


RESEARCH

Identification of tryptophan metabolism-related biomarkers for nonalcoholic fatty liver disease through network analysis

Cuihua Jiang^{1,*}, Jianqi Liang^{2,*}, Kaibo Hu², Yanqing Ye³, Jiajia Yang⁴, Xiaozhi Zhang⁴, Guilin Ye⁴, Jing Zhang⁵, Deju Zhang⁶, Bin Zhong⁷, Peng Yu², Liefeng Wang^{4,8} and Bin Zeng^{4,8} 

¹Department of Pain Management, The Affiliated Ganzhou Hospital of Nanchang University, Ganzhou, China

²Department of Endocrinology and Metabolism, The Second Affiliated Hospital, Jiangxi Medical College, Nanchang University, Nanchang, China

³Department of Gastroenterology, First Affiliated Hospital of Gannan Medical University, Ganzhou, China

⁴School of Basic Medicine, Gannan Medical University, Ganzhou, China

⁵Department of Anesthesiology, The Second Affiliated Hospital, Jiangxi Medical College, Nanchang University, Nanchang, China

⁶Food and Nutritional Sciences, School of Biological Sciences, The University of Hong Kong, Hong Kong, China

⁷Department of Pharmacy, First Affiliated Hospital of Gannan Medical University, Ganzhou, China

⁸China Medical University, Shenyang, China

Correspondence should be addressed to P Yu: zhangyuteam2022@163.com or to L Wang: wlf0709@gmu.edu.cn or to B Zeng: fyzengb2020@163.com

*C Jiang and J Liang contributed equally to this work

Abstract

Background: Increasing evidence demonstrates that tryptophan metabolism is closely related to the development of nonalcoholic fatty liver disease (NAFLD). This study aimed to identify specific biomarkers of NAFLD associated with tryptophan metabolism and research its functional mechanism.

Methods: We downloaded NAFLD RNA-sequencing data from GSE89632 and GSE24807, and obtained tryptophan metabolism-related genes (TMRGs) from the MsigDB database. The R package limma and WGCNA were used to identify TMRGs–DEGs, and GO, KEGG and Cytoscape were used to analyze and visualize the data. Immune cell infiltration analysis was used to explore the immune mechanism of NAFLD and the biomarkers. We also validated extended levels of biomarkers.

Results: We identified 375 NAFLD differentially expressed genes (DEGs) and 85 TMRGs–DEGs. GO/KEGG analysis revealed that TMRGs–DEGs were mainly enriched in triglyceride and cholesterol metabolism. ROC curves identified CCL20 (AUC = 0.917), CD160 (AUC = 0.933) and CYP7A1 (AUC = 1) as biomarkers of NAFLD. Immune infiltration analysis showed significant differences in ten immune cells, and the activation of dendritic cells and mast cells were highly positively correlated with NAFLD. CCL20, CD160 and CYP7A1 were highly correlated with M2 macrophage, neutrophil and mast cells activation, respectively. Twenty-seven TMRGs correlated with hub genes, and gene set enrichment analysis demonstrated their function in tryptophan- and lysine-containing metabolic process. We identified 41 therapeutic drug matches which corresponded to two hub genes and four drugs which co-targeted CCL20 and CYP7A1. Finally, three hub genes were validated in our mouse model.

Conclusions: CCL20, CD160 and CYP7A1 are tryptophan metabolism-related biomarkers of NAFLD, related to glycerol ester and cholesterol metabolism. We screened four compounds which co-target CCL29 and CYP7A1 to provide potential experimental drugs for NAFLD.

Keywords: nonalcoholic fatty liver disease; tryptophan metabolism; biomarkers; machine learning; network analysis

Introduction

Nonalcoholic fatty liver disease (NAFLD) is a heterogeneous disease with a wide spectrum of pathologies, including nonalcoholic fatty liver (NAFL) characterized by simple steatosis (SS), and nonalcoholic steatohepatitis characterized by hepatic steatosis with inflammation and hepatocyte ballooning, which in advanced stages gradually develops into fibrosis, cirrhosis and hepatocellular carcinoma (1). At present, NAFLD is considered as the most common cause of chronic liver disease (CLD) worldwide, and the leading cause of morbidity and mortality of diseases associated with liver (2, 3). Non-alcoholic steatohepatitis (NASH) was the fastest growing cause of liver cancer deaths globally (4). The prevalence of NAFLD was estimated to be approximately 25% in 2018, and has been increasing in parallel with obesity and type 2 diabetes (5, 6). According to a systematic review and meta-analysis, the total global prevalence of NAFLD up to 2021 was estimated to be 32.4% (95% CI 29.9–34.9) (7). It is predicted that the prevalence of NAFLD in 2040 will be 55.7%, which means more than half of the adult population will suffer from this disease (3). In addition, the global burden of NAFLD is forecasted to continue to increase in the coming decades (8).

Furthermore, NAFLD has emerged as a multisystem disease and is tightly linked to the risks of progressing to extrahepatic complications, such as cardiovascular diseases, certain types of extrahepatic cancer and type 2 diabetes mellitus, which are also the most frequent causes of death in patients with NAFLD (9, 10). However, to date, the treatment of NAFLD has been based on lifestyle modifications with poor compliance and there are still no FDA-approved drug therapies (11, 12). As a consequence, it is vital to research molecular mechanisms, and identify specific biomarkers for the diagnosis and treatment of this disease.

At present, the pathogenesis of NAFLD is complex and remains unclear (13). The development and progression of NAFLD are recognized as a multiple-hit theory, in which the interaction between various factors, such as insulin resistance, adipose tissue dysfunction, diet, gut microbiota and genetic and epigenetic inheritance, leads to lipid accumulation, inflammation and fibrosis (14). For the past few years, increasing evidence has revealed the key role of gut microbiota and their metabolites, including bile acids, tryptophan breakdown products and branched-chain amino acids, in the pathogenesis of NAFLD (15, 16). Tryptophan is an essential amino acid and participates in protein synthesis in the body. In addition, the catabolism of tryptophan *in vivo* consists of three pathways, including the kynurenine pathway, the indole pathway and the serotonin pathway (17). About 95% of ingested tryptophan is metabolized through the kynurenine pathway, which is mediated by tryptophan 2,3-dioxygenase (TDO), indoleamine 2,3-dioxygenase 1 (IDO1) and indoleamine 2,3-dioxygenase 2 (IDO2), resulting in the generation of a variety of biologically active substances, such as

kynurenine, kynurenic acid (KA) and 3-hydroxyanthranilic acid (3HAA) (18). The indole pathway conducts about 4–6% of tryptophan metabolism, regulated by gut microbiota through the enzyme tryptophanase, which is produced by various Gram-negative and Gram-positive intestinal bacteria, transforming tryptophan into indole and various derivatives, including indoleacetic acid (IAA), indoleacrylic acid and indolepropionic acid (IPA) (18, 19). The serotonin pathway leads to the production of 5-hydroxytryptamine and melatonin controlled by tryptophan hydroxylase 1 (TPH1), in which approximately 1–2% of ingested tryptophan is degraded (18).

Although inflammatory molecules upregulate the kynurenine pathway through IDO1, the effect of IDO1 on the development of NAFLD is still unknown (20, 21). KA alleviated liver steatosis through inhibiting endoplasmic reticulum stress mediated by AMPK/autophagy and AMPK/ORP150 signaling in a NAFLD murine model fed with a high-fat diet (HFD) (22). Indole is an advantageous signal for strengthening the intestinal epithelial cell barrier, attenuates the activation of NF- κ B signaling induced by TNF- α , inhibits the expression of the proinflammatory chemokine IL-8 and elevates the production of the anti-inflammatory cytokine IL-10 (23). IAA exerts a hepatoprotective effect through attenuating insulin resistance and oxidative stress, reducing the infiltration of F4/80+ macrophages and decreasing the level of inflammation in NAFLD mice induced by a HFD (24). IPA can mitigate liver inflammation and injury by promoting intestinal barrier integrity with increased expression of tight junction proteins, suppressing the TLR4/NF- κ B pathway, decreasing the levels of proinflammatory molecules, and inhibiting the expression of genes related to fibrosis and collagen synthesis in HFD-induced steatohepatitis rats (25). Serotonin promotes the production of reactive oxygen species, which induces mitochondrial dysfunction leading to hepatocellular injury in a model of NASH established by a choline- and methionine-deficient diet (26). Melatonin exerts liver protection in NAFLD according to several research studies (27). Due to the strong association between tryptophan metabolites and the development of NAFLD, identifying targets related to tryptophan metabolism may provide novel diagnostic and therapeutic methods for NAFLD.

In this study, we downloaded two NAFLD datasets from the Gene Expression Omnibus (GEO) database and identified differentially expressed genes (DEGs) using the limma package. Tryptophan metabolism-related genes (TMRGs) were obtained from the Molecular Signatures Database (MSigDB) and then used for screening key module genes through weighted gene co-expression network analysis (WGCNA). Subsequently, via the application of bioinformatic methods, including functional enrichment analysis, protein–protein interaction (PPI) network construction,

machine learning (least absolute shrinkage and selection operator (LASSO) and support vector machine-recursive feature elimination (SVM-RFE)) and receiver operating characteristic (ROC) curves and nomogram assessment, the crucial biomarkers related to TMRGs in NAFLD were identified. Next, immune cell infiltration analysis and gene set enrichment analysis (GSEA) were performed to demonstrate differences in immune cell distribution between NAFLD and control groups, and explore functions associated with biomarkers. We also analyzed diseases related to these genes, projected potential targeted therapeutic drugs, built a competing endogenous RNA (ceRNA) network and validated the expression levels of biomarkers. These biomarkers may bring new insights into the role of tryptophan metabolism in the pathogenesis of NAFLD, and provide potential targets for early diagnosis and drug treatment.

Materials and methods

Data collection

Two NAFLD datasets (GSE89632 and GSE24807) encompassing gene expression data were acquired from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). GSE89632 contains 39 NAFLD patients and 24 controls as the training set for analysis. GSE24807 contains 12 NAFLD patients and 5 controls, used for ROC curve and gene expression level evaluation. TMRGs were collected from the Molecular Signatures Database (MSigDB, <http://www.gsea-msigdb.org/gsea/msigdb>) through the keywords 'tryptophan metabolism' and 'tryptophan metabolic process', and after removing duplication, 47 genes were included.

DEG selection in NAFLD samples and controls

DEGs were identified with the use of the 'limma' R package, and the screening criterion of $|\log_2 FC| > 1$ and $P_{\text{adjust}} < 0.05$ between the NAFLD group and the control group in GSE89632, then visualized by volcano plot using the 'ggplot2' package. A heat map of the 20 most significant upregulated and downregulated genes was obtained through applying the 'pheatmap' package.

WGCNA and screening of TMRG-related module genes

A gene co-expression network was established to select the module genes related to TMRGs by using the 'WGCNA' package in R. First, using the 'ssGSEA' algorithm based on the 'GSVA' package, the scores of TMRGs in the dataset GSE89632 were calculated with TMRGs as the background gene set. Second, cluster analysis was performed on the samples to remove the outliers and ensure the accuracy of later analysis. Third, by taking the expression matrix of all genes as input data to compute, the most suitable soft

threshold was set to ensure that the interaction between genes conformed to a scale-free distribution. Next, adjacency was calculated based on the optimal soft threshold, the intergenic similarity was determined according to the adjacency, the coefficient of dissimilarity between genes was derived, and the systematic clustering tree between genes was obtained. Fourth, the gene co-expression modules were separated through the hybrid dynamic tree cutting algorithm; the minimum number of genes in each module criterion was 150. Finally, Pearson correlation analysis was performed between the modules and two characteristics, including disease status (yes or no) and TMRG scores, and the key module genes with highest relevance and statistical significance ($P < 0.05$) were selected for further analysis.

Identification of TMRGs–DEGs and functional enrichment analysis

The intersection genes of the key module genes related to TMRGs and DEGs for NAFLD were identified via the 'eulerr' package, and defined as TMRGs–DEGs. Functional enrichment analysis of TMRGs–DEGs, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG), was conducted through the use of the 'clusterProfiler' package in R software, with a P value criterion of < 0.05 . The results are shown through a lollipop chart and a chord diagram using the R 'ggpubr' and 'Goplot' packages, respectively.

Construction of a PPI network

In order to investigate the interactions between TMRGs and DEGs, a PPI network was constructed by applying the STRING database (<http://string.embl.de/>) and setting a confidence score greater than 0.4. Genes that had interactions with others were acquired and visualized with the utilization of the Cytoscape software, and screened out for the following analysis.

Machine learning for selecting candidate hub genes

To screen crucial biomarkers related to NAFLD TMRGs, two machine learning algorithms were utilized. LASSO is a method with increased prediction accuracy through variable selection and regularization. LASSO regression analysis was conducted with 3-fold cross-validation by the 'glmnet' package in R software. SVM-RFE, an algorithm based on SVM, was applied to select candidate genes using the 'e1071' package through the process in which, first, the importance and ranking of each gene in TMRGs–DEGs and the error and accuracy rate of each iteration combination were obtained, then the combination with the lowest error rate was selected as the best combination, and finally, the genes contained in the best combination were taken out. The intersection of genes screened by the two machine learning algorithms

was shown by a Venn diagram. In addition, intersection genes were determined as candidate hub genes.

ROC curve evaluation and nomogram construction

ROC curves were drawn with the application of the 'pROC' package in the GSE89632 NAFLD training dataset and the GSE24807 NAFLD external validation dataset to evaluate the diagnostic value of candidate hub genes. The area under the curve (AUC) was calculated and set to be greater than 0.8 to acquire ideal diagnostic prediction accuracy. In addition, the 'rms' package in R software was utilized to plot nomograms based on hub genes validated by the ROC curves. Scores were given according to the expression levels of genes, and all the scores of the above genes were added to be the total score, which was used to predict the diagnostic probability of NAFLD. The higher the total score, the greater the probability of NAFLD diagnosis. Then, based on the nomogram prediction model, the corresponding calibration curve was conducted to assess the predictive ability of the nomogram. In addition, a decision curve analysis (DCA) curve was constructed to integrate clinical utility in the analysis.

GSEA and correlation analysis

To further explore the functions of hub genes, GSEA was performed through the 'clusterProfiler' package based on the background gene sets of BP, CC, MF and KEGG in MSigDB. Moreover, Spearman correlation analysis was conducted between the expression level of every hub gene and all genes using the 'psych' package, and the correlation coefficient was used as the ranking standard. The screening criterion of single-gene GSEA was $|NES| > 1$ and $P_{\text{adjust}} < 0.05$.

Immune cell infiltration analysis and correlation analysis

The proportion of 22 types of immune cells in NAFLD samples and control samples was calculated with the application of the CIBERSORT algorithm (version 1.03), and the gene set of immune cell LM22. The results were shown using a bar graph. Then, a boxplot was drawn to show the infiltration proportion of 22 types of immune cells between the control and NAFLD groups, and the Wilcoxon rank sum test was performed. Subsequently, Spearman correlation calculations were conducted for the immune cells, for which significant differences in infiltration proportion were observed using the 'psych' package in R software, and a heat map was plotted to display the correlation between immune cells using the 'ggcor' package. Next, correlations between hub genes and infiltrating immune cells with significant differences were calculated by Spearman correlation analyses and visualized by the lollipop chart applying the 'ggplot2' package.

Analysis of hub gene-related diseases

The CTD database (<http://ctdbase.org>), which is publicly available, was utilized to determine the co-expression scores of liver and cardiovascular diseases (28) associated with hub genes.

Competing endogenous RNA regulatory network construction and functional analysis

The starBase (<http://starbase.sysu.edu.cn/>) and the TargetScan (<http://www.targetscan.org>) databases were applied to screen for miRNAs related to hub genes. In addition, miRNAs screened from the starBase and TargetScan databases were intersected, and utilized to predict corresponding lncRNAs in the starBase database with the selection criterion of clipExpNum > 15 . Then, lncRNA-miRNA-mRNA networks were established. Pearson correlation analysis was performed to analyze the co-expression relationship between TMRGs and hub genes with the threshold set as $|r| > 0.3$ and $P < 0.05$. In addition, functional enrichment analysis of TMRGs and hub genes was conducted using Metascape as an annotation tool.

Prediction of drug targeting hub genes

Therapeutic drug targeting hub genes were obtained from the CTD database with the screening condition of reference count > 2 , and the interaction network between targeted drugs and hub genes was visualized through Cytoscape software.

Validation of expression of hub genes

The differences in expression levels of hub genes were visualized and compared between control groups and NAFLD groups in the GSE89632 training dataset and the GSE24807 external validating dataset. Furthermore, t-tests were performed between different groups. Twelve male C57 mice (6–7 weeks old) were purchased from the Laboratory Animal Center of Nanchang University (Nanchang, China). After purchasing, the mice were adaptively reared for 7 days and then six mice were fed with a HFD (basic diet + supplementary ingredients (2% cholesterol + 10% lard + 0.2% propylthiouracil + 0.5% sodium cholate)) for 8 weeks, and after 14 days of high-fat feeding, intraperitoneal injection of 5% CCl₄ edible oil was performed, two times a week. The other six mice were fed with a normal diet. Oil Red O and hematoxylin-eosin (H&E) staining of liver tissue sections was performed. Moreover, RT-qPCR was performed for expression validation at the animal level; primers and sequences are provided in Supplementary Table 1 (see section on [Supplementary materials](#) given at the end of the article).

Statistical analysis

R software was applied for all analyses related to bioinformatic methods. The Wilcoxon rank sum test was performed to compare the immune cell infiltration proportion, and Student’s *t*-tests were conducted to make a comparison with the different expression of hub genes between the control group and the NAFLD group. Pearson correlation analysis was utilized to investigate the correlations between different variables. *P*-value <0.05 was regarded as having statistical significance.

Results

Identification of NAFLD DEGs

In total, 375 DEGs were identified by the ‘limma’ package between the control group and the NAFLD group in the

GSE89632 dataset, encompassing 126 upregulated genes and 249 downregulated genes (Fig. 1A). NAFLD DEGs are depicted in a volcano plot, and the top ten most significant upregulated and downregulated DEGs are shown by heat map (Fig. 1B).

Identification of key module genes related to TMRGs through WGCNA

WGCNA was performed to screen the crucial module genes related to TMRGs. The optimal soft threshold was determined as $\beta = 15$ (scale-free R2 = 0.9) to ensure the interaction network of genes in accord with a scale-free distribution to the maximum (Fig. 1C and D). The result of a clustering analysis of samples is shown by a clustering dendrogram with two outlier samples, with GSM2385767 and GSM2385782 removed. Through the utilization of a hybrid dynamic tree cutting algorithm, ten gene modules were identified and are represented in different colors.

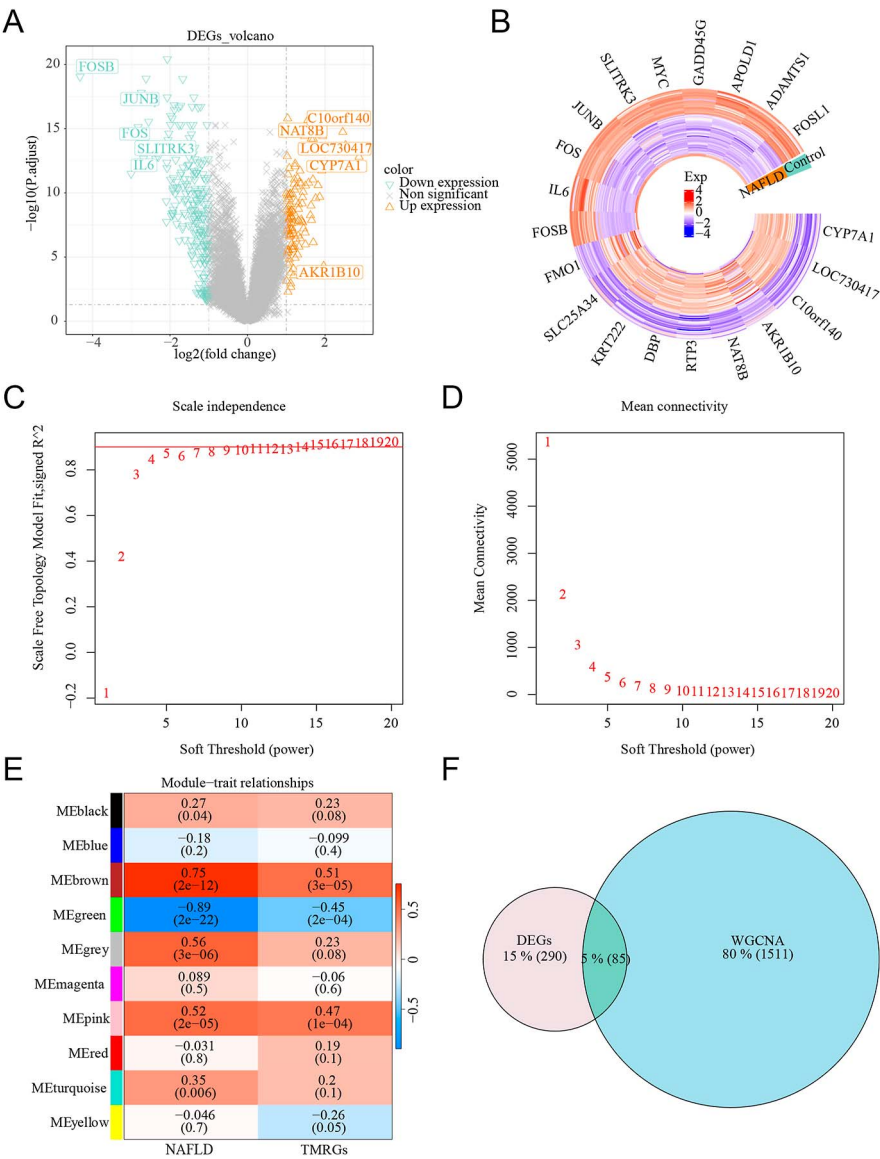


Figure 1 Identification of TMRGs-DEGs in NAFLD. (A) A volcano map indicating DEGs in the control and NAFLD groups. (B) A circular heat map depicting expression of the top ten overexpressed DEGs and top ten underexpressed DEGs. (C & D) Selection of the soft threshold for WGCNA. (E) Heat map illustrating correlation coefficients and *P* values of each module in the WGCNA results. (F) Venn diagram revealing final TMRG-DEG selection.

The relevance between the gene modules and two traits encompassing disease state (yes or no) and TMRG scores was obtained by the heat map, and the brown module, which contains 1,596 genes and has the highest correlation with NAFLD (correlation coefficient = 0.75, $p = 2 \times 10^{-12}$) and TMRGs (correlation coefficient = 0.51, $p = 3 \times 10^{-5}$), was considered as the key module for subsequent analysis (Fig. 1E).

Identification of TMRGs–DEGs and functional enrichment analysis

A total of 85 genes obtained from the intersection of 375 DEGs and the 1,596 module genes were regarded as TMRGs–DEGs and are shown by a Venn plot (Fig. 1F). Next, functional enrichment analyses including GO and KEGG were conducted to explore and understand the biological functions and related pathways for the intersection genes. GO enrichment analysis indicated

that a total of 244 items were significantly enriched, including 233 items of biological processes (BP), 8 items of cellular components (CC) and 13 items of molecular functions (MF). ‘Positive regulation of cell adhesion’, ‘glycerolipid biosynthetic process’, ‘apical plasma membrane’, ‘mitotic spindle’ and ‘O-acyltransferase activity’ were the items in which the intersection genes were mainly enriched. The top eight items of BP, CC and MF ordered by *P*-value from small to large are visualized by a lollipop graph, and the top eight items overall are shown in a chord diagram. KEGG analysis demonstrated that the intersection genes were chiefly enriched in seven pathways, including ‘glycerolipid metabolism’, ‘PD-L1 expression and PD-1 checkpoint pathway in cancer’, ‘fat digestion and absorption’, ‘cholesterol metabolism’, ‘cytokine–cytokine receptor interaction’, ‘bile secretion’ and the ‘IL-17 signaling pathway’ (Fig. 2A, B, C).

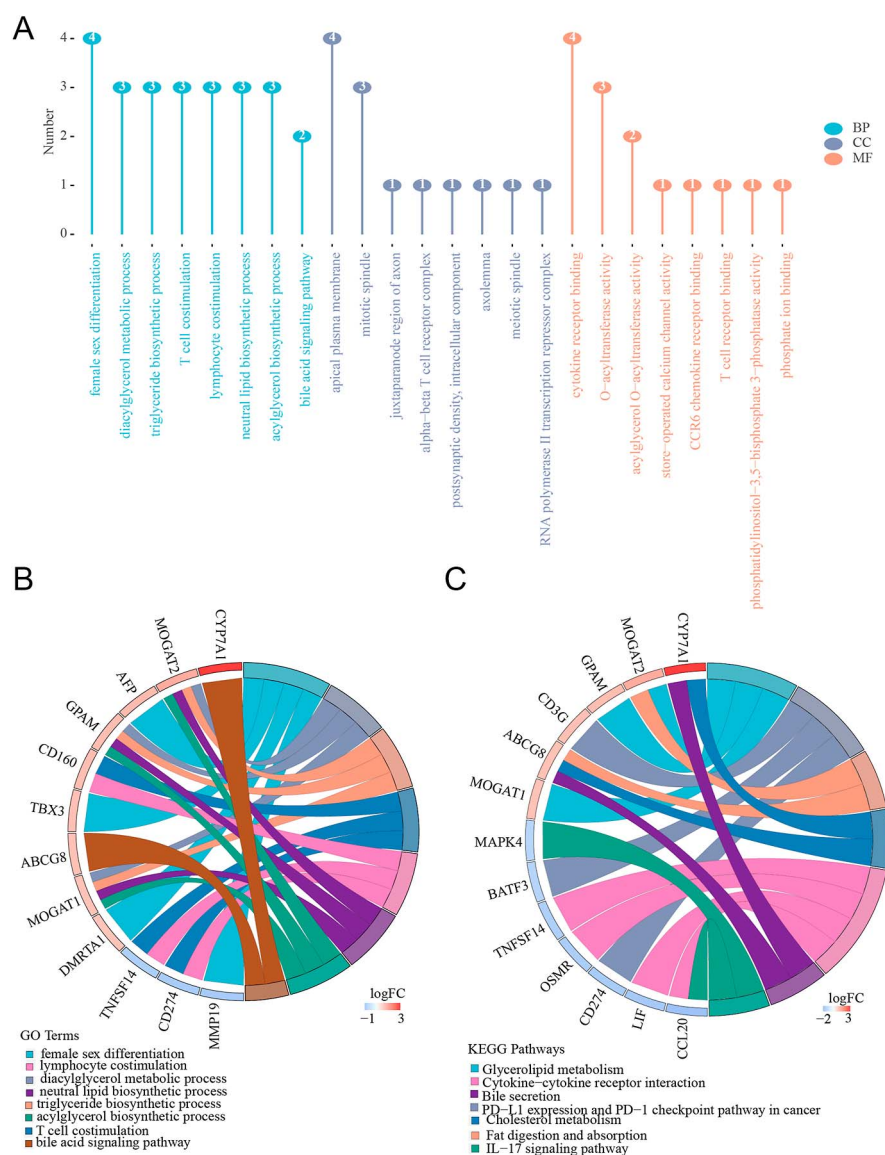
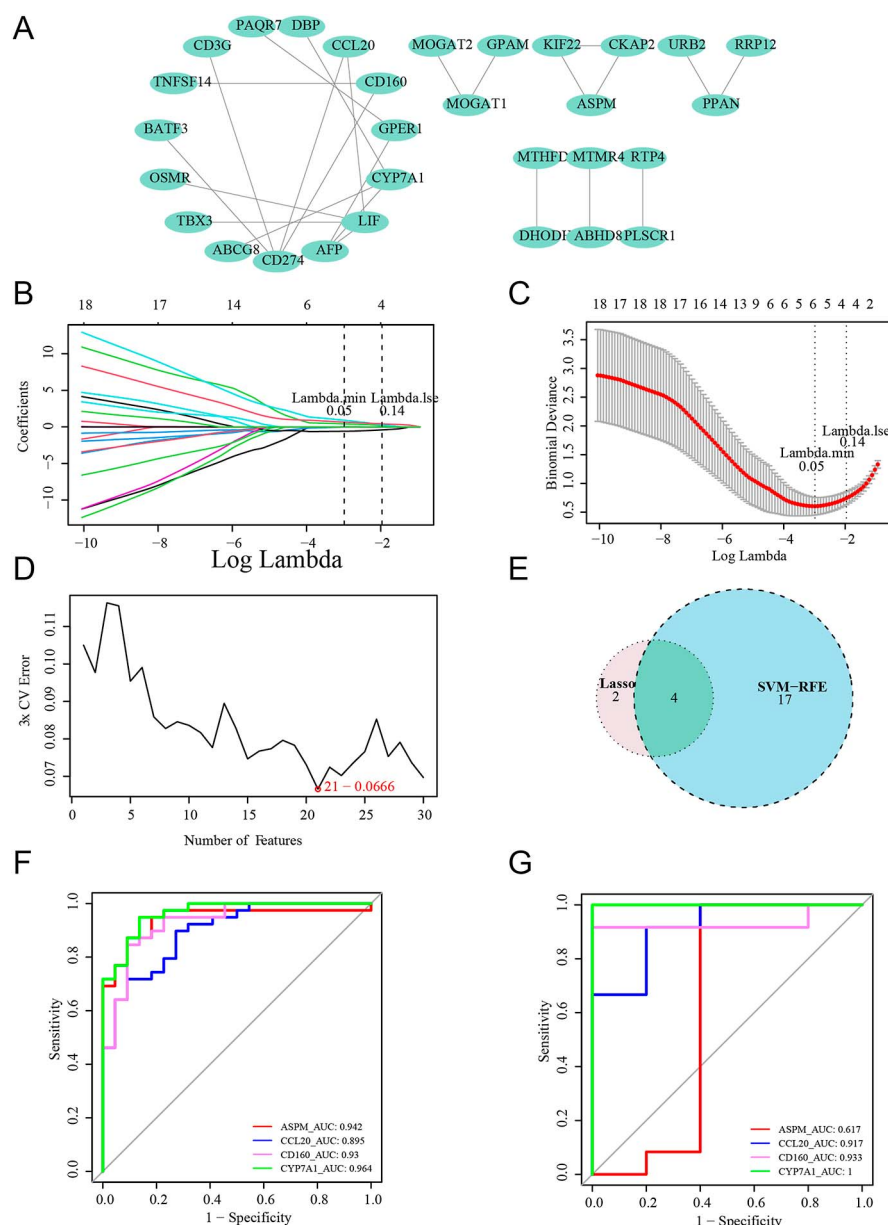


Figure 2

Functional background of the TMRGs–DEGs. (A) Lollipop chart of the general results of the GO enrichment analysis. (B & C) Chord diagram of several significant pathways and the corresponding genes in the results of (B) GO and (C) KEGG enrichment analysis.

**Figure 3**

Selection of the hub genes. (A) PPI network of the TMRGs-DEGs. (B & C) Results of selecting hub genes via LASSO regression. (D) Results of selecting hub genes via SVM-RFE. (E) Venn diagram revealing the hub gene final selection. (F & G) ROC curve demonstrating the diagnostic role of the four hub genes in the (F) training set and (G) validation set.

Construction of PPI network

A PPI network was established to explore the interaction relationship between TMRGs and DEGs. Thirty genes that have interaction with other genes were screened out of the following analysis (Fig. 3A).

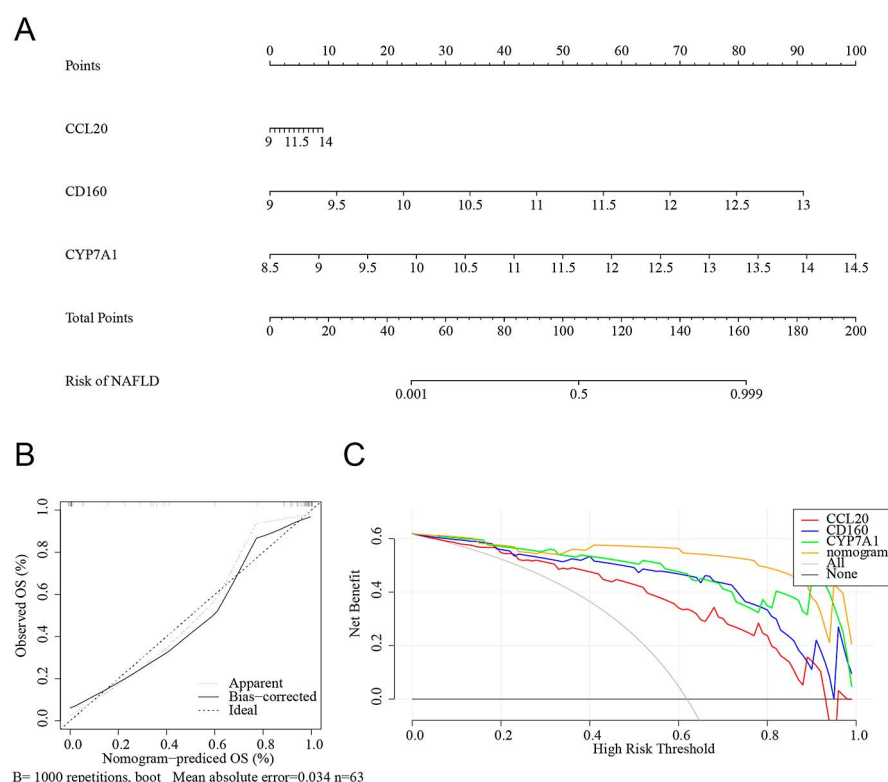
Screening of candidate hub genes through machine learning

Two machine learning algorithms, including LASSO and SVM-RFE, were utilized to identify the candidate hub genes. Six candidate genes were obtained via LASSO regression analysis: ASPM, CCL20, CD160, CYP7A1, PPA1 and RTP4 (Fig. 3B and C). As shown in the figure of SVM-RFE analysis, when the number of genes was 21,

the error rate was the lowest at 0.0666 and 21 genes were selected as candidate genes: CD160, CYP7A1, ABHD8, CKAP2, ASPM, RRP12, PLSCR1, URB2, GPAM, MOGAT1, CCL20, KIF22, PAQR7, OSMR, DHODH, CD3G, AFP, LIF, BATF3, ABCG8 and TBX3 (Fig. 3D). Then, four candidate hub genes (ASPM, CCL20, CD160 and CYP7A1) were determined by the intersection of the six genes from the LASSO analysis and the 21 genes from the SVM-RFE analysis (Fig. 3E).

Identification of hub genes related to TMRGs via ROC curves and nomogram construction

In order to evaluate the capability of the four candidate hub genes to distinguish NAFLD samples from control

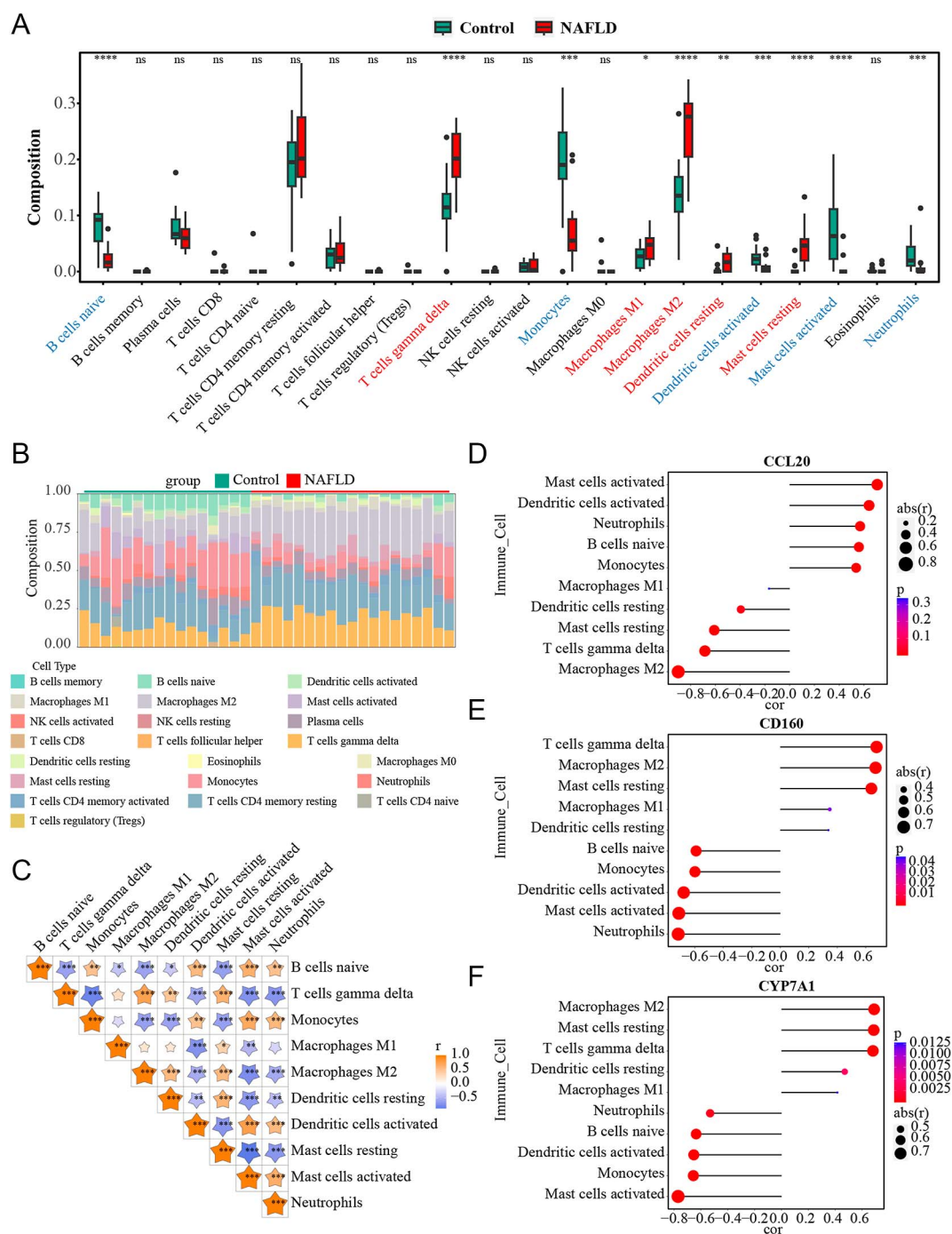
**Figure 4**

Construction of the diagnostic model via the four hub genes. (A) A nomogram consisting of four hub genes. (B) Calibration curve of the nomogram. (C) DCA curve of the nomogram and the four hub genes.

samples, a ROC curve was plotted in the training dataset; the results demonstrated that the AUCs of all four candidate hub genes were greater than 0.8. The details were as follows: ASPM (AUC: 0.942), CCL20 (AUC: 0.895), CD160 (AUC: 0.93) and CYP7A1 (AUC: 0.964) (Fig. 3F). The results of further ROC analysis for four genes in the external validation dataset indicated that the AUCs of three genes, CCL20, CD160 and CYP7A1, were greater than 0.8. The details were as follows: ASPM (AUC: 0.617), CCL20 (AUC: 0.917), CD160 (AUC: 0.933) and CYP7A1 (AUC: 1) (Fig. 3G). As a consequence, CCL20, CD160 and CYP7A1 were determined as the hub genes, and defined as biomarkers for NAFLD related to TMRGs. Subsequently, a nomogram was drawn for the three biomarkers with regression equation $y = 56.392 + (0.198) \times \text{CCL20} + (2.751) \times \text{CD160} + (2.014) \times \text{CYP7A1}$ (Fig. 4A). In addition, the corresponding calibration curve was plotted based on the nomogram prediction model (Fig. 4B). The slope of the calibration curve was close to 1, which demonstrated that the accuracy of the nomogram model and the diagnostic value of the three biomarkers was good. Moreover, the DCA curve was plotted according to the three biomarkers and the nomogram model, a multi-factor logistic regression model (Fig. 4C). The results indicated that the benefit value of all three biomarkers was high and the benefit value of the nomogram model was higher than that of a single biomarker. The utilization of the DCA curve demonstrated the clinical feasibility of the nomogram model.

Immune cell infiltration analysis and correlation analysis

The results of the Wilcoxon rank sum test demonstrated that the infiltration proportion of ten immune cells in the control group and the NAFLD group was significantly different, including naive B cells, gamma delta T cells, monocytes, M1 macrophages, M2 macrophages, resting dendritic cells, activated dendritic cells, resting mast cells, activated mast cells and neutrophils. In comparison with the control group, the levels of gamma delta T cells, M1 macrophages, M2 macrophages, resting dendritic cells and resting mast cells were elevated, and the levels of naive B cells, monocytes, activated dendritic cells, activated mast cells and neutrophils were lower in the NAFLD group, which is shown by a boxplot (Fig. 5A). A bar graph shows the proportion of 22 types of immune cells in each sample of the control group and the NAFLD group (Fig. 5B). Correlations between ten immune cells are displayed by a heat map, which demonstrates the highest negative correlation between activated mast cells and resting mast cells ($r = -0.865$, $P < 0.001$), and the greatest positive correlation between activated dendritic cells and activated mast cells ($r = 0.752$, $P < 0.001$) (Fig. 5C). A lollipop chart depicts correlations between the expression levels of three hub genes and the ratios of ten immune cells. CCL20, CD160 and CYP7A1 had the highest correlation with M2 macrophages ($r = -0.899$, $P < 0.0001$), neutrophils ($r = -0.716$, $P < 0.0001$) and activated mast cells ($r = -0.770$, $P < 0.0001$), respectively (Fig. 5D, E, F).

**Figure 5**

Immune microenvironment mediated via the four hub genes in NAFLD. (A) Differences in infiltration of 22 immune cell types between control and NAFLD groups. (B) The general landscape depicts the proportion of each immune cell in the control and NAFLD groups. (C) A correlation heatmap for the 22 immune cells. (D–F) Lollipop chart indicating correlation between the infiltration score of immune cells and the expression of (D) CCL20, (E) CD160 and (F) CYP7A1.

Analysis of diseases related to three hub genes

Bar plots depict the co-expression scores of three hub genes with various liver and cardiovascular diseases, and demonstrates that all three hub genes target these diseases; moreover, the scores of CCL20 with heart diseases, CD160 with hypertension and CYP7A1 with fatty liver were the highest (Fig. 6A, B, C).

ceRNA regulatory network construction for three hub genes and functional analysis with TMRGs

A total of 20 miRNAs were selected by the intersection of 41 miRNAs in the starBase database and 1,103 miRNAs in the TargetScan database, based on the prediction of three hub genes. Then, 53 lncRNAs corresponding to 20 miRNAs were obtained from the starBase database. The ceRNA network, including 3 mRNAs, 20 miRNAs and 53 lncRNAs, was constructed as 'lncRNA-miRNA-mRNA' with 166 pairs of action relationships (Fig. 6D). In addition, 27 TMRGs were selected through Pearson correlation analysis according to the correlations of TMRGs with three hub genes. Functional analysis was performed based on 3 hub genes and 27 TMRGs, which demonstrated that these genes were enriched in 12 clusters, including 'tryptophan metabolism', 'indole-containing compound metabolic process' and 'lysine degradation'.

Drug prediction related to three hub genes

A total of 41 therapeutic drugs corresponding to two hub genes (CCL20 and CYP7A1) were obtained from the CTD database, and compounds with potential therapeutic effects associated with CD160 were discovered. The relationship between the 2 hub genes and 41 targeted drugs was visualized by a network diagram, which demonstrates that four compounds, including sodium arsenite, cyclosporine, benzo(a)pyrene and tetrachlorodibenzodioxin, were the common targeted drugs for CCL20 and CYP7A1 (Fig. 6E).

Validation of the expression of three hub genes

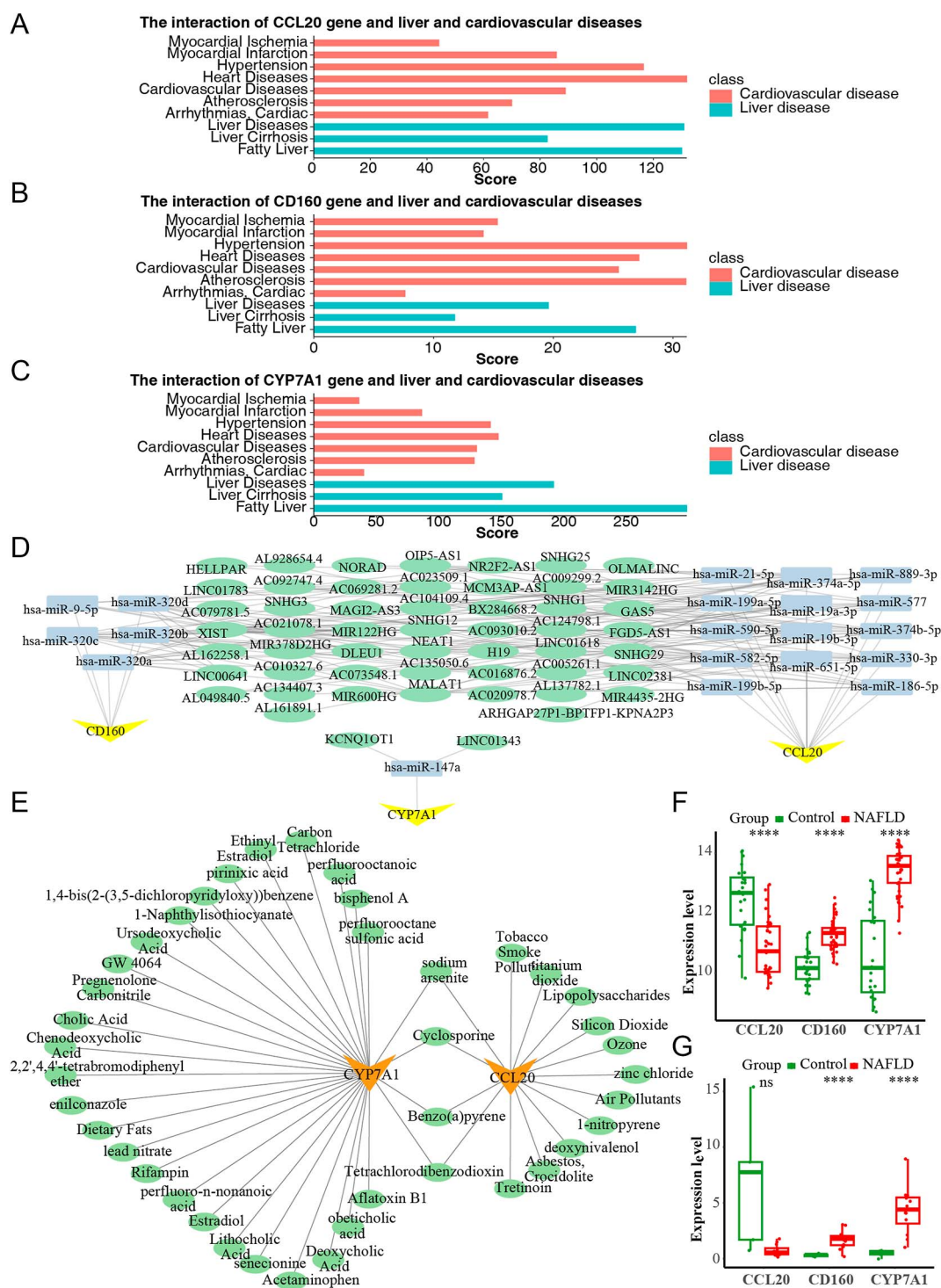
As shown in the figure, compared with the control group, the expression levels of CD160 and CYP7A1 in the NAFLD group were significantly elevated in both the training dataset (GSE89632) and the external validating dataset (GSE24807). In contrast, the expression of CCL20 in the NAFLD group was lower compared with the control group in the two NAFLD datasets; however, while the change in the training dataset was statistically significant, the change was not statistically significant in the external validating dataset (Fig. 6F and G). Moreover, we constructed a mouse model to verify our three hub genes at the animal level. HE staining and Oil Red O staining further revealed the successful establishment

of our NAFLD model (Fig. 7A). In addition, expression of the three hub genes was further validated via qPCR, in which CD160 and CYP7A1 exhibited the same significant trend in NAFLD, while the expression difference of CCL20 was insignificant and opposite (Fig. 7B and C).

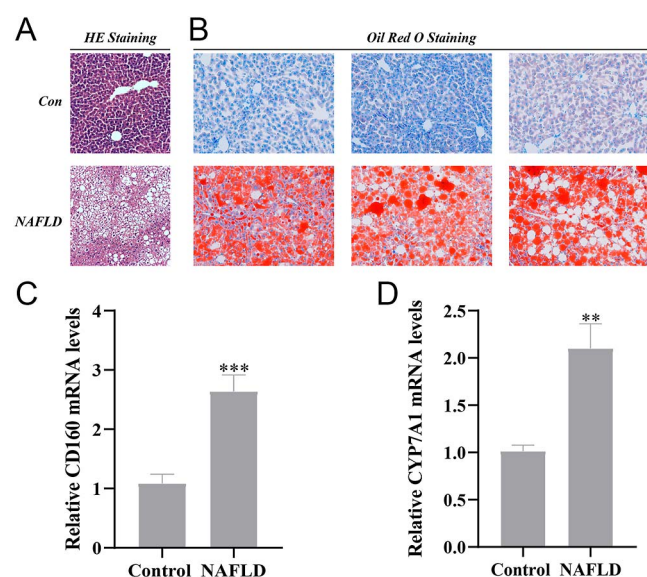
Discussion

NAFLD is regarded as not only the most common cause of CLD worldwide but also a multisystem disease and a growing healthcare burden due to NAFLD is projected in the coming decades (2, 8, 9). Increasing evidence has demonstrated the potential contributory role of tryptophan metabolites mediated by the gut microbiota in the development of NAFLD (16). Identifying tryptophan metabolism-related biomarkers is important to further explore and understand the pathogenesis of NAFLD. In this study, the key module genes associated with TMRGs determined by WGCNA and DEGs, obtained by differential expression analysis of genes, were intersected and functional enrichment analysis was performed on the intersection genes defined as TMRGs-DEGs; these analyses indicated that lipid metabolism and inflammatory pathways were mainly enriched. Then, three hub genes related to TMRGs, including CCL20, CD160 and CYP7A1, were screened out as biomarkers through the construction of a PPI network, the application of two machine learning algorithms and the evaluation of ROC curves, which displayed ideal diagnostic values for all three biomarkers in the training and external validation datasets. In addition, validation of biomarker expression in the training set and the external validation set demonstrated that, compared with the control group, CD160 and CYP7A1 were significantly upregulated in the NAFLD group. These biomarkers may provide novel targets and bring new methods for diagnosing and treating NAFLD.

The GSE89632 dataset was derived from a cross-sectional study at the University Health Network in Toronto, Canada, in which 39 NAFLD samples, including 20 SS and 19 NASH samples, were obtained from patients aged older than 18, diagnosed by liver biopsy after excluding other liver diseases and comorbidities from March 2007 to November 2011, and 24 healthy control samples were obtained from normal patients without hepatic steatosis or fibrosis on imaging or histologic examination (29). The 12 NASH samples of the GSE24807 dataset were obtained from patients aged 9–19 years and diagnosed by liver biopsy at the State University of New York, USA, between July 2010 and September 2013, and the five control samples were obtained from healthy donor livers used for liver transplantation from Admet Technologies (30, 31). In this study, GSE89632 was used as a training set to screen out DEGs in NAFLD versus controls, while GSE24807 was used for ROC curves and expression validation analysis. It is clear that the GSE89632 dataset

**Figure 6**

Further exploration of the role of hub genes. (A–C) Relationship between the cardiovascular or liver diseases and (A) CCL20, (B) CD160 and (C) CYP7A1. (D) ceRNA network for the three hub genes. (E) Drug-mRNA network of the three hub genes. (F & G) Expression difference of the three hub genes in the (F) training set and (G) validation set.

**Figure 7**

Verification of the three hub genes in the NAFLD animal model. (A) HE staining and Oil Red O staining confirmed the successful establishment of the NAFLD mouse model. (B & C) Expression difference of (C) CD160 and (D) CYP7A1 at the mRNA level via qPCR.

is concerned about adult patients in Canada, while GSE24807 is from children in the USA; neither dataset had a large sample size. However, strict inclusion and exclusion criteria, and data quality control ensured the accuracy of genetic information. These datasets are not representative of global NAFLD patients, and further research should at least concentrate on populations in Asia and Africa.

LASSO and SVM-RFE are two commonly used machine learning algorithms in bioinformatics. LASSO is a regularized linear logistic regression analysis method that minimizes the overfitting that may occur with high-dimensional data (32, 33). However, when analyzing data in the biomedical field, especially genomic data, there are problems of unstable handling of correlated variables and inconsistent selection of linear dependent variables, which can be solved by using Precision Lasso, a LASSO variant (34). Based on the nonlinear discriminative property, SVM-RFE recursively eliminates the least important variables, models the remaining variables and then compares them to filter out the best combination of variables (35, 36). Biomedical research has a specific need to filter the variables most relevant to the outcome and understand the strength of their association, and the SVM-RFE algorithm leaves room for further research and performance improvement in this field (37). In this study, prediction accuracy was enhanced for the intersection of genes screened by both LASSO and SVM-RFE methods.

NAFLD is also known as metabolism-associated fatty liver disease (MAFLD). In a study to identify metabolism-related biomarkers in NAFLD, four genes were screened out: AMDHD1, FMO1, LPL and P4HA1 (38). FMO1 and P4HA1 are hub genes for NAFLD, AMDHD1 is associated with liver regeneration, of which expression was downregulated in NAFLD, and LPL is associated with lipid metabolism (38). Although similar bioinformatic methods were used, the four genes related to metabolism in the study mentioned above have no intersection with genes screened out in this study, which may be the result of the difference between cell metabolism and tryptophan catabolism mediated by gut microbiota. Another bioinformatic study exploring biomarkers for different stages of NAFLD demonstrated that ACS2, PCSK9 and CYP7A1 were candidate genes for SS, while ANGPTL3, CD36, CYP51A1, FASN, FAS, FDFT1 and LSS were candidate genes for NASH (39). In this study, CYP7A1 was also identified as a hub gene of NAFLD, which was consistent with the results of the above study, indicating that CYP7A1 may be a diagnostic biomarker for early-stage NAFLD. In a study analyzing hepatic transcriptome data using robust rank aggregation and bioinformatic methods, the top ten genes in NASH patients (SPP1, FBLN5, CHI3L1, CCL20, CD24, FABP4, GPNMB, VCAN, EFEMP1 and CXCL10) were upregulated and their functions were focused on extracellular matrix-related pathways or immune-related pathways, which are transcriptomic features associated with NASH progression and fibrosis severity (40). In this study, the expression level of CCL20 was downregulated in NAFLD compared with the control group, which was not consistent with the results mentioned in the above study and needs further research to verify. Chronic lipid overaccumulation in hepatocytes and disruption of lipid metabolism create a lipotoxic environment, leading to hepatocyte death and release of ALT; thus, elevated circulating ALT levels in patients with MAFLD were observed and ALT was regarded as a risk marker for MAFLD (41, 42, 43). However, in a retrospective cohort study, a significant proportion of MAFLD patients had normal circulating levels of ALT, which may restrict the appliance of ALT in clinical settings (44). Currently, the gold standard for NAFLD diagnosis is liver biopsy, an invasive test with limited clinical application. The identification of NAFLD-related markers can not only contribute to diagnose and intervene in NAFLD as early as possible, to avoid progression to cirrhosis or hepatocellular carcinoma, but also to be conducive to improve the understanding of the pathogenesis of NAFLD and search for effective drug therapy targets.

Cysteine–cysteine motif chemokine ligand 20 (CCL20), also known as macrophage inflammatory protein 3α (MIP-3α), exodus-1 and liver- and activation-regulated chemokine (LARC), is a low molecular weight cytokine and a member of the CC chemokine family, which exerts its effect by acting specifically on C–C chemokine receptor

6 (CCR6) (45). The SCYA20 (small inducible cytokine family A (Cys–Cys), member 20) gene, localized to chromosome 2q33–q372, encodes the CCL20 protein, which consists of 70 amino acids (46, 47). The primary cell types producing CCL20 in the liver are macrophages and hepatic stellate cells (48). It was found that the expression of miR-590-5p was low in biopsy samples of patients with NAFLD fibrosis, resulting in attenuated inhibition of the expression of CCL20, consistent with the observation of elevated CCL20 in the progression of NAFLD and the aggravation of NAFLD histological changes (49, 50). Serving as an inflammatory chemokine, CCL20 is upregulated in hepatic stellate cells and mediates the recruitment of immature dendritic cells, which secrete various inflammatory molecules involved in the progression of the fibrosis process in NAFLD (51). CCL20, associated with the development of NAFLD fibrosis, may serve as a potential target for NAFLD (51).

Cluster of differentiation 160 (CD160) is a glycoprotein anchored on the cell membrane by glycosylphosphatidylinositol (GPI), and a member of the immunoglobulin superfamily expressed on the surface of natural killer cells and cytotoxic T-cell subsets (52, 53). The CD160 gene is localized to human chromosome 1q21.1, contains six exons and encodes the CD160 protein composed of 181 amino acid residues (53, 54). CD160 interacts with major histocompatibility complex class Ia and Ib molecules in a low-affinity way, which triggers the cytotoxic function mediated by circulating natural killer cells and the secretion of cytokines, such as interferon gamma (IFN- γ), tumor necrosis factor alpha (TNF- α) and interleukin 6 (IL-6) (55, 56). Serving as a coinhibitory molecule, CD160 binds with HVEM on the surface of antigen-presenting cells to provide a negative regulatory signal to the activation of CD4(+) T cells (57, 58). CD160, interacting with HVEM, exerted BTLA-dependent inhibition of NKT cells, which may serve as a potential target for treating acute inflammatory diseases in the liver (59). However, at present, no studies have reported the association of CD160 with the development of NAFLD and more research studies need to be conducted to reveal the role of CD160 in the pathogenesis of NAFLD.

Cholesterol 7 α -hydroxylase (CYP7A1) is a member of the cytochrome P450 family 7 subfamily A1 and is located in the microsome of hepatocytes, which is a substrate-specific and rate-limiting enzyme catalyzing the conversion of cholesterol to bile acids in the classic pathway (60). The CYP7A1 gene, also known as human 7 α -hydroxylase gene (CYP70), is located on chromosome 8q11–q12, contains five introns and six exons, and encodes the CYP7A1 protein consisting of 503 amino acid residues (61, 62). The FGF19/FGFR4/Erk1/2 signaling pathway in hepatocytes may mediate the inhibition of bile acid synthesis via downregulated CYP7A1 (63). The transcriptional factor hepatocyte

growth factor suppresses the expression of CYP7A1 by the c-Jun, JNK1/2 and ERK1/2 signaling pathways (64). In addition, miR-122, which is considered as the liver-specific miRNA, was demonstrated to inhibit CYP7A1 via lowering mRNA stability and the expression of miR-122 was reduced in NASH (65, 66, 67). CYP7A1 was shown to be negatively regulated by miR-17 and may be involved in the process of hepatocellular steatosis, which was consistent with the results of a bioinformatic research study (39, 68). CYP7A1 may serve as an early diagnostic biomarker for NAFLD intervention and a potential therapeutic target by regulating cholesterol metabolism for NAFLD.

In this study, a ceRNA network was established comprising of 3 mRNAs, 20 miRNAs and 53 lncRNAs, some of which were shown to be associated with the pathogenesis of NAFLD. Hsa-miR-21-5p plays a crucial role in the development of NAFLD from SS- to NASH-associated liver cancer (69). It was observed that the expression of lncRNA MIR22HG and IGF1 was downregulated, while miR-9-3p was upregulated in NASH patients, and it was also demonstrated that the overexpression of MIR22HG attenuated the emergence and progression of NASH through inhibiting fibrosis and pyroptosis, and promoting autophagy via the miR-9-3p/IGF1 pathway (70). lncRNA NEAT1, binding to microRNA-140, exerts a key role in the progression of NAFLD by inactivating the AMPK/SREBP-1 signaling pathway and the reduction of NEAT1 attenuated NAFLD through mTOR/S6K1 signaling (71, 72). lncRNA MALAT1 facilitated lipid accumulation in the liver mediated by the PPAR α /CD36 pathway in NAFLD, by regulating the miR-206/ARNT axis (73). lncRNA H19 facilitates liver lipogenesis through directly regulating the miR-130a/PPAR γ pathway, and promotes hepatic steatosis by increasing the transcription of MLXIPL and activating the mTORC1 signaling pathway in NAFLD (74, 75). Further validations are needed to explore the interactions revealed by ceRNA.

It was demonstrated that higher serum levels of CCL20 compared with healthy people were related to ischemic heart disease (76), and overexpressed CCL20 was determined in atherosclerotic lesions (77). It was indicated that cardiac fibroblasts secrete CCL20 to chemotactic Th17 cells infiltrating the heart muscle (78). CCL20 was associated with various heart diseases, which was consistent with the results of the analysis in our study. Although our study found a strong association between CD160 and hypertension, at present, the role of CD160 in the development of heart diseases has not been reported. The CYP7A1 protein was a rate-limiting enzyme in the synthesis of bile acids, whose change in both liver and intestine was demonstrated to promote the development of NAFLD in relevant preclinical models (79). Thus, CYP7A1 may be a potential therapeutic target in NAFLD through regulating synthesis of bile acids.

In this study, a total of 41 drugs targeting two biomarkers (CCL20 and CYP7A1) were found. Cyclosporine A (CsA) and its analogs, which were considered as potent and typical cyclophilin inhibitors, exerted protective effects in NASH by preventing mitochondrial dysfunction, reducing ROS production and interdicting cell death mediated by the mitochondrial permeability transition (80). It has been demonstrated that the level of plasma cholesterol increased in CsA-treated rats, which was associated with the significant reduction of CYP7A1 (81). At present, the relationship between CsA and the expression of CCL20 has not been reported. Ursodeoxycholic acid (UDCA) inhibited apoptosis and induced autophagy through regulating the interaction between the Bcl-2/Beclin-1 complex and the Bcl-2/Bax complex, with the activation of AMPK in a rat model of NAFLD (82). It was reported that the administration of UDCA in a mouse model of NASH had beneficial effects, including improving gut microbiota and intestinal barrier integrity, and reducing the level of inflammation in the liver (83). It was indicated that the mRNA level of CYP7A1 was unaffected by UDCA treatment (84). GW4064, through activating farnesoid X receptor (FXR), mitigated inflammation in the liver, and downregulated the levels of apoptosis and proinflammatory molecules induced by endotoxin in a NAFLD mouse model (85). The FXR agonist GW4064 elevated the expression of adipokines and corresponding receptors in preadipocytes and HepG2 cells, which may be a novel way to prevent the progression of NAFLD (86). Estradiol, which is considered as a potent endogenous antioxidant, was demonstrated to attenuate hepatic steatosis and inhibit liver fibrosis in animal models (87, 88). It was indicated that the expression of CYP7A1 was upregulated and the activity of cholesterol 7 α -hydroxylase was elevated with the administration of the natural estrogen 17 β -estradiol (E2) under physiological doses, which disappeared with high-dose treatment (89). Obeticholic acid, which is regarded as a selective FXR agonist that has liver protective properties, has been studied for the treatment of NASH. Obeticholic acid exerts its effects through a novel mechanism, where lipid accumulation in the liver is inhibited through the direct suppression of the activation of the NLRP3 inflammasome in macrophages; its combination with lipid peroxidative inhibitors improved the antifibrotic effect (90, 91, 92). A diet with less saturated fat and increased unsaturated fat is supported for the treatment of NAFLD (93). It was demonstrated that the activity of cholesterol 7 α -hydroxylase was increased in animals fed with a diet rich in unsaturated fatty acids compared with those with saturated fatty acid feed, and the mRNA level of CYP7A1 was upregulated after the administration of a monounsaturated fatty acid diet (94).

Compared with the control group, the expression levels of CD160 and CYP7A1 in the NAFLD group were significantly increased in both the training and external validation datasets, while the expression of CCL20 in the NAFLD

group was decreased in the two datasets, although the decrease was not statistically significant in the external validation dataset. However, it was demonstrated that CCL20 was elevated in NAFLD, which is not consistent with the verification results in the two datasets (51). At present, the role of CD160 in the pathogenesis of NAFLD has not been reported. The expression of CYP7A1 was upregulated in NAFLD, which contributed to the synthesis of bile acids and is consistent with the validation in the two datasets (95).

The CCR6 ligand CCL20 is regarded as the marker gene of M2 macrophages. Secretion of the cytokine CCL20 is closely linked to M2 polarization, and the chemokine axis CCR6/CCL20 is involved in cancer progression in a variety of tumors, such as prostate cancer (96), melanoma (97) and hepatocellular carcinoma (98). It was observed that the level of CCL20 elevated with gradual worsening lesions of liver histology in NAFLD patients (50). The TM4SF5-dependent chemokine secretion of hepatocytes, including CCL20 and CXCL10, induced macrophage reprogramming from proinflammatory M1-type macrophages to pro-fibrotic M2-type macrophages, leading to the progression of NASH to hepatic fibrosis (99). In addition, visfatin upregulated the expression of CCL20 through activating the NF- κ B and MKK3/6-p38 signaling pathways in macrophages, promoting expression of fibrosis markers in hepatic stellate cells, which demonstrated that the secretion of CCL20 by macrophages might play an important role in the progression of liver fibrosis (100).

CD160 is a molecular marker on the surface of cytotoxic NK cells, it exists as a GPI or transmembrane protein, and it is predominantly expressed in the cytotoxic subset of NK cells (CD56dim/CD16+). In NK cells, CD160 enhances cell activation and cytotoxicity, and activates the secretion of cytokines IFN- γ , IL-6, IL-8 and TNF- α . It has been shown that NKT-CD160 and marginal zone B clusters were decreased in both convalescent moderate and severe cases (101). The functions and mechanisms of the interactions between CD160 and neutrophils have not yet been reported in the literature.

CYP7A1 is the key regulatory molecule in the synthesis of bile acids from cholesterol. It was demonstrated that the expression level and secretion of IL-1 β and TNF- α in macrophage cell lines and Kupffer cells were elevated through TGR5 activation, which suppressed the expression of CYP7A1 in hepatocytes (102). In this study, the highest correlation between CYP7A1 and activated mast cells was identified through immune cell infiltration analysis. However, the interaction of CYP7A1 and mast cells in the pathogenesis and progression of NAFLD has not been reported.

Supplementary materials

This is linked to the online version of the paper at
<https://doi.org/10.1530/EC-24-0470>.

Declaration 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.

Funding

This research was funded by the Science and Technology Research Project of Jiangxi Provincial Department of Education (grant number GJJ201520 to BZ), the National Natural Science Foundation of China (82100869 to PY), the Natural Science Foundation of Jiangxi Province (grant numbers 20212BAB216051 to JZ and 20212BAB216047 and 202004BCJL23049 to PY) and the 'Thousand Talents Plan' for Introducing and Cultivating High Level Talents in Innovation and Entrepreneurship of Jiangxi Province (grant number jxsq2023201105 to PY).

Author contribution statement

BZ and KH were responsible for data acquisition and analysis. KH, GS and SH conducted the human experiments. LX, RH and KH drafted the manuscript. GY, JZ and DZ revised the manuscript. PY and LW designed this study. All authors read and approved the final manuscript for submission.

Data availability

The original data used in this project can be downloaded from the GEO public database (<https://www.ncbi.nlm.nih.gov/geo/>). The data analyzed during the human experiments are available from the corresponding author on reasonable request.

Acknowledgments

We sincerely acknowledge the contribution of the SangerBox public platform (<http://vip.sangerbox.com/home.html>).

References

- 1 Paternostro R & Trauner M. Current treatment of non-alcoholic fatty liver disease. *J Intern Med* 2022 **292** 190–204. (<https://doi.org/10.1111/joim.13531>)
- 2 Paik JM, Golabi P, Younossi Y, *et al.* Changes in the global burden of chronic liver diseases from 2012 to 2017: the growing impact of NAFLD. *Hepatology* 2020 **72** 1605–1616. (<https://doi.org/10.1002/hep.31173>)
- 3 Le MH, Yeo YH, Zou B, *et al.* Forecasted 2040 global prevalence of nonalcoholic fatty liver disease using hierarchical bayesian approach. *Clin Mol Hepatol* 2022 **28** 841–850. (<https://doi.org/10.3350/cmh.2022.0239>)
- 4 Huang DQ, Singal AG, Kono Y, *et al.* Changing global epidemiology of liver cancer from 2010 to 2019: NASH is the fastest growing cause of liver cancer. *Cell Metab* 2022 **34** 969–977.e2. (<https://doi.org/10.1016/j.cmet.2022.05.003>)
- 5 Araújo AR, Rosso N, Bedogni G, *et al.* Global epidemiology of non-alcoholic fatty liver disease/non-alcoholic steatohepatitis: what we need in the future. *Liver Int* 2018 **38** (Supplement 1) 47–51. (<https://doi.org/10.1111/liv.13643>)
- 6 Henry L, Paik J & Younossi ZM. Review article: the epidemiologic burden of non-alcoholic fatty liver disease across the world. *Aliment Pharmacol Ther* 2022 **56** 942–956. (<https://doi.org/10.1111/apt.17158>)
- 7 Riaz K, Azhari H, Charette JH, *et al.* The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* 2022 **7** 851–861. ([https://doi.org/10.1016/s2468-1253\(22\)00165-0](https://doi.org/10.1016/s2468-1253(22)00165-0))
- 8 Estes C, Anstee QM, Arias-Loste MT, *et al.* Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016–2030. *J Hepatol* 2018 **69** 896–904. (<https://doi.org/10.1016/j.jhep.2018.05.036>)
- 9 Targher G, Tilg H & Byrne CD. Non-alcoholic fatty liver disease: a multisystem disease requiring a multidisciplinary and holistic approach. *Lancet Gastroenterol Hepatol* 2021 **6** 578–588. ([https://doi.org/10.1016/s2468-1253\(21\)00020-0](https://doi.org/10.1016/s2468-1253(21)00020-0))
- 10 Cho JY & Sohn W. The growing burden of non-alcoholic fatty liver disease on mortality. *Clin Mol Hepatol* 2023 **29** 374–376. (<https://doi.org/10.3350/cmh.2023.0084>)
- 11 Ferguson D & Finck BN. Emerging therapeutic approaches for the treatment of NAFLD and type 2 diabetes mellitus. *Nat Rev Endocrinol* 2021 **17** 484–495. (<https://doi.org/10.1038/s41574-021-00507-z>)
- 12 Makri E, Goulas A & Polyzos SA. Epidemiology, pathogenesis, diagnosis and emerging treatment of nonalcoholic fatty liver disease. *Arch Med Res* 2021 **52** 25–37. (<https://doi.org/10.1016/j.arcmed.2020.11.010>)
- 13 Guo X, Yin X, Liu Z, *et al.* Non-alcoholic fatty liver disease (NAFLD) pathogenesis and natural products for prevention and treatment. *Int J Mol Sci* 2022 **23** 15489. (<https://doi.org/10.3390/ijms232415489>)
- 14 Buzzetti E, Pinzani M & Tsochatzis EA. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Metabolism* 2016 **65** 1038–1048. (<https://doi.org/10.1016/j.metabol.2015.12.012>)
- 15 Wu J, Wang K, Wang X, *et al.* The role of the gut microbiome and its metabolites in metabolic diseases. *Protein Cell* 2021 **12** 360–373. (<https://doi.org/10.1007/s13238-020-00814-7>)
- 16 Ni Y, Ni L, Zhuge F, *et al.* The gut microbiota and its metabolites, novel targets for treating and preventing non-alcoholic fatty liver disease. *Mol Nutr Food Res* 2020 **64** e2000375. (<https://doi.org/10.1002/mnfr.202000375>)
- 17 Lee H, Zandkarimi F, Zhang Y, *et al.* Energy-stress-mediated AMPK activation inhibits ferroptosis. *Nat Cell Biol* 2020 **22** 225–234. (<https://doi.org/10.1038/s41556-020-0461-8>)
- 18 Taleb S. Tryptophan dietary impacts gut barrier and metabolic diseases. *Front Immunol* 2019 **10** 2113. (<https://doi.org/10.3389/fimmu.2019.02113>)
- 19 Roager HM & Licht TR. Microbial tryptophan catabolites in health and disease. *Nat Commun* 2018 **9** 3294. (<https://doi.org/10.1038/s41467-018-05470-4>)
- 20 Teunis C, Nieuwdorp M & Hanssen N. Interactions between tryptophan metabolism, the gut microbiome and the immune system as potential drivers of non-alcoholic fatty liver disease (NAFLD) and metabolic diseases. *Metabolites* 2022 **12** 514. (<https://doi.org/10.3390/metabo12060514>)
- 21 Arora A, Tripodi GL, Kareinen I, *et al.* Genetic deficiency of indoleamine 2,3-dioxygenase aggravates vascular but not liver disease in a nonalcoholic steatohepatitis and atherosclerosis comorbidity model. *Int J Mol Sci* 2022 **23** 5203. (<https://doi.org/10.3390/ijms23095203>)
- 22 Pyun DH, Kim TJ, Kim MJ, *et al.* Endogenous metabolite, kynurenine acid, attenuates nonalcoholic fatty liver disease via AMPK/autophagy- and AMPK/ORP150-mediated signaling. *J Cell Physiol* 2021 **236** 4902–4912. (<https://doi.org/10.1002/jcp.30199>)

- 23 Bansal T, Alaniz RC, Wood TK, *et al.* The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. *Proc Natl Acad Sci U S A* 2010 **107** 228–233. (<https://doi.org/10.1073/pnas.0906112107>)
- 24 Ji Y, Gao Y, Chen H, *et al.* Indole-3-acetic acid alleviates nonalcoholic fatty liver disease in mice via attenuation of hepatic lipogenesis, and oxidative and inflammatory stress. *Nutrients* 2019 **11** 2062. (<https://doi.org/10.3390/nu11092062>)
- 25 Zhao ZH, Xin FZ, Xue Y, *et al.* Indole-3-propionic acid inhibits gut dysbiosis and endotoxin leakage to attenuate steatohepatitis in rats. *Exp Mol Med* 2019 **51** 1–14. (<https://doi.org/10.1038/s12276-019-0304-5>)
- 26 Nocito A, Dahm F, Jochum W, *et al.* Serotonin mediates oxidative stress and mitochondrial toxicity in a murine model of nonalcoholic steatohepatitis. *Gastroenterology* 2007 **133** 608–618. (<https://doi.org/10.1053/j.gastro.2007.05.019>)
- 27 Terziev D & Terzieva D. Experimental data on the role of melatonin in the pathogenesis of nonalcoholic fatty liver disease. *Biomedicines* 2023 **11** 1722. (<https://doi.org/10.3390/biomedicines11061722>)
- 28 Chu Y, Yu F, Wu Y, *et al.* Identification of genes and key pathways underlying the pathophysiological association between nonalcoholic fatty liver disease and atrial fibrillation. *BMC Med Genomics* 2022 **15** 150. (<https://doi.org/10.1186/s12920-022-01300-1>)
- 29 Arendt BM, Comelli EM, Ma DW, *et al.* Altered hepatic gene expression in nonalcoholic fatty liver disease is associated with lower hepatic n-3 and n-6 polyunsaturated fatty acids. *Hepatology* 2015 **61** 1565–1578. (<https://doi.org/10.1002/hep.27695>)
- 30 Liu W, Baker SS, Baker RD, *et al.* Upregulation of hemoglobin expression by oxidative stress in hepatocytes and its implication in nonalcoholic steatohepatitis. *PLoS One* 2011 **6** e24363. (<https://doi.org/10.1371/journal.pone.0024363>)
- 31 Feng Q, Baker SS, Liu W, *et al.* Increased apolipoprotein A5 expression in human and rat non-alcoholic fatty livers. *Pathology* 2015 **47** 341–348. (<https://doi.org/10.1097/pat.0000000000000251>)
- 32 Wu Z, Chen H, Ke S, *et al.* Identifying potential biomarkers of idiopathic pulmonary fibrosis through machine learning analysis. *Sci Rep* 2023 **13** 16559. (<https://doi.org/10.1038/s41598-023-43834-z>)
- 33 Atabaki-Pasdar N, Ohlsson M, Viñuela A, *et al.* Predicting and elucidating the etiology of fatty liver disease: a machine learning modeling and validation study in the IMI DIRECT cohorts. *PLoS Med* 2020 **17** e1003149. (<https://doi.org/10.1371/journal.pmed.1003149>)
- 34 Wang H, Lengerich BJ, Aragam B, *et al.* Precision lasso: accounting for correlations and linear dependencies in high-dimensional genomic data. *Bioinformatics* 2019 **35** 1181–1187. (<https://doi.org/10.1093/bioinformatics/bty750>)
- 35 Wu LD, Li F, Chen JY, *et al.* Analysis of potential genetic biomarkers using machine learning methods and immune infiltration regulatory mechanisms underlying atrial fibrillation. *BMC Med Genomics* 2022 **15** 64. (<https://doi.org/10.1186/s12920-022-01212-0>)
- 36 Wang H, Cheng W, Hu P, *et al.* Integrative analysis identifies oxidative stress biomarkers in non-alcoholic fatty liver disease via machine learning and weighted gene co-expression network analysis. *Front Immunol* 2024 **15** 1335112. (<https://doi.org/10.3389/fimmu.2024.1335112>)
- 37 Sanz H, Valim C, Vegas E, *et al.* SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinf* 2018 **19** 432. (<https://doi.org/10.1186/s12859-018-2451-4>)
- 38 Jiang H, Hu Y, Zhang Z, *et al.* Identification of metabolic biomarkers associated with nonalcoholic fatty liver disease. *Lipids Health Dis* 2023 **22** 150. (<https://doi.org/10.1186/s12944-023-01911-2>)
- 39 Reyes-Avendaño I, Villaseñor-Altamirano AB, Reyes-Jimenez E, *et al.* Identification of key markers for the stages of nonalcoholic fatty liver disease: an integrated bioinformatics analysis and experimental validation. *Dig Liver Dis* 2024 **56** 1887–1896. (<https://doi.org/10.1016/j.dld.2024.05.010>)
- 40 He W, Huang C, Zhang X, *et al.* Identification of transcriptomic signatures and crucial pathways involved in non-alcoholic steatohepatitis. *Endocrine* 2021 **73** 52–64. (<https://doi.org/10.1007/s12020-021-02716-y>)
- 41 Mota M, Banini BA, Cazanave SC, *et al.* Molecular mechanisms of lipotoxicity and glucotoxicity in nonalcoholic fatty liver disease. *Metabolism* 2016 **65** 1049–1061. (<https://doi.org/10.1016/j.metabol.2016.02.014>)
- 42 Musso G, Cassader M, Paschetta E, *et al.* Bioactive lipid species and metabolic pathways in progression and resolution of nonalcoholic steatohepatitis. *Gastroenterology* 2018 **155** 282–302.e8. (<https://doi.org/10.1053/j.gastro.2018.06.031>)
- 43 Wang D, Zhou BY, Xiang L, *et al.* Alanine aminotransferase as a risk marker for new-onset metabolic dysfunction-associated fatty liver disease. *World J Gastroenterol* 2024 **30** 3132–3139. (<https://doi.org/10.3748/wjg.v30.i25.3132>)
- 44 Chen JF, Wu ZQ, Liu HS, *et al.* Cumulative effects of excess high-normal alanine aminotransferase levels in relation to new-onset metabolic dysfunction-associated fatty liver disease in China. *World J Gastroenterol* 2024 **30** 1346–1357. (<https://doi.org/10.3748/wjg.v30.i10.1346>)
- 45 Kadomoto S, Izumi K & Mizokami A. The CCL20-CCR6 Axis in cancer progression. *Int J Mol Sci* 2020 **21** 5186. (<https://doi.org/10.3390/ijms21155186>)
- 46 Zlotnik A & Yoshie O. Chemokines: a new classification system and their role in immunity. *Immunity* 2000 **12** 121–127. ([https://doi.org/10.1016/s1074-7613\(00\)80165-x](https://doi.org/10.1016/s1074-7613(00)80165-x))
- 47 Hieshima K, Imai T, Opdenakker G, *et al.* Molecular cloning of a novel human CC chemokine liver and activation-regulated chemokine (LARC) expressed in liver. Chemotactic activity for lymphocytes and gene localization on chromosome 2. *J Biol Chem* 1997 **272** 5846–5853. (<https://doi.org/10.1074/jbc.272.9.5846>)
- 48 Affò S, Morales-Ibanez O, Rodrigo-Torres D, *et al.* CCL20 mediates lipopolysaccharide induced liver injury and is a potential driver of inflammation and fibrosis in alcoholic hepatitis. *Gut* 2014 **63** 1782–1792. (<https://doi.org/10.1136/gutjnl-2013-306098>)
- 49 Leti F, Malenica I, Doshi M, *et al.* High-throughput sequencing reveals altered expression of hepatic microRNAs in nonalcoholic fatty liver disease-related fibrosis. *Transl Res* 2015 **166** 304–314. (<https://doi.org/10.1016/j.trsl.2015.04.014>)
- 50 Hanson A, Piras IS, Wilhelmsen D, *et al.* Chemokine ligand 20 (CCL20) expression increases with NAFLD stage and hepatic stellate cell activation and is regulated by miR-590-5p. *Cytokine* 2019 **123** 154789. (<https://doi.org/10.1016/j.cyto.2019.154789>)
- 51 Chu X, Jin Q, Chen H, *et al.* CCL20 is up-regulated in non-alcoholic fatty liver disease fibrosis and is produced by hepatic stellate cells in response to fatty acid loading. *J Transl Med* 2018 **16** 108. (<https://doi.org/10.1186/s12967-018-1490-y>)
- 52 Oumeslakht L, Aziz AI, Bensussan A, *et al.* CD160 receptor in CLL: current state and future avenues. *Front Immunol* 2022 **13** 1028013. (<https://doi.org/10.3389/fimmu.2022.1028013>)

- 53 Piotrowska M, Spodzieja M, Kuncewicz K, *et al.* CD160 protein as a new therapeutic target in a battle against autoimmune, infectious and lifestyle diseases. Analysis of the structure, interactions and functions. *Eur J Med Chem* 2021 **224** 113694. (<https://doi.org/10.1016/j.ejmech.2021.113694>)
- 54 Schmitt C, Ghazi B, Bellier F, *et al.* Identification and analysis of the human CD160 promoter: implication of a potential AML-1 binding site in promoter activation. *Genes Immun* 2009 **10** 616–623. (<https://doi.org/10.1038/gene.2009.52>)
- 55 Agrawal S, Marquet J, Freeman GJ, *et al.* Cutting edge: MHC class I triggering by a novel cell surface ligand costimulates proliferation of activated human T cells. *J Immunol* 1999 **162** 1223–1226. (<https://doi.org/10.4049/jimmunol.162.3.1223>)
- 56 Barakonyi A, Rabot M, Marie-Cardine A, *et al.* Cutting edge: engagement of CD160 by its HLA-C physiological ligand triggers a unique cytokine profile secretion in the cytotoxic peripheral blood NK cell subset. *J Immunol* 2004 **173** 5349–5354. (<https://doi.org/10.4049/jimmunol.173.9.5349>)
- 57 Kojima R, Kajikawa M, Shiroishi M, *et al.* Molecular basis for herpesvirus entry mediator recognition by the human immune inhibitory receptor CD160 and its relationship to the cosignaling molecules BTLA and LIGHT. *J Mol Biol* 2011 **413** 762–772. (<https://doi.org/10.1016/j.jmb.2011.09.018>)
- 58 Cai G, Anumanthan A, Brown JA, *et al.* CD160 inhibits activation of human CD4+ T cells through interaction with herpesvirus entry mediator. *Nat Immunol* 2008 **9** 176–185. (<https://doi.org/10.1038/ni1554>)
- 59 Kim TJ, Park G, Kim J, *et al.* CD160 serves as a negative regulator of NKT cells in acute hepatic injury. *Nat Commun* 2019 **10** 3258. (<https://doi.org/10.1038/s41467-019-10320-y>)
- 60 Chiang JYL & Ferrell JM. Up to date on cholesterol 7 alpha-hydroxylase (CYP7A1) in bile acid synthesis. *Liver Res* 2020 **4** 47–63. (<https://doi.org/10.1016/j.livres.2020.05.001>)
- 61 Cohen JC, Cali JJ, Jelinek DF, *et al.* Cloning of the human cholesterol 7 alpha-hydroxylase gene (CYP7) and localization to chromosome 8q11-q12. *Genomics* 1992 **14** 153–161. ([https://doi.org/10.1016/s0888-7543\(05\)80298-8](https://doi.org/10.1016/s0888-7543(05)80298-8))
- 62 Li YC, Wang DP & Chiang JY. Regulation of cholesterol 7 alpha-hydroxylase in the liver. Cloning, sequencing, and regulation of cholesterol 7 alpha-hydroxylase mRNA. *J Biol Chem* 1990 **265** 12012–12019. ([https://doi.org/10.1016/s0021-9258\(19\)38501-1](https://doi.org/10.1016/s0021-9258(19)38501-1))
- 63 Song KH, Li T, Owsley E, *et al.* Bile acids activate fibroblast growth factor 19 signaling in human hepatocytes to inhibit cholesterol 7 α -hydroxylase gene expression. *Hepatology* 2009 **49** 297–305. (<https://doi.org/10.1002/hep.22627>)
- 64 Song KH, Ellis E, Strom S, *et al.* Hepatocyte growth factor signaling pathway inhibits cholesterol 7 α -hydroxylase and bile acid synthesis in human hepatocytes. *Hepatology* 2007 **46** 1993–2002. (<https://doi.org/10.1002/hep.21878>)
- 65 Song KH, Li T, Owsley E, *et al.* A putative role of micro RNA in regulation of cholesterol 7 α -hydroxylase expression in human hepatocytes. *J Lipid Res* 2010 **51** 2223–2233. (<https://doi.org/10.1194/jlr.m004531>)
- 66 Cheung O, Puri P, Eicken C, *et al.* Nonalcoholic steatohepatitis is associated with altered hepatic microRNA expression. *Hepatology* 2008 **48** 1810–1820. (<https://doi.org/10.1002/hep.22569>)
- 67 Esau C, Davis S, Murray SF, *et al.* miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. *Cell Metab* 2006 **3** 87–98. (<https://doi.org/10.1016/j.cmet.2006.01.005>)
- 68 Gong R, Lv X & Liu F. MiRNA-17 encoded by the miR-17-92 cluster increases the potential for steatosis in hepatoma cells by targeting CYP7A1. *Cell Mol Biol Lett* 2018 **23** 16. (<https://doi.org/10.1186/s11658-018-0083-3>)
- 69 Rodrigues PM, Afonso MB, Simão AL, *et al.* miR-21-5p promotes NASH-related hepatocarcinogenesis. *Liver Int* 2023 **43** 2256–2274. (<https://doi.org/10.1111/liv.15682>)
- 70 Chen X, Zhou S, Chen Y, *et al.* LncRNA MIR22HG/microRNA-9-3p/IGF1 in nonalcoholic steatohepatitis, the ceRNA network increases fibrosis by inhibiting autophagy and promoting pyroptosis. *Clin Nutr* 2024 **43** 52–64. (<https://doi.org/10.1016/j.clnu.2023.11.004>)
- 71 Sun Y, Song Y, Liu C, *et al.* LncRNA NEAT1-MicroRNA-140 axis exacerbates nonalcoholic fatty liver through interrupting AMPK/SREBP-1 signaling. *Biochem Biophys Res Commun* 2019 **516** 584–590. (<https://doi.org/10.1016/j.bbrc.2019.06.104>)
- 72 Wang X. Down-regulation of lncRNA-NEAT1 alleviated the non-alcoholic fatty liver disease via mTOR/S6K1 signaling pathway. *J Cell Biochem* 2018 **119** 1567–1574. (<https://doi.org/10.1002/jcb.26317>)
- 73 Xiang J, Deng YY, Liu HX, *et al.* LncRNA MALAT1 promotes PPAR α /CD36-mediated hepatic lipogenesis in nonalcoholic fatty liver disease by modulating miR-206/ARNT axis. *Front Bioeng Biotechnol* 2022 **10** 858558. (<https://doi.org/10.3389/fbioe.2022.858558>)
- 74 Liu J, Tang T, Wang GD, *et al.* LncRNA-H19 promotes hepatic lipogenesis by directly regulating miR-130a/PPAR γ axis in non-alcoholic fatty liver disease. *Biosci Rep* 2019 **39** BSR20181722. (<https://doi.org/10.1042/bsr20181722>)
- 75 Wang H, Cao Y, Shu L, *et al.* Long non-coding RNA (lncRNA) H19 induces hepatic steatosis through activating MLXIPL and mTORC1 networks in hepatocytes. *J Cell Mol Med* 2020 **24** 1399–1412. (<https://doi.org/10.1111/jcmm.14818>)
- 76 Safa A, Rashidinejad HR, Khalili M, *et al.* Higher circulating levels of chemokines CXCL10, CCL20 and CCL22 in patients with ischemic heart disease. *Cytokine* 2016 **83** 147–157. (<https://doi.org/10.1016/j.cyto.2016.04.006>)
- 77 Calvayrac O, Rodríguez-Calvo R, Alonso J, *et al.* CCL20 is increased in hypercholesterolemic subjects and is upregulated by LDL in vascular smooth muscle cells: role of NF- κ B. *Arterioscler Thromb Vasc Biol* 2011 **31** 2733–2741. (<https://doi.org/10.1161/ATVBAHA.111.235721>)
- 78 Yu M, Hu J, Zhu MX, *et al.* Cardiac fibroblasts recruit Th17 cells infiltration into myocardium by secreting CCL20 in CVB3-induced acute viral myocarditis. *Cell Physiol Biochem* 2013 **32** 1437–1450. (<https://doi.org/10.1159/000356581>)
- 79 Gillard J, Clerbaux LA, Nachit M, *et al.* Bile acids contribute to the development of non-alcoholic steatohepatitis in mice. *JHEP Rep* 2022 **4** 100387. (<https://doi.org/10.1016/j.jhepr.2021.100387>)
- 80 Ure DR, Trepanier DJ, Mayo PR, *et al.* Cyclophilin inhibition as a potential treatment for nonalcoholic steatohepatitis (NASH). *Expert Opin Investig Drugs* 2020 **29** 163–178. (<https://doi.org/10.1080/13543784.2020.1703948>)
- 81 Vaziri ND, Liang K & Azad H. Effect of cyclosporine on HMG-CoA reductase, cholesterol 7 α -hydroxylase, LDL receptor, HDL receptor, VLDL receptor, and lipoprotein lipase expressions. *J Pharmacol Exp Ther* 2000 **294** 778–783. ([https://doi.org/10.1016/s0022-3565\(24\)39134-7](https://doi.org/10.1016/s0022-3565(24)39134-7))
- 82 Wu P, Zhao J, Guo Y, *et al.* Ursodeoxycholic acid alleviates nonalcoholic fatty liver disease by inhibiting apoptosis and improving autophagy via activating AMPK. *Biochem Biophys Res Commun* 2020 **529** 834–838. (<https://doi.org/10.1016/j.bbrc.2020.05.128>)

- 83 Li H, Wang Q, Chen P, *et al.* Ursodeoxycholic acid treatment restores gut microbiota and alleviates liver inflammation in non-alcoholic steatohepatic mouse model. *Front Pharmacol* 2021 **12** 788558. (<https://doi.org/10.3389/fphar.2021.788558>)
- 84 Shefer S, Kren BT, Salen G, *et al.* Regulation of bile acid synthesis by deoxycholic acid in the rat: different effects on cholesterol 7 α -hydroxylase and sterol 27-hydroxylase. *Hepatology* 1995 **22** 1215–1221. (<https://doi.org/10.1002/hep.1840220429>)
- 85 Yao J, Zhou CS, Ma X, *et al.* FXR agonist GW4064 alleviates endotoxin-induced hepatic inflammation by repressing macrophage activation. *World J Gastroenterol* 2014 **20** 14430–14441. (<https://doi.org/10.3748/wjg.v20.i39.14430>)
- 86 Xin XM, Zhong MX, Yang GL, *et al.* GW4064, a farnesoid X receptor agonist, upregulates adipokine expression in preadipocytes and HepG2 cells. *World J Gastroenterol* 2014 **20** 15727–15735. (<https://doi.org/10.3748/wjg.v20.i42.15727>)
- 87 Shimizu I & Ito S. Protection of estrogens against the progression of chronic liver disease. *Hepatol Res* 2007 **37** 239–247. (<https://doi.org/10.1111/j.1872-034x.2007.00032.x>)
- 88 Kamada Y, Kiso S, Yoshida Y, *et al.* Estrogen deficiency worsens steatohepatitis in mice fed high-fat and high-cholesterol diet. *Am J Physiol Gastrointest Liver Physiol* 2011 **301** G1031–G1043. (<https://doi.org/10.1152/ajpgi.00211.2011>)
- 89 Parini P, Angelin B, Stavr  us-Evers A, *et al.* Biphasic effects of the natural estrogen 17 β -estradiol on hepatic cholesterol metabolism in intact female rats. *Arterioscler Thromb Vasc Biol* 2000 **20** 1817–1823. (<https://doi.org/10.1161/01.atv.20.7.1817>)
- 90 Huang S, Wu Y, Zhao Z, *et al.* A new mechanism of obeticholic acid on NASH treatment by inhibiting NLRP3 inflammasome activation in macrophage. *Metabolism* 2021 **120** 154797. (<https://doi.org/10.1016/j.metabol.2021.154797>)
- 91 Zhuge A, Li S, Yuan Y, *et al.* Microbiota-induced lipid peroxidation impairs obeticholic acid-mediated antifibrotic effect towards nonalcoholic steatohepatitis in mice. *Redox Biol* 2023 **59** 102582. (<https://doi.org/10.1016/j.redox.2022.102582>)
- 92 Abenavoli L, Falalyeyeva T, Boccutto L, *et al.* Obeticholic acid: a new era in the treatment of nonalcoholic fatty liver disease. *Pharmaceuticals* 2018 **11** 104. (<https://doi.org/10.3390/ph11040104>)
- 93 Yki-J  rvinen H, Luukkonen PK, Hodson L, *et al.* Dietary carbohydrates and fats in nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol* 2021 **18** 770–786. (<https://doi.org/10.1038/s41575-021-00472-y>)
- 94 Bravo E, Flora L, Cantafora A, *et al.* The influence of dietary saturated and unsaturated fat on hepatic cholesterol metabolism and the biliary excretion of chylomicron cholesterol in the rat. *Biochim Biophys Acta* 1998 **1390** 134–148. ([https://doi.org/10.1016/s0005-2760\(97\)00174-4](https://doi.org/10.1016/s0005-2760(97)00174-4))
- 95 Jiao N, Baker SS, Chapa-Rodr  guez A, *et al.* Suppressed hepatic bile acid signalling despite elevated production of primary and secondary bile acids in NAFLD. *Gut* 2018 **67** 1881–1891. (<https://doi.org/10.1136/gutjnl-2017-314307>)
- 96 Xie T, Fu DJ, Li ZM, *et al.* CircSMARCC1 facilitates tumor progression by disrupting the crosstalk between prostate cancer cells and tumor-associated macrophages via miR-1322/CCL20/CCR6 signaling. *Mol Cancer* 2022 **21** 173. (<https://doi.org/10.1186/s12943-022-01630-9>)
- 97 Samaniego R, Guti  rrez-Gonz  lez A, Guti  rrez-Seijo A, *et al.* CCL20 expression by tumor-associated macrophages predicts progression of human primary cutaneous melanoma. *Cancer Immunol Res* 2018 **6** 267–275. (<https://doi.org/10.1158/2326-6066.cir-17-0198>)
- 98 Chen Y, Zhou P, Deng Y, *et al.* ALKBH5-mediated m(6)A demethylation of TIRAP mRNA promotes radiation-induced liver fibrosis and decreases radiosensitivity of hepatocellular carcinoma. *Clin Transl Med* 2023 **13** e1198. (<https://doi.org/10.1002/ctm2.1198>)
- 99 Kim E, Um H, Park J, *et al.* TM4SF5-dependent crosstalk between hepatocytes and macrophages to reprogram the inflammatory environment. *Cell Rep* 2021 **37** 110018. (<https://doi.org/10.1016/j.celrep.2021.110018>)
- 100 Heo YJ, Choi SE, Lee N, *et al.* CCL20 induced by visfatin in macrophages via the NF- κ B and MKK3/6-p38 signaling pathways contributes to hepatic stellate cell activation. *Mol Biol Rep* 2020 **47** 4285–4293. (<https://doi.org/10.1007/s11033-020-05510-7>)
- 101 Zhao P, Zou J, Zhou F, *et al.* Immune features of COVID-19 convalescent individuals revealed by a single-cell RNA sequencing. *Int Immunopharmacol* 2022 **108** 108767. (<https://doi.org/10.1016/j.intimp.2022.108767>)
- 102 Lou G, Ma X, Fu X, *et al.* GPBAR1/TGR5 mediates bile acid-induced cytokine expression in murine Kupffer cells. *PLoS One* 2014 **9** e93567. (<https://doi.org/10.1371/journal.pone.0093567>)