker genes using the GSE20681 dataset. The Immunology Database and Analysis Portal (immPort) database provided the 1793 IRGs (Supplementary Table 1) used in this study (16). Drugs that target marker genes were predicted using the Drug-Gene Interaction Database (DGIdb) (17). Additionally, the DrugBank database was used to collect structural data on drugs that target marker genes.

2.2. Identification of DE-IRGs

We screened for DE-IRGs between CAD and controls using the "LIMMA" package in the R software using the combined datasets ($P < 0.05$). Genes were categorized as being up-regulated or down-regulated based on their log₂FC values (18), which are displayed on a Volcano plot.

2.3. Functional enrichment analysis

Many related human pathologies, including cancer, neurodegenerative diseases, and infectious diseases, are the result of abnormal protein-protein interactions (PPIs) that alter molecular recognition mechanisms and binding partner affinity under given conditions. Highly complex interactome diagrams are increasingly being used to decipher disease-specific protein associations and characterize the effects of splicing and genetic variation on these systems (19). We confined the confidence (combined score) > 0.4 as a condition when examining the PPI to establish the validity of this interaction.

The Gene Ontology (GO) project provides a comprehensive source for functional genomics. There are three separate aspects: (i) development and maintenance of the ontology, (ii) annotation of gene products, and (iii) development and continuous improvement of tools and training that facilitate the creation, maintenance, and use of the ontologies (20). The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a knowledge base for the systematic analysis of gene functions, linking genomic information with higher order functional information. The PATHWAY database represents the higher order functions in terms of the network for the interacting molecules (21). As a result, with the help of the "ClusterProfiler" package in R, GO was used to identify characteristic biological attributes, and

KEGG pathway enrichment analysis was performed to identify functional attributes. Significance was set at $P < 0.05$.

2.4. Identifying CAD-related OFGs

In biomedical research, it is important to select the variables most associated with the studied outcome and to determine the strength of this association. Least absolute shrinkage and selection operator (LASSO) Cox regression analyses and support vector machines-recursive feature elimination (SVM-RFE) analyses are powerful tools to analyze data with a number of predictors approximately equal or larger than the number of observations. The "glmnet" package was used to obtain the LASSO algorithm to minimize the dimensionality of the data. By locating distinct genetic biomarkers of CAD using the LASSO technique, which were compared by the average misjudgment rates of their 10-fold cross-validations, models of DE-IRGs between CAD patients and normal samples were created (22). A SVM-RFE model was created at the same time using the "SVM" package, and its average false positive rate was also compared using a 10-fold cross-validation (23). To create OFGs for CAD, genes from both algorithms were overlapped. By determining the receiver operating characteristic (ROC) curve and computing the area under the curve (AUC), accuracy, sensitivity, and selectivity, the diagnostic potency of OFGs was evaluated. Similarly, the diagnostic power of the logistic regression model was evaluated using ROC curves.

2.5. Single sample gene set enrichment analysis (ssGSEA) and single-gene gene set variation analysis (ssGSVA)

Gene set enrichment analysis (GSEA) is a popular framework for condensing information in gene expression profiles into pathways or signature summaries. While GSEA is often considered the endpoint of bioinformatics analysis, gene set variation analysis (GSVA) has an increased ability to detect subtle pathway activity changes in sample populations compared to GSEA, which constitutes the starting point for building pathway-centric biological models (24). The background gene set for this study was the KEGG pathway set, and each marker gene underwent GSVA analysis using the "GSVA" package in the R software. Simultaneously, we utilized the "LIMMA" package to examine the variations in GSVA scores between the marker gene groups with high and low expression. Based on the filtering criteria of $|t| > 2$ and $P < 0.05$, if $t > 0$, we regarded the pathway to be activated in the high expression group, and if $t < 0$, we considered the pathway to be activated in the low expression group. Furthermore, with the help of the "LIMMA" package, we performed Single sample

gene set enrichment analysis (ssGSEA) to further understand the functions of the OFGs.

2.6. Evaluation of CAD immune cell infiltration

CIBERSORT is a type of deconvolution algorithm, by which the normalized gene expression matrix can be transformed into the composition of infiltrating immune cells. Through the CIBERSORT package, we examined the proportion of 28 infiltrating immune cell types in CAD patients and healthy controls. We presented them as risk heat maps and violin plots to better understand the infiltration of immune cells in the samples in the dataset (25). Additionally, we examined the immune cell infiltration associated with each gene and visualized it as a correlation heatmap (26).

2.7. Statistical analysis

R version 4.1.3 was used to complete all analyses. Using Student's *t*-tests, comparisons between the two groups were made. The association between DE-IRGs was discovered using Pearson correlation analysis. Targeted drugs were displayed using Cytoscape. Differences were deemed significant at $P < 0.05$.

3. Results

3.1. Recognition of DE-IRGs

The network is described in **Figure 1**. With the help of the "LIMMA" package, 82 of the 1,793 IRGs, which consisted of 35 up-regulated and 47 down-regulated genes, were differentially expressed between CAD and healthy controls according to differential analysis of the GSE20680 data (**Supplementary Table 2**). To show the differential expression of the 82 genes between CAD and control samples, we generated a volcano plot to show genetic differences and marked the three genes that were the most significantly up-regulated and down-regulated ($P < 0.05$) (**Figure 2A**).

3.2. Functional analysis of the DE-IRGs

We built a PPI network to investigate the connection between the IRGs and the biological processes engaged in the regulation process. Of the 82 IRGs, 15 were unrelated, thus, we constructed a PPI network based on the remaining 67 IRGs. These genes were tightly connected at the protein level, according to PPI analysis (**Figure 2B**). IRGs involved

in molecular functions, biological processes, and cellular components were identified by GO analysis; IRGs principally focused on the activation of signaling receptors in the molecular function component, whereas in the biological processes component, they were concentrated on the control of the positive response to stimuli from the environment. Additionally, IRGs were found to be primarily abundant in the outer layer of the plasma membrane in the cellular component aspect (**Figure 2C**). The GO results were also validated by KEGG; the pathway enrichment map revealed that the primary enriched pathways involved receptor ligand activity, signaling receptor activator activity, and positive response regulation to external stimulus, indicating a strong correlation between the inflammatory response and CAD (**Figure 2D**).

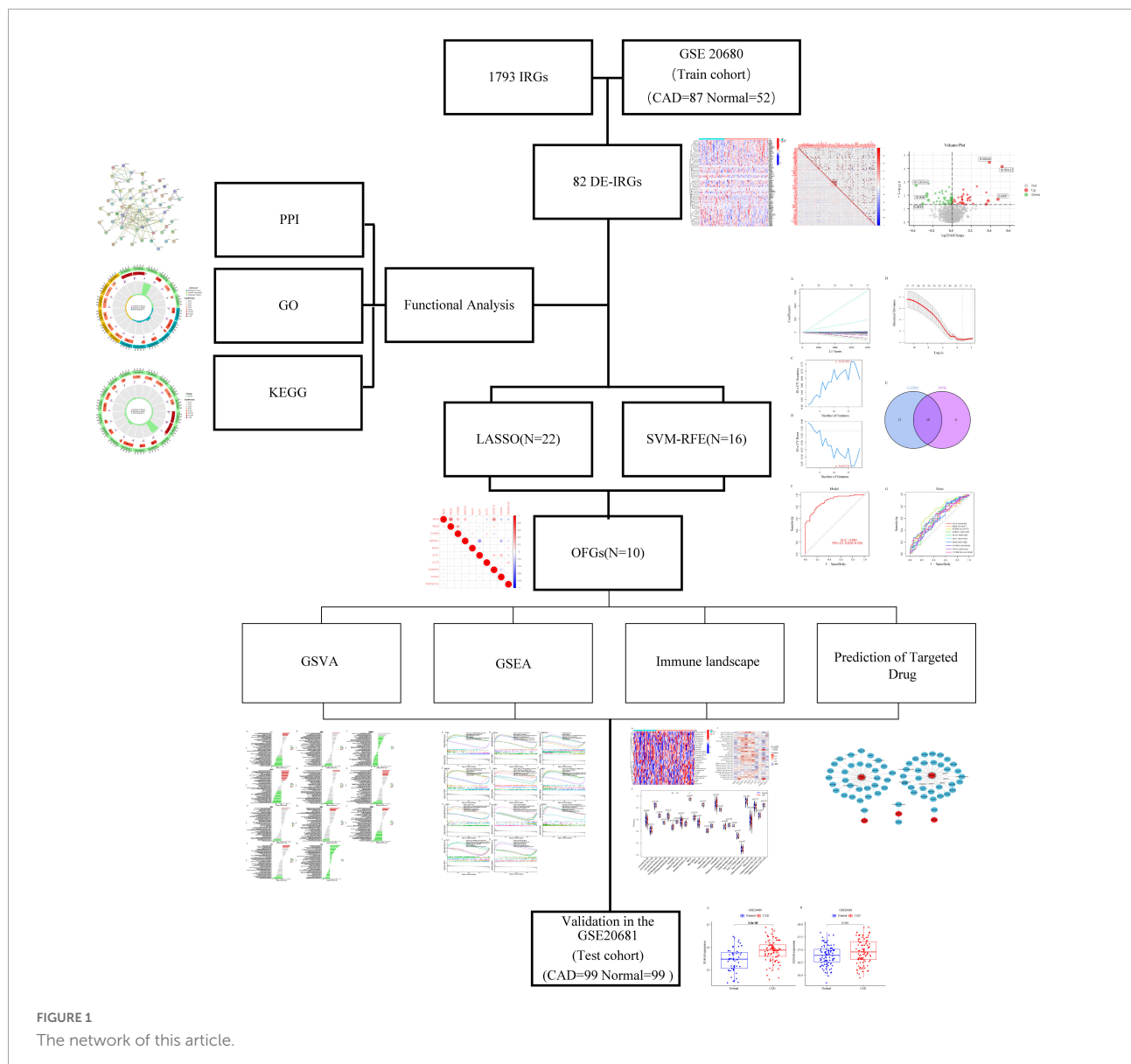
3.3. Identification of OFGs for CAD

To reduce the contingency of the DE-IRGs, we created IRG models sequentially using the LASSO and SVM-RFE approaches, and chose the intersection genes of the two as OFGs to screen out DE-IRGs between CAD patients and normal controls. Based on the average false positive rate of 10 times cross-validation, we created a model consisting of 22 DE-IRGs by applying the LASSO logistic regression algorithm and tweaking the penalty value using 10-fold cross-validation (**Figures 3A, B**). Then, using the SVM-RFE approach, which was compared by the average misjudgment rates of their 10-fold cross-validations, we created a model including 16 DE-IRGs, which was then used to determine the optimal combination of eigengenes (maximal accuracy = 0.726, minimal RMSE = 0.274) (**Figures 3C, D**). Finally, we combined the marker genes from the LASSO and SVM-RFE models to identify 10 marker genes (MICB, RELB, S100A8, MAP2K1, IGHA1, SLIT1, SLIT2, TNFSF15, FGFR3, and TNFRSF13C) for further investigation (**Figure 3E**).

Using the "glm" package, we created a logistic regression model based on the 10 OFGs mentioned above, which could distinguish between normal and CAD samples with an AUC of 0.889 [95% confidence interval (CI): 0.830–0.938], according to the ROC curve (**Figure 3F**). Furthermore, ROC curves for the 10 marker genes were created to clarify how well a single gene can distinguish between samples with and without CAD; all genes exhibited an AUC > 0.6, with the S100A8 gene having the best predictive power (AUC = 0.710) (**Figure 3G**). Logistic regression models are more accurate and specific than individual marker genes at differentiating CAD samples from healthy controls.

3.4. ssGSVA and ssGSEA

Based on the levels of expression of each marker gene paired with GSVA, we found differentially activated



pathways across groups with high and low expression. We began by running GSVA analysis on OFGs; CAD patients exhibited lower expression in “taste transduction,” “olfactory transduction,” and “retinol metabolism” compared to normal controls, and increased expression in “circadian rhythm mammalian,” “limonene and pinene degradation,” and “DNA replication” (Figure 4A). Based on the degree of each marker gene’s expression in combination with GSVA, we could identify differences between CAD and normal controls. The expression of the IGH1 gene was mainly up-regulated for “limonene and pinene degradation,” whereas it was down-regulated for “DNA replication.” Patients with CAD exhibited upregulation of the MAK2P1 gene in “linoleic acid metabolized” and downregulation in “glycosphingolipid biosynthetic Globo series.” Patients with CAD had higher

levels of “pentose and glucuronic acid interconversion” and lower levels of “other glycan degradation” for the MICB gene. For the RELB gene, “limonene and pinene degradation” and “folate biosynthesis” were the two key areas where it was up-regulated. In addition to being up-regulated in “limonene and pinene degradation,” the S100A8 gene was also down-regulated in “pantothenic acid and COA biosynthesis.” RNA degradation and tyrosine metabolism for the SILT1 gene were both increased in CAD patients. TNFSF15 was downregulated in “taste conduction” but upregulated in “mammalian circadian rhythm.” Additionally, TNFRSF13C was exclusively up-regulated in “sulfur metabolism,” “biosynthesis of unsaturated fatty acids,” “non-homologous end joining,” and “pantothenic acid and COA biosynthesis.” The values were different for other aspects but were not statistically significant.

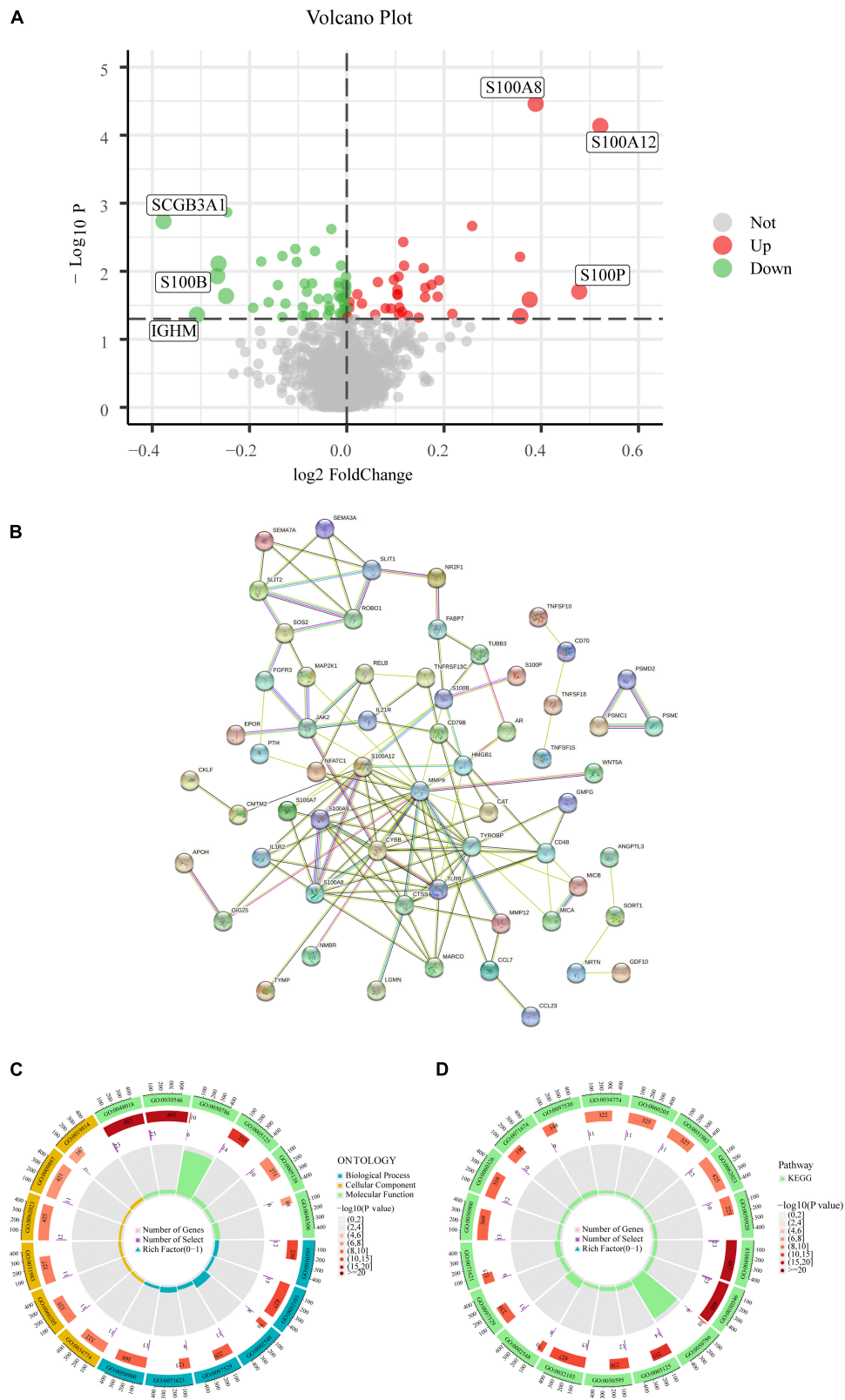


FIGURE 2 Analysis of DE-IRGs. (A) Volcano plot of the DE-IRGs in CAD. (B) PPI network. (C) GO enrichment analysis. (D) KEGG enrichment analysis.

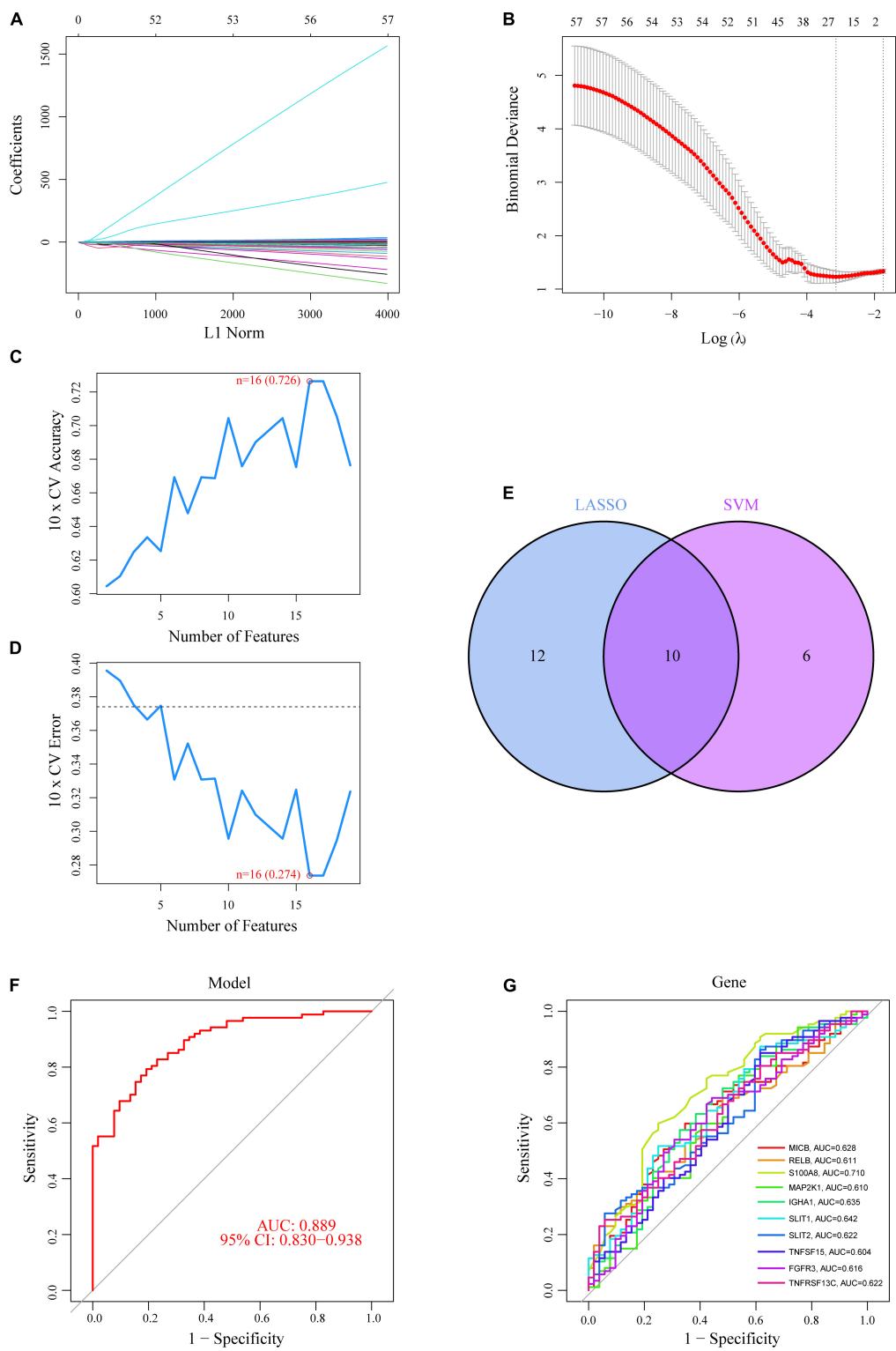
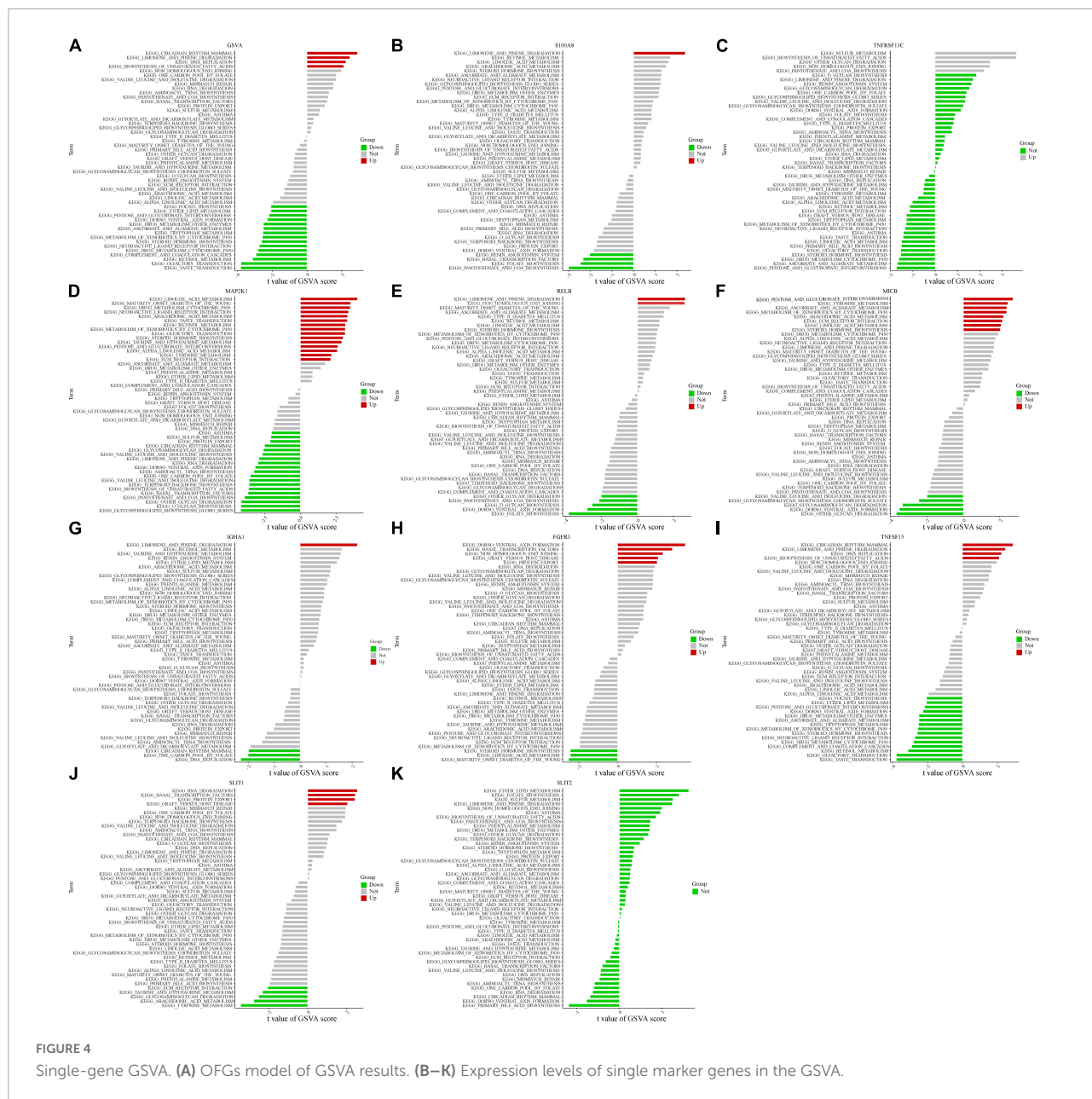


FIGURE 3

Ten DE-IRGs were shown to be CAD diagnostic genes. (A,B) Using the LASSO logistic regression algorithm, with penalty parameter tuning conducted by 10-fold cross-validation, 22 CAD-related features were selected. (C,D) The SVM-RFE algorithm was used to filter the 10 DE-IRGs to identify the optimal combination of feature genes. Finally, 16 genes (maximal accuracy = 0.726, minimal RMSE = 0.274) were identified as the OFGs. (E) The 10 marker genes obtained from the LASSO and SVM-RFE models. (F) The AUC for disease samples was determined using a logistic regression model. (G) ROC curves for the 10 marker genes.



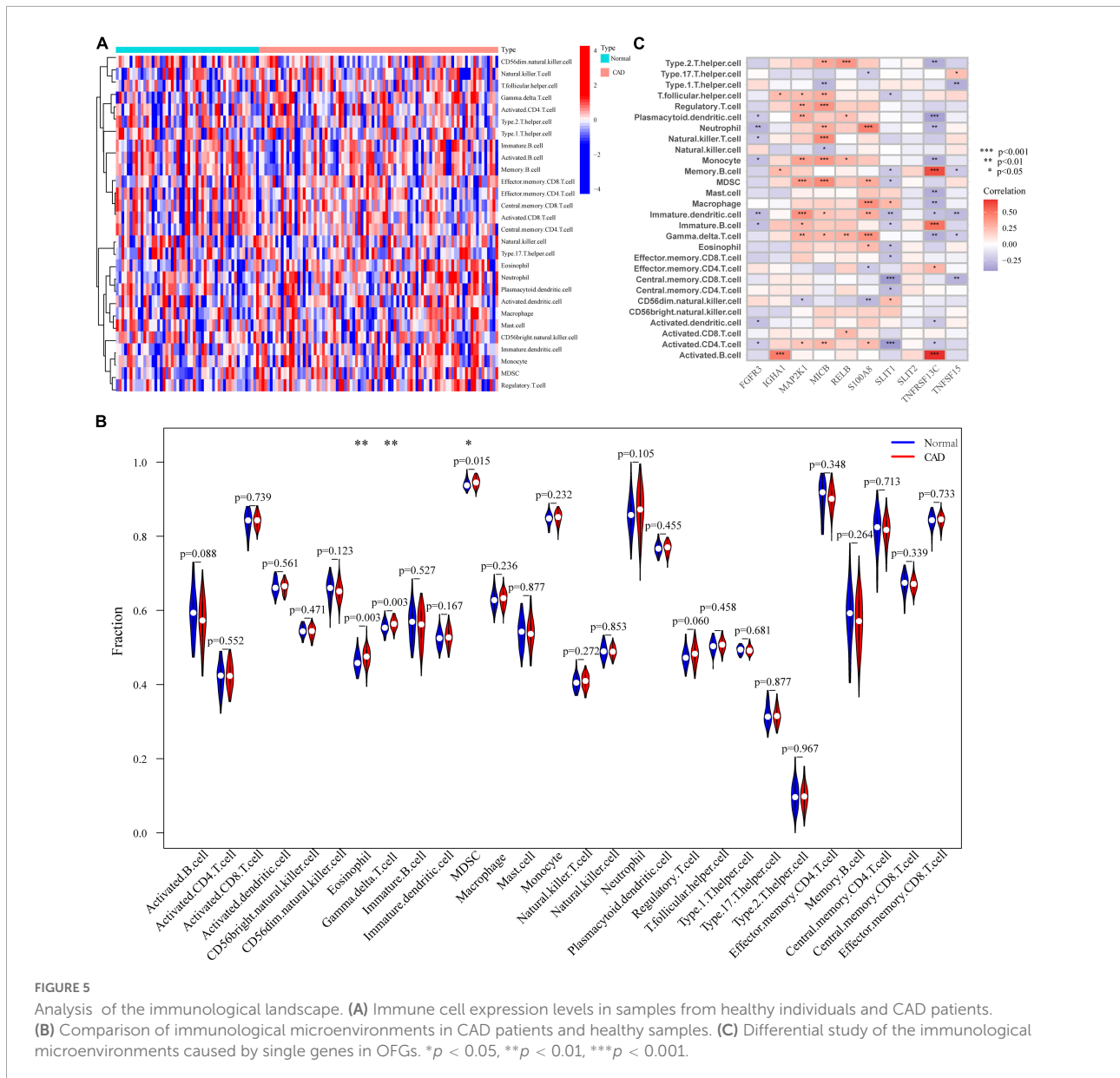
Surprisingly, there was no statistically significant distinction for the *SILT2* gene between CAD patients and healthy individuals (Figures 4B–K). From the above GSEA analysis, it is not difficult to conclude that the higher expression of these genes in coronary heart disease patients is related to cell cycle activity and energy metabolism, which to a certain extent reflects the correlation between coronary heart disease and the chronic inflammatory response.

To further explore the potential function of marker genes that can distinguish diseased samples from normal samples, we performed GSEA analysis on the OFGs model and drew the enrichment curve for the top six signaling pathways. The results showed that for CAD patients, the “B

cell receptor signaling pathway,” “chronic myeloid leukemia,” “endometrial cancer,” and “lysosome” pathways were up-regulated, while the “neuroactive ligand-receptor interaction” and “olfactory conduction” pathways were down-regulated (Supplementary Figure 1).

3.5. Immune landscape

There is increasing data to demonstrate how closely related CAD and the immune microenvironment are. To investigate changes in the immune microenvironment between CAD patients and normal controls, we utilized the CIBERSORT

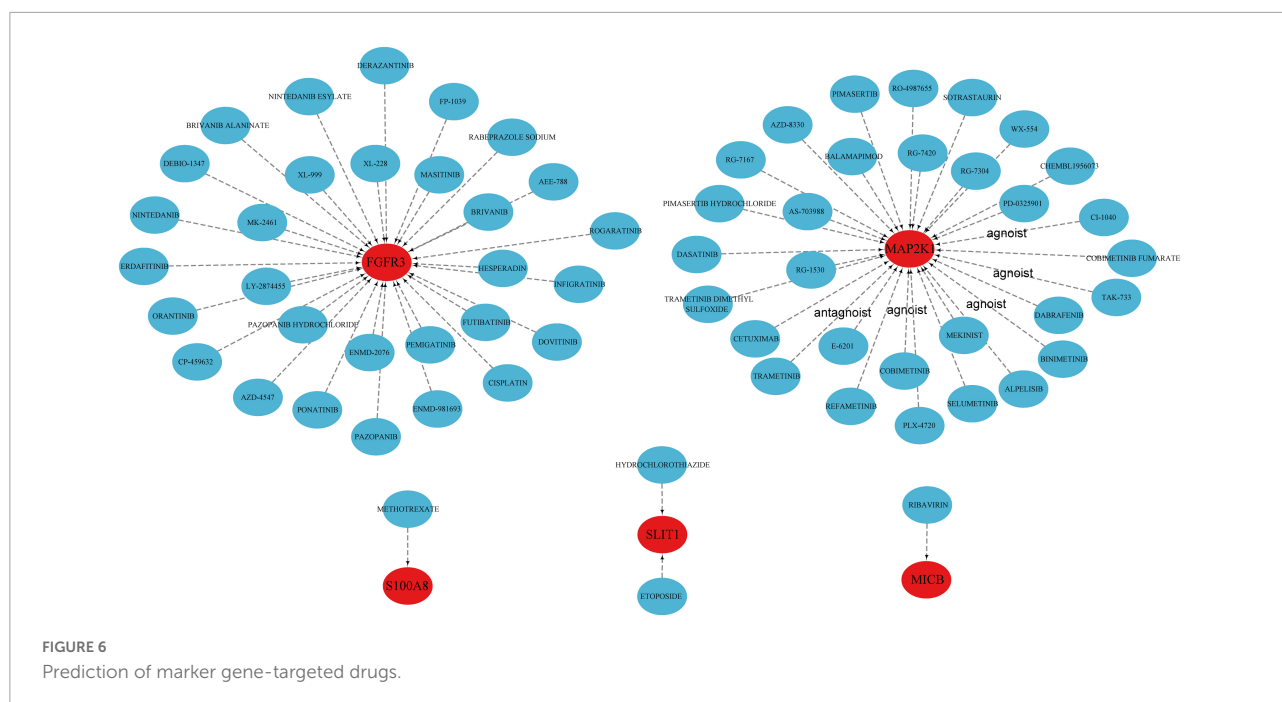


method; after examining the expression levels of 28 types of immune cells in various patient populations 13 were comparatively up-regulated and 15 were down-regulated in CAD patients. The findings were shown as a heat map (Figure 5A). In addition, we categorized the genes by single-gene immune cell infiltration; IGHA1, MAP2K1, MICB, RELB, and S100A8 were primarily linked to up-regulation of immune cell expression, whereas FGFR3, SILT1, SILT2, TNFRSF13C, and TNFSF15 down-regulated immune cell expression (Figure 5B). We analyzed the difference in the expression level of immune cells in 28 patients with CAD and normal controls, and the results are shown as violin diagrams. Eosinophils, gamma delta T cells, and myeloid-derived suppressor cells were up-regulated in CAD patients, and their distinctions were then shown using

correlation analysis with the 28 immune cells. Although other immune cells were also found to differ statistically, it was not significant (Figure 5C).

3.6. Prediction of targeted drugs

We further revealed the drugs that may target marker genes through the DGIdb database. In total, 64 drugs that were linked to five genes (S100A8, SLIT1, MICB, FGFR3, MAP2K1) were analyzed, including 42 targeted inhibitor drugs, six inhibitor and allosteric modulator drugs, one inhibitor and antagonist drug with controversial effects, and 15 as yet undetermined drugs (Supplementary Table 3). Unfortunately, for the other



five genes, we did not find relevant drugs. The results were visualized by the Cytoscape software (Figure 6). Methotrexate (MTX) appears to affect S100A8; however, the precise impact still needs to be validated.

3.7. Expression of marker genes in the validation set

In the GSE20681 dataset, we verified the expression of the marker genes. We used differential analysis to examine S100A8 gene expression in the validation set; the results revealed that S100A8 expression was higher in CAD patients than in healthy controls ($P = 0.019$) (Figure 7).

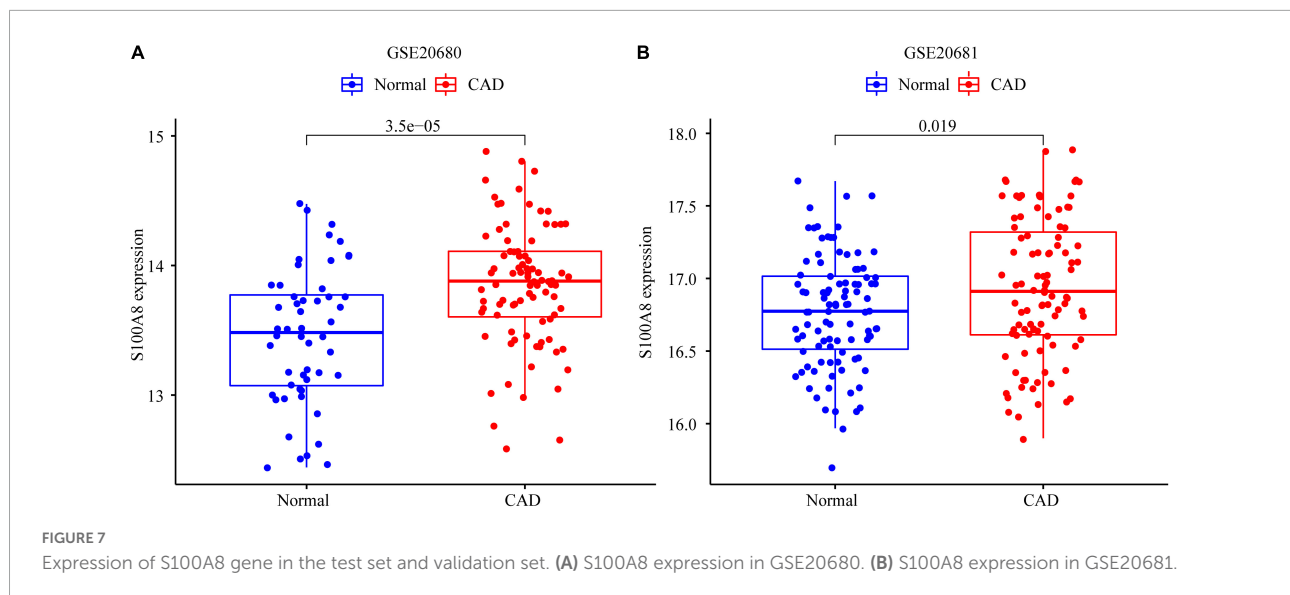
4. Discussion

Coronary atherosclerosis is a complex, persistent, and progressive inflammatory disease that is the leading cause of death in both developed and developing countries (2). Immune cells maintain homeostasis in the heart, and various immune cells that reside or penetrate the heart tissue play an important role in the repair process following tissue damage. Many previous studies have suggested that CAD and immunoreaction are inseparable and that inflammation is closely related to CAD; in particular, monocytes and mast cells play a crucial role in the process of the chronic inflammatory response. However, it is worth noting that the specific relationship between coronary atherosclerosis and immune genes has not been reported in

detail. Therefore, it is of clinical significance to establish an immune-related prognostic model. Therefore, we used the GSE20680 gene set as original data to analyze the DEGs between CAD patients and normal controls. Next, we found common functional pathways through bioinformatics enrichment, and analyzed the differences in the expression of IRGs between CAD patients and controls, which is of great significance in exploring the pathogenesis of IRGs in CAD patients.

An individual's risk of CAD is controlled by interactions between hereditary and lifestyle factors, as is the case with most complicated diseases (27). There are several well-acknowledged risk factors for CAD, such as tobacco use, hypertension, hyperlipidemia, diabetes, and family history; at least one of these risk variables is present in 80–90% of patients with documented CAD (28). The study of gene expression, mRNA, and miRNA has opened up new possibilities for the pursuit of diagnostic and therapeutic approaches, widening the boundaries of genetic research as laboratory technology and informatics advance (7). Metabolic syndrome is defined as central obesity plus hypertriglyceridemia, low high-density lipoprotein cholesterol, hypertension, or high fasting glucose levels. Grayson et al. have identified gene expression patterns associated with metabolic syndrome, which includes many genes involved in the innate immune response. As the condition develops into sequelae, such as CAD, which is characterized by an increase in genes related to macrophage activation and signaling, this pattern changes (29, 30). Therefore, immune expression research can be used to infer the stage and prognosis of CAD patients.

Numerous immunological components, such as cells (endothelial cells, macrophages, and lymphocytes), cytokines,



chemokines, acute-phase proteins, and adhesion molecules, are involved in atherosclerosis (31). Prospective epidemiological data is most consistent with the predictive usefulness of CRP, interleukin 6 (IL-6), and tumor necrosis factor- α (TNF- α), among a wide range of immune system parameters (32). In addition, immune system characteristics are closely related to CAD risk factors, including hypertension, smoking, dyslipidemia, etc. CAD risk factors indirectly induce inflammatory responses to promote the occurrence and development of coronary heart disease. The vicious cycle of inflammation and high-risk factors intertwines and promotes the occurrence of CAD. The initial inflammatory response to arterial injury is frequently advantageous; however, if the harmful agent continues to exist, it could cause permanent inflammation and change the way the body reacts initially to harmful inflammation. The inflammatory process becomes the first feature of the development of CAD, with subsequent progression of inflammation that eventually impairs arterial function (33, 34). Therefore, it is particularly important to identify the early inflammatory response in CAD.

In this study, we screened 10 IRGs, dominated by S100A8, by bioinformatics. Each of these 10 genes has an AUC value > 0.6 , indicating that they are accurate and specific at differentiating coronary heart disease samples from normal samples. The S100 family member S100A8 (also known as MRP8) typically appears as a heterodimer. As a Ca^{2+} sensor also involved in cytoskeletal reorganization and arachidonic acid metabolism, S100A8 is constitutively expressed in neutrophils and monocytes. In disorders connected to inflammation, S100A8 has been proposed as a biomarker for diagnosis and follow-up, and as a predictor of therapeutic response (35). S100A8 is actively produced during inflammation and is essential for the inflammatory response since it promotes leukocyte recruitment and triggers cytokine secretion (36, 37).

S100A8 controls the production of pro-inflammatory mediators, including cytokines, chemokines, reactive oxygen species, and nitric oxide, among others, to prevent tissue damage brought on by excessive inflammation (38). Therefore, S100A8 expression is a double-edged sword; on one hand, as an inflammatory mediator, it promotes the progression of local inflammation, whereas on the other hand, overexpression of S100A8 will inhibit the explosive spread of inflammation, and eventually lead to the chronic progressive development of inflammation. Clinical evidence suggests that the neutrophil-to-lymphocyte ratio is one of the most reliable predictors of death in acute coronary syndrome (ACS) patients, meaning that lowering the neutrophil count may have positive effects (39, 40). In ACS patients, neutrophils pass through the injury site and release a significant amount of S100A8, which acts as a chemoattractant to encourage the recruitment of other cell types, including monocytes, which determines the subsequent inflammatory response. Additionally, neutrophils are the first responders to aseptic injury (41, 42). Our study found that CAD patients have a high expression of S100A8, which supports the view that a chronic inflammatory response occurs in CAD; therefore, this model may provide a new therapeutic target for the diagnosis and treatment of CAD.

To our knowledge, this is the first study to use bioinformatics to comprehensively analyze the role of IRGs in CAD. In addition, most of the existing studies have constructed prognostic models through methods such as Weighted Gene Co-Expression Network Analysis (WGCNA) (42–45). Compared with them, we are the first to use LASSO and SVM-RFE dual methods to establish a prognosis model of coronary heart disease, which greatly improved the accuracy of the model, which is undoubtedly of certain value. Besides, on the basis of PPI network, GO and KEGG enrichment analysis of differential genes, we further analyzed the single

genes of the model prognosis genes, and further demonstrated the specific mechanism of action and related pathways of the model-related genes by using GSVA and ssGSEA methods, which is undoubtedly of constructive significance. Nonetheless, this investigation is not without its drawbacks. First, even though the OFGs had a significant prognostic impact in both the experimental and control groups, we cannot rule out the possibility that the prognostic effect was accidental, thus, it needs to be confirmed in a study with a larger sample size. However, there are currently no data sets with a large enough sample size and clinical prognostic information to be used for further verification. The interaction of other characteristics with predictive value was not considered in this study, even though this OFG model performed exceptionally well in terms of prognosis for patients with CAD. Third, since this prognostic signature was developed and tested using information from open sources, the model requires additional biological support, which will need to be confirmed in subsequent research.

5. Conclusion

We formed an immune-associated coronary heart disease prognostic model for a total of 10 genes (MICB, RELB, S100A8, MAP2K1, IGHA1, SLIT1, SLIT2, TNFSF15, FGFR3, and TNFRSF13C). Among them, S100A8 was the most closely related to CAD. Through bioinformatics correlation analysis, we confirmed that the cardiac immune microenvironment of CAD patients may be affected by this OFG model. In addition, we found that the relationship between immune cells and these key genes could have a significant impact on the development and progression of CAD, and that studies of these genes may shed new light on how to treat cardiovascular disease. Although gene expression levels may not always correlate to protein expression, their research importance is indisputable. To better understand the pathophysiology and management of coronary heart disease, we will continue to focus on these genes in future studies.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/>, GSE20680 and <https://www.ncbi.nlm.nih.gov/geo/>, GSE20681.

[ncbi.nlm.nih.gov/geo/](https://www.ncbi.nlm.nih.gov/geo/), GSE20680 and <https://www.ncbi.nlm.nih.gov/geo/>, GSE20681.

Author contributions

All authors were solely responsible for the content and writing of the manuscript and made significant contributions to the design, data collection and interpretation, and manuscript preparation and revision of this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2022.1055422/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Gene set enrichment analysis (GSEA). (A) OFGs model of GSEA results. (B–K) Expression levels of single marker genes in the GSEA.

SUPPLEMENTARY TABLE 1

1793 immune-related genes.

SUPPLEMENTARY TABLE 2

82 DEGs in GSE20680.

SUPPLEMENTARY TABLE 3

Marker gene-targeted drugs.

References

- Cassar A, Holmes D, Rihal C, Gersh B. Chronic coronary artery disease: diagnosis and management. *Mayo Clin Proc.* (2009) 84:1130–46. doi: 10.4065/mcp.2009.0391
- McCullough P. Coronary artery disease. *Clin J Am Soc Nephrol.* (2007) 2:611–6. doi: 10.2215/CJN.03871106
- Ma L, Chen W, Gao R, Liu L, Zhu M, Wang Y, et al. China cardiovascular diseases report 2018: an updated summary. *J Geriatr Cardiol.* (2020) 17:1–8.
- Roth G, Johnson C, Abajobir A, Abd-Allah F, Abera S, Abyu G, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol.* (2017) 70:1–25.

5. Kwok C, Satchithananda D, Mallen C. Missed opportunities in coronary artery disease: reflection on practice to improve patient outcomes. *Coron Artery Dis.* (2022) 33:233–8. doi: 10.1097/MCA.0000000000001075
6. Libby P. The changing nature of atherosclerosis: what we thought we knew, what we think we know, and what we have to learn. *Eur Heart J.* (2021) 42:4781–2. doi: 10.1093/eurheartj/ehab438
7. Weber C, Noels H. Atherosclerosis: current pathogenesis and therapeutic options. *Nat Med.* (2011) 17:1410–22.
8. Shea S, Navas-Acien A, Shimbo D, Brown E, Budoff M, Bancks M, et al. Spatially weighted coronary artery calcium score and coronary heart disease events in the multi-ethnic study of atherosclerosis. *Circ Cardiovasc Imaging.* (2021) 14:e011981.
9. Libby P. The changing landscape of atherosclerosis. *Nature.* (2021) 592:524–33. doi: 10.1038/s41586-021-03392-8
10. Adam C, Șalaru D, Priscaariu C, Marcu D, Sascău R, Stătescu C. Novel biomarkers of atherosclerotic vascular disease—latest insights in the research field. *Int J Mol Sci.* (2022) 23:4998. doi: 10.3390/ijms23094998
11. Xiao S, Zhou Y, Wu Q, Ma W, Chen M, Pan D. Uncovering the differentially expressed genes and pathways involved in the progression of stable coronary artery disease to acute myocardial infarction using bioinformatics analysis. *Eur Rev Med Pharmacol Sci.* (2021) 25:301–12. doi: 10.26355/eurrev_202101_24396
12. Du Y, Hu Y, Wen N, Fu S, Zhang G, Li L, et al. Abnormal expression of TGFB2, EGF, LRP10, and IQGAP1 is involved in the pathogenesis of coronary artery disease. *Rev Cardiovasc Med.* (2021) 22:947–58. doi: 10.31083/j.rcm2203103
13. Tibaut M, Caprnda M, Kubatka P, Sinković A, Valentova V, Filipova S, et al. Markers of atherosclerosis: part 2 - genetic and imaging markers. *Heart Lung Circ.* (2019) 28:678–89. doi: 10.1016/j.hlc.2018.09.006
14. Plesković A, Vraspir-Porenta O, Zorc-Plesković R, Petrovič D, Zorc M, Milutinović A. Deficiency of mast cells in coronary artery endarterectomy of male patients with type 2 diabetes. *Cardiovasc Diabetol.* (2011) 10:40. doi: 10.1186/1475-2840-10-40
15. Li Z, Song Y, Wang M, Shen R, Qin K, Zhang Y, et al. m6A regulator-mediated RNA methylation modification patterns are involved in immune microenvironment regulation of coronary heart disease. *Front Cardiovasc Med.* (2022) 9:905737. doi: 10.3389/fcvm.2022.905737
16. Christiansen M, Larsen S, Nyegaard M, Neergaard-Petersen S, Ajjan R, Würtz M, et al. Coronary artery disease-associated genetic variants and biomarkers of inflammation. *PLoS One.* (2017) 12:e0180365. doi: 10.1371/journal.pone.0180365
17. Wagner A, Coffman A, Ainscough B, Spies N, Skidmore Z, Campbell K, et al. DGIdb 2.0: mining clinically relevant drug-gene interactions. *Nucleic Acids Res.* (2016) 44:D1036–44. doi: 10.1093/nar/gkv1165
18. Qiu J, Peng B, Tang Y, Qian Y, Guo P, Li M, et al. CpG methylation signature predicts recurrence in early-stage hepatocellular carcinoma: results from a multicenter study. *J Clin Oncol.* (2017) 35:734–42. doi: 10.1200/JCO.2016.68.2153
19. Soleymani F, Paquet E, Viktor H, Michalowski W, Spinello D. Protein-protein interaction prediction with deep learning: a comprehensive review. *Comput Struct Biotechnol J.* (2022) 20:5316–41. doi: 10.1016/j.csbj.2022.08.070
20. Gene Ontology Consortium. Gene ontology consortium: going forward. *Nucleic Acids Res.* (2015) 43:D1049–56. doi: 10.1093/nar/gku1179
21. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* (2000) 28:27–30. doi: 10.1093/nar/28.1.27
22. Ye Y, Dai Q, Qi H. A novel defined pyroptosis-related gene signature for predicting the prognosis of ovarian cancer. *Cell Death Discov.* (2021) 7:71. doi: 10.1038/s41420-021-00451-x
23. Miao Y, Liu J, Liu X, Yuan Q, Li H, Zhang Y, et al. Machine learning identification of cuproptosis and necroptosis-associated molecular subtypes to aid in prognosis assessment and immunotherapy response prediction in low-grade glioma. *Front Genet.* (2022) 13:951239. doi: 10.3389/fgene.2022.951239
24. Liang L, Yu J, Li J, Li N, Liu J, Xiu L, et al. Integration of scRNA-Seq and bulk RNA-Seq to analyse the heterogeneity of ovarian cancer immune cells and establish a molecular risk model. *Front Oncol.* (2021) 11:711020. doi: 10.3389/fonc.2021.711020
25. Cheng Q, Wang L. LncRNA XIST serves as a ceRNA to regulate the expression of ASF1A, BRWD1M, and PFKFB2 in kidney transplant acute kidney injury via sponging hsa-miR-212-3p and hsa-miR-122-5p. *Cell Cycle.* (2020) 19:290–9. doi: 10.1080/15384101.2019.1707454
26. Zhang S, Zhang E, Long J, Hu Z, Peng J, Liu L, et al. Immune infiltration in renal cell carcinoma. *Cancer Sci.* (2019) 110:1564–72.
27. Khera A, Emdin C, Drake I, Natarajan P, Bick A, Cook N, et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N Engl J Med.* (2016) 375:2349–58.
28. Khot U, Khot M, Bajzer C, Sapp S, Ohman E, Brener S, et al. Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA.* (2003) 290:898–904. doi: 10.1001/jama.290.7.898
29. Freedman J, Larson M, Tanriverdi K, O'Donnell C, Morin K, Hakanson A, et al. Relation of platelet and leukocyte inflammatory transcripts to body mass index in the Framingham heart study. *Circulation.* (2010) 122:119–29. doi: 10.1161/CIRCULATIONAHA.109.928192
30. Grayson B, Wang L, Aune T. Peripheral blood gene expression profiles in metabolic syndrome, coronary artery disease and type 2 diabetes. *Genes Immun.* (2011) 12:341–51. doi: 10.1038/gene.2011.13
31. Naghavi M, Libby P, Falk E, Casscells S, Litovsky S, Rumberger J, et al. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: part I. *Circulation.* (2003) 108:1664–72.
32. Danesh J, Collins R, Peto R. Chronic infections and coronary heart disease: is there a link? *Lancet.* (1997) 350:430–6. doi: 10.1016/S0140-6736(97)03079-1
33. Kop W, Gottdiener J. The role of immune system parameters in the relationship between depression and coronary artery disease. *Psychosom Med.* (2005) 67(Suppl. 1):S37–41. doi: 10.1097/01.psy.0000162256.18710.4a
34. Ross R. Atherosclerosis—an inflammatory disease. *N Engl J Med.* (1999) 340:115–26. doi: 10.1056/NEJM199901143400207
35. Wang S, Song R, Wang Z, Jing Z, Wang S, Ma J. S100A8/A9 in inflammation. *Front Immunol.* (2018) 9:1298. doi: 10.3389/fimmu.2018.01298
36. Holzinger D, Nippe N, Vogl T, Marketon K, Mysore V, Weinlage T, et al. Myeloid-related proteins 8 and 14 contribute to monosodium urate monohydrate crystal-induced inflammation in gout. *Arthritis Rheumatol.* (2014) 66:1327–39. doi: 10.1002/art.38369
37. Vogl T, Tenbrock K, Ludwig S, Leukert N, Ehrhardt C, van Zoelen M, et al. Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock. *Nat Med.* (2007) 13:1042–9. doi: 10.1038/nm1638
38. Zhao J, Endoh I, Hsu K, Tedla N, Endoh Y, Geczy C. S100A8 modulates mast cell function and suppresses eosinophil migration in acute asthma. *Antioxid Redox Signal.* (2011) 14:1589–600. doi: 10.1089/ars.2010.3583
39. Soehnlein O, Steffens S, Hidalgo A, Weber C. Neutrophils as protagonists and targets in chronic inflammation. *Nat Rev Immunol.* (2017) 17:248–61. doi: 10.1038/nri.2017.10
40. Bhat T, Teli S, Rijal J, Bhat H, Raza M, Khoueiry G, et al. Neutrophil to lymphocyte ratio and cardiovascular diseases: a review. *Expert Rev Cardiovasc Ther.* (2013) 11:55–9.
41. Sreejit G, Abdel-Latif A, Athmanathan B, Annabathula R, Dhyani A, Noothi S, et al. Neutrophil-derived S100A8/A9 amplify granulopoiesis after myocardial infarction. *Circulation.* (2020) 141:1080–94. doi: 10.1161/CIRCULATIONAHA.119.043833
42. Ryckman C, Vandal K, Rouleau P, Talbot M, Tessier P. Proinflammatory activities of S100: proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. *J Immunol.* (2003) 170:3233–42. doi: 10.4049/jimmunol.170.6.3233
43. Liu K, Chen S, Lu R. Identification of important genes related to ferroptosis and hypoxia in acute myocardial infarction based on wgcna. *Bioengineered.* (2021) 12:7950–63. doi: 10.1080/21655979.2021.1984004
44. Liang W, Sun F, Zhao Y, Shan L, Lou H. Identification of susceptibility modules and genes for cardiovascular disease in diabetic patients using wgcna analysis. *J Diabetes Res.* (2020) 2020:4178639. doi: 10.1155/2020/4178639
45. Liu Y, Liu N, Liu Q. Constructing a cerna-immunoregulatory network associated with the development and prognosis of human atherosclerosis through weighted gene co-expression network analysis. *Aging.* (2021) 13:3080–100. doi: 10.18632/aging.202486