



OPEN ACCESS

EDITED BY

Marcelo R. S. Briones,
Federal University of São Paulo, Brazil

REVIEWED BY

Feng Zhang,
The Second Affiliated Hospital of Harbin
Medical University, China
Bo Xiong,
Baoshan Hospital Affiliated to Shanghai
University of Traditional Chinese Medicine,
China

*CORRESPONDENCE

Xiaohong Li,
✉ li_xiaoh2005@163.com
Jingmei Song,
✉ songjingmei1983@163.com

RECEIVED 10 February 2025

ACCEPTED 09 July 2025

PUBLISHED 17 July 2025

CITATION

Lin Y, Song J and Li X (2025) Identification of
biomarkers between coronary artery disease
and non-alcoholic steatohepatitis: a
combination of bioinformatics and
machine learning.
Front. Genet. 16:1573621.
doi: 10.3389/fgene.2025.1573621

COPYRIGHT

© 2025 Lin, Song and Li. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](#).
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.

Identification of biomarkers between coronary artery disease and non-alcoholic steatohepatitis: a combination of bioinformatics and machine learning

Yihong Lin¹, Jingmei Song^{2*} and Xiaohong Li^{3*}

¹The Second School of Clinical Medicine, Zhejiang Chinese Medical University, Hangzhou, China,

²School of Basic Medicine, Zhejiang Chinese Medical University, Hangzhou, China, ³School of
Pharmaceutical Sciences, Zhejiang Chinese Medical University, Hangzhou, China

Background: Non-alcoholic steatohepatitis (NASH) commonly complicates coronary artery disease (CAD), yet the interaction mechanism remains unclear. Our research seeks to investigate the common mechanisms and key signature genes between CAD and NASH.

Methods: RNA sequence information for CAD and NASH was screened from the GEO database. Weighted gene co-expression network analysis (WGCNA) and differentially expressed gene analysis identified key genes, followed by functional enrichment analysis of these shared genes. Three machine learning methods—LASSO, random forest, and SVM-RFE—were used to identify signature genes. Gene set enrichment analysis (GSEA) was then performed to explore potential mechanisms associated with the signature genes. In addition, single-sample gene set enrichment analysis (ssGSEA) evaluated immune infiltration in CAD and NASH and its correlation with the signature genes.

Results: WGCNA has revealed two key modules for CAD and NASH. The intersection of the CAD modules and their differential genes narrowed the key genes down to 2,808 shared genes. Finally, 44 shared genes were selected for both CAD and NASH. Kyoto Encyclopedia of Genes and Genomes analysis showed that these genes were primarily enriched in insulin resistance and inflammation pathways. Machine learning identified the signature genes BATF3, SOCS2, and GPER, all with ROC values above 0.7, validated in external datasets. GSEA revealed that these genes act through common mechanisms in CAD and NASH, regulating metabolic, inflammatory, and cardiovascular pathways. In addition, ssGSEA suggested their involvement in immune cell infiltration.

Conclusion: BATF3, SOCS2, and GPER have emerged as promising gene candidates that may serve as biomarkers or potential therapeutic targets for

CAD combined with NASH, linked to the regulation of metabolic, inflammatory, and cardiovascular pathways. We also identified insulin resistance and inflammation pathways as common mechanisms underlying both diseases.

KEYWORDS

coronary artery disease, non-alcoholic steatohepatitis, machine learning, WGCNA, bioinformatics

1 Introduction

Coronary artery disease (CAD) is a significant type of cardiovascular disease (CVD) and one of the leading causes of global mortality and the loss of disability-adjusted life years (Ralapanawa and Sivakanesan, 2021). There are several risk factors that can lead to CAD. These include family history, smoking, alcohol consumption, diabetes, age, obesity, and hypertension (Madhavan et al., 2018). The development of atherosclerotic plaques is the main pathological feature of CAD. This process involves the continuous deposition of excess cholesterol and cholesterol esters in the arterial intima, resulting in the proliferation of connective tissue, thickening and hardening of the arterial wall, and subsequent connective tissue necrosis (Chiu et al., 2018).

Non-alcoholic fatty liver disease (NAFLD) has emerged as the most common chronic liver condition in adults globally, affecting roughly 25% of the population (Powell et al., 2021). NAFLD represents a significant public health issue, contributing greatly to liver-related morbidity and mortality. In addition, it has become the most rapidly increasing cause of hepatocellular carcinoma. Non-alcoholic steatohepatitis (NASH) is a more severe form of NAFLD. It is characterized by the development of hepatocellular ballooning, lobular inflammatory infiltrates, and fibrosis developing atop simple hepatic steatosis (Książek et al., 2024). Current studies have found that the incidence of cirrhosis in NASH patients within 10–15 years is ten times higher than that in patients with simple steatosis, reaching 15%–25%. Moreover, the five-year and ten-year survival rates for patients with NASH are 67% and 59%, respectively (Wang et al., 2014; Fan et al., 2019). Therefore, NASH is the primary target for clinical interventions aimed at preventing the progression of NAFLD and represents a critical stage in the progression of liver disease in NAFLD patients. Patients with NAFLD exhibit cardiometabolic risks comparable to those observed in individuals with CAD, including inflammation, dyslipidemia, and endothelial dysfunction. The presence of NAFLD may serve as a considerable risk factor for the advancement of CAD (Zhang et al., 2020; Zhang and Fang, 2021). One study found that among the subtypes of NAFLD, patients with NASH pose a greater risk of dying from CVD than those with simple steatosis alone (Vanni et al., 2015).

A robust correlation exists between NASH and atherosclerotic cardiovascular disease. NASH has been demonstrated to contribute to the onset and progression of CVD. This is accomplished through inflammatory mediators released systemically (e.g., TNF- α , CRP, PCF) and oxidative stress markers (Targher et al., 2010). Nevertheless, the precise mechanisms through which NASH contributes to the development of atherosclerosis remain poorly understood. The molecular biology research on the shared mechanisms of the two diseases is limited. Genetic variants in PNPLA3, TM6SF2, and GCKR have been proposed as potential drivers of atherosclerosis in NAFLD patients, though this remains a topic of debate (Petta et al., 2013; Dongiovanni et al., 2015; Xia et al., 2016). Acute myocardial infarction constitutes an acute manifestation of CAD. A study in China found that elevated serum FSTL3 levels may increase the risk of fibrosis and acute myocardial infarction in type 2 diabetes patients with NAFLD (Duan et al., 2023). Bioinformatics studies have identified *PLCXD3*, *CCL19*, *PKD2* and *MMP9* as signature genes in atherosclerosis and NAFLD (Mo et al., 2023; Lv et al., 2024). Although NASH has received widespread attention as an important subtype of NAFLD, there is still a lack of bioinformatics studies directly focusing on the link between CAD and NASH.

This study aims to investigate the common pathogenic mechanisms of CAD and NASH by analyzing transcriptomic data sourced from the Gene Expression Omnibus (GEO) database. To this end, bioinformatics approaches and machine learning methods will be employed to explore common pathogenic pathways, signature genes, and immune infiltration profiles associated with CAD and NASH.

2 Materials and methods

2.1 GEO dataset download and process

The key terms “coronary artery disease” and “non-alcoholic steatohepatitis” were employed to search for gene expression profiles associated with CAD and NASH in the GEO database (Home - GEO - NCBI (nih.gov)). The datasets were considered eligible if they met the following criteria: high-throughput expression data, human tissue samples, and at least 10 samples per group to ensure the reliability of the weighted gene co-expression network analysis (WGCNA). Finally, the following datasets were chosen for analysis: GSE113079, GSE66360, GSE89632, and GSE135251. Of these, GSE113079 and GSE89632 were subjected to comprehensive analysis, while GSE66360 and GSE135251 were employed for validation purposes. The GSE113079 dataset comprises 93 CAD patients and 48 normal controls, while the GSE89632 dataset

Abbreviations: CAD, Coronary artery disease; NASH, Non-alcoholic steatohepatitis; WGCNA, Weighted gene co-expression network Analysis; GO, Gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; GSEA, Gene set enrichment analysis; ssGSEA, Single-sample gene set enrichment analysis; $\gamma\delta$ T Cells, gamma delta T cells.

contains 19 NASH patients and 24 normal controls. The GSE64566 dataset (GPL6947 platform, 26 CAD cases, and 20 controls) and the GSE135251 dataset (155 NASH cases and 10 controls) were employed to validate the results further. The “limma” package in R was used to normalize the data, ensuring appropriate corrections across datasets were implemented. Probe annotations have been mapped to gene symbols by means of the annotation files provided by the respective platforms. The resulting gene expression matrices were utilized for subsequent analyses, with samples as rows and gene symbols as columns.

2.2 Weighted gene co-expression network analysis

The modules associated with CAD and NASH were explored using WGCNA (Langfelder and Horvath, 2008). First, hierarchical clustering was performed using the “hclust” function in R to remove outliers. The “pickSoftThreshold” function in the WGCNA package was used to search for the best soft-threshold power and adjacency according to scale-free network criteria. The adjacency matrix was converted into a topological overlap matrix, and phase differences were calculated. A hierarchical clustering dendrogram was constructed, and genes exhibiting analogous expression patterns were grouped into discrete modules. The dynamic tree-cut method was employed to construct the co-expressed gene modules, with the minimum module size set to 50. Finally, the relationship between the gene module and the disease was assessed using the gene significance and module membership values to identify the core module.

2.3 Identifying differentially expressed genes between CAD samples and controls

Differential expression analysis was performed on the CAD dataset to further narrow down the number of genes identified in the key CAD module obtained from WGCNA. The R package “limma” (Ritchie et al., 2015) was used to identify differentially expressed genes (DEGs) in GSE113079 using the inclusion criteria $|\log_2 \text{Fold Change}| \geq 0.5$ and $p - \text{adjust} < 0.05$. Volcano plots were generated to characterize the DEGs, and the expression levels of the 50 most significantly represented genes were plotted using the heatmap. The intersection of key modules and differentially expressed genes in the CAD dataset is defined as CAD key genes. We employed the JVenn online tool to construct a Venn diagram (Philippe Bardou et al., 2014).

2.4 Identification and enrichment analysis of shared genes between CAD and NASH

Genes that are present in both the CAD and NASH key modules and have the same expression trend in both diseases are defined as shared genes. The visualization was carried out using Venn diagrams (Philippe Bardou et al., 2014). To investigate the potential functions of the genes shared between CAD and NASH, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses using the org.Hs.e.g.db, ggplot2, clusterProfiler, and enrichplot packages in R. GO analysis was

utilized to clarify the biological processes (BP), molecular functions (MF), and cellular components (CC) involved in the overlapping genes. KEGG pathway analysis was conducted, aiding in the identification of the underlying biological mechanisms within the shared genes.

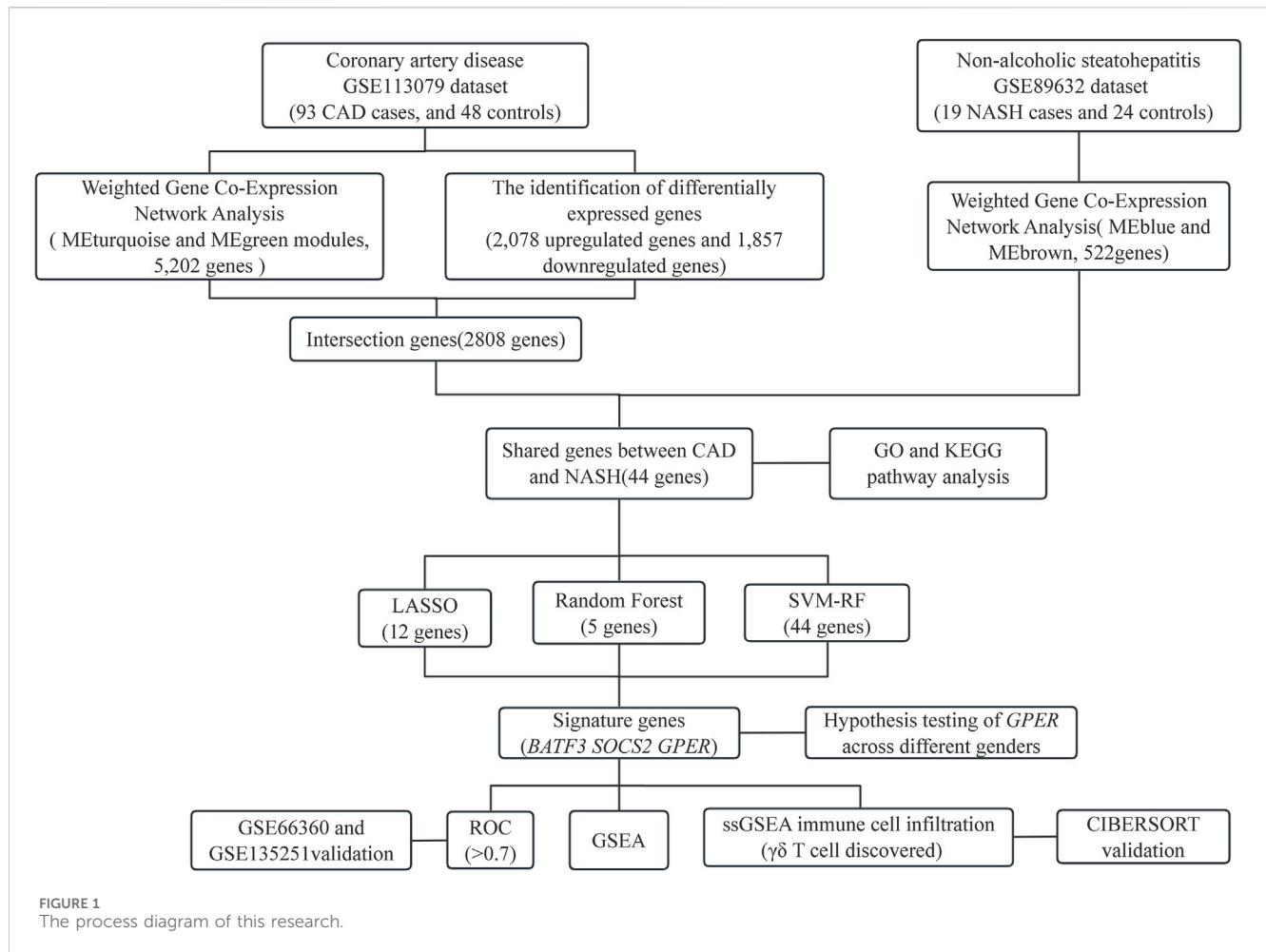
2.5 Machine learning and signature gene identification

The shared genes between CAD and NASH were investigated using three machine-learning methods. The goal was to find common signature genes using Least Absolute Shrinkage and Selection Operator (LASSO) Regression, random forest (RF), and support vector machine recursive feature elimination (SVM-RFE). LASSO regression is an improved form of linear regression which introduces an L1-norm penalty to the least squares objective function. This regularization shrinks some regression coefficients to zero, thereby enabling effective variable selection (Tibshirani, 1996; Ranstam and Cook, 2018). The LASSO analysis was performed using the “glmnet” package, with the penalty parameters selected through 10-fold cross-validation. RF is an ensemble learning method that enhances prediction stability and accuracy by aggregating the results of multiple decision trees (Cutler et al., 2012). In this study, gene importance scores were computed using the RF algorithm to identify candidate genes that may play a crucial role in the development and progression of disease. The “randomForest” package was used to perform the analysis. To identify the ideal number of trees for the RF algorithm executed by the “randomForest” package, the error rate was initially computed for a series of values ranging from 1 to 500 trees. Ultimately, the significance of each candidate signature gene was assessed using the “importance” function, and the top five genes were chosen according to the computed importance scores. SVM-RFE is a recursive feature elimination method based on support vector machines that iteratively trains the model and removes the least informative features to identify the most significant variables for classification (Guyon et al., 2002). The SVM-RFE algorithm was trained and evaluated using the “e1071” and “caret” packages.

The overlapping genes between the genes derived by the three machine learning algorithms were called signature genes. A Venn diagram was constructed to illustrate the overlap between the two groups. The diagnostic performance of these signature genes was evaluated using the area under the receiver operating characteristic (ROC) curve (AUC). An AUC value greater than 0.7 was considered indicative of excellent diagnostic performance, while values between 0.5 and 0.7 were considered indicative of satisfactory performance. To facilitate the visualization of gene expression in the GSE113079 and GSE89632 datasets, boxplots were generated using the R package “ggplot2”.

2.6 Validating key genes

Validation was conducted using the GSE66360 and GSE135251 datasets to confirm the reliability of the expression levels of these signature genes in CAD and NASH. We compared the expression levels of the signature genes between the patient and control cohorts to determine significant differences and generated



boxplots for visual representation. The diagnostic performance of these signature genes was then evaluated using the AUC.

2.7 Gene set enrichment analysis

To analyze the role of the signature genes in these two diseases, we used gene set enrichment analysis (GSEA) with GSEA software version 4.3.2. The gene set “c2. cp.kegg_legacy.v2023.2. Hs.symbols.gmt” was taken as the reference. The genes were then separated into two groups based on the expression levels in the software, and GSEA was conducted to identify significant pathways with a p-value below 0.05.

2.8 Immune infiltration analysis

The variations in immune cell infiltration in peripheral blood samples from CAD and NASH patients were compared with normal peripheral blood samples using single-sample gene set enrichment analysis (ssGSEA). We calculated Spearman's rank correlation coefficients in R for immune cells with differentially signature genes to analyze their relationship with specific genes. The CIBERSORT immunization algorithm was validated against immune cells associated with and expressing the same trend of signature genes in both diseases.

2.9 Statistical analysis

Statistical analyses were performed with R software (version 4.3.1). P-values below 0.05 were deemed statistically significant, and all P-values were computed as two-tailed. The flowchart in Figure 1 shows the methodology of the study.

3 Results

3.1 Identifying co-expression modules in CAD and NASH

WGCNA analysis of CAD and NASH-related genes was performed to identify key genetic modules. As a result, Spearman's correlation coefficient was employed to create heat maps of module-trait associations. Each color represents a different module (Figures 2A,B). Figure 2C shows that eight gene modules were detected in the GSE113079 dataset. The MEturquoise module demonstrated a robust positive correlation with the disease ($r = 0.79, p = 1e - 31$), whereas the MEgreen module exhibited a negative correlation ($r = -0.45, p = 2e - 08$). Accordingly, the MEturquoise and MEgreen modules were identified as CAD key modules, encompassing a total of 5,202 genes. In the GSE89632 dataset, three gene modules were identified

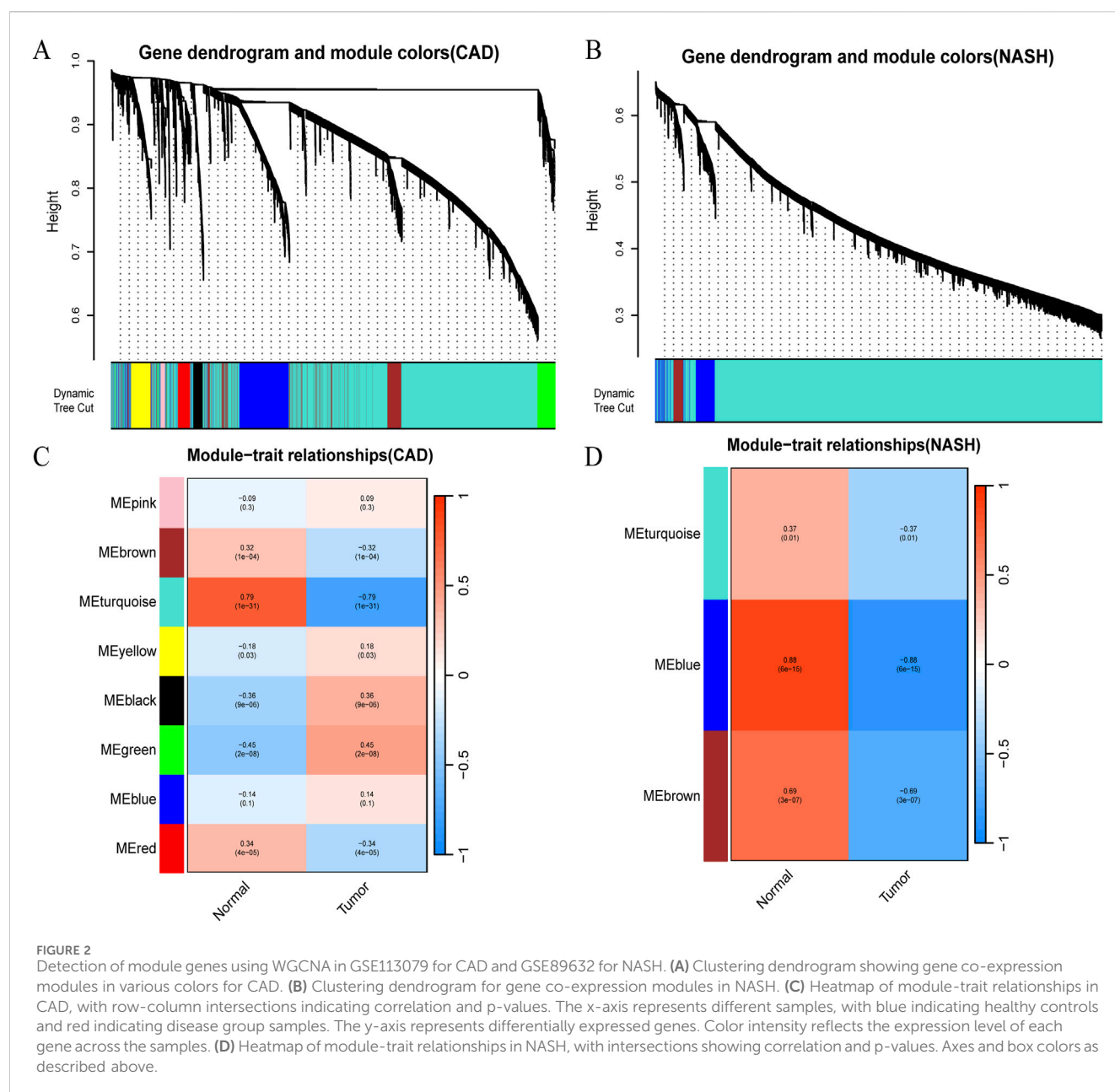


FIGURE 2

Detection of module genes using WGCNA in GSE113079 for CAD and GSE89632 for NASH. (A) Clustering dendrogram showing gene co-expression modules in various colors for CAD. (B) Clustering dendrogram for gene co-expression modules in NASH. (C) Heatmap of module-trait relationships in CAD, with row-column intersections indicating correlation and p-values. The x-axis represents different samples, with blue indicating healthy controls and red indicating disease group samples. The y-axis represents differentially expressed genes. Color intensity reflects the expression level of each gene across the samples. (D) Heatmap of module-trait relationships in NASH, with intersections showing correlation and p-values. Axes and box colors as described above.

(Figure 2D). The MEblue ($r = 0.88$, $p = 6e - 15$) and MEbrown ($r = 0.69$, $p = 3e - 07$) modules were selected as key modules for NASH, comprising 375 and 147 genes, respectively.

3.2 Identifying DEGs between CAD samples and controls

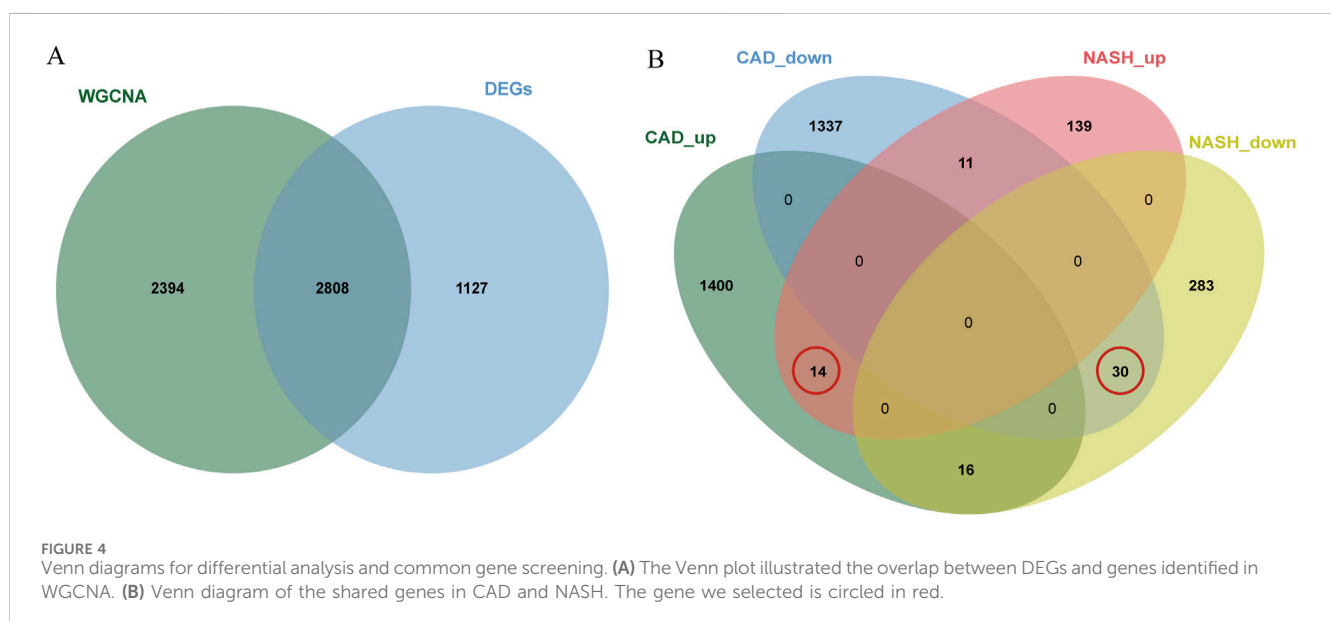
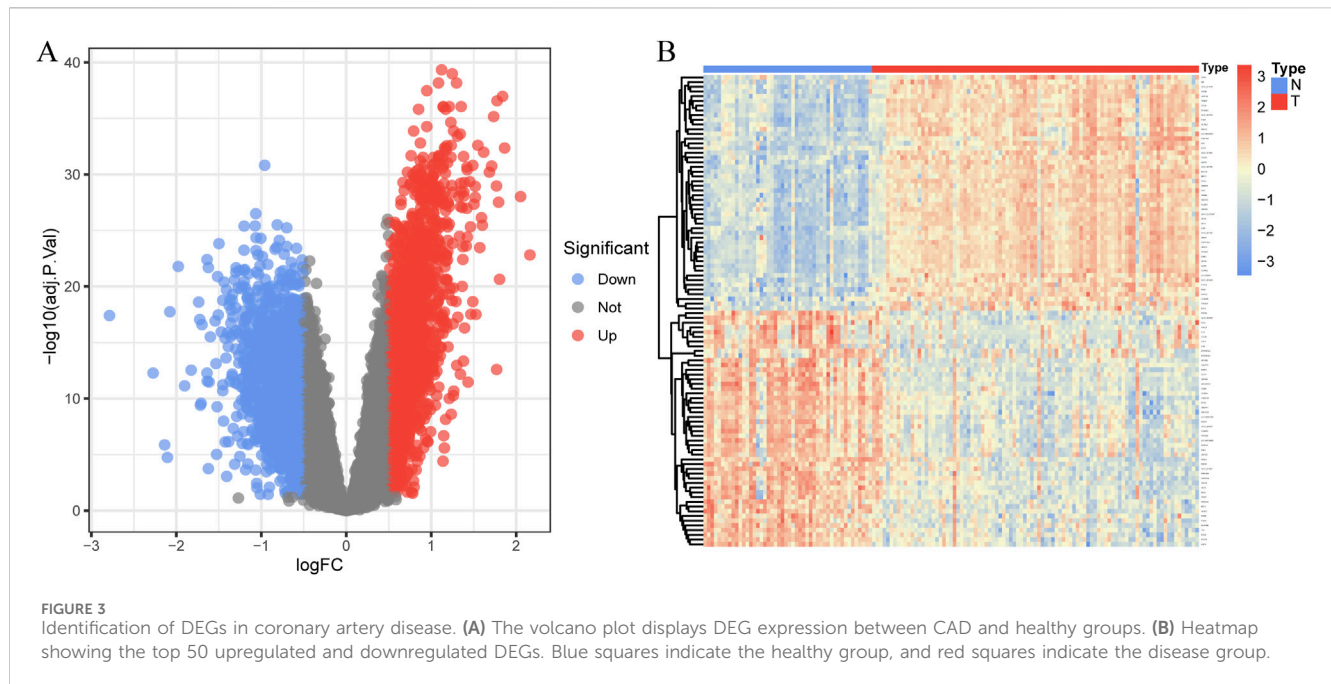
In total, 3,935 genes were screened, including 2,078 upregulated genes and 1,857 downregulated genes (Figure 3A). A heatmap displaying the top 50 upregulated and DEGs between CAD samples and the healthy group is shown in Figure 3B. The intersection of DEGs in CAD and the key modules in WGCNA was taken, resulting in a total of 2,808 genes, and a Venn diagram was drawn (Figure 4A).

3.3 Identifying of shared gene signatures in CAD and NASH

A total of 44 genes were recognized by the overlap of the previously identified key CAD genes and the genes from the blue and brown modules of NASH (Figure 4B).

3.4 GO and KEGG enrichment analysis of shared genes

The KEGG and GO analyses identified several enriched pathways relevant to the investigated diseases. In the KEGG analysis results, we displayed only the pathways related to our disease research, including apoptosis, FoxO signaling pathway,

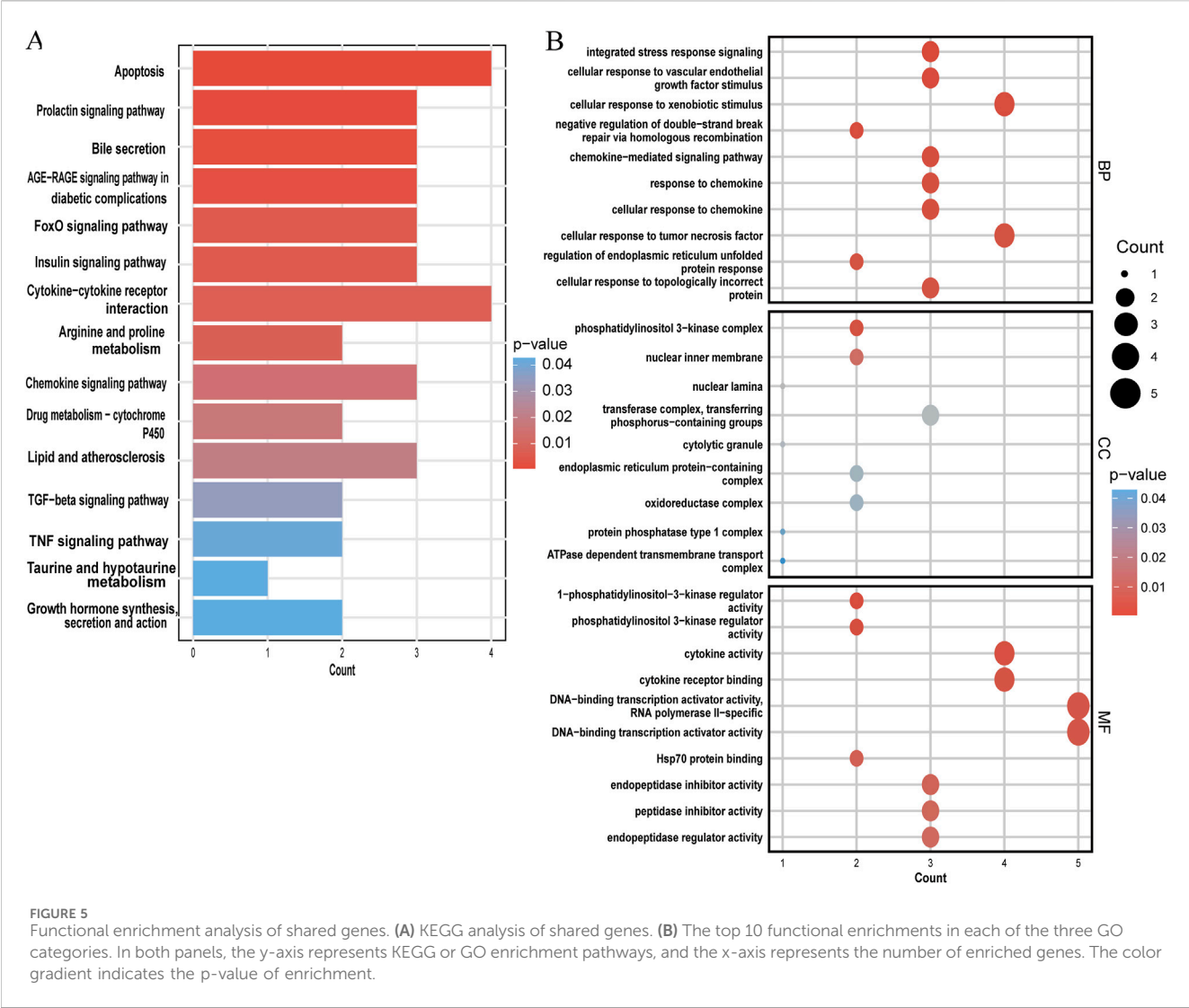


insulin signaling pathway, cytokine-cytokine receptor interaction, arginine and proline metabolism, lipid and atherosclerosis, TGF-beta signaling pathway, TNF signaling pathway, etc. and drew a bar graph (Figure 5A). Three categories make up the GO analysis: BP, CC, and MF. The BP analysis results revealed that the three most notable processes were integrated stress response signaling, cellular response to vascular endothelial growth factor stimulus, and cellular response to xenobiotic stimulus. In the CC analysis, the top three categories were identified as phosphatidylinositol 3-kinase complex, nuclear inner membrane, and nuclear lamina. In addition, 1-phosphatidylinositol 3-kinase regulator activity, cytokine activity

and cytokine receptor binding were observed to have a significant role in the MF category (Figure 5B).

3.5 Machine learning and identification of signature genes

The LASSO analysis revealed 12 signature genes (Figures 6A,B), whereas the RF analysis highlighted the top five most significant signature genes (Figures 6C,D). Genes selected based on LASSO are in Supplementary Table S1. The SVM-RFE analysis indicated that a



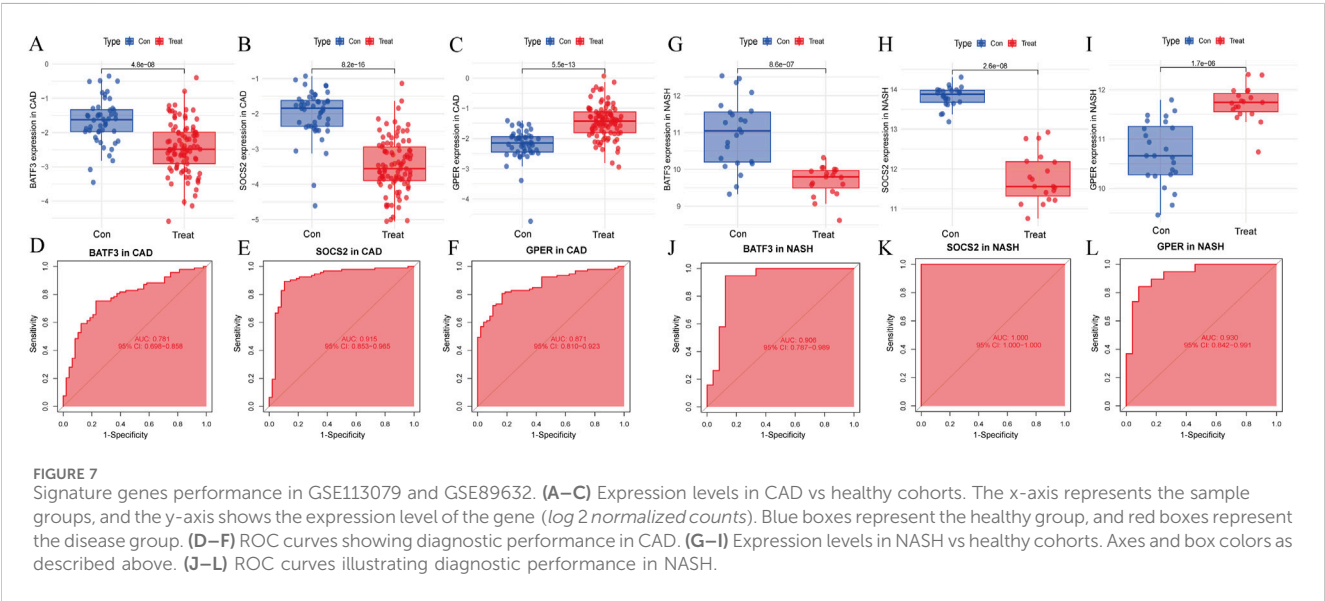
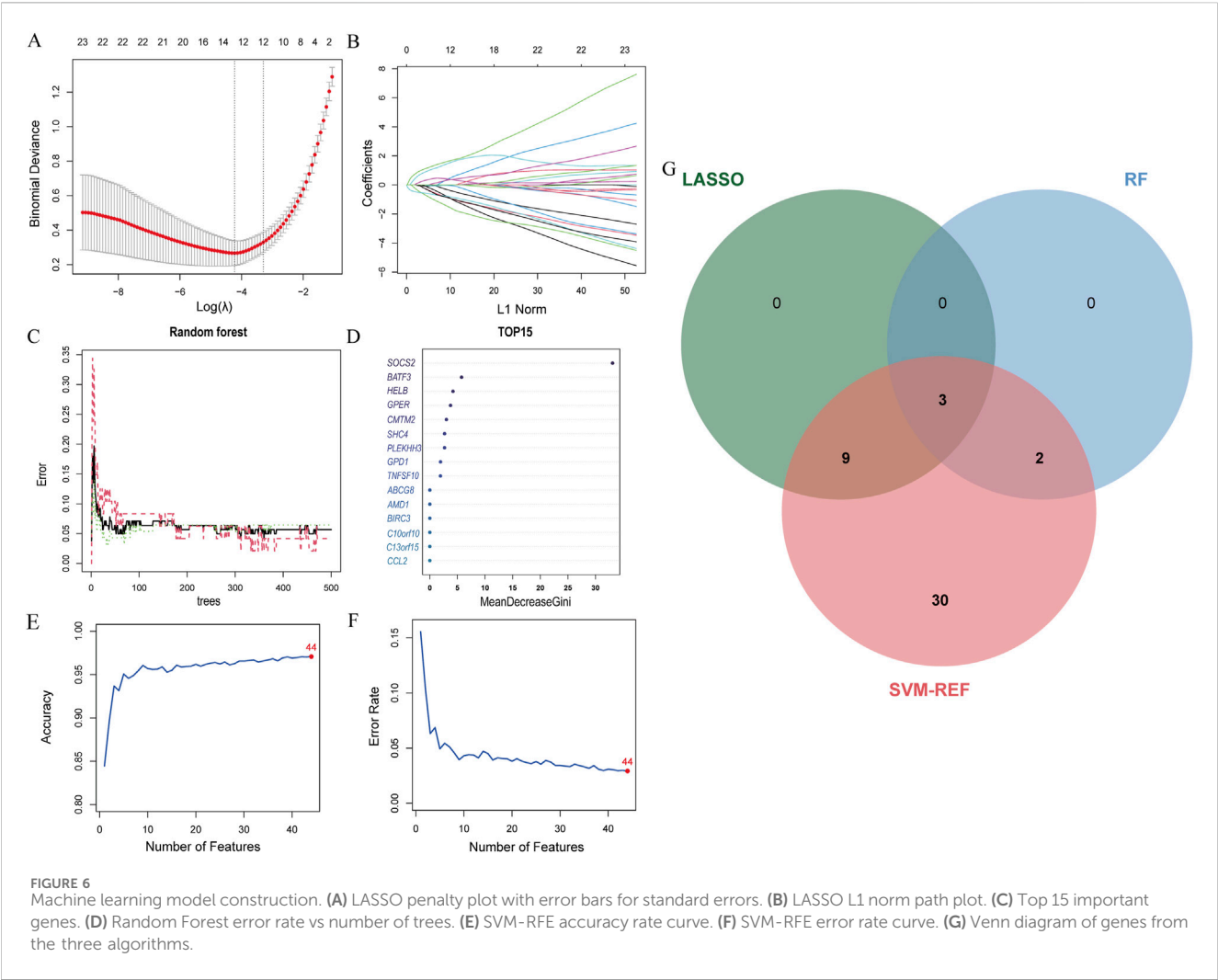
model containing 44 genes demonstrated the highest accuracy (Figures 6E,F). The genes obtained from the three algorithms were ultimately found to be in common, identifying three signature genes. The final signature genes were *BATF3*, *GPER*, and *SOCS2* (Figure 6G). Both *BATF3* and *SOCS2* were lowly expressed in CAD and NASH (Figures 7A,B; Figures 7G,H). This suggests that those genes are suppressed in the disease. In both conditions, the parameter test for these two genes yielded p-values less than 0.05, indicating a statistically significant difference in expression between the normal and disease groups. In patients with CAD, the area under the AUC for the signature genes *BATF3* and *SOCS2* were 0.781 and 0.915, respectively (Figures 7D,E). In NASH, the AUC values for *BATF3* and *SOCS2* were 0.906 and 1, respectively (Figures 7J,K). Both genes exhibited AUC values exceeding 0.7, indicative of robust predictive efficacy. Conversely, *GPER* was highly expressed in both CAD and NASH (Figures 7C,I), suggesting that this gene is upregulated significantly in both diseases. The area under the AUC for the signature gene *GPER* in CAD was 0.871 (Figure 7F). In NASH, it was 0.930 (Figure 7L). Once more, both AUC values exceeded 0.7, indicating robust predictive efficacy.

3.6 Hypothesis testing of *GPER* across different genders

Since *GPER* acts as an estrogen receptor, it functions mainly in women (Morán-Costoya et al., 2021). To avoid the bias caused by gender differences, we verified whether *GPER* still has expression differences in male patients by significance testing. Due to the absence of gender-specific data in the remaining datasets, the NASH dataset GSE89632 was the sole source available for this analysis. The samples were divided into two groups, one comprising male subjects and the other composed of female subjects. We then separately compared the differences between the two groups of patients and normal samples. Male and female patient groups showed statistically significant differences compared to normal subjects ($p < 0.001$) (Figure 8).

3.7 Validating key genes

Next, we assessed whether the signature genes were significantly different in diseased and normal tissues in external validation



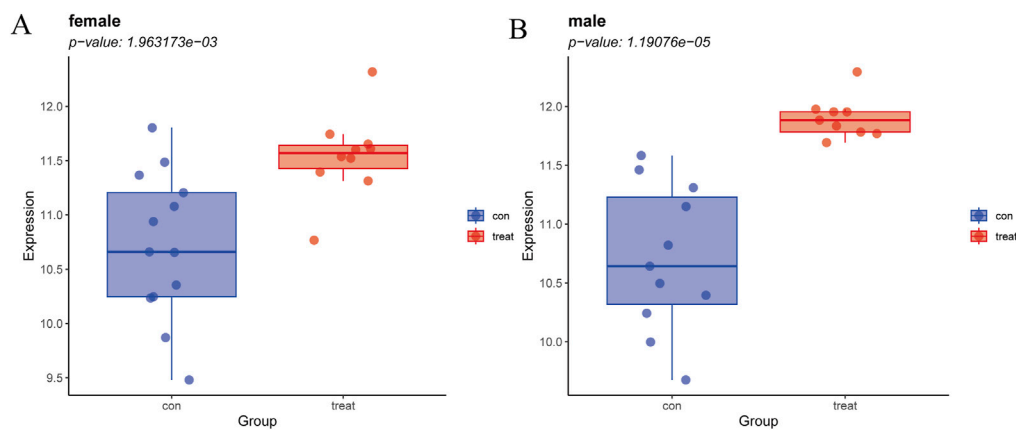


FIGURE 8
(A) Expression levels of GPER in NASH vs. healthy cohorts in females. (B) Expression levels of GPER in NASH vs. healthy cohorts in males.

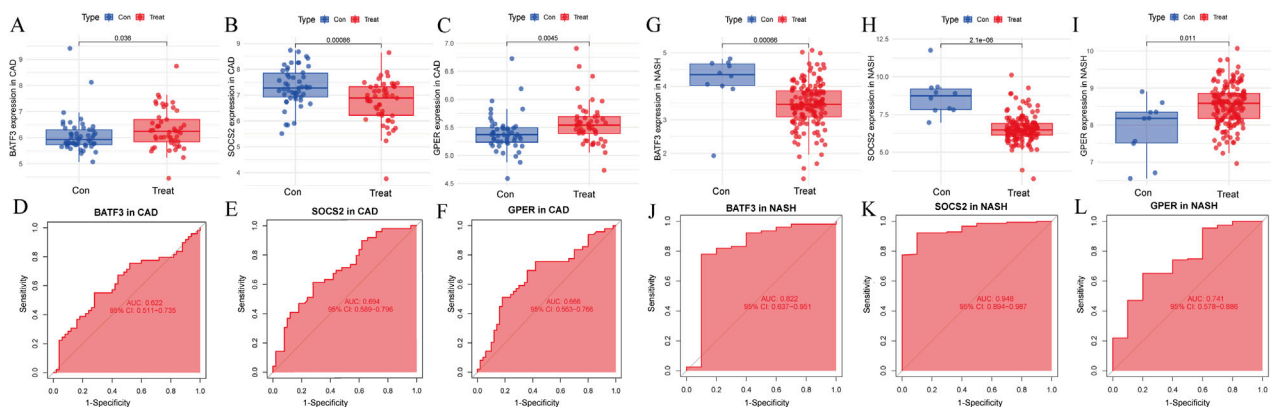


FIGURE 9
Evaluation of signature genes in GSE66360 and GSE135251. (A–C) Expression levels of signature genes in CAD vs. healthy cohorts. (D–F) ROC curves demonstrating the predictive power of signature genes in CAD. (G–I) Expression levels of signature genes in NASH vs. healthy cohorts. (J–L) ROC curves showcasing the diagnostic efficacy of the signature genes in NASH.

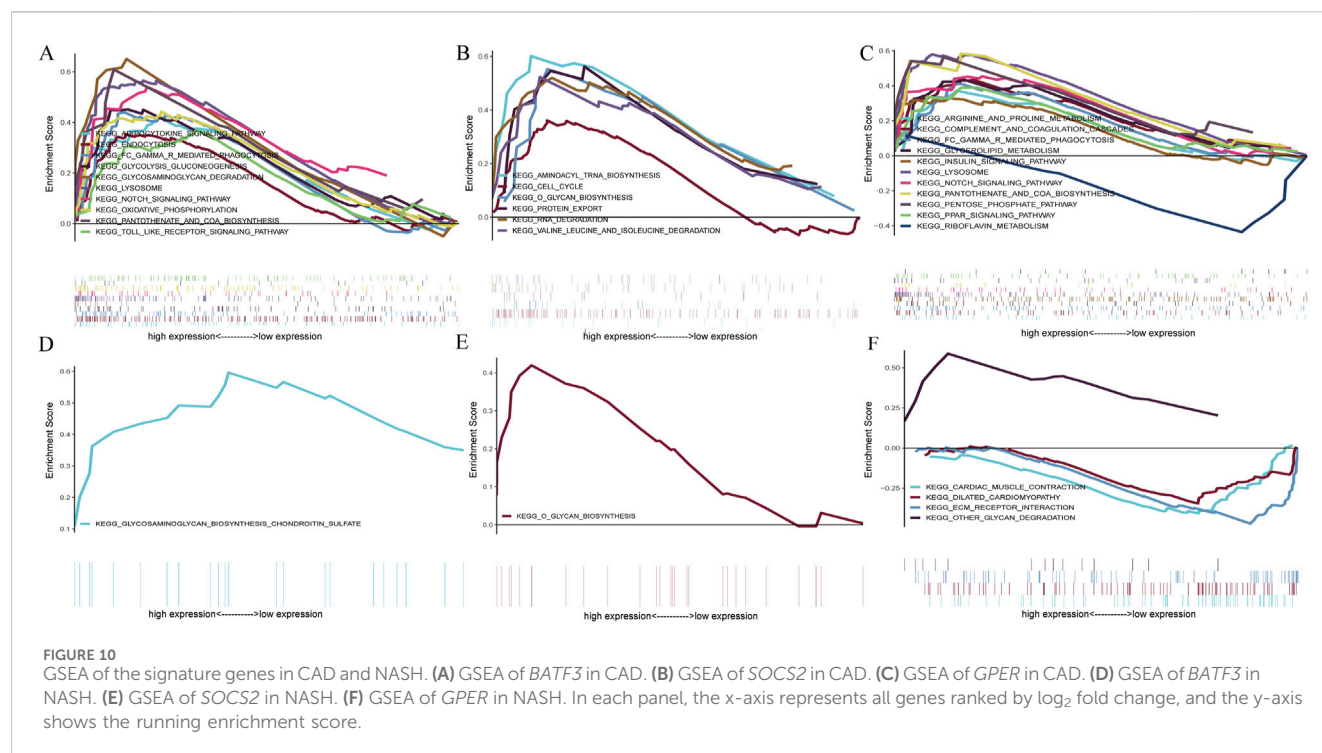
cohorts. We also evaluated the ROC curves of these genes to determine their diagnostic efficacy in predicting NASH and CAD. In the GSE66360(CAD validation cohort), the expression of *GPER* was significantly higher in the disease cohort than in the control cohort. The other genes showed lower expression levels, consistent with the findings in GSE113079 (Figures 9A–C). All three genes in this group were statistically different between the disease and normal groups ($p < 0.05$). The AUC values derived from the ROC curves for *BATF3*, *SOCS2*, and *GPER* were 0.622, 0.695, and 0.666, respectively (Figures 9D–F). All AUC values were greater than 0.5 but less than 0.7, indicating moderate predictive ability.

In the GSE135251 (NASH validation cohort), *GPER* was once more identified as being highly expressed relative to the control group. In contrast, the other genes displayed lower expression levels, consistent with the findings from GSE89632 (Figures 9G–I). All three genes showed statistically significant differences between the disease and normal groups ($p < 0.05$). The AUC values for *BATF3*, *SOCS2* and *GPER* were 0.822, 0.741 and 0.948, respectively. Each

AUC value was greater than 0.7, demonstrating robust predictive ability (Figures 9J–K).

3.8 GSEA analysis

GSEA analysis assessed the pathways related to the signature genes to identify their relevance to the disease. In the context of CAD, The upregulation of *BATF3* was found to have statistically significant relevance to several pathways, including adipocytokine signaling pathway, endocytosis, Fc gamma R-mediated phagocytosis, glycolysis gluconeogenesis, glycosaminoglycan degradation, lysosome, Notch signaling pathway, oxidative phosphorylation, pantothenate and CoA biosynthesis, and Toll-like receptor signaling pathway (Figure 10A). The upregulation of *SOCS2* was significantly associated with aminoacyl-tRNA biosynthesis, cell cycle, O-glycan biosynthesis, protein export, RNA degradation, and valine leucine and isoleucine degradation (Figure 10B). The upregulation of *GPER* is significantly associated



with pathways related to arginine and proline metabolism, glycerolipid metabolism, insulin signaling pathway, etc. Its downregulation was significantly associated with riboflavin metabolism (Figure 10C). In the context of NASH, the upregulation of *BATF3* was significantly associated with glycosaminoglycan biosynthesis of chondroitin sulfate (Figure 10D). The upregulation of *SOCS2* was significantly associated with O-glycan biosynthesis (Figure 10E). The upregulation of *GPER* was significantly associated with other glycan degradation, whereas the downregulation was associated with ECM-receptor interaction, cardiac muscle contraction, and dilated cardiomyopathy (Figure 10F).

In conclusion, *BATF3* is involved in the metabolic pathways related to both diseases. *SOCS2* has been shown to activate the O-glycan biosynthesis pathway in both CAD and NASH. *GPER*'s upregulation is associated with pathways related to glycolipid metabolism in both conditions, while its downregulation is linked to pathways concerning cardiovascular function. This suggests that *GPER* may be important in regulating metabolism and maintaining cardiovascular health.

3.9 Immune cell infiltration analysis

For CAD, CAD patients showed higher infiltration levels of CD56bright natural killer cells, gamma delta T cells ($\gamma\delta$ T cells), immature B cells, neutrophils, regulatory T cells, etc., compared to the normal group. Conversely, infiltration levels of activated B cells, activated CD4 T cells, activated CD8 T cells, CD56dim natural killer cells, myeloid-derived suppressor cells (MDSC), etc. were lower (Figure 11A). For NASH, the patients showed higher infiltration levels of activated CD8 T cells, CD56bright natural killer cells, $\gamma\delta$ T cells, immature B cells, and monocytes compared to the normal

group. In contrast, activated CD4 T cells, activated dendritic cells, CD56dim natural killer cells, MDSCs, etc., showed lower infiltration levels (Figure 11B).

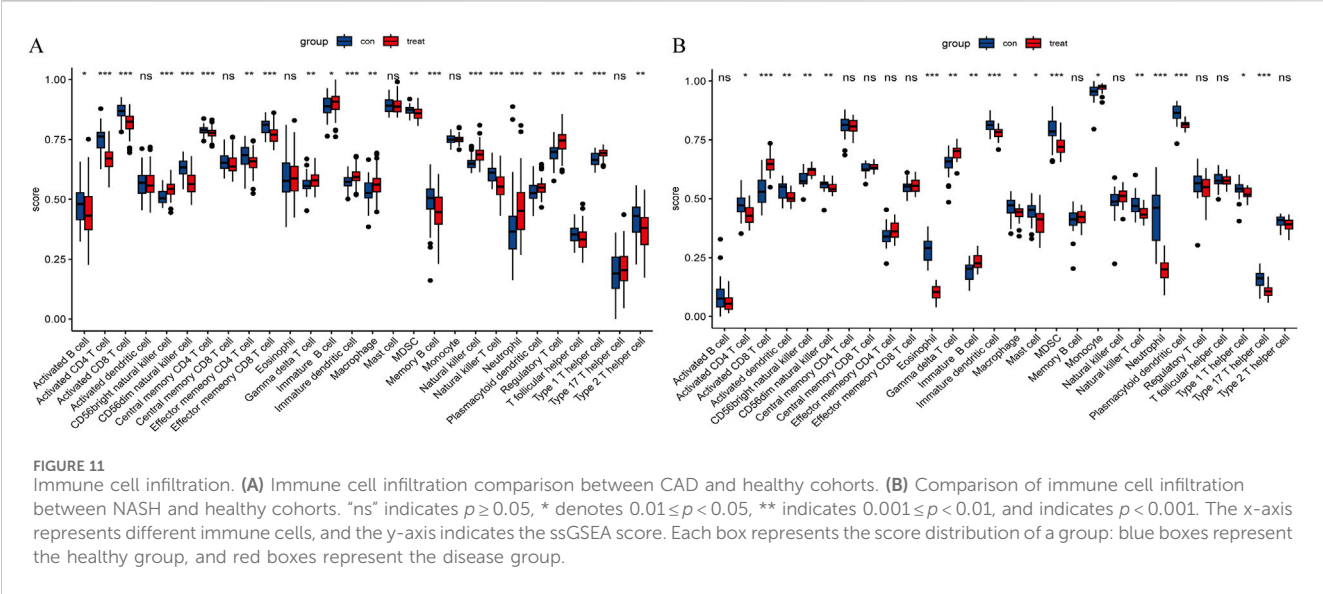
Notably, in both diseases, the infiltration levels of CD56bright natural killer cells, $\gamma\delta$ T cells, and immature B cells were significantly elevated compared to the normal group, whereas the infiltration levels of activated CD4 T cells, CD56dim natural killer cells, and MDSCs were reduced.

3.10 The expression of *BATF3* shows a positive correlation with $\gamma\delta$ T cells

The Spearman rank correlation coefficients between the differential immune cell populations and the feature genes were evaluated (Figures 12A,B). The results of our analysis reveal that there is a positive correlation between $\gamma\delta$ T cells and the expression levels of *BATF3* in both disease states examined. As previously stated, the number of $\gamma\delta$ T cells is increased in both conditions. To validate the expression trend of $\gamma\delta$ T cells in both diseases, we employed the CIBERSORT algorithm. The results obtained from this analysis were consistent with those derived from ssGSEA (Figures 12C,D).

4 Discussion

NASH and CAD are common diseases that significantly affect the health of patients. In recent years, increasing epidemiologic, genetic, pathologic, and clinical observational evidence suggests that NAFLD may be an independent risk factor for CAD (Zhang D. et al., 2023). NASH is a subtype of NAFLD with more severe damage and is an important intermediate link in the development from simple



fatty degeneration of hepatocytes to liver fibrosis and liver cancer. An epidemiologic study on NASH showed a global prevalence of NASH of more than 15% as of 2016 (Dong et al., 2024). In addition, approximately 20% of NASH patients will develop cirrhosis (Sheka et al., 2020). Moreover, higher risk of CAD in NASH patients than in those with simple steatosis, a Japanese study shows (Niikura et al., 2020). Therefore, further research into the incidence and development mechanism of NASH is especially important. This study is designed to discover the common genetic markers and pathogenesis of NASH and CAD and to evaluate the common immune microenvironment of the two diseases in pursuit of more effective preventive measures, diagnosis, and treatment approaches for these diseases. In this study, WGCNA and differential gene selection were performed on the two diseases to find out the common key gene sets. By performing GO and KEGG functional analysis on 44 potential target genes related to the two diseases, the various biological processes and pathways associated with these genes were uncovered, thereby determining their potential modes of action in the disease. In addition, three machine learning algorithms were applied to identify three signature genes, *BATF3*, *SOCS2*, and *GPB*, and their respective roles in the disease were explored. Finally, ssGSEA was employed to assess the shared immune infiltration in CAD and NASH and to analyze its correlation with signature genes.

At present, the recognized common risk factors for NAFLD and CAD include metabolic syndrome, insulin resistance, dyslipidemia, as well as obesity (Brouwers et al., 2019). While our KEGG pathway analysis confirmed the involvement of shared pathways such as the insulin signaling and FoxO signaling pathways, a deeper understanding of their functional implications is warranted. FoxO transcription factors are downstream mediators of insulin signaling. Their activity is regulated by serine/threonine phosphorylation catalyzed by protein kinase B and other kinases. Prior research has demonstrated their substantial contribution to lipid homeostasis in diet-induced fatty liver disease (Pan et al., 2017), thereby indicating a mechanistic association with NASH pathogenesis. Based on the analysis as mentioned above and

conclusions, we postulate that FoxO may influence the pathogenesis of CAD by regulating mechanisms such as insulin signaling, oxidative stress, and inflammatory responses. In light of these findings, the FoxO signaling pathway may serve as a viable intervention target for CAD in conjunction with NASH.

In addition, our study demonstrated that the development of both diseases is also significantly influenced by inflammatory pathways, including cytokine-cytokine receptor interaction, TNF signaling pathway and TGF-beta signaling pathway. Endothelial dysfunction, a hallmark of early-stage CAD, initiates an inflammatory cascade by promoting the infiltration of macrophages and the release of pro-inflammatory mediators, including cytokines and chemokines, thereby amplifying local vascular inflammation (Zouridakis et al., 2004; Ebadi et al., 2022). Studies have demonstrated that levels of inflammatory factors are statistically significantly increased in the epicardial tissue of CAD patients (Zhou et al., 2011). Similarly, NASH is characterized by hepatic lipid accumulation, hepatocyte injury, and fibrosis, all of which are closely linked to an enhanced inflammatory milieu (Thibaut et al., 2022). Notably, TNF- α not only speeds up the progression of NASH but also prevents hepatocyte apoptosis by inducing TNF- α -induced protein 8-like 1. This action helps reduce steatosis and fibrosis, offering protection against NASH (Wu et al., 2021). In conclusion, inflammatory signaling pathways are of great importance in the pathogenesis of both CAD and NASH and may serve as common biomarkers for these diseases.

Basic leucine zipper ATF-like transcription factor 3 (*BATF3*) is a member of the basic leucine zipper transcription factor family and plays a part in controlling gene expression. This process is facilitated by the formation of heterodimers with multiple transcription factors. *BATF3* is a key transcription factor involved in generating CD8 α^+ dendritic cells within lymphoid tissues and CD103 $^+$ dendritic cells in the skin. Our study shows that *BATF3* levels are decreased in the peripheral blood of patients with CAD and NASH and that this expression is positively correlated with $\gamma\delta$ T-cell infiltration in these conditions. Furthermore, elevated levels of $\gamma\delta$ T cell expression were observed in both diseases. This finding

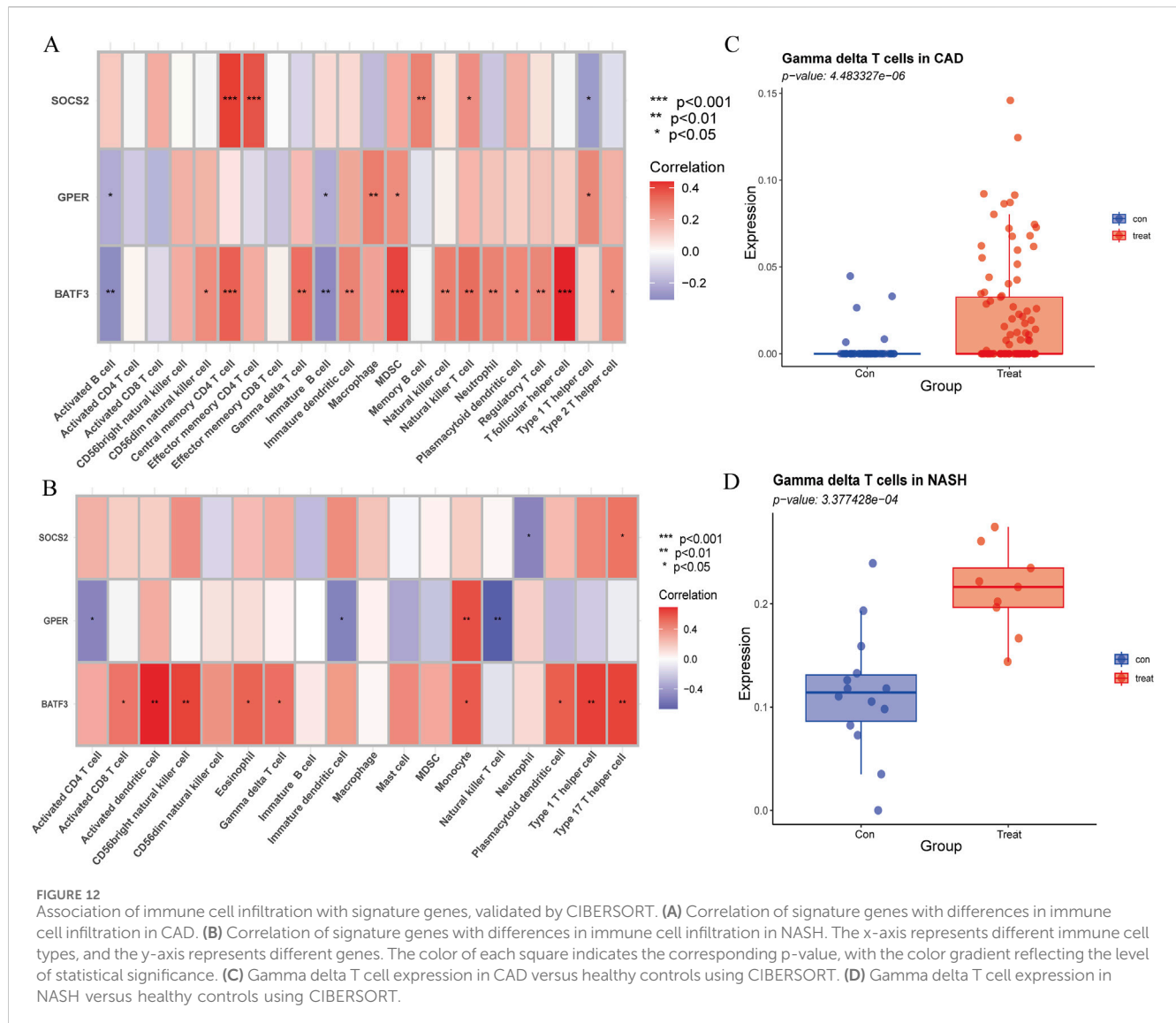


FIGURE 12

Association of immune cell infiltration with signature genes, validated by CIBERSORT. (A) Correlation of signature genes with differences in immune cell infiltration in CAD. (B) Correlation of signature genes with differences in immune cell infiltration in NASH. The x-axis represents different immune cell types, and the y-axis represents different genes. The color of each square indicates the corresponding p-value, with the color gradient reflecting the level of statistical significance. (C) Gamma delta T cell expression in CAD versus healthy controls using CIBERSORT. (D) Gamma delta T cell expression in NASH versus healthy controls using CIBERSORT.

suggests that *BATF3* might contribute to immune regulation in both diseases. A study on patients with CAD demonstrated elevated $\gamma\delta$ T cell expression in atherosclerotic plaque samples, which is consistent with our findings (Liu et al., 2024). However, another study reported a decreasing trend in the proportion of $\gamma\delta$ T cells in patients with CAD (Han et al., 2021). This discrepancy may be attributed to differences in sample sources, as the study in question used data from the GSE40231 dataset, which was derived from carotid lesion tissues, whereas our samples were obtained from peripheral blood. Furthermore, a study from China demonstrated that in patients with acute myocardial infarction, certain $\gamma\delta$ T-cell subsets exhibited restricted expression, while others were markedly elevated (Chen et al., 2018). This finding suggests that $\gamma\delta$ T cells in CAD patients may not be uniformly decreased, but rather exhibit fluctuating expression patterns that are closely associated with specific subsets. Despite the limitations of the GSEA approach employed in our analysis, which precludes the capacity to discern between distinct $\gamma\delta$ T-cell subpopulations, our findings align with those of the aforementioned study, collectively suggesting that $\gamma\delta$ T cells may have a pivotal function in the development and

progression of CAD. Nevertheless, the role of $\gamma\delta$ T cells in NAFLD remains underexplored. Our study demonstrates that $\gamma\delta$ T cells are upregulated in NASH. $\gamma\delta$ T cells can detect pathogens and stimulate dendritic cell maturation, functional activation, migration, and antigen presentation. There has been evidence that mature dendritic cells exhibit elevated levels of costimulation molecules in the context of CD1-restricted T cells and reduced antigen uptake when there are CD1-restricted T cells around (Leslie et al., 2002). It is postulated that the downregulation of *BATF3* may impair dendritic cells function, thereby affecting $\gamma\delta$ T cells and resulting in aberrant $\gamma\delta$ T cell infiltration in NASH and CAD. The precise mechanisms remain unclear and require further investigation.

Suppressor of Cytokine Signaling 2 (SOCS2) is a critical protein that plays a significant role in modulating cytokine responses. It acts as a classic negative regulator of the JAK/STAT signaling pathway (Cabrera-Galván et al., 2023). SOCS2 can mitigate inflammation by attenuating inflammasome signaling pathways, including inhibiting the NF- κ B pathway to decrease macrophage apoptosis. This action helps reduce inflammation and slows the progression of NASH (Xiao et al., 2023). The observed reduction of SOCS2 in liver tissue

under high-fat diet conditions and its inverse correlation with NASH severity support its protective role in liver homeostasis (Li S et al., 2021; Zhang Z. et al., 2023). It is consistent with our results. Research suggests that SOCS2 may interact with the JAK/STAT3 pathway to cause coronary CAD in diabetic patients (Sheng et al., 2017). While the JAK/STAT3 pathway typically exerts cardioprotective effects by inhibiting apoptosis, preserving mitochondrial integrity, and promoting angiogenesis, its overactivation may trigger proinflammatory cascades under certain conditions (Yasukawa et al., 2012). For instance, in ApoE^{-/-} mice, JAK2/STAT2 activation has been shown to exacerbate macrophage-mediated inflammation in response to homocysteine (Xu et al., 2022). Therefore, we speculate that downregulation of SOCS2 may reduce its inhibitory effect on the JAK2/STAT2 pathway, leading to an enhanced macrophage inflammatory response in CAD.

Our GSEA analysis of SOCS2 demonstrated that SOCS2 expression is associated with O-glycosylation biosynthesis in both diseases. Glycosylation is among the most common post-translational modifications in eukaryotic cells and is considered a key factor in the progression of NAFLD (Zhan et al., 2016). The glycosylation of apolipoprotein B, extensive N-linked glycosylation of fatty acid translocase, and the interaction between carbohydrate response element-binding protein and O-GlcNAc transferase have all been closely associated with metabolic dysregulation and inflammatory responses in the progression of NAFLD (Ihara et al., 1998; Guinez et al., 2011). Additionally, previous studies have shown that patients with NASH exhibit a distinct glycosylation pattern characterized by an increase in core-fucosylated biantennary glycans and a decrease in galactosylated, non-fucosylated biantennary glycans, which can help distinguish between NAFLD and NASH (Verhelst et al., 2020). In CAD, leukocyte recruitment is a critical step in the formation of arterial plaques, and the function of adhesion molecules and chemokine receptors involved in leukocyte recruitment is determined by proper post-translational glycosylation (Lowe, 2002; Frommhold et al., 2008). For example, the inhibition of α 2,3-sialyltransferase IV, α 2,3-fucosyltransferase, and core 2 β 1,6-galactosyltransferase I has been shown to suppress the development of atherosclerosis (Pu and Yu, 2014). Our study suggests that SOCS2 may influence the progression of both diseases through the O-glycosylation biosynthesis pathway and serve as a potential molecular link between them. However, this remains a hypothesis based solely on bioinformatics predictions and requires further validation through *in vivo* and *in vitro* experiments to determine whether and how SOCS2 is involved in the regulation of this pathway.

G-protein coupled estrogen receptor 1 (GPER) is a seven-transmembrane domain receptor that interacts with estrogen. It acts mainly by activating several downstream signaling pathways. (Revankar et al., 2005; Prossnitz and Barton, 2011). Emerging evidence suggests that GPER plays a critical role in the progression of NASH by activating the AMPK signaling pathway. This activation modulates lipid metabolism by promoting lipogenesis while simultaneously inhibiting lipolysis in both male and female mice. The net effect is a reduction in hepatic fat accumulation and the downregulation of fibrosis-related genes, ultimately alleviating liver fibrosis (Li et al., 2024). GPER is also associated with immune cell infiltration, fibrosis, and the release of inflammatory factors in the liver. Additionally, it is correlated with HDL and LDL levels in the blood. In endothelial cells, the activation of GPER by estrogen

decreases the transcytosis of LDL cholesterol into these cells, thereby providing indirect protection to the vasculature (Ghaffari et al., 2018). GPER activation promotes relaxation of the coronary arteries and reduces smooth muscle cell proliferation and migration. This helps prevent or reverse the progression of coronary atherosclerotic disease by increasing coronary blood flow in the affected arteries (Barton and Prossnitz, 2015).

However, the incidence of both NASH and CAD is sexually dimorphic. The incidence and severity of CVD and NAFLD are lower in premenopausal women than in men of the same age and postmenopausal women (Mouat et al., 2018; Lonardo et al., 2019). Higher levels of oestrogen may explain the lower prevalence in women. And GPER is the primary receptor that mediates estrogen levels (Li et al., 2024). However, studies report no significant sex-based differences in GPER expression across major tissues such as the aorta, heart, and kidney in *Sprague-Dawley* rats, suggesting estrogen may exert similar signaling efficiency in both sexes (Hutson et al., 2019). In addition, male GPER-deficient mice exhibit an overall increase in body fat, insulin resistance and a pro-inflammatory phenotype (Sharma et al., 2013), indicating that GPER may also play a crucial protective role in males. These suggest that GPER also plays a protective role in males. Our study indicates that GPER expression is significantly associated with the development of NASH in both men and women, reinforcing its importance in men as well. Thus, GPER may serve not only as a sex-independent signature gene but also as a potential therapeutic target. Future research is needed to clarify the mechanistic roles GPER plays specifically in males and how its modulation might be leveraged therapeutically.

We need to acknowledge the limitations of our study. First, the data were obtained from a public database with a relatively limited sample size, which may have some impact on the performance of the machine learning model. The small sample size may lead to an unstable model training process and increase the risk of overfitting. In addition, the small sample size prevents ensemble methods such as RF from fully exploiting their benefits. In addition, the data in public databases may have certain biases, such as selection bias and population bias. The CAD and NASH populations in this study were from China and Canada, respectively. However, we obtained good validation results in external datasets from the United States and the United Kingdom. Nonetheless, the study remains a bioinformatics-based investigation without supporting *in vivo* or *in vitro* experimental validation. Therefore, further molecular biology experiments and clinical studies are strongly recommended to validate the biological functions and clinical significance of the signature genes identified in this study.

5 Conclusion

This study identified three signature genes, *BATF3*, *SOCS2*, and *GPER*, shared at the transcriptional level in both NASH and CAD. Additionally, we demonstrated that insulin resistance-related and inflammation-related pathways represent common mechanisms underlying these two diseases. Moreover, we examined the infiltration of immune cells in CAD and NASH and their relationship with the characteristic genes, offering new perspectives on the role of immunity in CAD complicated by NASH. Nevertheless, further research is needed to help us

understand the mechanisms underlying the common pathways and to address the potential limitations of this study.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: GSE113079: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE113079> GSE66360: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66360> GSE89632: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89632> GSE135251: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE135251> GEO database: <https://www.ncbi.nlm.nih.gov/geo/>.

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

YL: Methodology, Validation, Writing – original draft. JS: Funding acquisition, Supervision, Writing – review and editing. XL: Funding acquisition, Project administration, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by grants from the Guangxi Science and Technology Program (Grant No. AB23026018) and the National Natural Science Foundation of China (Grant No. 82460338).

References

- Barton, M., and Prossnitz, E. R. (2015). Emerging roles of GPER in diabetes and atherosclerosis. *Trends Endocrinol. and Metabolism* 26 (4), 185–192. doi:10.1016/j.tem.2015.02.003
- Brouwers, M. C., Simons, N., Stehouwer, C. D., Koek, G. H., Schaper, N. C., and Isaacs, A. (2019). Relationship between nonalcoholic fatty liver disease susceptibility genes and coronary artery disease. *Hepatology* 3 (4), 587–596. doi:10.1002/hep4.1319
- Cabrera-Galván, J. J., Araujo, E., de Mirecki-Garrido, M., Pérez-Rodríguez, D., Guerra, B., Aranda-Tavío, H., et al. (2023). SOCS2 protects against chemical-induced hepatocellular carcinoma progression by modulating inflammation and cell proliferation in the liver. *Biomed. and Pharmacother.* 157, 114060. doi:10.1016/j.biopha.2022.114060
- Chen, X.-m., Zhang, T., Qiu, D., Feng, J.-y., Jin, Z.-y., Luo, Q., et al. (2018). Gene expression pattern of TCR repertoire and alteration expression of IL-17A gene of $\gamma\delta$ T cells in patients with acute myocardial infarction. *J. Transl. Med.* 16 (1), 189–10. doi:10.1186/s12967-018-1567-7
- Chiu, M. H., Heydari, B., Batulan, Z., Maarouf, N., Subramanya, V., Schenck-Gustafsson, K., et al. (2018). Coronary artery disease in post-menopausal women: are there appropriate means of assessment? *Clin. Sci.* 132 (17), 1937–1952. doi:10.1042/CS20180067
- Cutler, A., Cutler, D. R., and Stevens, J. R. (2012). Random forests. *Ensemble Machine Learning: Methods and Applications, nsemble Mach. Learn. Methods Appl.*, 157–175. doi:10.1007/978-1-4419-9326-7_5
- Dong, X., Li, J.-M., Lu, X.-L., Lin, X.-Y., Hong, M.-Z., Weng, S., et al. (2024). Global burden of adult non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) has been steadily increasing over the past decades and is expected to persist in the future. *Transl. Gastroenterology Hepatology* 9, 33. doi:10.21037/tgh-23-118
- Dongiovanni, P., Petta, S., Maglio, C., Fracanzani, A. L., Pipitone, R., Mozzi, E., et al. (2015). Transmembrane 6 superfamily member 2 gene variant disentangles nonalcoholic steatohepatitis from cardiovascular disease. *Hepatology* 61 (2), 506–514. doi:10.1002/hep.27490
- Duan, W., Shi, R., Yang, F., Zhou, Z., Wang, L., Huang, Z., et al. (2023). FSTL3 partially mediates the association of increased nonalcoholic fatty liver disease fibrosis risk with acute myocardial infarction in patients with type 2 diabetes mellitus. *Cardiovasc. Diabetol.* 22 (1), 297. doi:10.1186/s12933-023-02024-x
- Ebadi, N., Ghafouri-Fard, S., Taheri, M., Arsang-Jang, S., and Omrani, M. D. (2022). Expression analysis of inflammatory response-associated genes in coronary artery disease. *Archives physiology Biochem.* 128 (3), 601–607. doi:10.1080/13813455.2019.1708953
- Fan, J. G., Wei, L., Zhuang, H., Cai, W., Feng Chen, D., Duan, Z. P., et al. (2019). Guidelines of prevention and treatment of nonalcoholic fatty liver disease (2018, China). *J. Dig. Dis.* 20 (4), 163–173. doi:10.1111/1751-2980.12685
- Frommhold, D., Ludwig, A., Bixel, M. G., Zarbock, A., Babushkina, I., Weissinger, M., et al. (2008). Sialyltransferase ST3Gal-IV controls CXCR2-mediated firm leukocyte arrest during inflammation. *J. Exp. Med.* 205 (6), 1435–1446. doi:10.1084/jem.20070846

Acknowledgments

We extend our appreciation to the GEO database for granting open access to the data, and to the authors of the GSE113079, GSE66360, GSE89632, and GSE135251 datasets for their significant contributions.

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

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/fgene.2025.1573621/full#supplementary-material>

- Ghaffari, S., Naderi Nabi, F., Sugiyama, M. G., and Lee, W. L. (2018). Estrogen inhibits LDL (low-density lipoprotein) transcytosis by human coronary artery endothelial cells via GPER (G-protein-coupled estrogen receptor) and SR-BI (scavenger receptor class B type 1). *Arteriosclerosis, thrombosis, Vasc. Biol.* 38 (10), 2283–2294. doi:10.1161/ATVBAHA.118.310792
- Guiney, C., Filhoulaud, G., Rayah-Benhamed, F., Marmier, S., Dubuquoy, C., Dentin, R., et al. (2011). O-GlcNAcylation increases ChREBP protein content and transcriptional activity in the liver. *Diabetes* 60 (5), 1399–1413. doi:10.2337/db10-0452
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46 (1), 389–422. doi:10.1023/a:1012487302797
- Han, H., Du, R., Cheng, P., Zhang, J., Chen, Y., and Li, G. (2021). Comprehensive analysis of the immune infiltrates and aberrant pathways activation in atherosclerotic plaque. *Front. Cardiovasc. Med.* 7, 602345. doi:10.3389/fcvm.2020.602345
- Hutson, D. D., Gurrall, R., Ogola, B. O., Zimmerman, M. A., Mostany, R., Satou, R., et al. (2019). Estrogen receptor profiles across tissues from Male and female *Rattus norvegicus*. *Biol. sex Differ.* 10 (1), 4–13. doi:10.1186/s13293-019-0219-9
- Ihara, Y., Yoshimura, M., Miyoshi, E., Nishikawa, A., Sultan, A. S., Toyosawa, S., et al. (1998). Ectopic expression of N-acetylglucosaminyltransferase III in transgenic hepatocytes disrupts apolipoprotein B secretion and induces aberrant cellular morphology with lipid storage. *Proc. Natl. Acad. Sci.* 95 (5), 2526–2530. doi:10.1073/pnas.95.5.2526
- Książek, E., Goluch, Z., and Bochniak, M. (2024). Vaccinium spp. berries in the prevention and treatment of non-alcoholic fatty liver disease: a comprehensive update of preclinical and clinical research. *Nutrients* 16 (17), 2940. doi:10.3390/nu16172940
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9 (12), 559–13. doi:10.1186/1471-2105-9-559
- Leslie, D. S., Vincent, M. S., Spada, F. M., Das, H., Sugita, M., Morita, C. T., et al. (2002). CD1-mediated gamma/delta T cell maturation of dendritic cells. *J. Exp. Med.* 196 (12), 1575–1584. doi:10.1084/jem.20021515
- Li, L., Yao, Y., Wang, Y., Cao, J., Jiang, Z., Yang, Y., et al. (2024). G protein-coupled estrogen receptor 1 ameliorates nonalcoholic steatohepatitis through targeting AMPK-Dependent signaling. *J. Biol. Chem.* 300 (3), 105661. doi:10.1016/j.jbc.2024.105661
- Li, S., Han, S., Jin, K., Yu, T., Chen, H., Zhou, X., et al. (2021). SOCS2 suppresses inflammation and apoptosis during NASH progression through limiting NF- κ B activation in macrophages. *Int. J. Biol. Sci.* 17(15), 4165–4175. doi:10.7150/ijbs.63889
- Liu, J.-W., Zhang, Z.-H., Lv, X.-S., Xu, M.-Y., Ni, B., He, B., et al. (2024). Identification of key pyroptosis-related genes and microimmune environment among peripheral arterial beds in atherosclerotic arteries. *Sci. Rep.* 14 (1), 233. doi:10.1038/s41598-023-50689-x
- Lonardo, A., Nascimbeni, F., Ballestri, S., Fairweather, D., Win, S., Than, T. A., et al. (2019). Sex differences in nonalcoholic fatty liver disease: state of the art and identification of research gaps. *Hepatology* 70 (4), 1457–1469. doi:10.1002/hep.30626
- Lowe, J. B. (2002). Glycosylation in the control of selectin counter-receptor structure and function. *Immunol. Rev.* 186 (1), 19–36. doi:10.1034/j.1600-065x.2002.18603.x
- Lv, Q., Han, Q., Wen, Z., Pan, Y., and Chen, J. (2024). The association between atherosclerosis and nonalcoholic fatty liver disease. *Medicine* 103 (1), e36815. doi:10.1097/MD.00000000000036815
- Madhavan, M. V., Gersh, B. J., Alexander, K. P., Granger, C. B., and Stone, G. W. (2018). Coronary artery disease in patients ≥ 80 years of age. *J. Am. Coll. Cardiol.* 71 (18), 2015–2040. doi:10.1016/j.jacc.2017.12.068
- Mo, S., Wang, Y., Yuan, X., Wu, W., Zhao, H., Wei, H., et al. (2023). Identification of common signature genes and pathways underlying the pathogenesis association between nonalcoholic fatty liver disease and atherosclerosis. *Front. Cardiovasc. Med.* 10 (1), 1142296. doi:10.3389/fcvm.2023.1142296
- Morán-Costoya, A., Proenza, A. M., Gianotti, M., Lladó, I., and Valle, A. (2021). Sex differences in nonalcoholic fatty liver disease: estrogen influence on the liver–adipose tissue crosstalk. *Antioxidants and Redox Signal.* 35 (9), 753–774. doi:10.1089/ars.2021.0044
- Mouat, M. A., Coleman, J. L., and Smith, N. J. (2018). GPCRs in context: sexual dimorphism in the cardiovascular system. *Br. J. Pharmacol.* 175 (21), 4047–4059. doi:10.1111/bph.14160
- Niikura, T., Imajo, K., Ozaki, A., Kobayashi, T., Iwaki, M., Honda, Y., et al. (2020). Coronary artery disease is more severe in patients with non-alcoholic steatohepatitis than fatty liver. *Diagnostics* 10 (3), 129. doi:10.3390/diagnostics10030129
- Pan, X., Zhang, Y., Kim, H.-G., Liangpunsakul, S., and Dong, X. C. (2017). FOXO transcription factors protect against the diet-induced fatty liver disease. *Sci. Rep.* 7 (1), 44597. doi:10.1038/srep44597
- Petta, S., Valenti, L., Marchesini, G., Di Marco, V., Licata, A., Camma, C., et al. (2013). PNPLA3 GG genotype and carotid atherosclerosis in patients with non-alcoholic fatty liver disease. *PloS one* 8 (9), e74089. doi:10.1371/journal.pone.0074089
- Philippe Bardou, J. M., Escudié, F., Djemiel, C., Klopp, C., and Klopp, C. (2014). Jvenn: an interactive venn diagram viewer. *BMC Bioinforma.* 15, 293. doi:10.1186/1471-2105-15-293
- Powell, E. E., Wong, V. W. S., and Rinella, M. (2021). Non-alcoholic fatty liver disease. *Lancet* 397 (10290), 2212–2224. doi:10.1016/s0140-6736(20)32511-3
- Prossnitz, E. R., and Barton, M. (2011). The G-protein-coupled estrogen receptor GPER in health and disease. *Nat. Rev. Endocrinol.* 7 (12), 715–726. doi:10.1038/nrendo.2011.122
- Pu, Q., and Yu, C. (2014). Glycosyltransferases, glycosylation and atherosclerosis. *Glycoconj. J.* 31 (9), 605–611. doi:10.1007/s10719-014-9560-8
- Ralapanawa, U., and Sivakanesan, R. (2021). Epidemiology and the magnitude of coronary artery disease and acute coronary syndrome: a narrative review. *J. Epidemiol. Glob. health* 11 (2), 169–177. doi:10.2991/jegh.k.201217.001
- Ranstam, J., and Cook, J. A. (2018). LASSO regression. *J. Br. Surg.* 105 (10), 1348. doi:10.1002/bjs.10895
- Revankar, C. M., Cimino, D. F., Sklar, L. A., Arterburn, J. B., and Prossnitz, E. R. (2005). A transmembrane intracellular estrogen receptor mediates rapid cell signaling. *Science* 307 (5715), 1625–1630. doi:10.1126/science.1106943
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-Sequencing and microarray studies. *Nucleic acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Sharma, G., Hu, C., Brigman, J. L., Zhu, G., Hathaway, H. J., and Prossnitz, E. R. (2013). GPER deficiency in Male mice results in insulin resistance, dyslipidemia, and a proinflammatory state. *Endocrinology* 154 (11), 4136–4145. doi:10.1210/en.2013-1357
- Sheka, A. C., Adeyi, O., Thompson, J., Hameed, B., Crawford, P. A., and Ikramuddin, S. (2020). Nonalcoholic steatohepatitis: a review. *Jama* 323 (12), 1175–1183. doi:10.1001/jama.2020.2298
- Sheng, M., Huang, Z., Pan, L., Yu, M., Yi, C., Teng, L., et al. (2017). SOCS2 exacerbates myocardial injury induced by ischemia/reperfusion in diabetic mice and H9c2 cells through inhibiting the JAK-STAT-IGF-1 pathway. *Life Sci.* 188, 101–109. doi:10.1016/j.lfs.2017.08.036
- Targher, G., Day, C. P., and Bonora, E. (2010). Risk of cardiovascular disease in patients with nonalcoholic fatty liver disease. *N. Engl. J. Med.* 363 (14), 1341–1350. doi:10.1056/NEJMra0912063
- Thibaut, R., Gage, M. C., Pineda-Torra, I., Chabrier, G., Venteclef, N., and Alzaid, F. (2022). Liver macrophages and inflammation in physiology and pathophysiology of non-alcoholic fatty liver disease. *FEBS J.* 289 (11), 3024–3057. doi:10.1111/febs.15877
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 58 (1), 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x
- Vanni, E., Marengo, A., Mezzabotta, L., and Bugianesi, E. (2015). “Systemic complications of nonalcoholic fatty liver disease: when the liver is not an innocent bystander,” in *Seminars in liver disease* (Thieme Medical Publishers), 236–249.
- Verhelst, X., Dias, A. M., Colombel, J.-F., Vermeire, S., Van Vlierberghe, H., Callewaert, N., et al. (2020). Protein glycosylation as a diagnostic and prognostic marker of chronic inflammatory gastrointestinal and liver diseases. *Gastroenterology* 158 (1), 95–110. doi:10.1053/j.gastro.2019.08.060
- Wang, F. S., Fan, J. G., Zhang, Z., Gao, B., and Wang, H. Y. (2014). The global burden of liver disease: the major impact of China. *Hepatology* 60 (6), 2099–2108. doi:10.1002/HEP.27406
- Wu, H., Xu, X., Zheng, A., Wang, W., Mei, L., Chen, Y., et al. (2021). TNF- α -Induce protein 8-Like 1 inhibits hepatic steatosis, inflammation, and fibrosis by suppressing polyubiquitination of apoptosis signal-regulating kinase 1. *Hepatology* 74 (3), 1251–1270. doi:10.1002/hep.31801
- Xia, M. F., Ling, Y., Bian, H., Lin, H. D., Yan, H. M., Chang, X. X., et al. (2016). I148M variant of PNPLA3 increases the susceptibility to non-alcoholic fatty liver disease caused by obesity and metabolic disorders. *Alimentary Pharmacol. and Ther.* 43 (5), 631–642. doi:10.1111/apt.13521
- Xiao, Z., Liu, M., Yang, F., Liu, G., Liu, J., Zhao, W., et al. (2023). Programmed cell death and lipid metabolism of macrophages in NAFLD. *Front. Immunol.* 14 (1), 1118449. doi:10.3389/fimmu.2023.1118449
- Xu, L., Zhang, H., Wang, Y., Yang, A., Dong, X., Gu, L., et al. (2022). FABP4 activates the JAK2/STAT2 pathway via Rap1a in the homocysteine-induced macrophage inflammatory response in ApoE $^{-/-}$ mice atherosclerosis. *Lab. Invest.* 102 (1), 25–37. doi:10.1038/s41374-021-00679-2
- Yasukawa, H., Nagata, T., Oba, T., and Imaizumi, T. (2012). SOCS3: a novel therapeutic target for cardioprotection. *Jak-stat* 1 (4), 234–240. doi:10.4161/jkst.22435
- Zhan, Y.-T., Su, H.-Y., and An, W. (2016). Glycosyltransferases and non-alcoholic fatty liver disease. *World J. Gastroenterology* 22 (8), 2483–2493. doi:10.3748/wjg.v22.i8.2483
- Zhang, D., Mi, Z., Peng, J., Yang, T., Han, Y., Zhai, Y., et al. (2023). Nonalcoholic fatty liver disease as an emerging risk factor and potential intervention target for atherosclerotic cardiovascular diseases. *J. Cardiovasc. Pharmacol.* 81 (5), 327–335. doi:10.1097/FJC.0000000000001418

- Zhang, L., She, Z.-G., Li, H., and Zhang, X.-J. (2020). Non-alcoholic fatty liver disease: a metabolic burden promoting atherosclerosis. *Clin. Sci.* 134 (13), 1775–1799. doi:10.1042/CS20200446
- Zhang, Y. X., and Fang, X. M. (2021). Hepatocardiac or cardiohepatic interaction: from traditional Chinese medicine to Western medicine. *Evidence-Based Complementary Altern. Med.* 2021 (3), 6655335. doi:10.1155/2021/6655335
- Zhang, Z., Wang, S., Zhu, Z., and Nie, B. (2023). Identification of potential feature genes in non-alcoholic fatty liver disease using bioinformatics analysis and machine learning strategies. *Comput. Biol. Med.* 157 (5), 106724. doi:10.1016/j.compbiomed.2023.106724
- Zhou, Y., Wei, Y., Wang, L., Wang, X., Du, X., Sun, Z., et al. (2011). Decreased adiponectin and increased inflammation expression in epicardial adipose tissue in coronary artery disease. *Cardiovasc. Diabetol.* 10, 2. doi:10.1186/1475-2840-10-2
- Zouridakis, E., Avanzas, P., Arroyo-Espiguero, R., Fredericks, S., and Kaski, J. C. (2004). Markers of inflammation and rapid coronary artery disease progression in patients with stable angina pectoris. *Circulation* 110 (13), 1747–1753. doi:10.1161/01.CIR.0000142664.18739.92