



OPEN ACCESS

EDITED BY

Senling Feng,
Third Affiliated Hospital of Guangzhou
Medical University, China

REVIEWED BY

Chenyuan Gong,
Shanghai University of Traditional Chinese
Medicine, China
Haiming Chen,
Guangzhou University of Chinese Medicine,
China

*CORRESPONDENCE

Lihong Wang
✉ syfighting007@yeah.net

†These authors have contributed equally to
this work

RECEIVED 15 October 2023

ACCEPTED 12 March 2024

PUBLISHED 28 March 2024

CITATION

Wang L, He W, Wang X, Wang J, Wei X, Wu D
and Wu Y (2024) Potential diagnostic
markers shared between non-alcoholic fatty
liver disease and atherosclerosis determined
by machine learning and bioinformatic
analysis.
Front. Med. 11:1322102.
doi: 10.3389/fmed.2024.1322102

COPYRIGHT

© 2024 Wang, He, Wang, Wang, Wei, Wu and
Wu. 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.

Potential diagnostic markers shared between non-alcoholic fatty liver disease and atherosclerosis determined by machine learning and bioinformatic analysis

Lihong Wang^{1*†}, Wenhui He^{2†}, Xilin Wang¹, Jianrong Wang¹,
Xiaojuan Wei¹, Dongzhi Wu² and Yundan Wu³

¹Department of Pharmacy, Fuzhou Second General Hospital, Fuzhou, China, ²Department of Orthopedic Research Institute, Fuzhou Second General Hospital, Fuzhou, China, ³Department of Pharmacy, The Third Affiliated Hospital of Fujian University of Traditional Chinese Medicine, Fuzhou, China

Background: Evidence indicates that chronic non-alcoholic fatty liver disease (NAFLD) can increase the risk of atherosclerosis (AS), but the underlying mechanism remains unclear.

Objective: This study is intended for confirming key genes shared between NAFLD and AS, and their clinical diagnostic value to establish a foundation for searching novel therapeutic targets.

Methods: We downloaded the Gene Expression Omnibus (GEO) datasets, GSE48452 and GSE89632 for NAFLD and GSE100927, GSE40231 and GSE28829 for AS. The progression of NAFLD co-expression gene modules were recognized via weighted gene co-expression network analysis (WGCNA). We screened for differentially expressed genes (DEGs) associated with AS and identified common genes associated with NAFLD and AS using Venn diagrams. We investigated the most significant core genes between NAFLD and AS using machine learning algorithms. We then constructed a diagnostic model by creating a nomogram and evaluating its performance using ROC curves. Furthermore, the CIBERSORT algorithm was utilized to explore the immune cell infiltration between the two diseases, and evaluate the relationship between diagnostic genes and immune cells.

Results: The WGCNA findings associated 1,129 key genes with NAFLD, and the difference analysis results identified 625 DEGs in AS, and 47 genes that were common to both diseases. We screened the core *RPS6KA1* and *SERPINA3* genes associated with NAFLD and AS using three machine learning algorithms. A nomogram and ROC curves demonstrated that these genes had great clinical meaning. We found differential expression of *RPS6KA1* in patients with steatosis and NASH, and of *SERPINA3* only in those with NASH compared with normal individuals. Immune infiltration findings revealed that macrophage and mast cell infiltration play important roles in the development of NAFLD and AS. Notably, *SERPINA3* correlated negatively, whereas *RPS6KA1* correlated positively with macrophages and mast cells.

Conclusion: We identified *RPS6KA1* and *SERPINA3* as potential diagnostic markers for NAFLD and AS. The most promising marker for a diagnosis of NAFLD and AS might be *RPS6KA1*, whereas *SERPINA3* is the most closely related gene for NASH and AS. We believe that further exploration of these core genes will reveal the etiology and a pathological relationship between NAFLD and AS.

KEYWORDS

non-alcoholic fatty liver disease, atherosclerosis, *RPS6KA1*, *SERPINA3*, bioinformatics

Introduction

Non-alcoholic fatty liver disease (NAFLD) is the most common chronic liver disease worldwide, and its incidence has increased from 25% in 2005 to 32% today (1), rendering it a leading chronic condition. In a study of abdominal ultrasound and coronary computed tomography angiography of 5,121 patients with neither history of coronary artery disease nor chronic alcohol consumption revealed a close correlation between the fatty liver index, NAFLD fibrosis scores, and non-calcified plaque (2). This proved that NAFLD is closely associated with noncalcified plaques, which are prone to sudden and unexpected cardiac events. In other words, NAFLD may induce the occurrence of cardiovascular diseases for instance atherosclerosis (3). Furthermore, cardiovascular complications correlate with increased risk for NAFLD (4). More focus on NAFLD is needed to reduce the risk of atherosclerotic events. The clinicopathological syndrome NAFLD is characterized by diffuse hepatocellular bulla fat, alcohol consumption, and other specific liver damaging factors, including simple hepatic steatosis and the evolution of non-alcoholic steatohepatitis (NASH), cirrhosis, and liver cancer (5, 6). The main pathological feature of liver steatosis is hepatocyte steatosis (> 5%), while NASH is inflammation and fibrosis (7). A meta-analysis has revealed significant correlation between AS and any degree of fibrosis, the severity of which amplifies this association (8). The risk of atherosclerosis increases when NAFLD progresses from simple steatosis to NASH (9).

Atherosclerosis is the deposition of apolipoprotein in the inner walls of blood vessels. This recruited various immune cells, resulting in endothelial dysfunction, which leads to a chronic inflammatory response (10). A close association between NAFLD and AS has been identified, and the increased abundance of systemic inflammatory factors such as IL6, IL-1 β , TNF- α leads to further endothelial dysfunction and enhanced vascular plaque formation in patients NAFLD (11). In addition, NAFLD and AS are manifestations of metabolic syndrome in end-organ damage, as well as diseases caused by abnormal fat metabolism. Therefore, systemic inflammation, endothelial dysfunction, and altered lipid metabolism might be mechanisms through which NAFLD increases the risk of AS. With the wide development and application of bioinformatics, WGCNA and machine learning algorithms (MLAs) can help to explore potential biological diagnostic markers between NAFLD with different stages and AS and assess immune characteristics between them using immune cell infiltration. This would contribute to the prevention and

early treatment of both diseases. Here, we used WGCNA analysis and MLAs to explore potential biological diagnostic markers that differ between NAFLD with different progression and AS, and are common to both. We also assessed immune signatures between potential markers using immune cell infiltration. Our findings should contribute to the prevention and early treatment of both diseases.

Materials and methods

Data processing and analysis

We downloaded the GEO datasets¹ GSE48452, GSE89632, GSE100927, GSE130970, GSE135251, GSE58979, GSE40231, GSE28829, GSE97210, GSE57691, and GSE163154 to determine gene expression and clinical information regarding patients with NAFLD and AS. The GSE40231 and GSE28829 were merged into one combined AS dataset through the SVA package (Supplementary Figure 1). Liver transcriptome data were derived from 32 NAFLD patients (steatosis, n = 14; NASH, n = 18) and 14 normal subjects in the GSE48452 dataset (GPL11532 platform). The GSE100927 dataset (GPL17077 platform) included 104 human peripheral artery specimens from 69 patients with AS and 35 healthy individuals. We verified the specific expression of gene signatures using the GSE89632, the combined AS dataset, GSE130970, GSE135251, GSE58979, GSE97210, GSE57691, GSE163154 datasets. All datasets were processed and analyzed using RStudio version 4.2.1. Detailed information on the GEO dataset can be found in Supplementary Table 1. Figure 1 displays the flowchart of the study. Table 1 displays the Abbreviations of the study.

Gene expression analyzed using WGCNA

Gene expression profiles in many samples can be analyzed using the WGCNA (12–14). It can cluster genes with similar expression profiles and analyze dependence between modules and particular traits or phenotypes. It is widely applied to phenotypic trait and gene association analyses among multiple groups (15). Both GSE100927 and GSE57691 datasets contain

¹ <https://www.ncbi.nlm.nih.gov/geo/>

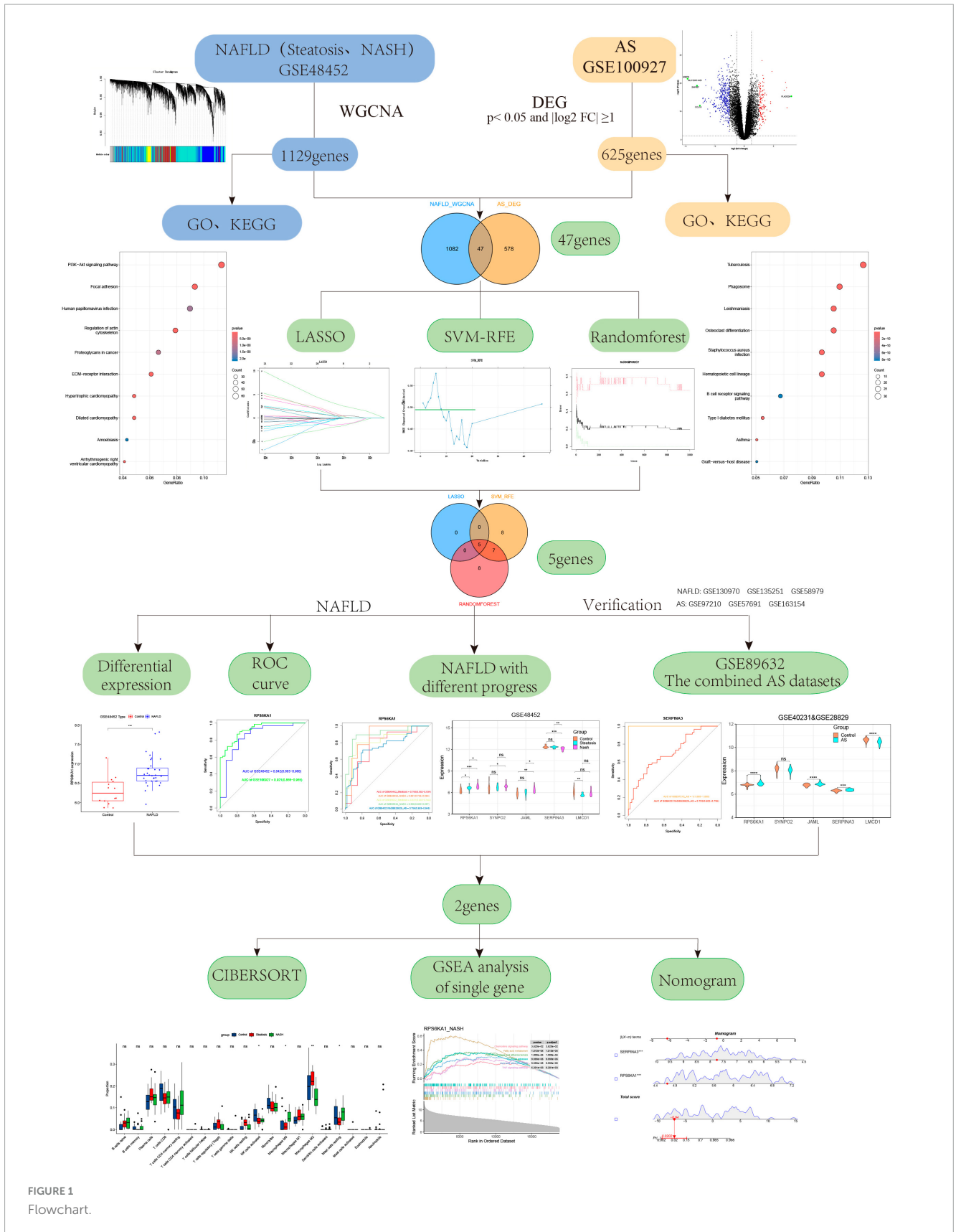


FIGURE 1
Flowchart.

clinical information related to the progression of NAFLD. We constructed a gene co-expression network for NAFLD by the WGCNA version 1.72.1 package in R version 4.3.1. The flash cluster

function in R was used analyze hierarchical clustering in NAFLD samples to identify and eliminate outliers. We then calculated an appropriate soft threshold using the pick Soft Threshold algorithm

TABLE 1 Displays the Abbreviations of the study.

Abbreviations	Full name	Abbreviations	Full name
AS	atherosclerosis	GSEA	Gene Set Enrichment Analysis
AUC	calculated the area	KEGG	Kyoto Encyclopedia of Genes and Genomes
CCR2	chemokine receptor type 2	KLF2-eNOS	Kruppel like factor 2- endothelial nitric oxide synthase
CXCR4	C-X-C chemokine receptor type 4	LASSO	Minimum absolute contraction and selection operator
DCA	Decision curve analysis	MLAs	machine learning algorithms
DEGs	Differentially expressed genes	NAFLD	Non-alcoholic fatty liver disease
ECM	Extracellular matrix	NASH	Non-alcoholic steatohepatitis
ECs	Endothelial cells	ROC	Receiver operating characteristic
ERK5	Extracellular signal-regulated kinase 5	SVM-RFE	Support vector machine recursive feature elimination
GEO	Gene expression omnibus	VCXAM-1	Vascular cell adhesion molecule 1
GO	Gene ontology	WGCNA	Weighted gene co-expression network analysis

in the WGCNA package. A scale-free network was constructed to obtain various gene modules related to NAFLD using the one-step network construction function in the WGCNA package. Divide the samples into control, steatosis, and NASH, perform module phenotype correlation analysis on each module, draw a module phenotype correlation heatmap, and determine the module with the highest correlation with the target phenotype. Key NAFLD-associated gene modules (NAFLD_WGCNA), were screened by association analysis of the module data with the clinical characteristics of NAFLD.

Difference analysis

Using R, we standardized and corrected all gene expression profile microarray data and annotated gene names using limma version 3.56.2 (16). DEGs in AS (AS_DEGs) were also obtained by limma package analysis of gene expression between AS patients and control cases. The gene threshold of AS_DEGs was $p < 0.05$ and $|\log_2 FC| \geq 1$. We visualized differential data as volcano plots. Genes that were common to NAFLD_WGCNA and AS_DEGs were identified using ggvenn version 0.1.10.

Functional enrichment analysis

We aimed to identify NAFLD and AS co-morbidities and reveal the potential biological significance of the core genes of these diseases. We therefore analyzed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways to determine the functional enrichment of NAFLD_WGCNA and AS_DEGs, respectively using the ClusterProfiler package (17) in R. Functional enrichment was confirmed at $p < 0.05$.

Key genes with diagnostic value screened by MLAs

Based on the common genes in NAFLD and AS, we further screened core genes using the LASSO regression (18), SVM-RFE

(19), and random forest (20). LASSO regression was applied using the glmnet package version 4.1.7 with parameters adjusted by 10-fold cross-validation. Using caret version 6.0–86, e1071 version 1.7–9, and kernlab version 0.9–29, the SVM-RFE algorithm was applied. Through five-fold cross-validation, the accuracy of the model algorithm was improved. The random forest algorithm was executed using randomForest package (version 4.7.1.1). The overlapping results of the three algorithms represented the diagnostic markers of NAFLD-associated AS. The core diagnostic markers of NAFLD-associated AS were further determined from analyses of ROC curves and gene expression.

Verification of core genes

We verified the differential expression of key genes in control and focal tissues by the NAFLD and AS datasets GSE89632, GSE130970, GSE135251, GSE58979, and the combined AS dataset, GSE97210, GSE57691, GSE163154 datasets respectively. We established ROC curves through the pROC (21) version 1.18.4 package, calculated the AUC value, and evaluated the ability of core genes to diagnose steatosis, NASH, and AS, respectively.

Construction of NAFLD-related AS diagnostic model

Expression profiles of the core diagnostic genes were obtained from the GSE100927 dataset. We constructed a nomogram diagnostic model (22, 23) based on gene expression, AS, and control samples to confirm the diagnostic value of these genes.

GSEA of individual genes

We assessed core genes using GSEA (24) and the ClusterProfiler package in R. Samples of steatosis, NASH and AS were classified based on high or low expression of core genes. A series of core DEGs was determined using the limma package. We then compared KEGG pathways among the three diseases using GSEA.

Correlations between core genes and immune cells

We explored the relationship between immune cell types with different degrees of infiltration and diagnostic markers of NAFLD and AS using CIBERSORT (25) and the preprocessCore version 1.62.1, parallel version 4.3.1, and e1071 packages in R.

Statistical analysis

Statistical analysis of all data was performed using R and Wilcoxon tests were used for analyzing differential gene expression. Differences among the control, steatosis, and NASH groups were compared by measuring immune infiltration and analysis of variance. More detailed statistical methods for data processing were introduced in the previous section. P values greater than 0.05 were considered statistically significant.

Results

Identification of co-expressed gene modules in NAFLD

We obtained gene expression profiles from 14 patients with steatosis, 18 with NASH, and 14 healthy individuals in the NAFLD datasets. The samples were assigned to steatosis, NASH and control groups. We ensured that the constructed scale-free network had biological significance as follows. We removed the outlier sample GSM1179010 from the NAFLD dataset, and selected nine as the soft threshold power β for the NAFLD dataset on the basis of a scale-free R^2 (0.8). We also calculated GS (correlation between a gene and a sample) and MM (correlation between gene expression and genes with a modular signature; ME) to correlate these modules with clinical features. Figures 2A-D shows that the NAFLD dataset genes were clustered into 27 modules visualized in different colors. Among them, light green, dark turquoise, light cyan, yellow, and white co-expressed gene modules were closely associated with NAFLD. The light green and dark turquoise modules were most associated with NASH, and light cyan and white were most associated with steatosis. The light green, dark turquoise, light cyan, and yellow modules were related to inflammation and fibrosis. The dark turquoise, light cyan were related to fat and nas. Therefore, we selected these five NAFLD-related gene modules as targets and found 1,129 genes (NAFLD_WGCNA), which is shown in Supplementary Table 2.

Identification of DEGs in AS

We compared DEGs between patients with AS and healthy individuals using GSE100927 datasets. With the difference analysis, 625 DEGs (Supplementary Table 2) were selected between AS and healthy controls based on combination and normalization of the microarray data ($p < 0.05$ and $|\log_2 FC| \geq 1$; Figure 2E, volcano plot; Supplementary Figure 2, heatmap). The Venn diagram

revealed 47 intersecting genes between AS and NAFLD (Figure 2F). Thus, these two diseases have 47 genes in common.

Functional enrichment analysis

Functional enrichment of AS_DEGs and NAFLD_WGCNA was analyzed. Figures 3A-D shows that the NAFLD_WGCNA genes were concentrated in the PI3K-AKT signaling pathway, focal adhesion, ECM-receptor interaction, and other pathways. The AS_DEGs were primarily enriched in the activation of immune responses, leukocyte-mediated immunity, B cell receptor signaling, and other pathways. Collectively, these diseases appear to be mediated by inflammation and immunity.

Machine learning

We extracted expression data for 47 genes from the common gene expression profiles of NAFLD and AS. Five, 20, and 20 genes were respectively selected as biomarkers associated with NAFLD and AS using the LASSO, SVM-RFE, and random forest algorithms (Figures 4A-D, Supplementary Table 3 and Supplementary Figure 3). Overlapping genes derived from the three algorithms were considered diagnostic for NAFLD-AS comorbidities, and we identified them as *RPS6KA1*, *SERPINA3*, *JAML*, *SYNP02*, and *LMCD1* (Figure 4E).

Differentially expressed genes and ROC curves

We investigated the diagnostic value of the five common genes screened in clinical practice, by analyzing differences among them between the NAFLD and AS datasets. Figures 5A-J shows that three diagnostic genes significantly differed between the NAFLD and AS datasets whereas *SERPINA3* and *SYNP02* did not. The *LMCD1* gene correlated negatively, whereas *RPS6KA1* and *JAML* correlated positively with NAFLD and AS. We then plotted ROC curves for the NAFLD and AS datasets. Figures 5K-O shows that all AUC values were > 0.6 , indicating that these five genes have clinical diagnostic value.

Five diagnostic genes and NAFLD progression

We further explored relationships between these diagnostic genes and NAFLD progression. We found that *RPS6KA1* expression differed between the steatosis and NASH groups, with all AUCs being > 0.7 , whereas *LMCD1* expression differed only in the steatosis group. The expression of *SYNP02*, *JAML*, and *SERPINA3* differed only in the NASH group, with all AUCs being > 0.7 , while there was no difference in the steatosis group, with their AUC values all less than 0.7 (Figures 6A-F, Supplementary Figure 4). Therefore, we selected the *RPS6KA1* and *LMCD1* genes as diagnostic for steatosis and AS, and the *RPS6KA1*, *SYNP02*, *JAML* and *SERPINA3* genes as diagnostic for NASH and AS.

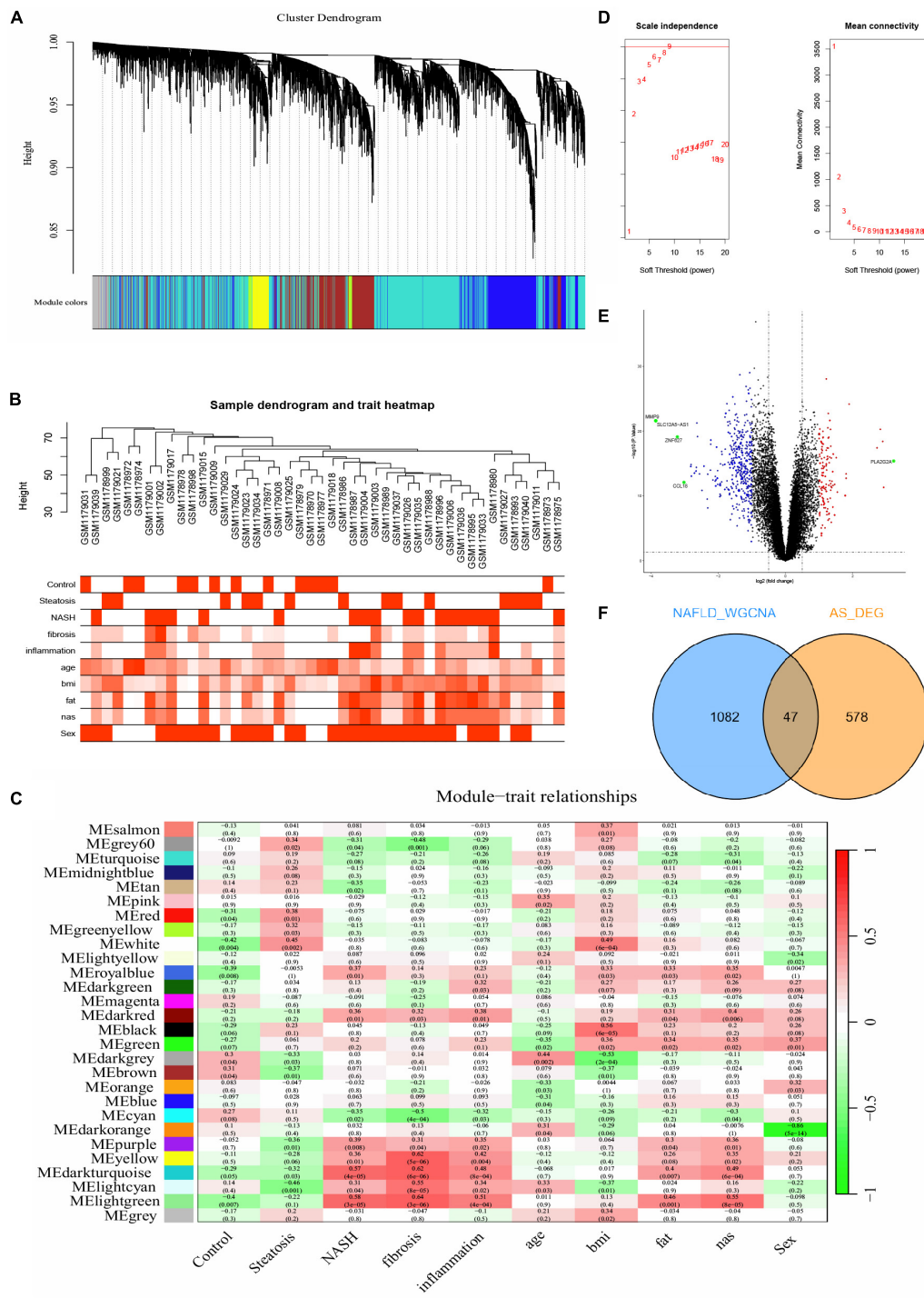
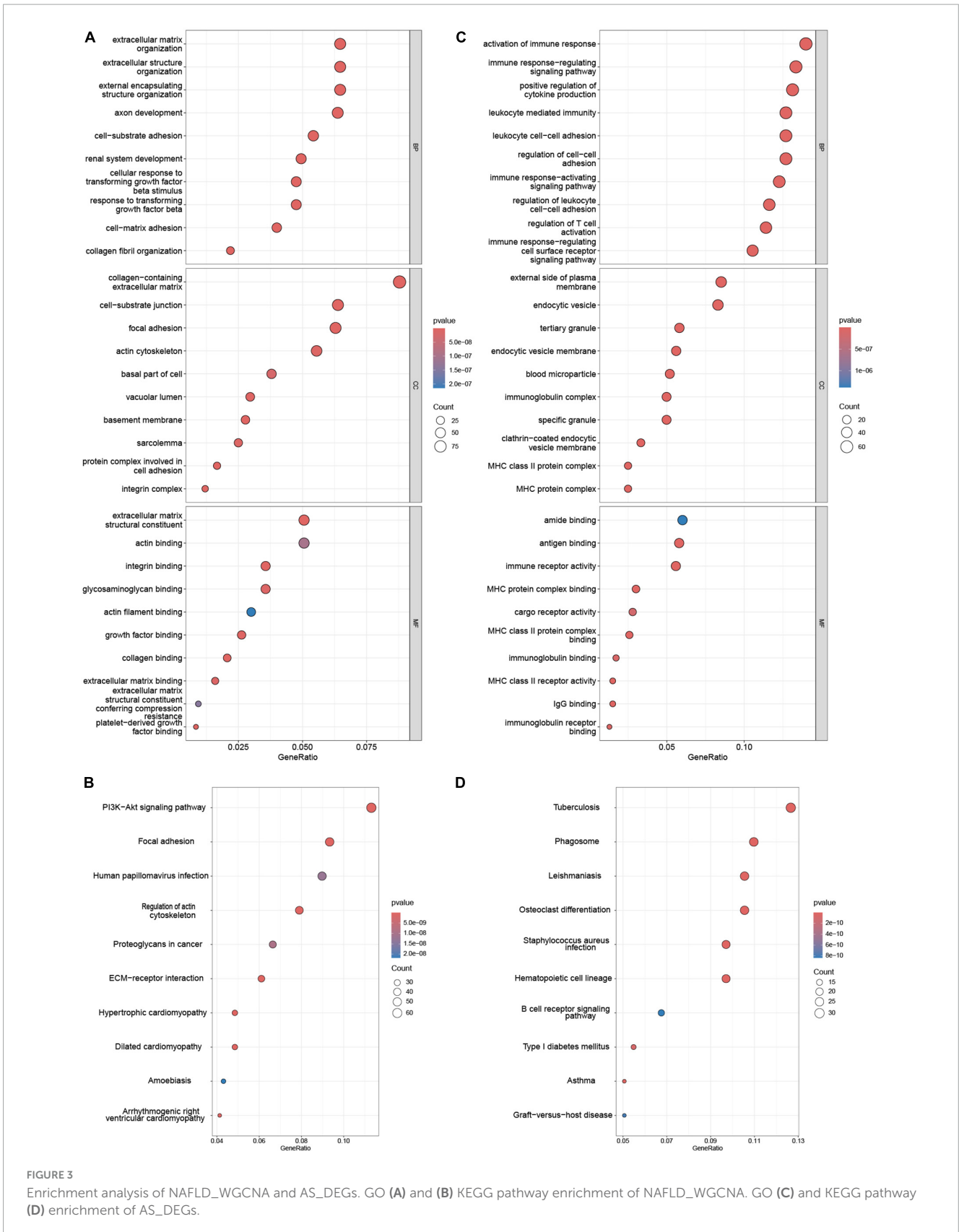


FIGURE 2 Difference analysis and WGCNA. (A-C,E) Construction of WGCNA co-expression network in GSE48452. (A) Sample clustering dendrogram. (B) Sample dendrogram and trait heatmap. (C) Soft threshold $\beta = 9$ and scale-free topological fit index (R2). (D) Volcano plot of DEGs in GSE100927. Red nodes indicate upregulated DEGs, blue nodes indicate downregulated DEGs, green nodes indicate $p < 0.05$ and $|\log_2 FC| > 3$, and black nodes indicate genes that are not significantly differentially expressed. (E) Heat map of module-trait correlations. (F) Common genes overlap in WGCNA_NAFLD and AS_DEGs.

Verification of core genes

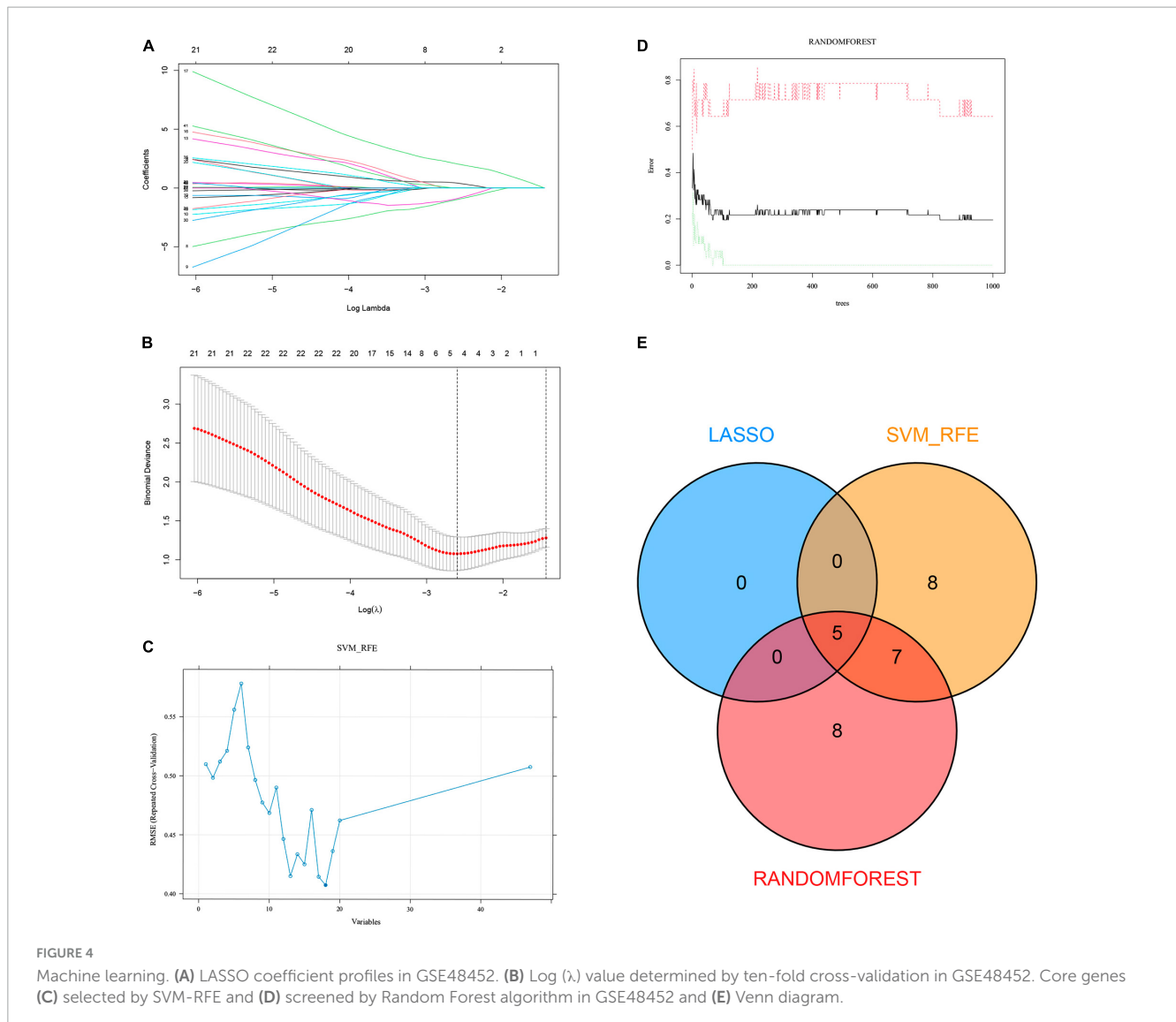
We validated the accuracy of five common genes using DEGs from GSE89632 and GSE57691. **Figures 6A-H** and **Supplementary Figure 4** shows that the expression of *RPS6KA1*, *SERPINA3*,

JAML, and *LMCD1* differed in the combined AS datasets and the ROC curve results showed that they had significant value for a diagnosing AS. Similarly, we have also seen similar results on other AS datasets. The expression of *RPS6KA1* and *SERPINA3* differed in the steatosis and NASH groups in the GSE89632 and



GSE58979 dataset, with AUCs being > 0.7. These results indicated that *RPS6KA1* and *SERPINA3* have significant value for diagnosing NAFLD (steatosis and NASH). Therefore, *RPS6KA1* might be

a promising marker for diagnosing NAFLD and AS, whereas *RPS6KA1* and *SERPINA3* might be promising for diagnosing NASH and AS.



Construction of diagnostic model for NAFLD-related AS

We evaluated the performance of the prediction model by basing a nomogram on the diagnostic genes, *RPS6KA1* and *SERPINA3*, and using calibration, decision curve analysis (DCA), and clinical impact curves (Figures 7A-D). The ROC curve in Figure 7E indicates that the nomogram has powerful diagnostic value for NAFLD-associated AS. The calibration curves in Figure 7C indicated that the prediction probability of the nomogram diagnostic model was similar to that of the ideal model, and the DCA in Figure 7D showed that NAFLD-associated AS diagnoses were more beneficial when based on the nomogram model rather than on a single gene.

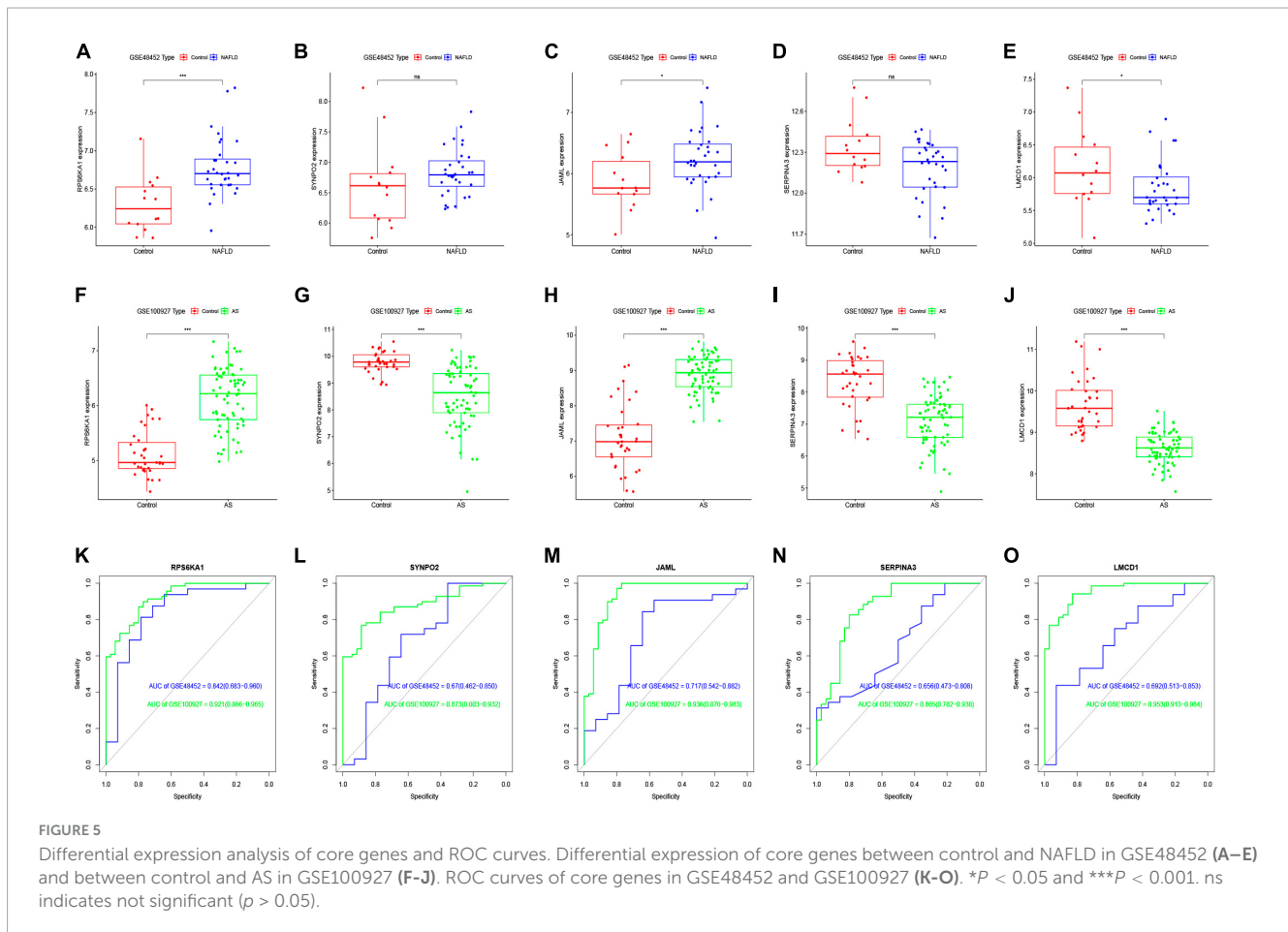
GSEA analysis of single gene

To further discuss the potential role of the core genes *RPS6KA1* and *SERPINA3* in steatosis, NASH and AS, we analyzed pathways

associated with a single gene using GSEA-KEGG. In the group with high expression, *RPS6KA1* might activate fatty acid metabolism, lipid and atherosclerosis, and TNF signaling, chemokine signaling, and focal adhesion pathways (Figures 8A-F), whereas *SERPINA3* might activate NF- κ B signaling, Toll-like receptor signaling, and FoxO signaling pathways, apoptosis, fatty acid metabolism, lipid and atherosclerosis.

Analysis of immune cell infiltration

The results of immune cell infiltration in patients with using CIBERSORT revealed differences in the expression of activated NK cells, M0 macrophages, and resting mast cells among the steatosis, NASH, NAFLD, and AS groups compared with controls (Figures 9A, B). This trend was consistent, with significantly increased M0 macrophages and significantly decreased NK cell activation and resting mast cells. The results of correlation analyses between diagnostic *RPS6KA1* and *SERPINA3* gene expression and immune cells showed that *RPS6KA1* correlated positively



with macrophages and *SERPINA3* correlated negatively with macrophages and mast cells in patients with NAFLD (Figure 9C), whereas *RPS6KA1* and *SERPINA3* respectively correlated positively and negatively with macrophages and mast cells in patients with AS (Figure 9D).

Discussion

The principal purpose of this study was to determine crosstalk genes between NAFLD and AS using bioinformatics. We identified *RPS6KA1* and *SERPINA3* as the core crosstalk genes in NAFLD and AS, with far higher AUC values than numerous genes. Inflammatory activation participated in crosstalk between NAFLD and AS due to the co-expression of macrophages and mast cells, and that hub genes correlated to a greater degree with different immune cells.

The important identified gene, *RPS6KA1*, also called *P90RSK*, belongs to the RSK family of serine/threonine kinases (26). Upon activation induced by H_2O_2 , *P90RSK* inhibits transcriptional activation of ERK5 and adjusts the subsequent expression of KLF2-eNOS and VCAM-1, thus inhibiting endothelial cells (ECs) inflammation and modulating the function of the ECs, which is responsible for atherosclerosis. However, this inhibitory effect was reversed by the *P90RSK* inhibitor (27). Blood flow disorder can easily progress to endothelial cell dysfunction, stimulate

P90RSK phosphorylation, induce SENP2 nuclear export, diminish SENP2 activity, enhance the sumoylation of ERK5 and tumor protein P53, and increase the expression of inflammatory and pro-apoptotic factors in ECs, which in turn decrease apoptosis and inflammatory responses that lead to their dysfunction; this ultimately results in the formation of atherosclerotic plaques (28–30). However, *P90RSK1* and *P90RSK2* gene knockout completely suppressed the appearance of such plaques. The mouse strain *WTP90RSK* -MTG that overexpresses *P90RSK* in bone marrow cells has far more atherosclerotic lesions and necrotic cores than non-transgenic littermate controls mice. In contrast, *DNP90RSK* -MTG mice expressed dominant-negative *P90RSK* in myeloid cells and have fewer obvious plaques and less developed necrotic cores. These results indicated that the activation of *RPS6KA1* (*P90RSK*) accelerates the formation of atherosclerotic plaques by influencing ECs function (31).

Direct investigations of *RPS6KA1* and NAFLD are limited. Currently, *RPS6KA1* plays an indeterminate role in chronic liver disease and might be involved in lipid metabolism and NAFLD. Inhibiting *RPS6KA1* and *STAT3* genes can impact the selective autophagic degradation of KLF3 mediated by CEBPB hat weakens adipogenesis (32).

Serine proteinase inhibitor A3 (*SERPINA3*) is also a key gene that belongs to the serpin superfamily (33). Most studies of *SERPINA3* have focused on tumors, and few have investigated NAFLD and AS, whereas the mouse homologue *SERPINA3C* is related to two diseases in more studies. Plasma levels of

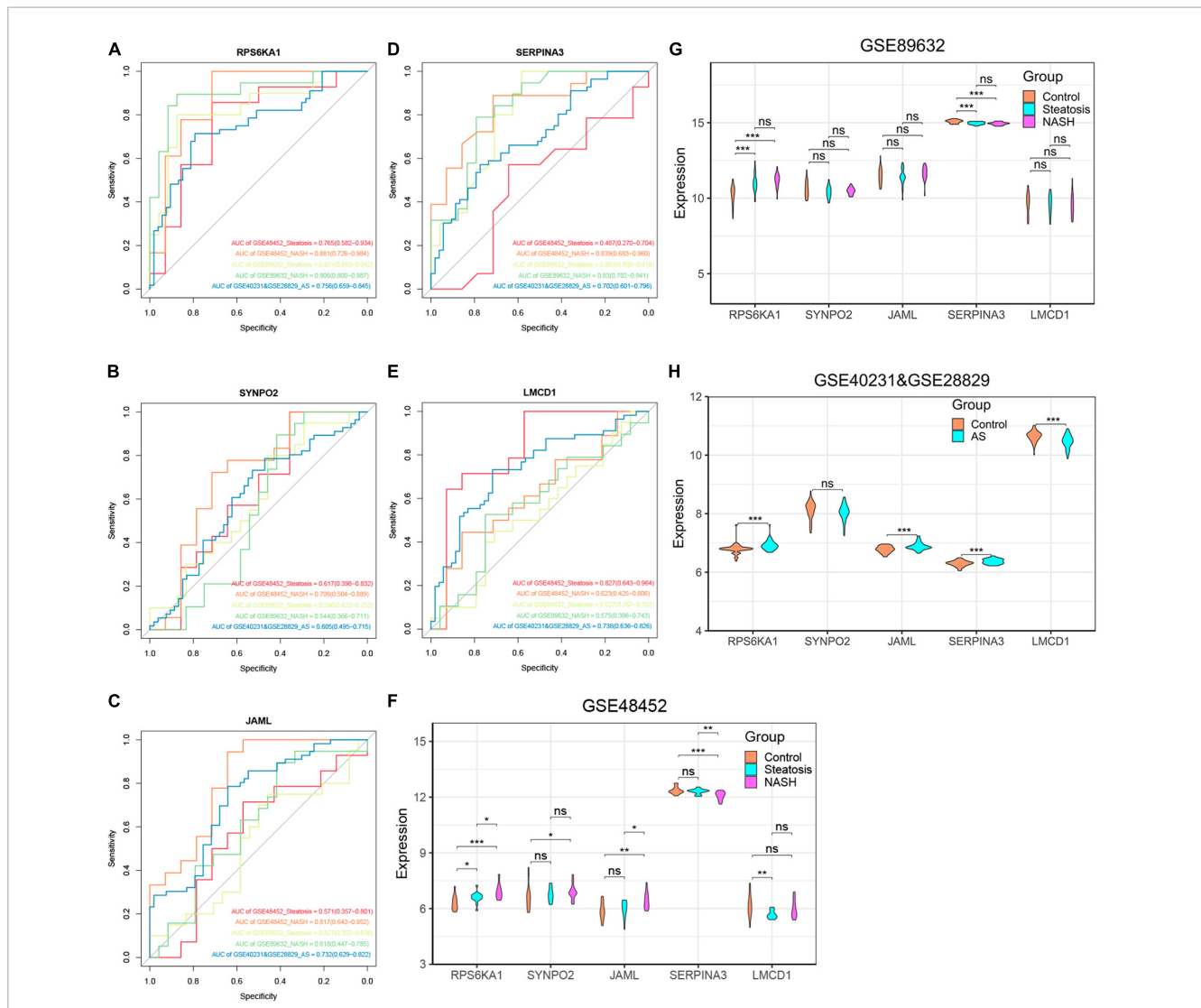
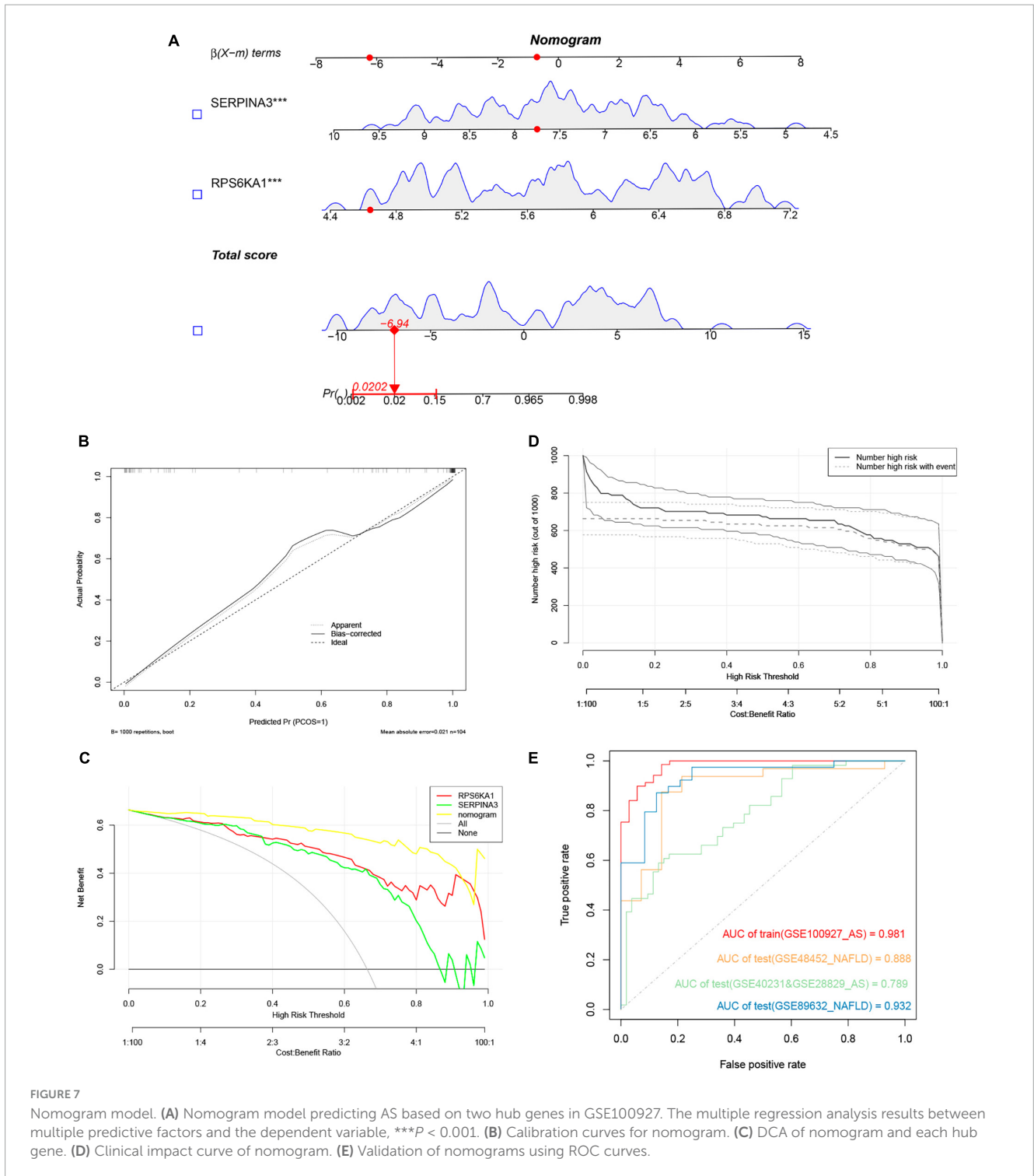


FIGURE 6 Verification of core genes. (A-E) Values for ROC and AUC of each hub gene among control, steatosis and NASH in GSE48452, GSE89632 and between control and AS in the combined AS dataset; (F-H) Differentially expressed core genes among control, steatosis and NASH in GSE48452, GSE89632 and between control and AS in the combined AS dataset. * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$. ns indicates not significant ($p > 0.05$).

the acute-phase protein *SERPINA3* increase substantially during inflammation (34). However, we found obviously diminished *SERPINA3* expression in patients compared with controls in the NAFLD and AS datasets, suggesting that *SERPINA3* has an anti-NAFLD (35) and anti-AS (36) effects. The anti-inflammatory role of *SERPINA3* is consistent with the results of its mouse homologue *SERPINA3C* (human homologue is *SERPINA3*) in NAFLD and AS.

Double knockout of *APO* and *SERPINA3C* substantially increased lipogenic *SREBP1* and *SCD1* gene expression in the livers of mice fed with a high-fat diet. These mice developed lipid accumulation, and steatosis compared with only *APO* knockout. Serum AST and ALT levels are significantly increased in double-knockout mice, and macrophage infiltration and ECM content expression are also increased in their liver tissues (37). The increased expression of JNK and P65 proteins in the liver activates the JNK/NF- κ B signaling pathway, which leads to enhanced expression of adhesion molecules and proinflammatory factor.

All factors cause further liver inflammation and fibrosis, and significantly aggravate liver injury. Additionally, *SERPINA3C* can also resist hepatocyte necrotic apoptosis by inhibiting the β -catenin/Foxo1/TLR4 signaling pathway (37). A deficiency of *SERPINA3C* strengthens liver sensitivity to lipid toxicity and promotes necrotic apoptosis, which, in turn, promotes inflammation and fibrosis in NASH (38). In summary, *SERPINA3C* can inhibit the expression of genes associated with lipogenesis, curb lipogenesis and hepatocyte necrotic apoptosis through the JNK/NF- κ B and Wnt/ β -catenin signaling pathways, diminish inflammation and fibrosis in the liver, and inhibit the development of more severe NASH. This is consistent with our findings of significantly reduced *SERPINA3* expression in NASH in GSE89632, GSE48452, GSE58979 NAFLD datasets but did not significantly change in steatosis. Although we did not see similar results in the GSE130970 and GSE135251 validation sets, considering that there are more control samples in the GSE48452 and other datasets, and



only 4 control samples in GSE130970, we also consider it as a core gene. Thus, *SERPINA3* might be a promising marker for a diagnosis of NASH syndrome and AS.

Otherwise, *SERPINA3C* has also been implicated in the development of AS. Thrombin activity is inhibited by *SERPINA3C*, which in turn weakens the inhibition of intracellular ERK1/2 and JNK phosphorylation, thereby declining the abnormal proliferation and migration of VSMCs. Excessive VSMC proliferation generates new intima and plaques, and facilitates the progression of

atherosclerosis (39). These indicate that *SERPINA3C* is a novel thrombin inhibitor that protects against the development of atherosclerosis.

Atherosclerosis and NAFLD are the results of chronic inflammatory reactions, and pathological development mainly involves macrophages and mast cells (40). Immune infiltration analysis revealed an increased abundance of macrophages and mast cells in patients with AS and those with NAFLD. Accumulated mast cells in plaques promotes the progression of atherosclerosis

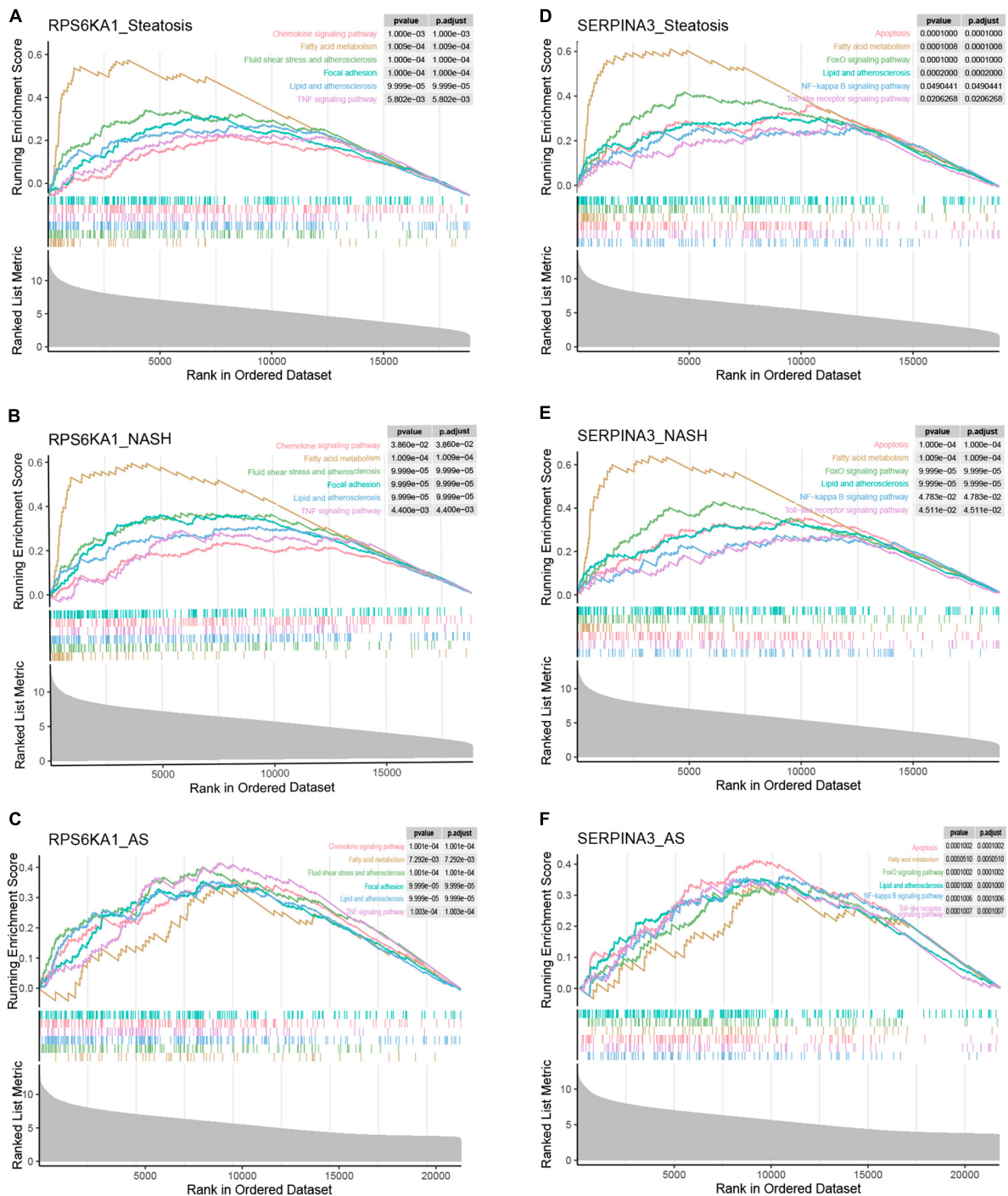
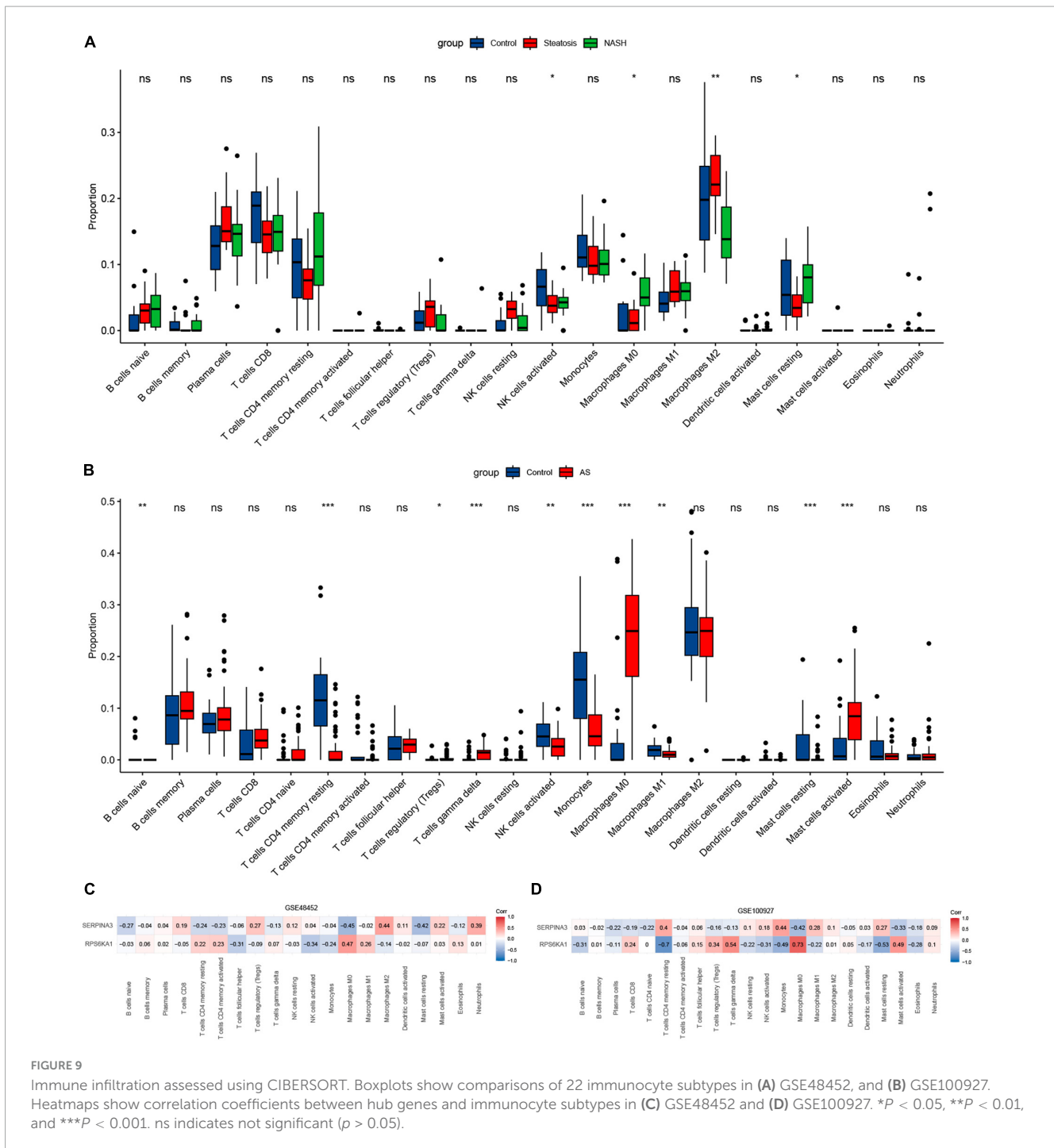


FIGURE 8 Single-gene GSEA of *RPS6KA1* and *SERPINA3*. Signaling pathways of *RPS6KA1* determined from single-gene GSEA in (A) steatosis, (B) NASH and (C) AS. Signaling pathways of *SERPINA3* determined from single-gene GSEA of in (D) steatosis, (E) NASH, and (F) AS.

by secreting pro-inflammatory factors and chemokines (41). Macrophages absorb low-density lipoprotein particles that have been modified by lipases and proteases and are trapped in the intima of blood vessels to form foam cells, which comprise the hallmark of early plaque formation (42). The continual accumulation of foam cells in the vascular intima aggravates local inflammatory responses by constantly secreting inflammatory factors and chemokines. Activation of *P90RSK* in macrophages

guides the expression of pro-inflammatory genes like *TNF-α*, which in turn facilitates atherosclerosis. Moreover, *P90RSK* overexpression in macrophages is involved in atherosclerotic and necrotic core formation.

Our results indicated that lipid metabolism, endothelial dysfunction, and inflammatory responses play vital roles in NAFLD and AS pathogenesis. Endothelial cell dysfunction is an initial stage of atherosclerosis (43, 44). The abundance of ECs expressing



cell adhesion molecules such as E-selectin, increases upon contact with a multitude of hazardous factors for atherosclerosis. These adhesion molecules then release chemokines that induce the recruitment and infiltration of monocytes and lymphocytes. The adhesion of monocytes and lymphocytes to ECs constitutes a proportion of atherosclerotic lesions, augments the recruitment of white blood cells and activates proinflammatory factors in ECs. This leads to inflammation that ultimately exacerbates atherosclerosis.

During the progression from simple steatosis to more severe NASH, Endothelial dysfunction and inflammation are

important role (10, 45). Hepatic sinusoidal endothelial cells activate NF-κB signaling through TLR9 to reinforce secretion of CCR2 and CXCR4, recruit macrophages to inflammatory sites, and release inflammatory factors during this period. These inflammatory factors might cause vascular endothelial cell dysfunction, which in turn results in atherosclerosis (46, 47).

As both NAFLD and AS are manifestations of end-organ damage in metabolic syndrome, they have several similarities. However, the common mechanisms of NAFLD with different stages and AS has not been explored until now. Based on the WGCNA

and MLAs, we identified characteristic genes that will help to further elucidate the mechanisms of action that are common to NAFLD with different stages and AS of progression. Our study was retrospective, and further experimental and clinical data are required to confirm our results. We will collect cases of NAFLD combined with AS in clinical practice, which will be our future research. The expression and diagnostic models based on two signature genes functioned well in the external validation sets for NAFLD with different stages and AS of progression, which increased confidence in our results. We revealed new biomarkers of an association between NAFLD with different stages and AS of progression.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Author contributions

LW: Writing—original draft, Data curation, Formal analysis, Funding acquisition, Methodology, Software, Visualization. WH: Writing—original draft, Formal analysis, Methodology. XW: Formal analysis, Writing—original draft. JW: Writing—review and editing. XW: Writing—review and editing. DW: Writing—review and editing. YW: Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of the article. This study

References

- Riazi K, Azhari H, Charette J, Underwood F, King J, Afshar E, et al. The prevalence and incidence of NAFLD worldwide: A systematic review and meta-analysis. *Lancet Gastroenterol Hepatol*. (2022) 7:851–61. doi: 10.1016/S2468-1253(22)00165-0
- Lee S, Park G, Lee J, Lee B, Park J, Kim B, et al. Association between non-alcoholic fatty liver disease and subclinical coronary atherosclerosis: An observational cohort study. *J Hepatol*. (2018) 68:1018–24. doi: 10.1016/j.jhep.2017.12.012
- Mantovani A, Morieri M, Palmisano L, Masulli M, Cossu E, Baroni M, et al. Hepatic steatosis with significant fibrosis is associated with an increased 10-year estimated risk of cardiovascular disease in adults with type 1 diabetes mellitus. *Cardiovasc Diabetol*. (2023) 22:204. doi: 10.1186/s12933-023-01945-x
- VanWagner L, Ning H, Allen N, Ajmera V, Lewis C, Carr J, et al. Alcohol use and cardiovascular disease risk in patients with nonalcoholic fatty liver disease. *Gastroenterology*. (2017) 153:1260–72.e3. doi: 10.1053/j.gastro.2017.08.012
- Powell E, Wong V, Rinella M. Non-alcoholic fatty liver disease. *Lancet*. (2021) 7:2212–24. doi: 10.1016/S0140-6736(20)32511-3
- Barreby E, Chen P, Aouadi M. Macrophage functional diversity in NAFLD – more than inflammation. *Nat Rev Endocrinol*. (2022) 18:461–72. doi: 10.1038/s41574-022-00675-6
- Sheka A, Adeyi O, Thompson J, Hameed B, Crawford P, Ikramuddin S. Nonalcoholic steatohepatitis: A review. *JAMA*. (2020) 323:1175–83. doi: 10.1001/jama.2020.2298

was sponsored by the Natural Science Foundation of Fujian Province project (2023J011509 and 2022J011310), the Fuzhou Science and Technology Plan project (2022-S-052), and the Fuzhou Health Technology Project (2021-S-wq17).

Acknowledgments

We would like to thank Editage (www.editage.cn) for English language editing.

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/fmed.2024.1322102/full#supplementary-material>

- Jamalnia M, Zare F, Lankarani K. Systematic review and meta-analysis: Association between liver fibrosis and subclinical atherosclerosis in nonalcoholic fatty liver disease. *Aliment Pharmacol Ther*. (2023) 58:384–94. doi: 10.1111/apt.17617
- Cheng Y, Wang C, Kao J. Metabolic associated fatty liver disease better identifying patients at risk of liver and cardiovascular complications. *Hepatol Int*. (2023) 17:350–6. doi: 10.1007/s12072-022-10449-x
- Lee D, Chiu J. Atherosclerosis and flow: Roles of epigenetic modulation in vascular endothelium. *J Biomed Sci*. (2019) 26:56. doi: 10.1186/s12929-019-0551-8
- Franco S, van der Graaf D, Kwanten W. Non-alcoholic fatty liver disease and cardiovascular risk: Pathophysiological mechanisms and implications. *J Hepatol*. (2016) 65:425–43. doi: 10.1016/j.jhep.2016.04.005
- Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*. (2008) 9:559. doi: 10.1186/1471-2105-9-559
- Guo H, Wang M, Shang Y, Zhang B, Zhang S, Liu X, et al. Apoptosis-related prognostic biomarkers and potential targets for acute kidney injury based on machine learning algorithm and in vivo experiments. *Apoptosis*. (2023) 29:303–20. doi: 10.1007/s10495-023-01896-4
- Jiang Q, Chen Z, Meng F, Zhang H, Chen H, Xue J, et al. CD36-BATF2/MYB axis predicts anti-PD-1 immunotherapy response in gastric cancer. *Int J Biol Sci*. (2023) 19:4476–92. doi: 10.7150/ijbs.87635

15. Yuan Q, Zhou Q, Ren J, Wang G, Yin C, Shang D, et al. WGCNA identification of TLR7 as a novel diagnostic biomarker, progression and prognostic indicator, and immunotherapeutic target for stomach adenocarcinoma. *Cancer Med.* (2021) 10:4004–16. doi: 10.1002/cam4.3946
16. Ritchie M, Phipson B, Wu D, Hu Y, Law C, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007
17. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb).* (2021) 2:100141. doi: 10.1016/j.xinn.2021.100141
18. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw.* (2011) 39:1–13. doi: 10.18637/jss.v039.i05
19. Sanz H, Valim C, Vegas E, Oller J, Reverter F. SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinformatics.* (2018) 19:432. doi: 10.1186/s12859-018-2451-4
20. Xu D, Wang Y, Zhou K, Wu J, Zhang Z, Zhang J, et al. Identification of an extracellular vesicle-related gene signature in the prediction of pancreatic cancer clinical prognosis. *Biosci Rep.* (2020) 40:BSR20201087. doi: 10.1042/BSR20201087
21. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* (2011) 12:77. doi: 10.1186/1471-2105-12-77
22. Iasonos A, Schrag D, Raj G, Panageas K. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* (2008) 26:1364–70. doi: 10.1200/JCO.2007.12.9791
23. Jiang W, Yu X, Dong X, Long C, Chen D, Cheng J, et al. A nomogram based on collagen signature for predicting the immunoscore in colorectal cancer. *Front Immunol.* (2023) 14:1269700. doi: 10.3389/fimmu.2023.1269700
24. Subramanian A, Tamayo P, Mootha V, Mukherjee S, Ebert B, Gillette M, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
25. Chen B, Khodadoust M, Liu C, Newman A, Alizadeh A. Profiling Tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol.* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12
26. Yu D, Chen C, Liu X, Yao J, Li S, Ruan X. Dysregulation of miR-138-5p/RPS6KA1-AP2M1 is associated with poor prognosis in AML. *Front Cell Dev Biol.* (2021) 9:641629. doi: 10.3389/fcell.2021.641629
27. Le N, Heo K, Takei Y, Lee H, Woo C, Chang E, et al. A crucial role for p90RSK-mediated reduction of ERK5 transcriptional activity in endothelial dysfunction and atherosclerosis. *Circulation.* (2013) 127:486–99. doi: 10.1161/CIRCULATIONAHA.112.116988
28. Heo K, Le N, Cushman H, Giancursio C, Chang E, Woo C, et al. Disturbed flow-activated p90RSK kinase accelerates atherosclerosis by inhibiting SENP2 function. *J Clin Invest.* (2015) 125:1299–310. doi: 10.1172/JCI76453
29. Dehnavi S, Sadeghi M, Penson P, Banach M, Jamialahmadi T, Sahebkar A. The role of protein SUMOylation in the pathogenesis of atherosclerosis. *J Clin Med.* (2019) 8:1856. doi: 10.3390/jcm8111856
30. Heo K, Cushman H, Akaike M, Woo C, Wang X, Qiu X, et al. ERK5 activation in macrophages promotes efferocytosis and inhibits atherosclerosis. *Circulation.* (2014) 130:180–91. doi: 10.1161/CIRCULATIONAHA.113.005991
31. Singh M, Kotla S, Le N, Ae Ko K, Heo K, Wang Y, et al. Senescent phenotype induced by p90RSK-NRF2 signaling sensitizes monocytes and macrophages to oxidative stress in HIV-positive individuals. *Circulation.* (2019) 139:1199–216. doi: 10.1161/CIRCULATIONAHA.118.036232
32. Han J, Jang K, Myung C. *Garcinia cambogia* attenuates adipogenesis by affecting CEBPB and SQSTM1/p62-mediated selective autophagic degradation of KLF3 through RPS6KA1 and STAT3 suppression. *Autophagy.* (2022) 18:518–39. doi: 10.1080/15548627.2021.1936356
33. Soman A, Asha Nair S. Unfolding the cascade of SERPINA3: Inflammation to cancer. *Biochim Biophys Acta Rev Cancer.* (2022) 1877:188760. doi: 10.1016/j.bbcan.2022.188760
34. Li B, Lei Z, Wu Y, Li B, Zhai M, Zhong Y, et al. The association and pathogenesis of SERPINA3 in coronary artery disease. *Front Cardiovasc Med.* (2021) 8:56889. doi: 10.3389/fcvm.2021.756889
35. Qian L, Ji J, Jiang Y, Guo J, Wu Y, Yang Z, et al. Serpina3c deficiency induced necroptosis promotes non-alcoholic fatty liver disease through β -catenin/Foxo1/TLR4 signaling. *FASEB J.* (2022) 36:e22316. doi: 10.1096/fj.202101345RRR
36. Qian L, Ji J, Guo J, Wu Y, Ma G, Yao Y. Protective role of serpina3c as a novel thrombin inhibitor against atherosclerosis in mice. *Clin Sci (Lond).* (2021) 135:447–63. doi: 10.1042/CS20201235
37. Fan W, Morinaga H, Kim J, Bae E, Spann N, Heinz S, et al. FoxO1 regulates Tlr4 inflammatory pathway signalling in macrophages. *EMBO J.* (2010) 29:4223–36. doi: 10.1038/emboj.2010.268
38. Wu X, Poulsen K, Sanz-Garcia C, Huang E, McMullen M, Roychowdhury S, et al. MLKL-dependent signaling regulates autophagic flux in a murine model of non-alcohol-associated fatty liver and steatohepatitis. *J Hepatol.* (2020) 73:616–27. doi: 10.1016/j.jhep.2020.03.023
39. Libby P, Bornfeldt K, Tall A. Atherosclerosis: Successes, surprises, and future challenges. *Circ Res.* (2016) 118:531–4. doi: 10.1161/CIRCRESAHA.116.308334
40. Kuznetsova T, Prange K, Glass C, de Winther M. Transcriptional and epigenetic regulation of macrophages in atherosclerosis. *Nat Rev Cardiol.* (2020) 17:216–28. doi: 10.1038/s41569-019-0265-3
41. Pham L, Kennedy L, Baiocchi L, Meadows V, Ekser B, Kundu D, et al. Mast cells in liver disease progression: An update on current studies and implications. *Hepatology.* (2022) 75:213–8. doi: 10.1002/hep.32121
42. Barrett T. Macrophages in atherosclerosis regression. *Arterioscler Thromb Vasc Biol.* (2020) 40:20–33. doi: 10.1161/ATVBAHA.119.312802
43. Favero G, Paganelli C, Buffoli B, Rodella L, Rezzani R. Endothelium and its alterations in cardiovascular diseases: Life style intervention. *Biomed Res Int.* (2014) 2014:801896. doi: 10.1155/2014/801896
44. Esper R, Nordaby R, Vilariño J, Paragano A, Cacharrón J, Machado R. Endothelial dysfunction: A comprehensive appraisal. *Cardiovasc Diabetol.* (2006) 5:4. doi: 10.1186/1475-2840-5-4
45. Le N, Sandhu U, Quintana-Quezada R, Hoang N, Fujiwara K, Abe J. Flow signaling and atherosclerosis. *Cell Mol Life Sci.* (2017) 74:1835–58. doi: 10.1007/s00018-016-2442-4
46. Mitten E, Baffy G. Mechanotransduction in the pathogenesis of non-alcoholic fatty liver disease. *J Hepatol.* (2022) 77:1642–56. doi: 10.1016/j.jhep.2022.08.028
47. Nasiri-Ansari N, Androutsakos T, Flessa C, Kyrou I, Siasos G, Randeve H, et al. Endothelial cell dysfunction and nonalcoholic fatty liver disease (NAFLD): A concise review. *Cells.* (2022) 11:2511. doi: 10.3390/cells11162511