*Gastroenterology & Hepatology Research* **2026** vol. 8(1) | e00001
DOI: 10.14218/GHR.2026.00001

## Original Article

# Core Active Ingredients and Therapeutic Targets of Yiguanjian for Liver Fibrosis: A Computational and Genetic Inference-based Study

Yikun Jiang[#], Jiahui Wang[#], Lei Wang, Yang Zheng, Tiejian Zhao, Rongwu Zhang[*] and Huaye Xiao[*]

*Department of Medicine, Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine, Nanning, Guangxi, China*

**Received:** January 10, 2026 | **Revised:** March 12, 2026 | **Accepted:** April 29, 2026 | **Published online:** May 15, 2026

## Abstract

**Background and objectives:** Studies suggest that Yiguanjian (YGJ) may exert a therapeutic effect on liver fibrosis. However, the active components and molecular targets responsible for its action remain unclear. This study aimed to systematically evaluate the active ingredients and potential targets of YGJ in the treatment of liver fibrosis.

**Methods:** Active compounds and corresponding targets of YGJ were retrieved from the Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP) and the Encyclopedia of Traditional Chinese Medicine (ETCM) databases. Liver fibrosis-related datasets were obtained from the Gene Expression Omnibus (GEO) database and divided into training and validation sets. Differentially expressed genes (DEGs) from the training set were subsequently analyzed using network pharmacology, molecular dynamics simulations, and immune infiltration analysis. Three machine learning models were employed to screen for core targets, followed by Gene Set Enrichment Analysis (GSEA) and Mendelian randomization (MR) analysis. The validation set was used to assess the expression levels and diagnostic potential of core targets.

**Results:** A total of 2,887 liver fibrosis-related targets and 1,198 YGJ-related targets were identified. Three hundred and three putative targets for YGJ in the treatment of liver fibrosis were identified. Three machine learning methods further narrowed these down to five core targets. Immune infiltration analysis revealed an increase in effector B cells, resting $CD4^+$ memory T cells, γδ T cells, and M1 macrophages during liver fibrosis progression. MR analysis showed that all five core targets (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4) had odds ratios greater than 1, indicating that they function as risk factors. Expression analyses in both the training and validation sets consistently validated the MR results, demonstrating strong diagnostic potential. GSEA revealed that the core targets were enriched in key signaling pathways, including Wnt, PPAR, and MAPK. Molecular docking and molecular dynamics simulations showed that the active compounds of YGJ exhibited strong binding affinity and stability with the core targets.

**Conclusions:** YGJ exerts its potential antifibrotic effects by downregulating or antagonizing the risk-associated targets (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4). These findings provide new insights into the potential of YGJ for treating liver fibrosis, while offering a scientific reference for the prevention and treatment of chronic liver diseases.

**Keywords:** Network pharmacology; Machine learning; Mendelian randomization; Yiguanjian; Liver fibrosis; FABP4; MDM2; AKR1B1; PDGFRB; NR1H4.

[*]**Correspondence to**: Rongwu Zhang, Department of Medicine, Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine, Nanning, Guangxi 530222, China. ORCID: https://orcid.org/0009-0005-4448-1349. Tel: +86-19163768565, E-mail: ZRWZY112233@163.com; Huaye Xiao, Department of Medicine, Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine, Nanning, Guangxi 530222, China. ORCID: https://orcid.org/0009-0001-2194-6140. Tel: +86-15677167569, E-mail: 1532035137@qq.com

[#]These authors contributed equally to this work.

**How to cite this article:** Jiang Y, Wang J, Wang L, Zheng Y, Zhao T, Zhang R, *et al*. Core Active Ingredients and Therapeutic Targets of Yiguanjian for Liver Fibrosis: A Computational and Genetic Inference-based Study. *Gastroenterol Hepatol Res* 2026;8(1):e00001. doi: 10.14218/GHR.2026.00001.

## Introduction

Liver fibrosis is a pathological condition present in most chronic liver diseases, characterized by excessive proliferation and deposition of extracellular matrix (ECM) within liver tissue. This process leads to abnormal alterations in liver architecture and impairs normal liver function. Fundamentally, liver fibrosis represents an excessive repair response following chronic liver injury, which can become irreversible over time. Persistent fibrosis, accompanied by

the necrosis and apoptosis of normal hepatocytes, results in the accumulation of ECM. Eventually, this ECM forms scar tissue, progressively replacing the liver parenchyma and leading to cirrhosis, portal hypertension, or even liver cancer, ultimately resulting in liver failure.[1] Chronic liver diseases, including those caused by hepatitis viruses, alcohol, drugs, toxins, parasites, metabolic and genetic factors, cholestasis, and immune abnormalities, are defined by a disease course lasting more than six months. Long-term pathogenic stimulation, abnormal metabolism, and immune-inflammatory responses can damage hepatocytes and trigger the onset of liver fibrosis. Therefore, liver fibrosis is a common outcome in most chronic liver diseases,[2] and its further progression can lead to cirrhosis, severely impacting health and life expectancy. Although reliable data on the incidence and prevalence of liver fibrosis are lacking, epidemiological studies and clinical pathology suggest that liver fibrosis is prevalent globally. More than 1.5 billion people worldwide suffer from chronic liver diseases, with the most common causes being non-alcoholic fatty liver disease, hepatitis B virus (HBV), hepatitis C virus (HCV), and alcoholic liver disease.[3] Cirrhosis, a direct consequence of liver fibrosis, accounts for 1.82% of the global disease burden, resulting in 1.2 million deaths annually.[4] Studies have shown that etiological treatments in Western medicine, such as long-term antiviral therapy for HBV, can effectively suppress viral replication and reverse chronic hepatitis B-related liver fibrosis.[5] However, current treatments targeting the underlying etiology have limitations in fully suppressing inflammation, and once the mechanism of fibrosis is triggered, it tends to progress actively. Therefore, antifibrotic therapies targeting the proliferation and degradation of fibrous tissue are crucial and constitute an important therapeutic approach for chronic liver disease. In chronic liver diseases lacking effective etiological treatments, antifibrotic therapy becomes even more critical. Cirrhosis, the end stage of liver fibrosis, underscores the importance of preventing or reversing liver fibrosis as one of its foundational therapeutic strategies.[6]

Due to the pathological mechanisms of liver fibrosis, which involve a systemic reparative response following liver damage, drugs targeting single molecules often fail to show clinical efficacy. Consequently, there are currently no clearly effective chemical or biological drugs available for clinical use. Traditional Chinese medicine (TCM) is a comprehensive medical system with a robust theoretical framework derived from millennia of clinical practice. Recent decades of research and clinical application have demonstrated that TCM is effective in treating liver diseases, particularly in the prevention and management of liver fibrosis.[7] Although TCM did not originally recognize liver fibrosis as a distinct clinical entity, the pathological changes and clinical manifestations of liver fibrosis (including cirrhosis) have long been classified under terms such as "accumulation" or "pain in the hypochondrium." This understanding has been widely accepted in clinical practice for decades. The primary etiology of liver fibrosis varies, and clinical manifestations differ; however, the fundamental pathogenesis is often one of deficiency and stagnation, with long-term retention of toxins that impair liver meridians, causing qi stagnation and blood stasis, which can be summarized as "deficiency leading to accumulation." Liver fibrosis essentially results from damage to liver parenchyma and deficiency of the liver and kidney yin, leading to an inability to transform qi, resulting in impaired circulation of qi and blood, blood stasis, and the formation of accumulations.[8] The "deficiency" primarily manifests as spleen qi deficiency, liver qi deficiency, and liver-kidney yin deficiency. Qi deficiency reflects functional damage and decline, while liver-kidney yin deficiency denotes deeper pathological changes in liver parenchyma. Further basic research indicates that "deficiency" is mainly reflected in the reduced number and diminished function of liver parenchymal cells and damage to the liver sinusoidal walls, while "blood stasis" is characterized by excessive ECM deposition and capillarization of liver sinusoids.[9] In terms of treatment for liver-kidney yin deficiency, clinical manifestations include hypochondriac pain aggravated by exertion, soreness in the waist and knees, dry mouth and throat, restlessness, dizziness, insomnia, dry eyes, and a red tongue with scant white coating and minimal moisture. The pulse is string-like and fine. The therapeutic approach is to nourish the liver and kidney, with Yiguanjian (YGJ) commonly used in practice.[10] YGJ, referenced in the *Ancient Classical Famous Formulas Key Information Table (7 Formulas)* and originating from the *Yi Fang Jie Du*, consists of ingredients such as Glehniae Radix (Bei Sha Shen), Ophiopogon tuber (Mai Dong), *Angelica sinensis* (Dang Gui), Rehmannia (Sheng Di Huang), Goji berry (Gou Qi Zi), and *Melia toosendan* (Chuan Lian Zi). However, as a complex Chinese herbal formula, YGJ's components and targets are intricate, and the mechanism by which it treats liver fibrosis remains unclear.

Network pharmacology is a discipline that explores the relationships among drugs, targets, and diseases to virtually screen for potential active compounds, molecular targets, and signaling pathways, thus establishing "YGJ–target–disease" networks to predict the pharmacological basis and mechanisms of action of drugs.[11] Recent technological advances have revolutionized this field, particularly with the advent of machine learning. The application of machine learning enables comprehensive analysis of extensive biomedical datasets, leading to the development of more accurate diagnostic methods and individualized patient treatments.[12–14] Machine learning has enhanced our ability to discover disease biomarkers and elucidate underlying disease mechanisms.[15] Mendelian randomization (MR), a method combining instrumental variables (IVs) and Mendelian inheritance laws, overcomes the limitations of traditional randomized controlled trials by minimizing confounding factors. With advancements in Mendelian theory and its growing practical application, MR is becoming an effective tool for inferring the causal relationship between drug targets (e.g., antagonists, agonists) and disease risk.[16,17]

In this study, we propose a comprehensive multilevel *in silico* validation strategy to elucidate the potential mechanisms of YGJ in treating liver fibrosis. This strategy leverages the complementary nature of multiple computational approaches: network pharmacology and machine learning were utilized for unbiased target screening and feature extraction; GEO expression analysis provided transcriptomic validation of disease relevance; MR offered genetic inference to establish putative causal relationships minimizing confounding biases; and molecular docking alongside dynamics simulations yielded biophysical evidence of direct YGJ–target interactions. By integrating these orthogonal layers of evidence, we aimed to systematically evaluate the active ingredients and potential targets of YGJ in the treatment of liver fibrosis. Furthermore, we performed gene expression and diagnostic validation, immune cell infiltration analysis, and single-gene GSEA enrichment analysis, with the detailed workflow depicted in Figure 1.

## Materials and methods

### *Screening of core active ingredients and targets of TCM*

The six herbal ingredients of YGJ (Sheng Di Huang, Bei Sha

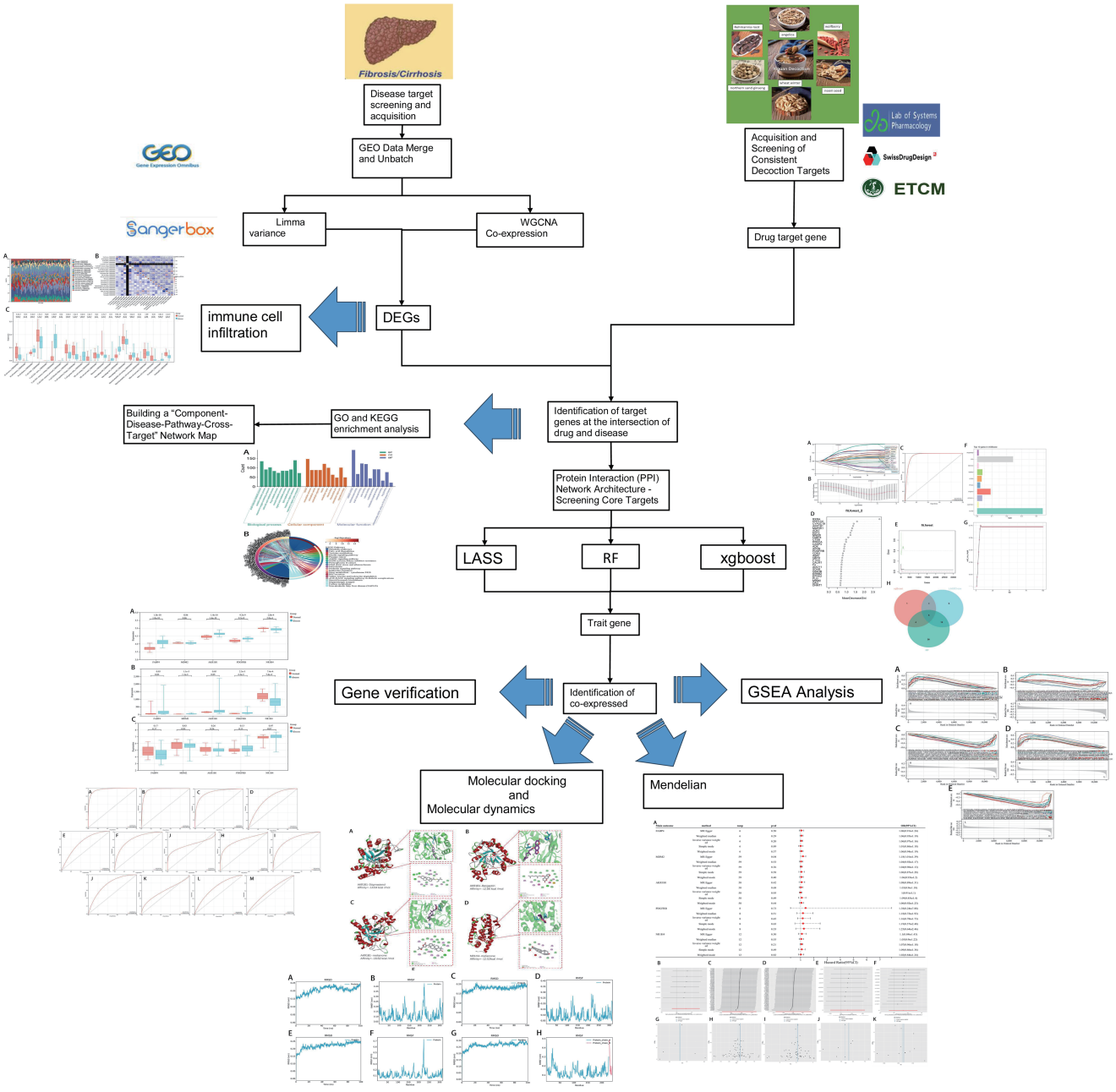


**Fig. 1. Overview flowchart of the entire process.** CI, confidence interval; DEGs, differentially expressed genes; ETCM, Encyclopedia of Traditional Chinese Medicine; GEO, Gene Expression Omnibus; GO, Gene Ontology; GSEA, Gene Set Enrichment Analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, Least Absolute Shrinkage and Selection Operator; MR, Mendelian randomization; WGCNA, weighted gene co-expression network analysis.

Shen, Mai Dong, Dang Gui, Gou Qi Zi, and Chuan Lian Zi) were input into the Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP, https://old.tcmsp-e.com/tcmsp.php) for screening. Active ingredients with oral bioavailability (OB) ≥ 30% and drug-like properties (DL) ≥ 0.18 were selected. The resulting data were merged and deduplicated to create a database of core active ingredients. These ingredients were then imported into Swiss Target Prediction (http://swisstargetprediction.ch/) for target prediction using a probability threshold of >0. The deduplicated target list was subsequently obtained. Additionally, active ingredients and targets were retrieved from the Encyclopedia of Traditional Chinese Medicine (ETCM) database (http://www.tcmip.cn/ETCM/index.php/Home/Index). After further deduplication, the final YGJ-target database was compiled.

### *Identification of disease targets*

Using "liver fibrosis" as the keyword, several relevant disease and control datasets were retrieved from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). These

datasets were merged, and batch effects were removed using the Sangerbox online platform (http://sangerbox.com/), followed by differential analysis using the limma package (differentially expressed genes were identified using thresholds of $|\log_2 FC| \geq 1.5$ and adjusted $P < 0.05$) and weighted gene co-expression network analysis (WGCNA).[18,19] The final disease target database was constructed by removing duplicate genes.

### *Immune cell infiltration analysis*

Immune cell infiltration was estimated using the TIMER2.0 platform (http://timer.cistrome.org/) via the CIBERSORT algorithm, applied to the batch-corrected datasets from multiple sources.[20] The CIBERSORT runtime parameters were set to their default values, and the sample filtering criterion was set to a CIBERSORT-calculated $P$-value < 0.05 to ensure the reliability of the estimated immune cell proportions.

### *Intersection of targets and construction of protein–protein interaction (PPI) network*

The intersection of the active ingredient targets and disease targets was identified to obtain the core pharmacological targets for treating liver fibrosis. These intersection targets were imported into the STRING database (https://version-11-5.string-db.org/cgi/input?sessionId=bC4qb8Lym4Qw&input_page_show_search=on), with "*Homo sapiens*" selected and a confidence threshold of >0.4. The original PPI network was generated, followed by optimization using Cytoscape 3.7.1. The MCODE plugin within Cytoscape 3.7.1 was employed to analyze the network, identifying subnetworks and feature gene sets based on the score values.[21]

### *Enrichment analysis*

Enrichment analysis was conducted using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The intersection target list was imported into the Sangerbox online platform (http://sangerbox.com/) for analysis. GO enrichment analysis included three categories: biological process (BP), molecular function (MF), and cellular component (CC). Results were filtered with a significance threshold of $P < 0.05$.

### *Feature gene selection using machine learning algorithms*

Machine learning techniques were employed to identify feature genes crucial to liver fibrosis. The dataset splitting strategy was as follows: the merged, batch-corrected dataset was randomly split into a training set and an internal validation set in a 7:3 ratio, and feature selection was performed only on the training set. To ensure robustness and comprehensive selection, Least Absolute Shrinkage and Selection Operator (LASSO) regression was used for variable selection and regularization, known for preventing overfitting by balancing model complexity and performance. The hyperparameter λ was selected using 10-fold cross-validation, and the optimal model was determined using the "1 standard error" criterion (λ = 1se) to identify genes with non-zero coefficients. The Random Forest (RF) algorithm, an ensemble learning method, was used to handle imbalanced data and estimate feature importance. It constructed 500 decision trees and evaluated feature importance using Recursive Feature Elimination (RFE), with the average decrease in the Gini coefficient serving as a measure of feature importance. Features with a score higher than 0.5 were considered significant. The XGBoost model, which integrates multiple weak classifiers (decision trees) to form a strong classifier, was also used. The hyperparameters for the XGBoost model were set as follows: the learning rate (eta) was 0.01, the maximum tree depth (max_depth) was 5, and the number of rounds (nrounds) was 500, with the optimal number of rounds determined via fivefold cross-validation. Feature importance was evaluated using the gain metric. Feature genes were identified from the intersection of the genes selected by LASSO logistic regression, XGBoost, and RF methods.

### *MR analysis*

A two-sample MR analysis was conducted to investigate the causal relationship between feature genes and liver fibrosis. Single nucleotide polymorphisms (SNPs) were used as IVs. Cis-eQTL data for the target genes were obtained from the Integrated Epidemiology Unit (IEU) database (https://gwas.mrcieu.ac.uk/) and GTEx (http://www.gtexportal.org/) to serve as exposure factors. The specific data sources were as follows: cis-eQTL data for FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4 were derived from pooled whole-blood tissue data in the GTEx v8 database. Pooled GWAS data on liver fibrosis served as the outcome factor. The data were sourced from the 11th release (R11) of the Finnish database (FinnGen), with the dataset ID finn-b-K11_FIBROCHIRLIV. This dataset includes case and control samples of liver fibrosis and cirrhosis from individuals of European descent. SNPs with a $P$-value $< 5 \times 10^{-8}$ were selected, and IVs with strong linkage disequilibrium (LD) ($r^2 < 0.3$, kb = 500) were excluded.[22] MR analysis was conducted using the "TwoSampleMR" package (version 0.6.8) in R (version 4.4.1). Methods used included inverse variance weighted (IVW), MR-Egger, weighted median (WM), simple mode, and weighted mode. Sensitivity analyses were performed to assess the reliability of MR results using heterogeneity, horizontal pleiotropy, and leave-one-out (LOO) analysis. Heterogeneity was assessed, with $P > 0.05$ indicating no heterogeneity.

### *In silico gene expression cross-validation*

To evaluate the generalizability of the selected signature genes, their differential expression and diagnostic performance were further validated on the training sets (GSE14323 and GSE164760) and independent validation sets (GSE6764 and GSE103580) to assess the model's stability and generalizability across different data sources. Differential gene expression of biomarkers was validated in multiple independent datasets. Receiver operating characteristic (ROC) curves were generated to evaluate the diagnostic accuracy and specificity of five feature genes for liver fibrosis. Analysis and visualization were performed using the Sangerbox platform (http://sangerbox.com/login.html).

## Gene Set Enrichment Analysis (GSEA)

GSEA was performed using the Sangerbox platform (http://sangerbox.com/login.html) and the KEGG database to identify significantly enriched pathways associated with the five feature genes and biomarkers in liver fibrosis. Samples were divided into high- and low-expression groups based on the median expression level of the feature gene, and 1,000 permutation tests were conducted, with a significance threshold set at $|NES| \geq 1.0$ and FDR < 0.25.

### *Molecular docking*

The 2D structures of active compounds were obtained from the PubChem database and stored in SDF format. The 3D structures of key targets were retrieved from the Protein Data Bank (PDB) in PDB format. Molecular docking was performed using Discovery Studio software, and the docking results were visualized with

PyMOL. A redocking process was performed at the original active site to validate the docking reliability.

### *Molecular dynamics simulation*

Molecular dynamics simulations were performed on the WeMol platform. The protein receptor was parameterized using the GROMACS 2023 software package and the CHARMM36 force field, and ligand topology files were generated using the CGenFF and GAFF force fields. The protein–ligand complex was placed in a TIP3P water box with 0.15 M NaCl to simulate physiological ionic strength. After energy minimization using the steepest descent method, a 100 ps NPT ensemble equilibrium was performed at 310 K (V-rescale thermostat) and 1 bar (Parrinello–Rahman pressure thermostat), followed by a 100 ns final molecular dynamics simulation. After the simulation, the trajectories were post-processed, and the root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) were calculated to assess the structural stability of the complex.

## Results

### *Screening of core active ingredients and targets of TCM*

A total of 59 active ingredients from five core herbs were identified from the TCMSP database after deduplication. These ingredients were then imported into the Swiss Target Prediction database, yielding 1,186 corresponding targets. After deduplication, 1,075 effective active ingredient–target pairs were obtained. Additionally, 735 targets were retrieved from the ETCM database (http://www.tcmip.cn/ETCM/index.php/Home/Index). After removing duplicates, the final YGJ target database contained 1,198 unique targets.

### *Identification of disease targets*

Four datasets (GSE14323, GSE164760, GSE6764, and GSE103580) were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), comprising a total of 260 samples from patients with chronic liver disease (covering various disease stages, including liver fibrosis, cirrhosis, and hepatocellular carcinoma) and 25 healthy control samples. Due to the lack of a unified disease staging standard in the original annotation information of each dataset and the limited sample size in some datasets, this study uniformly defined the above samples as the "disease group" to identify stable transcriptomic features associated with chronic liver disease status. It should be noted that the core objective of this study was to explore causal relationships between drug targets and diseases; subsequent MR analysis can provide causal inference evidence independent of sample composition. To mitigate batch effects, normalization was applied to the raw data of all datasets. The GSE14323 and GSE164760 datasets were combined to form the training cohort, while GSE6764 and GSE103580 were used as independent validation cohorts to confirm our findings. After batch effect removal and data merging via the Sangerbox platform (http://sangerbox.com/), the final dataset included 260 liver fibrosis samples and 25 control samples. Box plots revealed considerable distribution differences between the datasets before batch effect correction, indicating the presence of batch effects, whereas after batch effect removal, the data distributions aligned, with median values converging on the same line. Uniform Manifold Approximation and Projection (UMAP) analysis further confirmed the presence of batch effects before correction, with samples from each dataset clustering separately. However, after batch effect removal, samples from different datasets clustered together, indicating effective correction. Differentially expressed genes (DEGs) were identified using the limma package, with 144 upregulated genes and 54 downregulated genes. These results were visualized as volcano plots and heatmaps and organized into a matrix expression profile.

WGCNA was performed, and the soft threshold was determined to be 5 based on scale independence and average connectivity comparisons. Genes with similar expression profiles were classified into gene modules using hierarchical clustering based on TOM (topological overlap measure). The minimum module size for clustering was set to 30. Modules with a distance of less than 0.2 were merged, resulting in three co-expression modules. Notably, the black module, representing genes that could not be assigned to any module, was identified. Hub genes were filtered using the cutoff criterion of module membership (MM) > 0.3. After merging DEGs and hub genes and removing duplicates, the disease target database contained a total of 2,887 genes (see Fig. 2).

### *Immune cell infiltration*

To explore the role of immune cells in liver fibrosis pathogenesis, immune cell infiltration analysis was performed to determine the proportions of 22 immune cell types in each sample. As shown in Figure 3a, correlation analysis revealed strong associations between certain immune cell types. For instance, naive B cells were strongly correlated with memory B cells (r = 0.33), effector B cells with monocytes (r = 0.32), and effector B cells with neutrophils (r = 0.41). Cytotoxic T lymphocytes exhibited strong correlations with resting $CD4^+$ memory T cells (r = 0.58), follicular helper T cells (r = 0.42), and γδ T cells (r = 0.57). Resting $CD4^+$ memory T cells were strongly correlated with follicular helper T cells (r = 0.53), regulatory T cells (r = 0.57), γδ T cells (r = 0.48), resting NK cells (r = 0.33), and M1 macrophages (r = 0.40). Further analysis revealed significant correlations between various immune cells, including regulatory T cells and γδ T cells (r = 0.48), regulatory T cells and resting NK cells (r = 0.43), and resting $CD4^+$ memory T cells with neutrophils (r = 0.40), among others. Detailed correlations are shown in Figure 3b.

Box plot analysis further illustrated the differential expression of immune cells between the liver fibrosis and normal groups. Effector B cells, resting $CD4^+$ memory T cells, γδ T cells, and M1 macrophages showed significantly higher expression levels in the liver fibrosis group than in the normal group (Fig. 3c).

### *Intersection of targets and construction of the PPI network*

The intersection of the active ingredient–target library and the disease target library yielded 303 overlapping targets. These were imported into the STRING database to generate the original PPI network (Fig. 4). The PPI data were then analyzed using Cytoscape 3.7.1, calculating the degree, closeness, and betweenness values. The average degree value was 14.83, with closeness and betweenness values of 0.00127 and 507.64, respectively. Nodes with values lower than the average degree were excluded. Further analysis with the MCODE tool led to the identification of a subnetwork (Fig. 4).

### *Enrichment analysis*

A total of 2,420 entries were obtained from GO functional analysis. These included 1,965 BP entries, 193 CC entries, and 262 MF entries. The top 10 enriched terms, based on the number of enriched genes, were visualized (Fig. 5). KEGG pathway analysis identified 148 pathways, from which the top 20 were selected based on both the number of enriched genes and *P*-values. These pathways were primarily associated with non-alcoholic fatty liver disease, AMPK, AGE–RAGE signaling, and tyrosine metabolism,

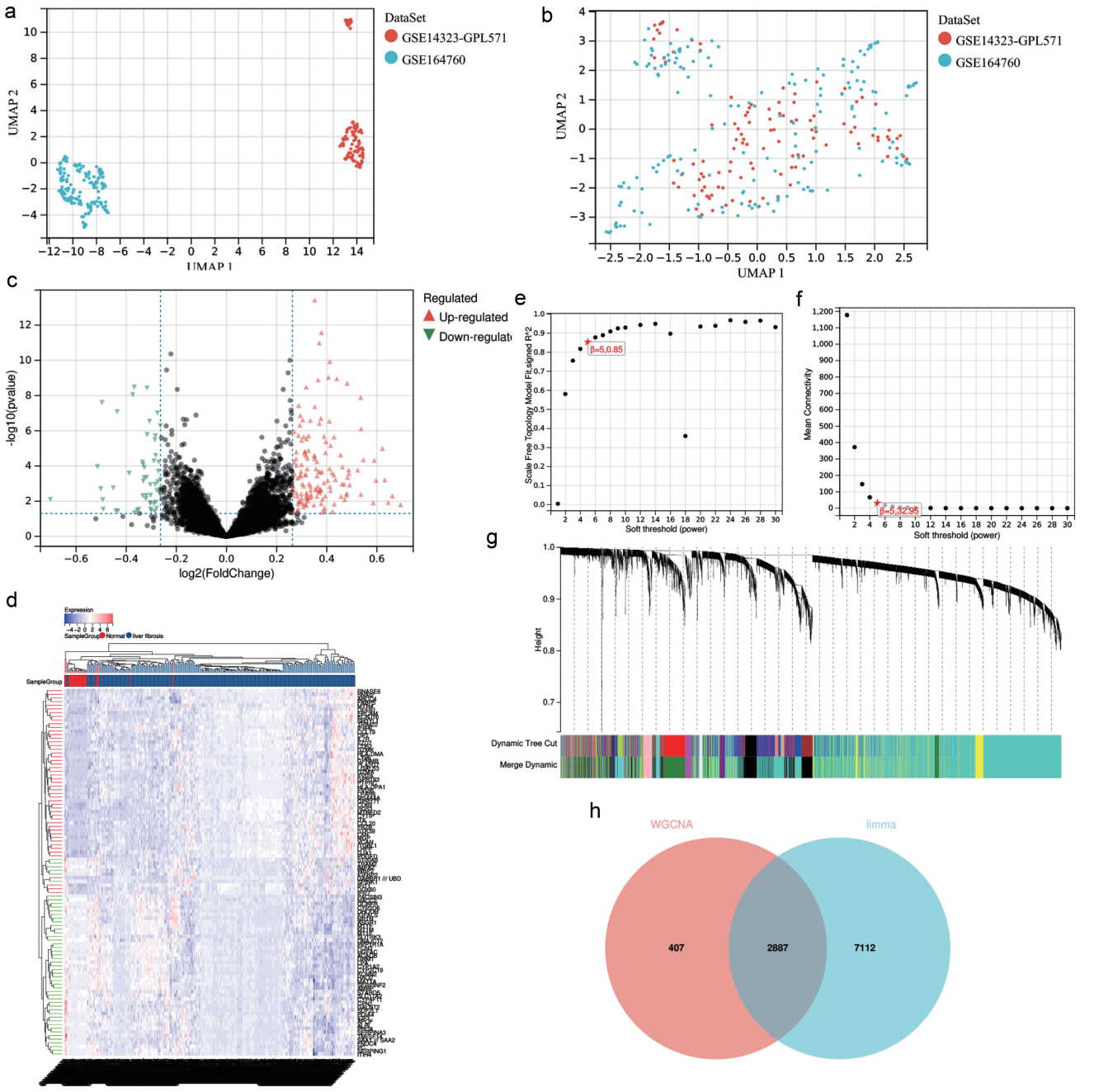


**Fig. 2. Screening of liver fibrosis–related targets.** (a) Principal component analysis (PCA) of the dataset before batch correction; (b) PCA of the dataset after batch correction; (c) Heatmap of differential analysis results; (d) Volcano plot of differential analysis results; (e) Scale-free topology plot from weighted gene co-expression network analysis (WGCNA); (f) Mean connectivity plot from WGCNA clustering analysis; (g) Gene network module plot from WGCNA clustering analysis; (h) Venn diagram illustrating the intersection between differentially expressed genes (DEGs) and hub genes.

among others. The enrichment bubble chart and bar chart illustrating these pathways are shown in Figure 5.

### *Feature gene selection using machine learning algorithms*

To identify the 303 key genes, we applied various machine learning algorithms. Specifically, LASSO regression identified 23 genes, yielding a ROC curve with an area under the curve (AUC) of 0.96. The XGBoost algorithm identified 10 genes, while the RF method identified 47 important genes. A Venn diagram was used to identify common genes across all three methods, resulting in the selection of five feature genes, as shown in Figure 6.

### *MR analysis*

MR analysis was performed to confirm the causal relationship between key target genes and liver fibrosis using methods such as IVW, MR-Egger, WM, simple mode, and weighted mode. The

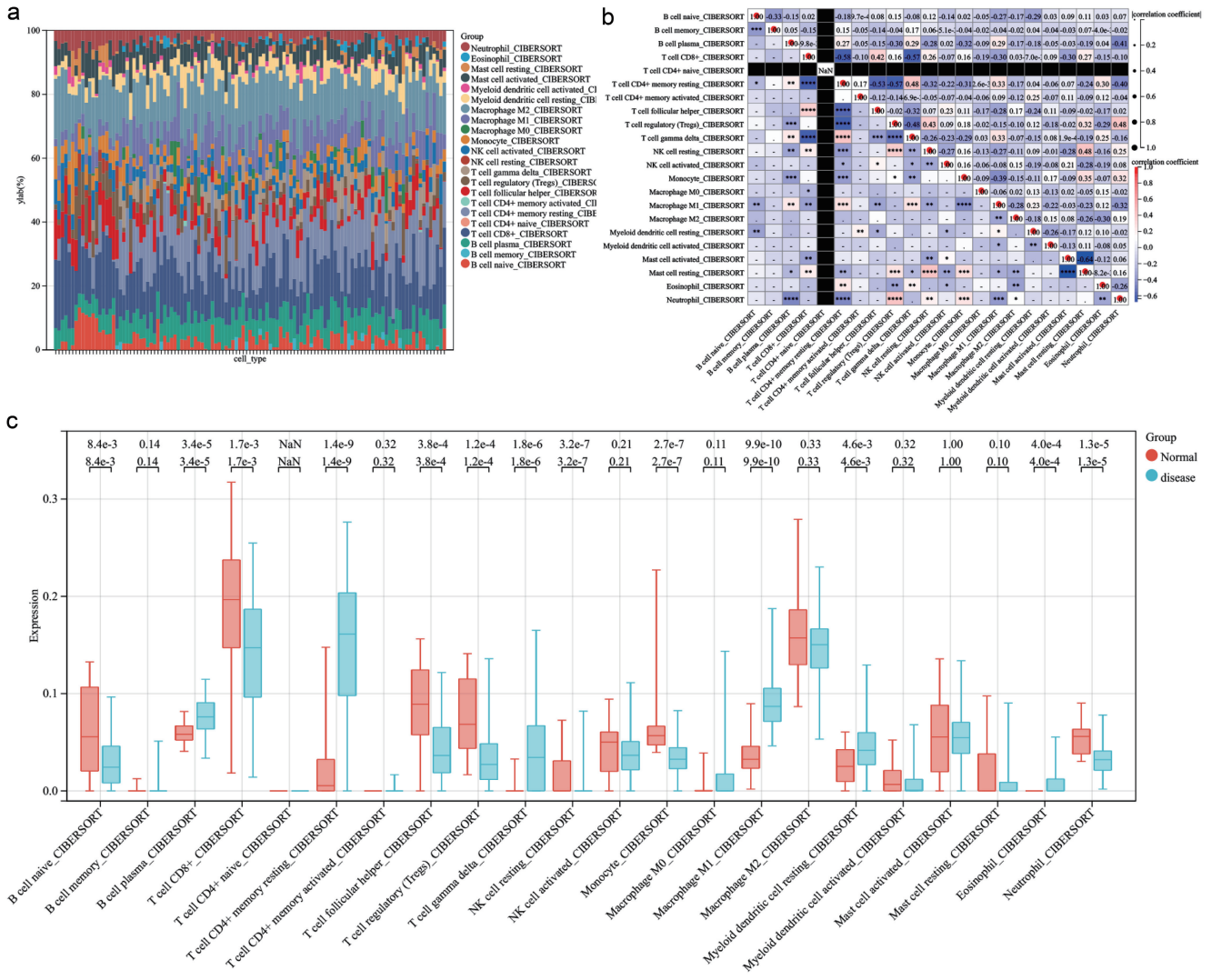


**Fig. 3. Immune cell infiltration analysis.** (a) Expression proportions of immune cells in the samples; (b) Correlation heatmap of immune cells; (c) Expression profiles of immune cells in the liver fibrosis group versus the normal group.

results, displayed in a forest plot, indicated that five feature genes (FABP4 (GTEx Database), MDM2 (eQTL-a-ENSG00000135679), AKR1B1 (eQTL-a-ENSG00000085662), PDGFRB (eQTL-a-ENSG00000085662), and NR1H4 (GTEx Database)) exhibited strong causal associations with liver fibrosis (odds ratio (OR) > 0.9), suggesting their potential as effective drug targets. Among these, the ORs for FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4 exceeded 1, indicating that their expression levels are positively correlated with the likelihood of liver fibrosis and classifying them as risk factors. Consequently, we hypothesize that the active compounds in YGJ may exert their therapeutic effects by acting as antagonists or downregulators of these risk-associated genes. Sensitivity analyses, including funnel plot evaluation, confirmed the robustness of these MR results. Leave-one-out analysis demonstrated consistent MR findings after systematically removing each SNP, further supporting the reliability of these results (Fig. 7).

### *Gene expression validation*

The expression of the five key genes (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4) was evaluated in both the training (GSE14323 and GSE164760) and independent validation cohorts. The expression trends for these genes were consistent across all datasets (Fig. 8). Specifically, compared with the control group, the expression levels of FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4 were higher in the liver fibrosis group, highlighting these core targets as potential risk factors. Furthermore, ROC curve analysis revealed AUC values greater than 0.5 for all five genes, confirming their potential diagnostic ability for liver fibrosis (Fig. 9). These results indicate that FABP4, AKR1B1, PDGFRB, and NR1H4 exhibit consistent differential expression in chronic liver diseases, suggesting their potential as biomarkers for chronic liver diseases. It is worth noting that, since the case cohort includes samples from different disease stages, the aforementioned diagnostic performance requires further validation in a more strictly defined cohort of early-stage liver fibrosis. It is also worth noting that MDM2 did not demonstrate outstanding diagnostic performance (with an AUC below 0.5) in the training set or in GSE82107. This may be due to differences in the tissues in which the same gene

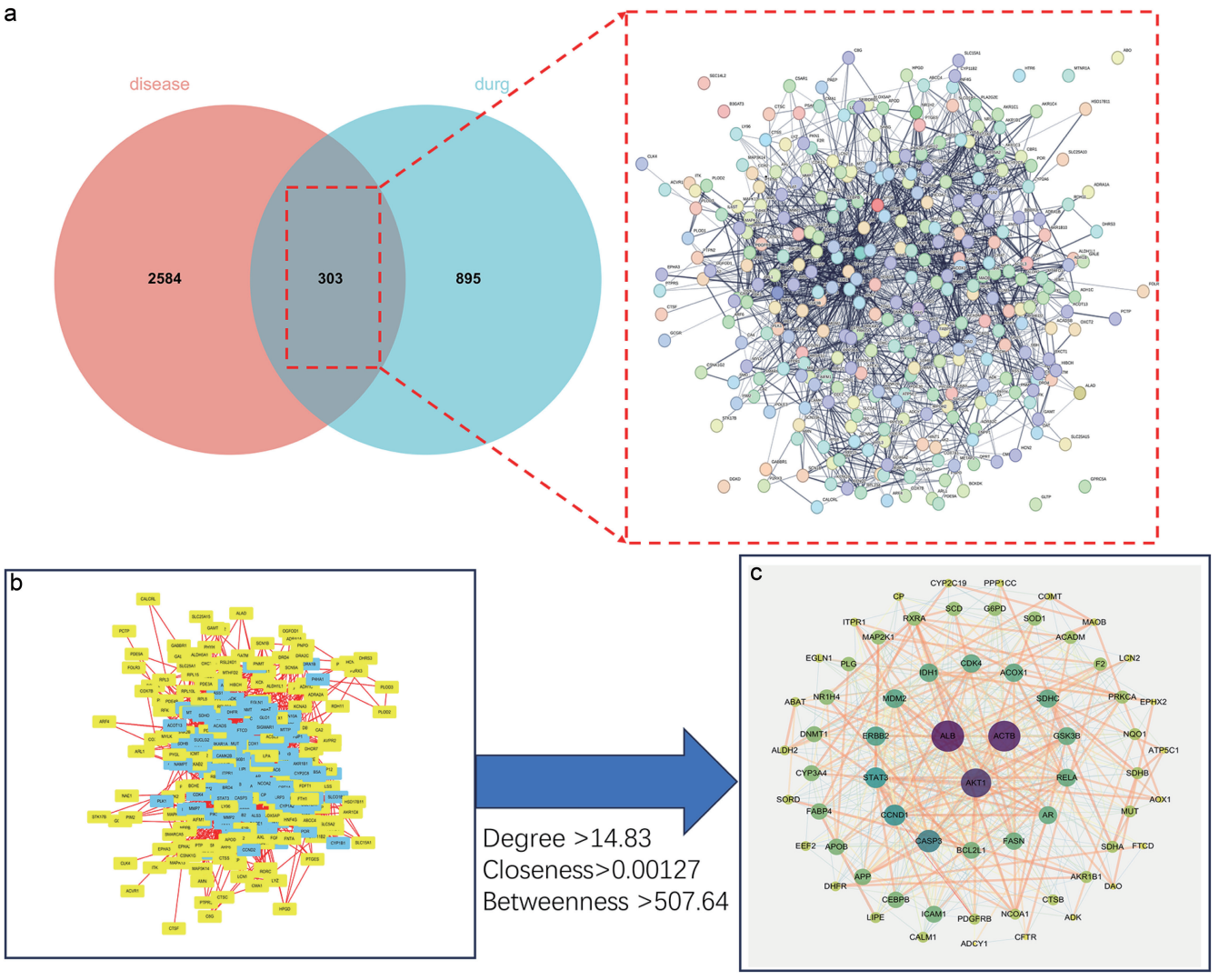


**Fig. 4. Protein–protein interaction (PPI) network visualization.** (a) Venn diagram of intersecting targets and the original PPI network diagram; (b) Diagram illustrating the analysis and screening of degree, closeness, and betweenness centrality values; (c) MCODE (module clustering) analysis results (note: the deeper the color and the larger the shape, the higher the score).

is expressed across GEO datasets; its diagnostic performance requires further research and validation.

### *GSEA*

To elucidate the underlying mechanisms, GSEA was conducted to analyze the significantly enriched pathways for the five core targets (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4). FABP4 was mainly enriched in actin cytoskeleton regulation pathways; MDM2 in apoptosis and Wnt signaling pathways; AKR1B1 in PPAR and pyruvate metabolism pathways; PDGFRB in ECM–receptor interaction and MAPK signaling pathways; and NR1H4 in alanine, aspartate, and glutamate metabolism and tryptophan metabolism pathways (Fig. 10).

### *Molecular docking and binding mode analysis*

The top 10 effective ingredients from the "Core TCM–Active Ingredients–Disease Pathways–Intersecting Targets" network, based on docking scores, were selected as core compounds: Cytidylic Acid B, Adenosine, Adenine Nucleoside, Stigmasterol, Cytidylic Acid A, Quercetin, Uridylic Acid, Bergaptin, Cnidilin, and Ethyl Linolenate. Molecular docking was performed with FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4. The highest binding scores were obtained for the following core compound–target pairs: AKR1B1–Stigmasterol, AKR1B1–Bergaptin, AKR1B1–Melianone, and NR1H4–Melianone, with average binding energies of −14.64, −12.86, −16.62, and −12.52 kcal/mol, respectively. Binding energies below −4.25, −5.0, or −7.0 kcal/mol indicate moderate, good, or strong binding affinity, respectively, suggesting that all four core compound–target pairs exhibit strong binding activity (Fig. 11).

### *Molecular dynamics simulation*

To verify binding stability, molecular dynamics (MD) simulations were performed on the docking complexes over a 100 ns period with a time step of 0.002 ns. RMSD was used to evaluate the stability of protein–ligand conformations, while RMSF was em-

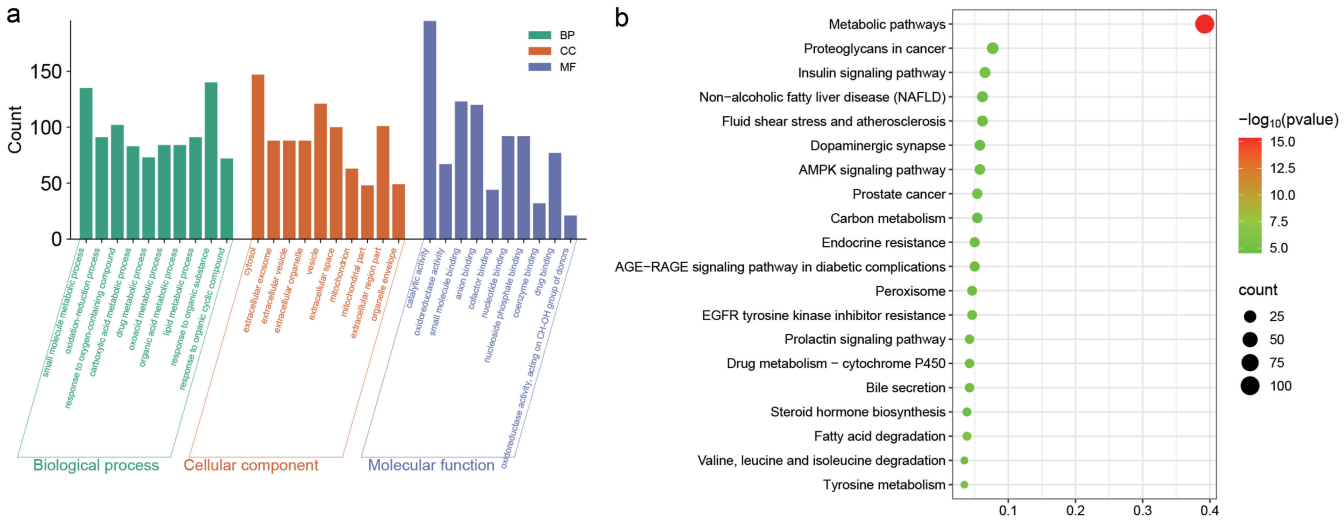


**Fig. 5. Enrichment analysis.** (a) Gene Ontology (GO) enrichment analysis results; (b) Kyoto Encyclopedia of Genes and Genomes (KEGG) circular plot.

ployed to analyze positional variations of proteins under specific pressure and temperature conditions. The results indicated that the RMSD curves for all pairs displayed good stability over the 100 ns simulation period. Specifically, AKR1B1–Stigmasterol stabilized after 25 ns, AKR1B1–Bergaptin after 20 ns, AKR1B1–Melianone after 10 ns, and NR1H4–Melianone after 20 ns. RMSF peaks indicated significant fluctuations in specific regions throughout the simulation. The RMSF values for the four core compound–target pairs ranged from 0.01 Å to 0.69 Å, suggesting that these protein–ligand complexes remained stable under the simulated conditions. Overall, these molecular dynamics simulations support the stability of the selected docking complexes (Fig. 12).

## Discussion

Through PPI network construction and machine learning-based screening, five core targets (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4) were computationally predicted as putative risk factors for liver fibrosis. Given their identification as risk factors in our MR analysis (OR > 1), we hypothesize that YGJ potentially exerts its antifibrotic effects by inhibiting or downregulating these specific targets. It should be noted that these findings are derived from bioinformatic inferences and require further experimental corroboration. FABP4 is a lipid factor secreted by adipocytes and macrophages, and studies have demonstrated its involvement in lipid metabolism and inflammatory responses.[23] MDM2, an E3 ubiquitin ligase, plays a crucial role in regulating the degradation of various proteins, including p53/TP53, which is implicated in hepatic stellate cell-mediated ferroptosis during liver fibrosis.[24,25] AKR1B1 encodes an aldose reductase enzyme, which is involved in cancer metabolism and intracellular fructose supply.[26] PDGFRB, a receptor tyrosine kinase, activates key signaling pathways such as PI3K/Akt, RAS/MAPK, and JAK/STAT, promoting cell proliferation and survival.[27] NR1H4, which encodes a bile acid receptor, is crucial for liver development and has been linked to cholestatic liver diseases.[28]

The findings suggest that these five key genes play distinct roles in the pathological progression of liver fibrosis. Furthermore, the "active ingredient–target network" constructed in this study indicates that YGJ exerts its therapeutic effects not through a single target or active ingredient but via a "multi-ingredient, multi-target" mechanism. This highlights the complex interactions between different active ingredients and targets in the treatment of liver fibrosis.

Immune infiltration analysis revealed an increased presence of effector B cells, resting CD4+ memory T cells, γδ T cells, and M1 macrophages during the development of liver fibrosis, suggesting that immune cells play a pivotal role in disease progression.

KEGG and GSEA enrichment analyses demonstrated that the key targets of YGJ in the treatment of liver fibrosis are enriched in multiple signaling pathways, including MAPK, AGE–RAGE, tyrosine metabolism, actin cytoskeleton regulation, apoptosis, Wnt signaling, PPAR signaling, pyruvate metabolism, ECM–receptor interactions, alanine, aspartate and glutamate metabolism, and tryptophan metabolism. MAPK signaling can induce mitochondrial-dependent apoptosis, promoting Bax translocation to mitochondria, inhibiting Bcl-2, and activating caspase-3, which may work synergistically with the Bcl-2/Bax signaling pathway in regulating liver fibrosis.[29] The AGE–RAGE signaling pathway has been identified as a key mechanism in the activation of hepatic stellate cells.[30] Wnt signaling regulates lactate dehydrogenase-A nuclear translocation, mediating glycolysis in hepatic stellate cells during the development of liver fibrosis.[31] Targeting PPAR-γ has been shown to inhibit hepatic stellate cell activation and improve liver fibrosis.[32] The role of pyruvate metabolism in driving hepatic stellate cell activation depends on its oxidative phosphorylation in the mitochondria or its metabolic processing in the cytoplasm.[33]

Furthermore, MR was employed to explore the potential causal relationships between the treatment targets (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4) and liver fibrosis. This approach leverages eQTL and liver fibrosis association data to establish causal links, thereby enhancing the predictive validity of the identified targets. MR analysis, employing five different methods to estimate causal effects, combined with sensitivity analysis and colocalization verification, provides robust causal estimates while minimizing the potential influence of pleiotropy in the IVs, further supporting the relevance of these targets in liver fibrosis.

In MR analysis, the five feature genes (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4) exhibited strong causal relationships with liver fibrosis (OR > 0.9), and all demonstrated sig-

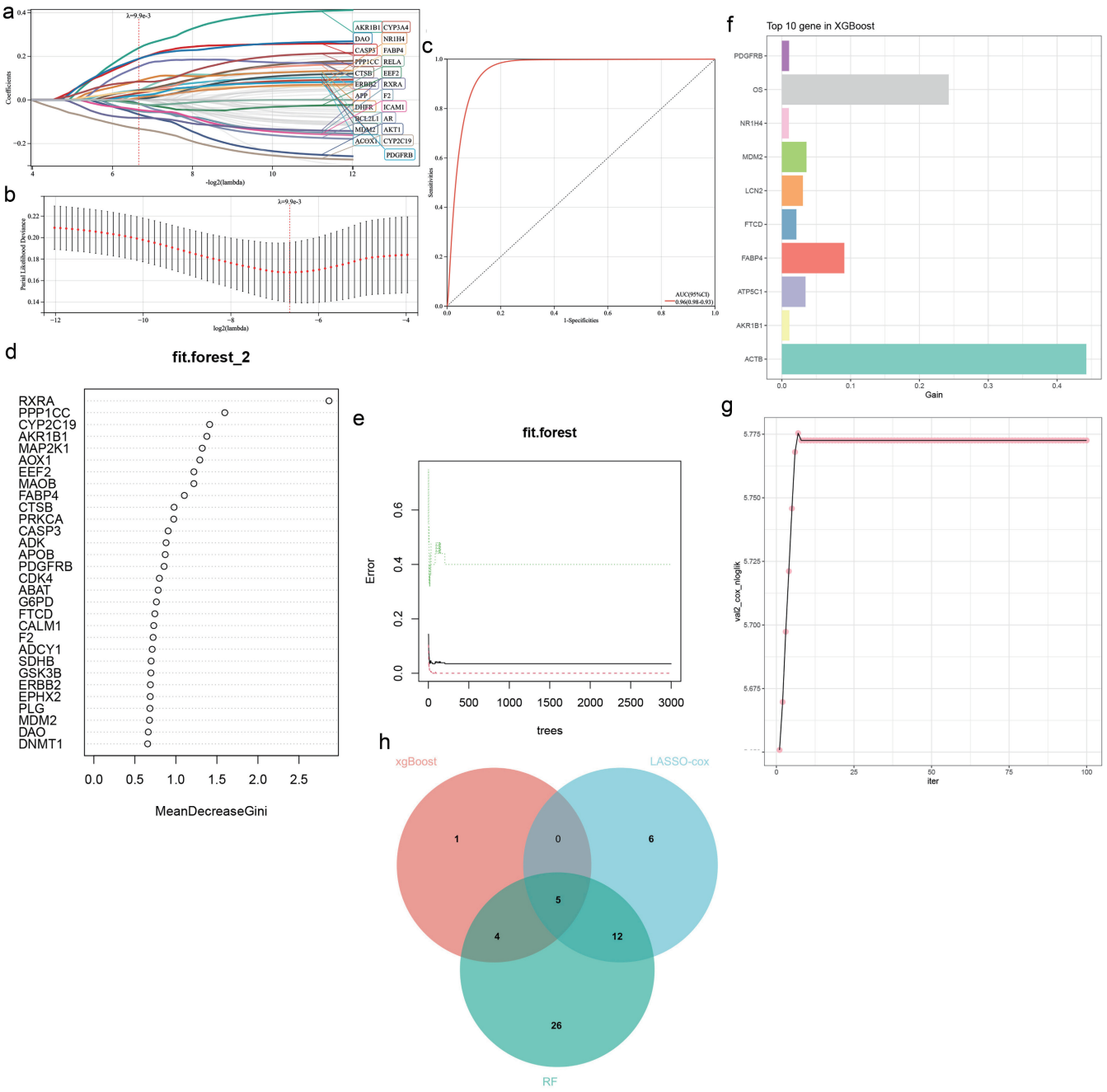


**Fig. 6. Feature gene selection using machine learning algorithms.** (a) Regression coefficients in the Least Absolute Shrinkage and Selection Operator (LASSO) regression algorithm; (b) Cross-validation curve in the LASSO regression algorithm; (c) Receiver operating characteristic (ROC) curve of the LASSO model; (d) Feature gene importance in the Random Forest (RF) algorithm; (e) Error rate versus number of trees in the RF algorithm; (f) Top 10 feature genes in the XGBoost model; (g) model evaluation curve for the XGBoost model; (h) Venn diagram of feature genes selected by LASSO, RF, and XGBoost.

nificant potential as drug targets. OR values greater than 1 suggest that increased expression of these genes correlates with a higher likelihood of liver fibrosis, positioning them as risk factors for the disease. From a pharmacological perspective, this directionality suggests that therapeutic interventions, such as YGJ, would need to inhibit or suppress these targets to mitigate disease progression. This aligns with the observed expression differences of these genes in liver fibrosis patients compared with control groups across all datasets. The consistency between MR analysis and differential expression analysis further corroborates the robust causal relationship of FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4 with liver fibrosis. ROC curves across all datasets confirm the diagnostic potential of these genes, with AUC values consistently exceeding 0.5, demonstrating predictive performance. At the same time, a major limitation of this study is that the case cohort in the GEO dataset encompasses a continuous disease spectrum ranging from

a

| Main outcome | method | nsnp | pval | OR(95%CI) |
|---|---|---|---|---|
| FABP4 | MR Egger | 6 | 0.50 | 1.06(0.91to1.24) |
| | Weighted median | 6 | 0.29 | 1.06(0.95to1.19) |
| | Inverse variance weighted | 6 | 0.20 | 1.06(0.97to1.16) |
| | Simple mode | 6 | 0.89 | 1.01(0.86to1.18) |
| | Weighted mode | 6 | 0.37 | 1.06(0.94to1.19) |
| MDM2 | MR Egger | 59 | 0.04 | 1.18(1.01to1.39) |
| | Weighted median | 59 | 0.53 | 1.04(0.92to1.17) |
| | Inverse variance weighted | 59 | 0.36 | 1.04(0.96to1.13) |
| | Simple mode | 59 | 0.58 | 1.06(0.87to1.28) |
| | Weighted mode | 59 | 0.40 | 1.06(0.93to1.2) |
| AKR1B1 | MR Egger | 58 | 0.42 | 1.08(0.89to1.31) |
| | Weighted median | 58 | 0.68 | 1.03(0.9to1.18) |
| | Inverse variance weighted | 58 | 0.95 | 1(0.91to1.1) |
| | Simple mode | 58 | 0.49 | 1.09(0.85to1.4) |
| | Weighted mode | 58 | 0.44 | 1.06(0.92to1.23) |
| PDGFRB | MR Egger | 8 | 0.73 | 1.38(0.24to7.88) |
| | Weighted median | 8 | 0.51 | 1.18(0.72to1.93) |
| | Inverse variance weighted | 8 | 0.45 | 1.16(0.78to1.73) |
| | Simple mode | 8 | 0.65 | 1.19(0.57to2.48) |
| | Weighted mode | 8 | 0.53 | 1.25(0.64to2.46) |
| NR1H4 | MR Egger | 12 | 0.50 | 1.1(0.84to1.43) |
| | Weighted median | 12 | 0.55 | 1.05(0.9to1.22) |
| | Inverse variance weighted | 12 | 0.21 | 1.07(0.96to1.18) |
| | Simple mode | 12 | 0.49 | 1.09(0.86to1.36) |
| | Weighted mode | 12 | 0.82 | 1.02(0.84to1.24) |

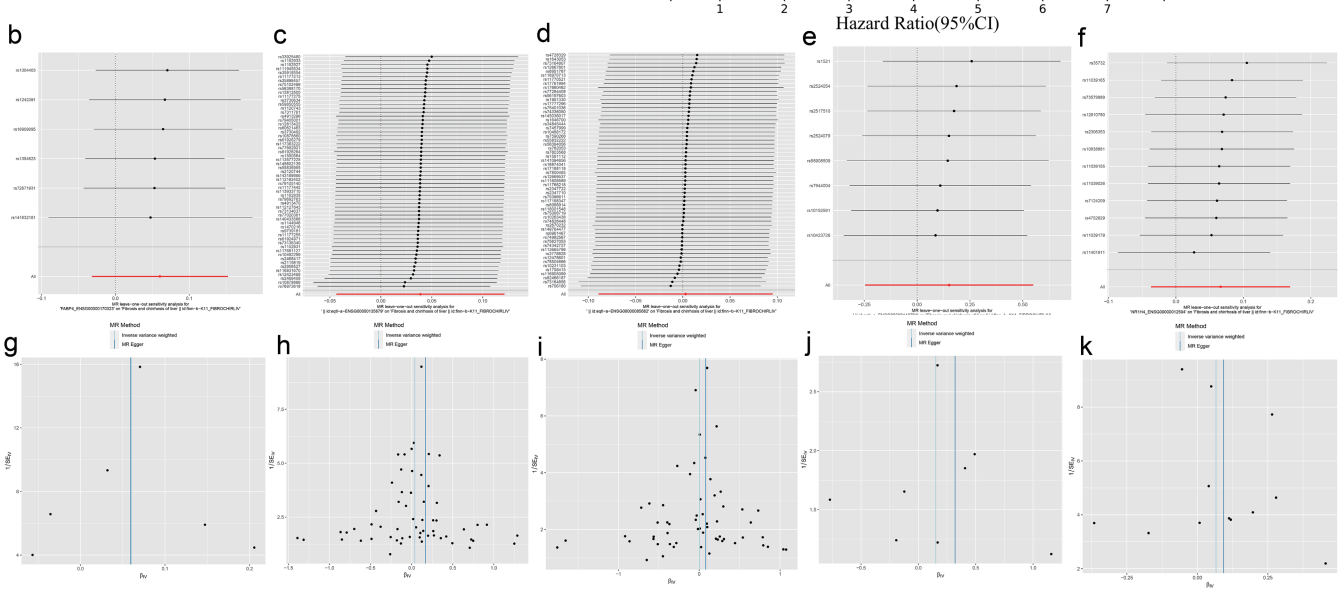


**Fig. 7. Mendelian randomization (MR) results for five feature genes.** (a) Forest plot of the overall MR results; (b) Sensitivity analysis for FABP4 using the leave-one-out method; (c) Sensitivity analysis for MDM2 using the leave-one-out method; (d) Sensitivity analysis for AKR1B1 using the leave-one-out method; (e) Sensitivity analysis for PDGFRB using the leave-one-out method; (f) Sensitivity analysis for NR1H4 using the leave-one-out method; (g) Funnel plot for FABP4; (h) Funnel plot for MDM2; (i) Funnel plot for AKR1B1; (j) Funnel plot for PDGFRB; (k) Funnel plot for NR1H4. CI, confidence interval; OR, odds ratio.

liver fibrosis to hepatocellular carcinoma, and precise stratification based on existing annotation information is not possible. However, the core findings of this study—namely, the causal relationships between FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4 and liver disease risk—are primarily based on MR analysis. MR analysis utilizes genetic variants as IVs, and its causal estimates are not influenced by the composition of the case cohort (since the association between exposure and outcome is based on population-level genetic data rather than transcriptomic data from the case cohort). Therefore, despite phenotypic heterogeneity in the transcriptomic-level validation, the MR results provide robust support for the key causal inferences of this study.

Lastly, molecular docking and molecular dynamics simulations were employed to validate the binding affinity of key active ingre-

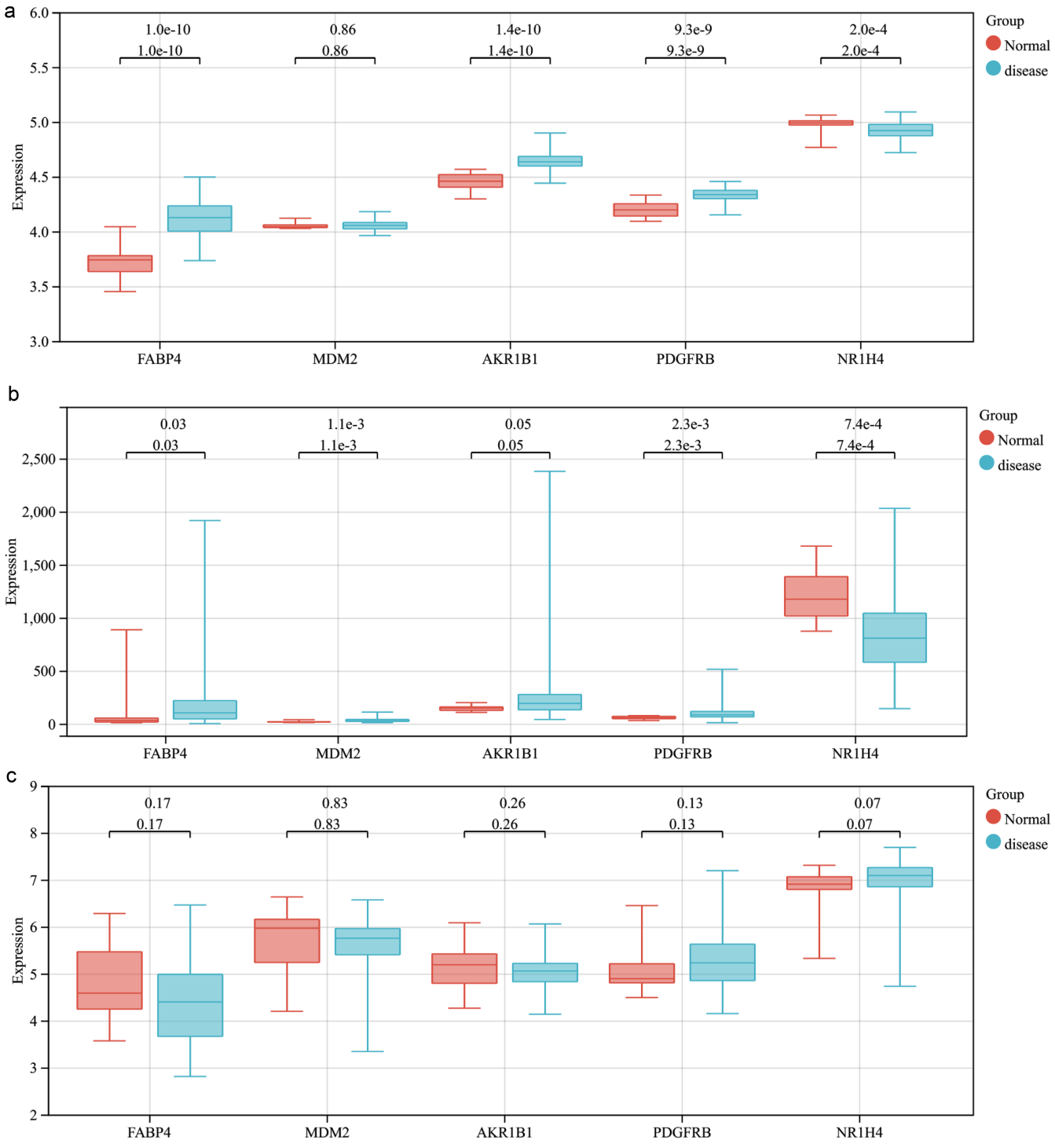


**Fig. 8. Gene validation.** (a) Differential gene expression of the five feature genes in the training cohort (GSE14323 and GSE164760); (b) Differential gene expression of the five feature genes in the GSE6764 dataset; (c) Differential gene expression of the five feature genes in the GSE103580 dataset.

dients in YGJ with the identified core targets. The binding energies between the core ingredients and core targets suggest strong binding interactions, providing further evidence of the synergistic therapeutic effects of these ingredients in treating liver fibrosis. However, further research is needed to explore whether these interactions extend to other active ingredients in YGJ.

In summary, our study successfully integrates network pharmacology and MR to establish a causal relationship between "YGJ–key components–core targets–liver fibrosis," identifying and validating FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4

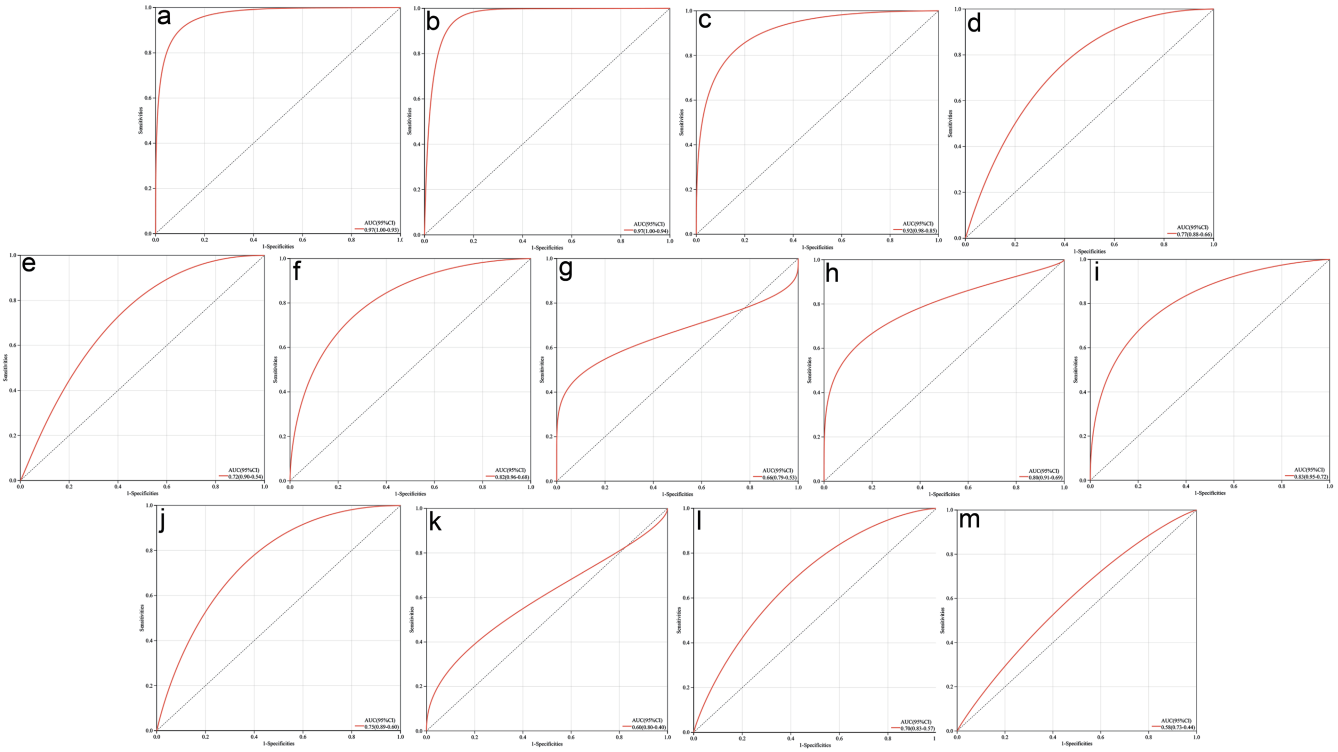


**Fig. 9. Receiver operating characteristic (ROC) curve analysis.** (a–e) Diagnostic performance analysis of the five feature genes (FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4) in the training cohort (GSE14323 and GSE164760); (f–i) Diagnostic performance analysis of the five feature genes in the GSE6764 dataset; (j–m) Diagnostic performance analysis of the five feature genes in the GSE103580 dataset.

as reliable biomarkers for liver fibrosis with good predictive performance. Functional investigations reveal significant interactions between these targets and immune-related pathways in liver fibrosis. These findings offer valuable insights into the complex molecular mechanisms of liver fibrosis and emphasize the potential for developing personalized diagnostic and therapeutic strategies, paving the way for targeted and effective interventions.

This study has several limitations:

- Due to the characteristics of TCM in syndrome differentiation and treatment, the dosage ratio of YGJ for treating various types of liver fibrosis (primarily targeting Liver-Kidney Yin Deficiency Syndrome) may vary, potentially influencing its efficacy. Insufficient clinical data were collected to further refine the liver fibrosis model and perform weighted analyses of different syndromes and dosages. Furthermore, Chinese herbal formulations involve complex "multi-ingredient, multi-target" actions, and quantifying synergistic interactions between active ingredients and their activation or inhibition of targets requires advanced network algorithms, which remains an area for future improvement in network pharmacology. With the advancement of artificial intelligence, network pharmacology, and molecular docking technologies, TCM and its formulations are likely to be more effectively integrated with tissue engineering for clinical application.
- In the MR analysis of drug targets, stringent criteria were used; therefore, other SNPs with causal relationships to liver fibrosis may still be unidentified. Due to sample limitations, stratification by disease type and patient sex was not performed, which may have introduced some bias. Further clinical validation of these key targets will require additional datasets.[34]
- The GEO dataset used in this study may present variability in gene expression across different tissues, which could potentially affect AUC values and other results.
- The GEO dataset used for transcriptomic validation includes samples of chronic liver disease at various stages of progression; however, due to limitations in annotation of the raw data, it was not possible to perform precise subgroup analyses or exclude samples with hepatocellular carcinoma. Consequently, ROC analysis results may be influenced by disease severity and cannot be directly extrapolated to the diagnostic context of early-stage liver fibrosis. However, the main conclusion of this study—the causal relationship between core genes and liver disease risk—is supported by MR analysis, which is independent of the aforementioned sample composition. Independent validation in a strictly defined early-stage fibrosis cohort remains necessary in the future.

Finally, it is important to emphasize that this is primarily a computational and genetic inference-based study. Although multiple algorithms and independent datasets were used for cross-validation, the precise mechanistic pathways lack direct *in vivo* or *in vitro* experimental validation. Therefore, the findings should be interpreted as predictive hypotheses intended to guide future experimental research. Comprehensive *in vitro* and *in vivo* studies are still required to empirically confirm these pharmacological targets and pathways.

## Conclusions

Network pharmacology and MR identified FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4 as potential risk biomarkers for liver fibrosis, suggesting that YGJ may exert therapeutic effects

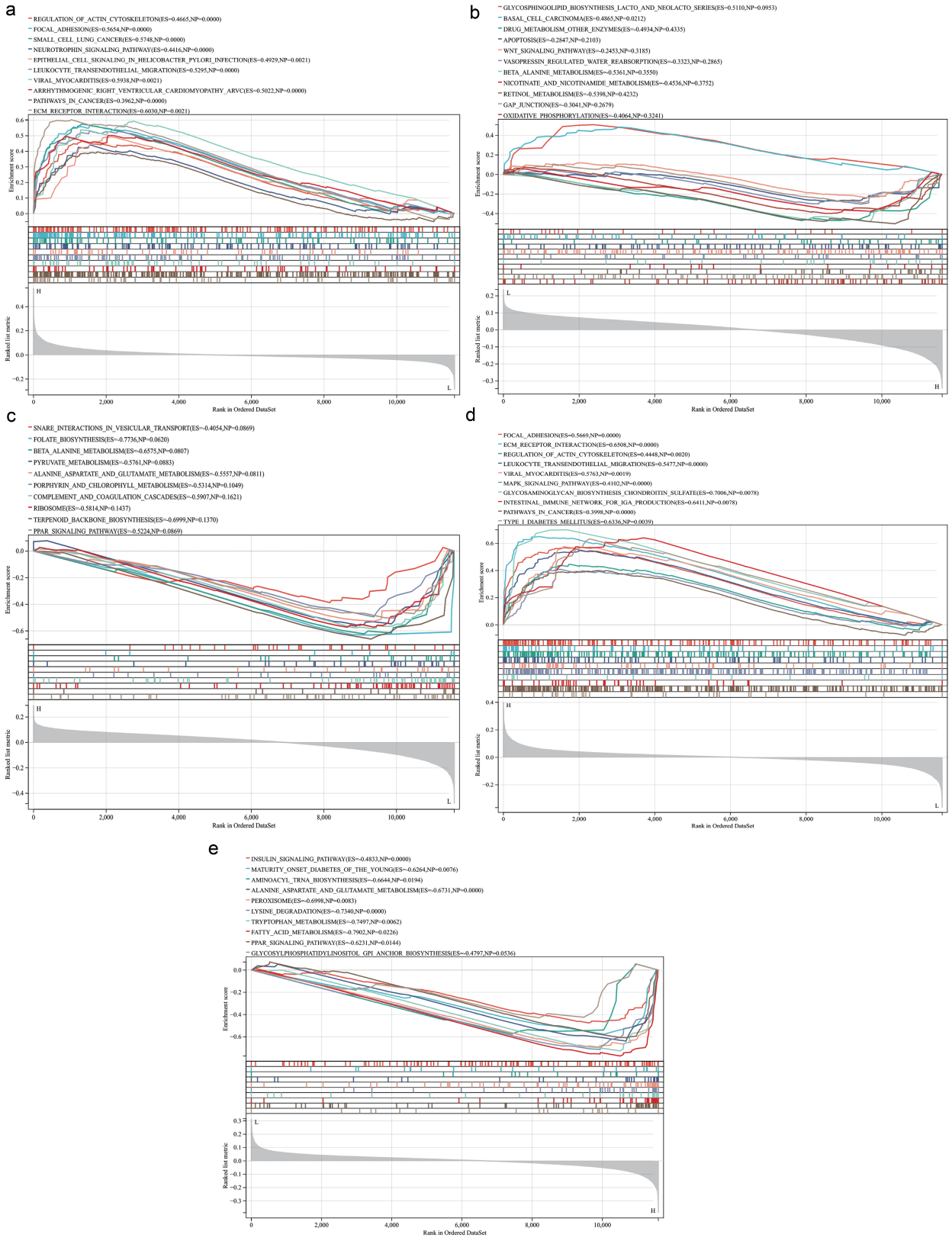


**Fig. 10. Gene Set Enrichment Analysis (GSEA).** (a) GSEA results for FABP4; (b) GSEA results for MDM2; (c) GSEA results for AKR1B1; (d) GSEA results for PDGFRB; (e) GSEA results for NR1H4.

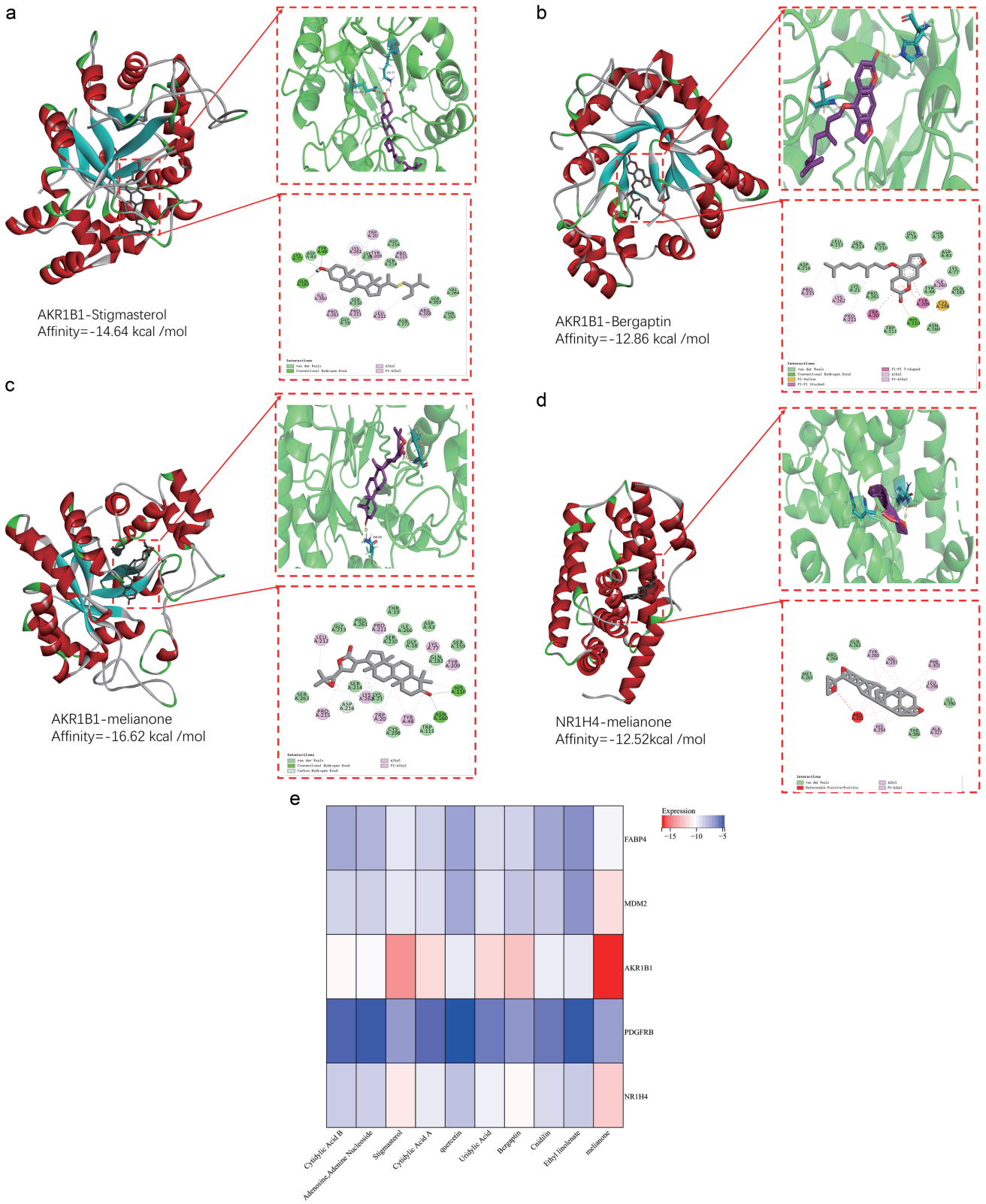


**Fig. 11. Molecular docking of core active ingredients with target molecules.** (a) Optimal conformation of AKR1B1–Stigmasterol docking; (b) Optimal conformation of AKR1B1–Bergaptin docking; (c) Optimized conformation of AKR1B1–Melianone docking; (d) Optimal conformation of NR1H4–Melianone docking; (e) Molecular docking results heatmap for core active ingredients with FABP4, MDM2, AKR1B1, PDGFRB, and NR1H4.

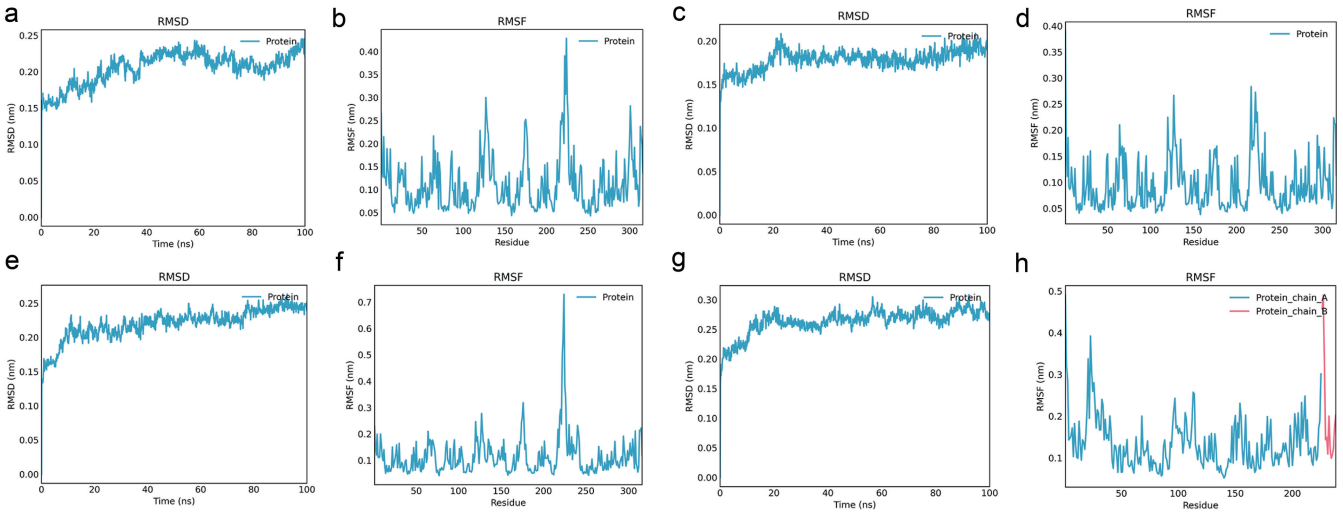


**Fig. 12. Molecular dynamics simulation results.** (a) Root-mean-square deviation (RMSD) curve for AKR1B1–Stigmasterol; (b) Root-mean-square fluctuation (RMSF) curve for AKR1B1–Stigmasterol; (c) RMSD curve for AKR1B1–Bergaptin; (d) RMSF curve for AKR1B1–Bergaptin; (e) RMSD curve for AKR1B1–Melianone; (f) RMSF curve for AKR1B1–Melianone; (g) RMSD curve for NR1H4–Melianone; (h) RMSF curve for NR1H4–Melianone.

by downregulating or antagonizing these targets. Functional analyses further revealed significant interactions between these targets and immune-related pathways. These findings offer mechanistic insights and potential strategies for targeted intervention, although direct experimental validation is still needed to confirm these in silico predictions.

## Acknowledgments

We thank the Guangxi University of Chinese Medicine (Guangxi, China) for providing laboratory equipment and technical support.

## Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 82204755), the Key Laboratory of Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine (No. 02: Basic Research on the Prevention and Treatment of Chronic Liver Diseases with Chinese Medicine and Zhuang-Yao Medicine), the Guangxi Natural Science Foundation (Grant Nos. 2023GXNSFBA026274 and 2024GXNSFAA010235), the Guangxi Zhuangyao Pharmaceutical Key Laboratory (Grant No. GXZYYKF2023-05), the Guangxi University of Traditional Chinese Medicine School-level Project Youth Fund (Grant No. 2022QN008), the Faculty of Chinese Medicine Science, Guangxi University of Chinese Medicine Research Project (Grant Nos. 2024ZZA002, 2024ZZA003, 2024ZZB006, 2024ZZB007, and 2024ZZB010), and the National Student Innovation and Entrepreneurship Training Programme of the Faculty of Chinese Medicine Science at the Guangxi University of Chinese Medicine (Grant No. 202513643017).

## Conflict of interest

The authors declare that there are no conflicts of interest.

## Author contributions

Study concept and design (YKJ, JHW), acquisition of data (LW), analysis and interpretation of data (JHW, YZ), drafting of the manuscript (YKJ), critical revision of the manuscript for important intellectual content (TJZ), administrative, technical, or material support (HYX), and study supervision (RWZ, HYX). All authors have made significant contributions to this study and have approved the final manuscript.

## Ethical statement

Not applicable.

## Data sharing statement

The data used and analyzed in this study are included within the article.

## References

[1] Bataller R, Brenner DA. Liver fibrosis. J Clin Invest 2005;115(2):209–218. doi:10.1172/JCI24282, PMID:15690074.
[2] Trautwein C, Friedman SL, Schuppan D, Pinzani M. Hepatic fibrosis: Concept to treatment. J Hepatol 2015;62(1 Suppl):S15–S24. doi:10.1016/j.jhep.2015.02.039, PMID:25920084.
[3] Moon AM, Singal AG, Tapper EB. Contemporary Epidemiology of Chronic Liver Disease and Cirrhosis. Clin Gastroenterol Hepatol 2020;18(12):2650–2666. doi:10.1016/j.cgh.2019.07.060, PMID:31401364.
[4] Wu Z, Wang W, Zhang K, Fan M, Lin R. Trends in the incidence of cirrhosis in global from 1990 to 2019: A joinpoint and age-period-cohort analysis. J Med Virol 2023;95(6):e28858. doi:10.1002/jmv.28858, PMID:37306296.
[5] Liaw YF. Reversal of cirrhosis: an achievable goal of hepatitis B antiviral therapy. J Hepatol 2013;59(4):880–881. doi:10.1016/j.jhep.2013.05.007, PMID:23673137.
[6] Friedman SL, Pinzani M. Hepatic fibrosis 2022: Unmet needs and a blueprint for the future. Hepatology 2022;75(2):473–488. doi:10.1002/hep.32285, PMID:34923653.
[7] Li Y, Lu Y, Nian M, Sheng Q, Zhang C, Han C, *et al*. Therapeutic potential and mechanism of Chinese herbal medicines in treating fibrotic liver disease. Chin J Nat Med 2023;21(9):643–657. doi:10.1016/S1875-5364(23)60443-1, PMID:37777315.

[8] Tan CY, Liu P. Theoretical Origin of Pathogenesis on Accumulation due to Deficiency of Liver Cirrhosis and Clinical Significance (in Chinese). Acad J Shanghai Univ Trad Chin Med 2010;24(4):25–28.
[9] Mu YP, Liu CH, Zhang H. Theory of "Asthenia and Deficiency Cause Accumulation" in Liver Cirrhosis —Brief Analysis on Academic System of Professor Liu Ping (in Chinese). Acad J Shanghai Univ Trad Chin Med 2013;27(2):1–4. doi:10.16306/j.1008-861x.2013.02.006.
[10] Wang LF, Zhao HB, Shao M, Xue BY. Syndrome Identification and Treatment Administration of Liver Cirrhosis with Liver-Kidney Yin Deficiency (in Chinese). J Trad Chin Med 2012;53(15):1318–1320.
[11] Zhang R, Zhu X, Bai H, Ning K. Network Pharmacology Databases for Traditional Chinese Medicine: Review and Assessment. Front Pharmacol 2019;10:123. doi:10.3389/fphar.2019.00123, PMID:30846939.
[12] Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. Nat Rev Mol Cell Biol 2022;23(1):40–55. doi:10.1038/s41580-021-00407-0, PMID:34518686.
[13] Krishnan R, Rajpurkar P, Topol EJ. Self-supervised learning in medicine and healthcare. Nat Biomed Eng 2022;6(12):1346–1352. doi:10.1038/s41551-022-00914-1, PMID:35953649.
[14] Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review. Brief Bioinform 2022;23(2):bbab569. doi:10.1093/bib/bbab569, PMID:35089332.
[15] Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. N Engl J Med 2023;388(13):1201–1208. doi:10.1056/NEJMra2302038, PMID:36988595.
[16] Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008;27(8):1133–1163. doi:10.1002/sim.3034, PMID:17886233.
[17] Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. Stat Methods Med Res 2007;16(4):309–330. doi:10.1177/0962280206077743, PMID:17715159.
[18] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47. doi:10.1093/nar/gkv007, PMID:25605792.
[19] 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, PMID:19114008.
[20] Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, *et al*. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol 2016;17(1):174. doi:10.1186/s13059-016-1028-7, PMID:27549193.
[21] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498–2504. doi:10.1101/gr.1239303, PMID:14597658.
[22] Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, *et al*. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018;7:e34408. doi:10.7554/eLife.34408, PMID:29846171.
[23] He Y, Li Y, Zhang S, Perry B, Zhao T, Wang Y, *et al*. Radicicol, a heat shock protein 90 inhibitor, inhibits differentiation and adipogenesis in 3T3-L1 preadipocytes. Biochem Biophys Res Commun 2013;436(2):169–174. doi:10.1016/j.bbrc.2013.05.068, PMID:23727383.
[24] Elkholi R, Abraham-Enachescu I, Trotta AP, Rubio-Patiño C, Mohammed JN, Luna-Vargas MPA, *et al*. MDM2 Integrates Cellular Respiration and Apoptotic Signaling through NDUFS1 and the Mitochondrial Network. Mol Cell 2019;74(3):452–465.e7. doi:10.1016/j.molcel.2019.02.012, PMID:30879903.
[25] Chen J, Zhang R, Li F, Lin S, Wang J. Integrated analysis and validation of TRIM23/p53 signaling pathway in hepatic stellate cells ferroptosis and liver fibrosis. Dig Liver Dis 2024;56(2):281–290. doi:10.1016/j.dld.2023.07.010, PMID:37495427.
[26] Zhao Q, Han B, Wang L, Wu J, Wang S, Ren Z, *et al*. AKR1B1-dependent fructose metabolism enhances malignancy of cancer cells. Cell Death Differ 2024;31(12):1611–1624. doi:10.1038/s41418-024-01393-4, PMID:39406918.
[27] Maleddu A, Pantaleo MA, Nannini M, Biasco G. The role of mutational analysis of KIT and PDGFRA in gastrointestinal stromal tumors in a clinical setting. J Transl Med 2011;9:75. doi:10.1186/1479-5876-9-75, PMID:21605429.
[28] Abifadel M, Varret M, Rabès JP, Allard D, Ouguerram K, Devillers M, *et al*. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet 2003;34(2):154–156. doi:10.1038/ng1161, PMID:12730697.
[29] Huang YK, Chang KC, Li CY, Lieu AS, Lin CL. AKR1B1 Represses Glioma Cell Proliferation through p38 MAPK-Mediated Bcl-2/BAX/Caspase-3 Apoptotic Signaling Pathways. Curr Issues Mol Biol 2023;45(4):3391–3405. doi:10.3390/cimb45040222, PMID:37185746.
[30] Yang C, Geng X, Wan G, Song L, Wang Y, Zhou G, *et al*. Transcriptomic and proteomic investigation of the ameliorative effect of total polyphenolic glycoside extract on hepatic fibrosis in Lamiophlomis rotata Kudo via the AGE/RAGE pathway. J Ethnopharmacol 2024;324:117720. doi:10.1016/j.jep.2024.117720, PMID:38211823.
[31] Wang F, Chen L, Kong D, Zhang X, Xia S, Liang B, *et al*. Canonical Wnt signaling promotes HSC glycolysis and liver fibrosis through an LDH-A/HIF-1α transcriptional complex. Hepatology 2024;79(3):606–623. doi:10.1097/HEP.0000000000000569, PMID:37733267.
[32] Guo C, Lai L, Ma B, Huang Q, Wang Z. Notoginsenoside R1 targets PPAR-γ to inhibit hepatic stellate cell activation and ameliorates liver fibrosis. Exp Cell Res 2024;437(1):113992. doi:10.1016/j.yexcr.2024.113992, PMID:38492634.
[33] Mejias M, Gallego J, Naranjo-Suarez S, Ramirez M, Pell N, Manzano A, *et al*. CPEB4 Increases Expression of PFKFB3 to Induce Glycolysis and Activate Mouse and Human Hepatic Stellate Cells, Promoting Liver Fibrosis. Gastroenterology 2020;159(1):273–288. doi:10.1053/j.gastro.2020.03.008, PMID:32169429.
[34] Larsson SC, Butterworth AS, Burgess S. Mendelian randomization for cardiovascular diseases: principles and applications. Eur Heart J 2023;44(47):4913–4924. doi:10.1093/eurheartj/ehad736, PMID:37935836.