Received: 25 June 2023 | Accepted: 25 August 2023

DOI: 10.1097/HC9.0000000000000298

ORIGINAL ARTICLE

OPEN

Multiscale integrative analyses unveil immune-related diagnostic signature for the progression of MASLD

Jing Bai[1] | Lun Zhu[1] | Wanqi Mi[1] | Zhengzheng Gao[2] | Minyue Ouyang[1] | Wanlu Sheng[2] | Lin Song[3] | Lidao Bao[4] | Yuheng Ma[2] | Yingqi Xu[1,2,4]

[1]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China

[2]College of Pharmacy, Inner Mongolia Medical University, Hohhot, China

[3]College of Mongolian Medicine, Inner Mongolia Medical University, Hohhot, China

[4]Hohhot Mongolian Medicine of Traditional Chinese Medicine Hospital, Hohhot, China

Correspondence
Yingqi Xu, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, 150081, China.
Email: xuyingqi1987@163.com

Yuheng Ma, College of Pharmacy, Inner Mongolia Medical University, Hohhot, 010110, China.
Email: myhstar@126.com

Lidao Bao, Hohhot Mongolian Medicine of Traditional Chinese Medicine Hospital, Hohhot, 010110, China.
Email: baolidao237@163.com

Abstract

Metabolic dysfunction–associated steatotic liver disease (MASLD) is a chronic liver disease prevalent worldwide, with an increasing incidence associated with obesity, diabetes, and metabolic syndrome. The progression of MASLD to metabolic dysfunction–associated steatohepatitis (MASH) poses a pressing health concern, highlighting the significance of accurately identifying MASLD and its progression to MASH as a primary challenge in the field. In this study, a systematic integration of 66 immune cell types was conducted. Comprehensive analyses were performed on bulk, single-cell RNA-Seq, and clinical data to investigate the immune cell types implicated in MASLD progression thoroughly. Multiple approaches, including immune infiltration, gene expression trend analysis, weighted gene coexpression network analysis, and 4 machine learning algorithms, were used to examine the dynamic changes in genes and immune cells during MASLD progression. C-X-C motif chemokine receptor 4 and dedicator of cytokinesis 8 have been identified as potential diagnostic biomarkers for MASLD progression. Furthermore, cell communication analysis at the single-cell level revealed that the involvement of C-X-C motif chemokine receptor 4 and dedicator of cytokinesis 8 in MASLD progression is mediated through their influence on T cells. Overall, our study identified vital immune cells and a 2-gene diagnostic signature for the progression of MASLD, providing a new perspective on the diagnosis and immune-related molecular mechanisms of MASLD. These findings have important implications for developing innovative diagnostic tools and therapies for MASLD.

Abbreviations: CXCR4, C-X-C motif chemokine receptor 4; DC, dendritic cell; DEG, differentially expressed gene; DOCK8, dedicator of cytokinesis 8; HLA-DRA, HLA class II histocompatibility antigen, DR alpha chain; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease; PPI, protein-protein interaction.

Jing Bai, Lun Zhu, and Wanqi Mi contributed equally to this work.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.hepcommjournal.com.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.
Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Association for the Study of Liver Diseases.

INTRODUCTION

Metabolic dysfunction–associated steatotic liver disease (MASLD) is the most common chronic liver disease worldwide, with a high prevalence rate, and is the second leading cause of end-stage liver disease.[1,2] Metabolic dysfunction–associated steatohepatitis (MASH) is an essential link in MASLD progression, which has the risk of further development into different degrees of liver fibrosis, even end-stage liver disease such as liver cirrhosis and liver cancer, and is positively correlated with an increase in mortality.[3] Despite its high prevalence and potential impact on public health, effective diagnostic methods for MASLD and its progression to MASH are lacking. The gold standard for diagnosing MASLD is liver biopsy, which is invasive and carries potential risks, including bleeding, infection, and pain. Therefore, there is an urgent need for diagnostic methods that can effectively identify MASLD progression to predict the prognosis and guide treatment.

Immune cell infiltration plays a crucial role in MASLD development and progression. The interaction between immune cells and hepatocytes leads to the production of proinflammatory cytokines and chemokines, which contribute to the development of hepatic steatosis, inflammation, and fibrosis.[4] As a significant immune organ, the liver contains a coordinated network of innate immune cells, including KCs, dendritic cells (DCs), and lymphocytes.[5] While MASLD primarily manifests as a metabolic disorder, it involves inflammatory processes mediated by diverse immune cells, particularly during the progression to MASH, where inflammation is essential.[6] Many studies have found that abnormalities in immune cells occur during the progression of MASLD. MASLD progression requires the release of neutrophil-specific components.[7] One study showed that Th1 cytokines, such as IL-6 and TNF-α, may promote Th1 cell differentiation and contribute to inflammation in MASLD.[8] In addition, B cell–activating factor levels are increased in MASLD progression, and a higher BAFF level in serum is considered a risk factor for MASH.[9] Merad et al found that the number of DCs increased in livers with MASH, and these DCs showed an activated phenotype.[10] Consequently, a systematic study of the characteristics of immune cells during MASLD progression will contribute to developing new treatments for MASLD and MASH.

With the development of microarray and high-throughput sequencing technologies, bioinformatics analysis can be used to identify diagnostic biomarkers for many diseases. For instance, a recent bioinformatics study demonstrated that the secreted protein TSP-2, encoded by THBS2, could be a noninvasive biomarker for identifying MASH and advanced fibrosis.[11]

This study aimed to investigate the dynamic changes in immune cell composition during MASLD progression and identify the diagnostic gene signature associated with this process. First, we analyzed transcriptomic data from liver biopsies of patients at different stages of MASLD, including Normal, MASLD, and MASH, to identify significant differences in immune cell infiltration during MASLD progression. Subsequently, we identified the differentially expressed immune-related genes associated with at least one clinical feature of MASLD. In addition, 4 machine learning methods were employed to construct an ensemble learning model. Finally, we elucidated the vital role of diagnostic genes and immune cells in MASLD progression at the single-cell level. By exploring the immune cell composition, functional changes, and molecular signaling pathways implicated in MASLD progression, as well as the interaction between immune cells and the pathogenesis of MASLD, this study provides a deeper understanding of the underlying mechanisms driving the progression of MASLD.

METHODS

Data collection and preprocessing

The microarray data sets E-MEXP-3291,[12] GSE33814,[13] GSE48452,[14] GSE61260,[15] GSE66676,[16] GSE89632,[17] and GSE164760;[18] the RNA-Seq data sets GSE126848,[19] GSE130970,[20] and GSE135251;[21] and the single-cell RNA-sequencing (scRNA-seq) data sets GSE174748[22] and GSE189175[23] were obtained from the ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) and Gene Expression Omnibus databases (https://www.ncbi.nlm.nih.gov/geo/), respectively. Clinical information associated with the data sets is presented in Supplemental Table S1, http://links.lww.com/HC9/A591.

The raw microarray data underwent preprocessing and normalization using the Robust Multiarray Average method implemented in the R package “affy.” For instances in which one gene corresponded to multiple probes, the average value was calculated. Additionally, the read counts of the RNA-seq data sets were converted to transcripts per million for further analysis.

Integration of immune cell gene sets

To identify immune cell signatures, 66 gene sets were collected from Wang and colleagues (Supplemental Table S2, http://links.lww.com/HC9/A592).[24–32] The Single-Sample Gene Set Enrichment Analysis method was used to assess the immune infiltration of each sample, which was computed based on the expression levels of immune cell signature genes. The R package “limma” was used to examine the differences in immune cell infiltration between the 2 comparisons (MASLD vs. Normal and MASH vs. MASLD). Pearson correlation test was conducted to evaluate the correlation between distinct immune cell infiltration, while R package

"corrplot" was used to visualize the correlation between immune cells.

Genes associated with MASLD progression

Six microarray data sets were merged into a single expression matrix after batch correction. The differentially expressed genes (DEGs) in 3 comparison groups (MASH vs. Normal, MASH vs. MASLD, and MASLD vs. Normal) were identified using the R package "limma." DEGs were defined as genes that met the threshold of $|\log2FC| > 0.3$ and p-value < 0.05. Overall, the DEGs from the 3 comparisons were combined.

Subsequently, all DEGs were clustered using the R package "Mfuzz" based on the order of MASLD progression (Normal, MASLD, and MASH), which clustered genes with similar expression patterns in the same group. The DEGs were divided into 10 clusters with a membership threshold of 0.25. Clusters of DEGs showing consistent expression patterns across different stages were merged and considered MASLD progression–related genes.

Functional enrichment analysis

To investigate the function and pathways of specific genes, we performed gene ontology and Kyoto Encyclopedia of Genes and Genomes pathway analyses using the R package "clusterProfiler." We screened for statistically significant gene ontology terms and signaling pathways with a threshold of $p < 0.05$. The results of the enrichment analysis were visualized using R packages "treemap" and "ggplot2." Gene set variation analysis was used to analyze functional differences in different cell populations. The hallmark gene set for gene set variation analysis was downloaded from the Molecular Signatures Database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp).

Immune-related coexpression networks

To identify coexpression modules with a high correlation to immune cell infiltration in MASLD, we employed the R package "weighted gene coexpression network analysis" to construct a weighted gene coexpression network and cluster all genes into different modules. We chose a scale-free R^2 threshold of 0.85 and a soft threshold of $\beta = 5$. The adjacency matrix was then used to calculate the topological overlap matrix. Hierarchical clustering was performed to identify modules containing a minimum of 50 genes. Finally, the Pearson correlation test was used to evaluate the association between modules and traits.

Construction of protein-protein interaction network

In this study, the protein-protein interaction (PPI) network was constructed using the Search Tool for the Retrieval of Interacting Genes/Proteins database (https://string-db.org/).[33] The minimum required interaction score was set to 0.4, and false discovery rate stringency was set to 0.05. The PPI network was visualized using the Cytoscape software. To identify hub genes, the cytoHubba plugin of Cytoscape was used, which scored each node gene using 11 algorithms, including Maximal Clique Centrality, Density of Maximum Neighborhood Component, Maximum Neighborhood Component, Degree, Edge-Percolated Component, BottleNeck, EcCentricity, Closeness, Radiality, Betweenness, and Stress. The top 50 node genes scored by each algorithm were screened for hub gene identification using Supervenn. Identification of clinically relevant hub genes was based on clinical features from the data set GSE48452.

Diagnostic signature for MASLD progression

To avoid overfitting, we used RNA-seq GSE126848 instead of integrated array data to construct the model. We performed the Boruta algorithm with maxRuns = 1000 to select genes that were more strongly associated with MASLD progression based on their importance. To establish an accurate diagnostic signature, we used an ensemble machine learning model based on 3 algorithms: random forest, Extreme Gradient Boosting, and decision tree. To avoid bias due to different methods, we used the R package "caretEnsemble" for this purpose.

To verify the diagnostic effectiveness of the identified diagnostic signature in different MASLD progression stages, we employed the data sets GSE130970, GSE135251, and GSE164760. Additionally, the diagnostic effectiveness of the identified diagnostic signature was evaluated using ROC analysis by R package "pROC," with the area under the ROC curve value used to determine the ability to distinguish between different stages of MASLD progression.

Building and evaluation of predictive nomogram

The R software package "rms" was employed to create a nomogram, which was evaluated by plotting calibration curves to compare the predicted diagnostic probabilities from the plots with the observed diagnostic probabilities. The net benefit of the nomogram was determined using the decision curve analysis curve.

scRNA-seq data preprocessing

We obtained scRNA-seq data sets for MASLD from the GEO database and conducted an analysis using Seurat.[34] To ensure the quality of our research, we first removed low-quality cells that exhibited fewer than 500 unique molecular identifiers and with mitochondrial gene expression levels exceeding 20%. We further refined our dataset to include only cells that expressed between 300 and 7000 genes, with genes expressed in at least three cells used for downstream analysis. Our analysis yielded a total of 9033 cells from healthy samples, 15,236 cells from patients with MASLD, and 8525 cells from patients with MASH. We conducted log normalization of raw unique molecular identifier counts and identified 2000 highly variable genes using FindVariableFeatures. Principal component analysis was performed using the variable genes, and dimensionality reduction was performed using RunPCA. To address the issue of batch effects and heterogeneity between samples, we used harmony to correct these effects. Finally, we conducted a second round of dimensionality reduction using t-Distributed Stochastic Neighbor Embedding and annotated the cell clusters using marker genes from CellMarker.[35] The R package "Scissor" used bulk phenotypes to identify cell subpopulations associated with the progression of MASLD.[36] The R package "AUCell" was employed to identify the active states of the diagnostic signature in T cells, categorizing cells in the top 25% of scores as the high group and the remaining cells as the low groups.

Ligand-receptor interactions between T cell and hepatic stellate cell (HSC)

We investigated ligand-receptor interactions between T cell subgroups and HSCs using the iTALK package. The top 20 ligand-receptor interactions exhibiting high expression levels were selected for further investigation.

RESULTS

The immune characteristics of MASLD progression

The present study aimed to investigate the immune infiltration characteristics at different stages of MASLD. To this end, 6 microarray data sets were selected for immune cell infiltration analysis, comprising 140 normal samples, 107 MASLD samples, and 95 MASH samples. The batch effect of the merged data was effectively mitigated, as demonstrated in Figure 1A, C, which displays the data before and after batch correction. The merged data sets were used for further analysis.

First, the Single-Sample Gene Set Enrichment Analysis algorithm was employed to calculate the infiltration scores of 66 immune cells in the 3 groups, which are presented in a heatmap (Figure 1D and Supplemental Table S3, http://links.lww.com/HC9/A593). Twelve immune cells were found to exhibit significant differences ($p < 0.05$) in the 3 comparisons (MASLD vs. Normal and MASH vs. MASLD), as depicted in a violin plot (Figure 1E). These immune cells include resting memory CD4 T cells, T cells, central memory CD4 T cells, naïve CD4 T cells, CD8 T cells, regulatory T cells, central memory CD8 T cells, mast cells, activated CD8 T cells, effector memory CD4 T cells, resting natural killer cells, and monocytic lineage. Furthermore, the infiltration of immune cells increases with the progression of MASLD. Finally, by analyzing single-cell data obtained from Normal, MASLD, and MASH samples, we successfully identified seven distinct cell types: hepatocytes, endothelial cells, cholangiocytes, HSCs, T cells, DCs, and macrophages (Figure 1F, G). Notably, hepatocytes constituted the highest proportion across all samples (Figure 1H). Among the immune cells, T cells accounted for the highest proportion in all sample groups. Furthermore, there was an escalation in the proportion of T cells, macrophages, and DCs from Normal to MASH groups. These findings suggest that the immune characteristics of liver tissues change with MASLD progression.

Characterizing gene expression alterations during MASLD progression

To identify genes exhibiting differential expression across distinct stages of MASLD, we performed differential expression analysis of 342 samples encompassing Normal, MASLD, and MASH. The analysis identified a set of 68 DEGs from the Normal-MASLD group, 385 DEGs from the MASLD-MASH group, and 537 DEGs from the Normal-MASH group, which met the criteria of $p < 0.05$ and $|\log2FC| > 0.3$ (Figure 2A-C and Supplemental Table S4, http://links.lww.com/HC9/A594). By integrating the three sets of DEGs, we ultimately identified 667 DEGs, which are visually represented in Figure 2D using a Venn diagram.

To determine the gene alteration characteristics during MASLD progression, we conducted Mfuzz cluster analysis to identify the expression trends of DEGs across the three stages (Figure 2E and Supplemental Table S5, http://links.lww.com/HC9/A595). The initial set of 667 DEGs identified was used as input for the Mfuzz analysis, resulting in the identification of 10 clusters that were categorized into 3 groups: upregulated (clusters 2, 4, and 5), downregulated (clusters 3 and 7), and discontinuous change (clusters 1, 6, 8, 9, and 10). Ultimately, we discovered 236 Mfuzz-DEGs belonging to the upregulated group and 117 Mfuzz-DEGs belonging to the downregulated group, demonstrating their

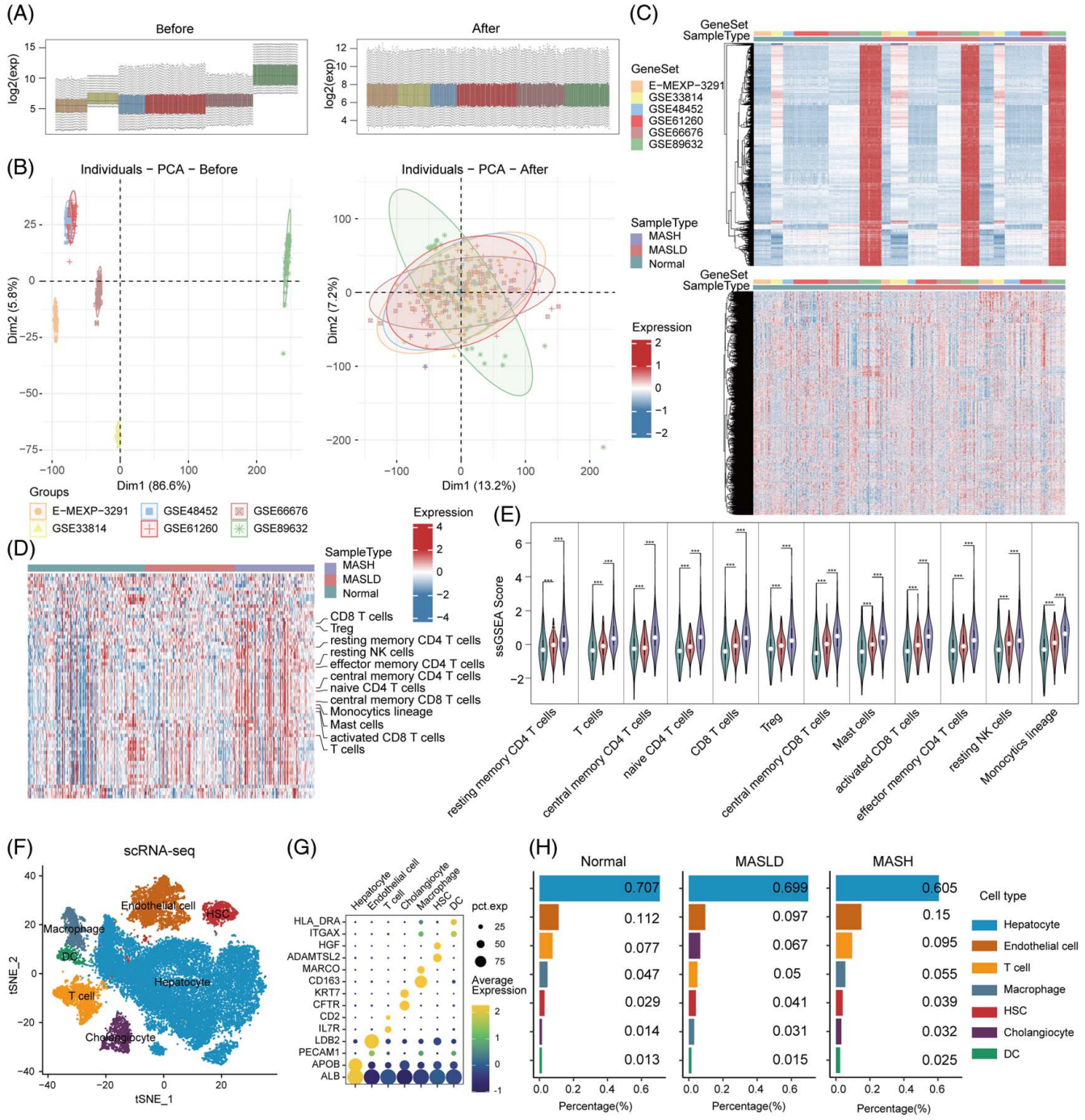


FIGURE 1 Boxplot and principal component analyses of before (left) and after (right) batch correction (A, B). Expression levels of the immune cell–related genes before (top) and after (bottom) batch correction (C). The infiltration score of 12 kinds of immune cells with significant differences in the 2 comparisons was shown in a heatmap and violin plot (D, E). ***$p \leq 0.001$. The main cell types in MASLD are illustrated by the t-SNE plot (F). Dot plot showing average expression of marker genes in each cell type (G). The proportions of the main cell types were ranked from high to low in Normal, MASLD, and MASH samples, respectively (H). Abbreviations: DC, dendritic cell; HLA-DRA, HLA class II histocompatibility antigen, DR alpha chain; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease; PCA, principal component analysis; scRNA-seq, single-cell RNA sequencing; ssGSEA, Single-sample Gene Set Enrichment Analysis; Treg, regulatory T cell; t-SNE, t-Distributed Stochastic Neighbor Embedding.

sustained changes during MASLD progression. We performed gene ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses using clusterProfiler to explore the mechanisms underlying MASLD progression. Our findings revealed a significant enrichment of upregulated Mfuzz-DEGs in immune-related functional pathways, whereas downregulated Mfuzz-DEGs were closely associated with metabolic processes. Additionally, T cell–related pathways were prominent among the top enriched pathways, implying

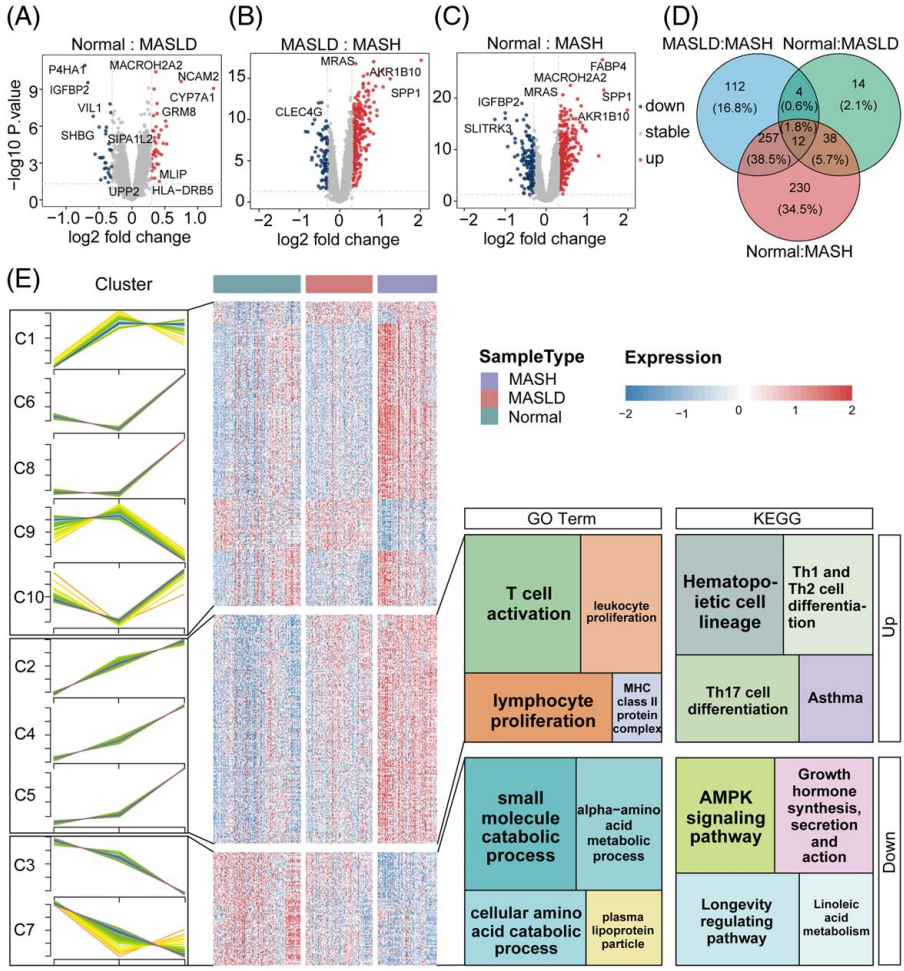


FIGURE 2 Differentially expressed gene in the 3 comparisons were visualized by the volcano map (A, B, C). A Venn diagram showed the combination of the differentially expressed genes (D). Mfuzz cluster analysis and heatmap illustrated the changes in the expression of differentially expressed genes, and the GO term and KEGG pathway enrichment analyses were performed on the upregulated and downregulated gene sets (E). Abbreviations: AMPK, adenosine 5'-monophosphate-activated protein kinase; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease.

the potential significant involvement of T cells during MASLD progression.

Construction of immune-related coexpression networks in MASLD progression

To investigate genes highly correlated with immune cell infiltration during MASLD progression, we constructed a coexpression network by considering immune cell types as phenotypes and identified phenotype-related genes. A power of $\beta = 5$ (scale-free $R^2 = 0.85$) was selected as the soft threshold to ensure a scale-free network (Figure 3A). Subsequently, hierarchical clustering and dynamic tree clipping were employed to identify ten modules (Figure 3B). We calculated correlation coefficients (Cor) to determine the correlation between the modules and immune cell types. The results in Figure 3C show that the black module was the most positively correlated module with T cells (Cor = 0.7, p = 6e − 49), whereas the dark red module exhibited the highest negative correlation with mast cells (Cor = −0.51, p = 8e − 23). Additionally, the black module displayed significant positive correlations with the other immune cell types. Therefore, the black and dark red modules were selected as crucial modules associated with immune cell infiltration for further analysis.

Identification of hub genes associated with clinical features in MASLD progression

An overlap of 157 genes was observed between the Mfuzz-DEGs and the genes in the crucial weighted gene coexpression network analysis modules (Figure 4A). To understand the functional mechanisms involved in MASLD progression by these overlapping genes, we constructed a PPI network comprising 157 overlapping genes. Using the

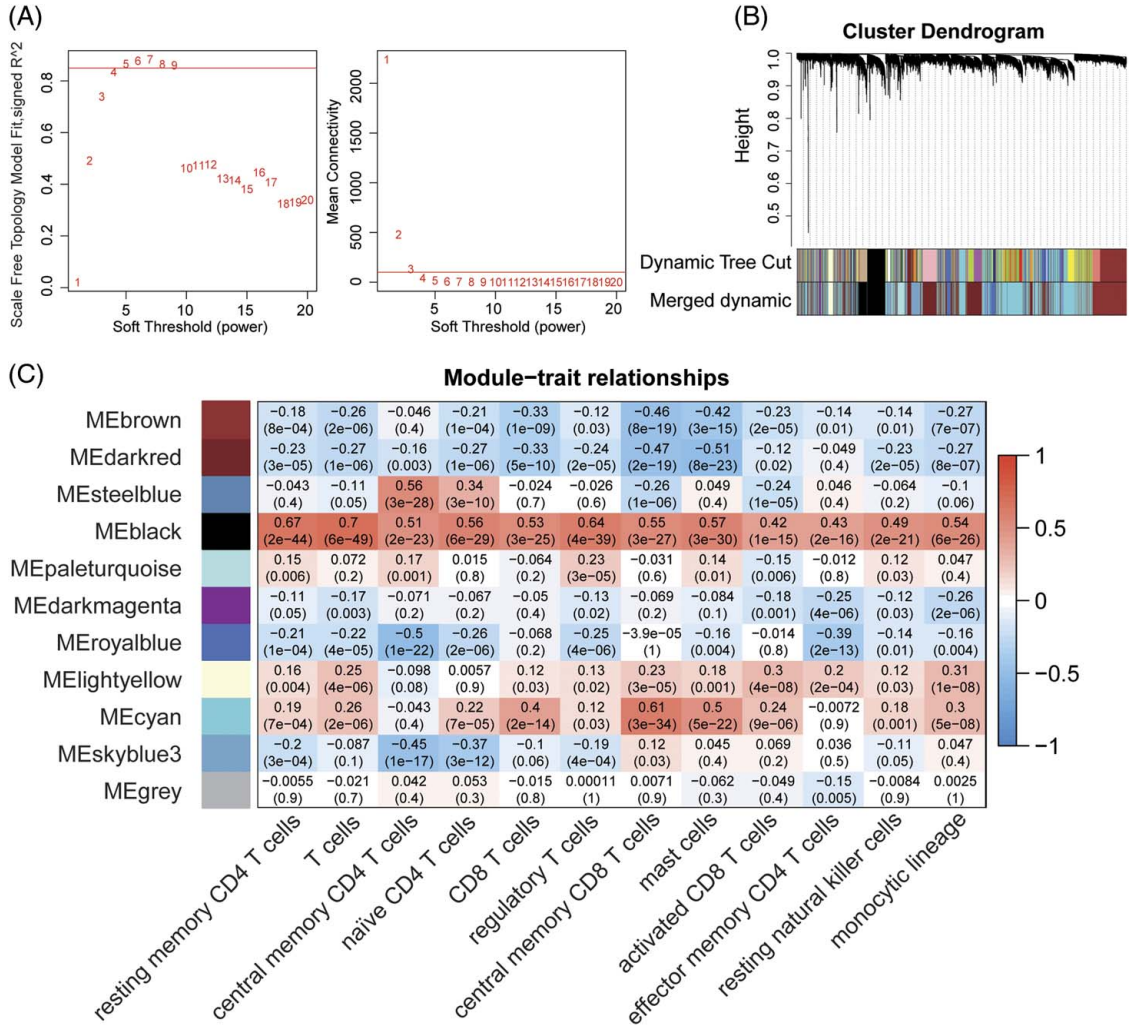


FIGURE 3 Calculating soft-thresholding power. Left: scale-free fit indices using different soft-thresholding powers (β); Right: mean connectivity using different soft-thresholding powers (A). The dendrogram is clustered by the Dynamic Tree Cut algorithm (B). The heatmap profiling the correlations between module eigengenes and the immune cell characteristics (C).

Search Tool for the Retrieval of Interacting Genes database, we set an interaction score threshold of 0.4 and removed isolated genes. Subsequently, the 113 filtered genes were classified into eight modules (immune response, fibrosis, lipid metabolism, inflammation, apoptosis, cell adhesion, oxidative stress, and others) (Figure 4B). The PPI network depicted the interactions among these functional gene modules. Employing 11 algorithms, we quantified the importance of each gene, and 16 hub genes were ultimately obtained (Figure 4C and Supplemental Table S6, http://links.lww.com/HC9/A596).

Furthermore, we investigated the correlation between 5 clinical features [body mass index, steatosis, inflammation, NAFLD activity score, and fibrosis] and the progression of MASLD. NAFLD activity score, a comprehensive scoring system for MASLD, provides an overall assessment based on 3 histological features: steatosis, ballooning, and lobular inflammation. The degree of fibrosis plays a crucial role in the progression of MASLD. Analyzing Normal to MASH samples, we observed an ascending trend in the 5 clinical feature indices, which exhibited significant positive correlations with each other (Figure 4D, E). To identify genes with clinical significance, we evaluated the correlation between 16 candidate genes and the 5 clinical indices, subsequently excluding those without significant associations. Finally, we successfully identified 14 genes that demonstrated substantial correlations with MASLD clinical features (Figure 4F, G).

Construction and validation of a 2-gene diagnostic signature for MASLD progression

To enhance the accuracy of diagnosing MASLD progression, we performed feature selection based on 14 hub genes with clinical significance separately in Normal-MASLD/MASH and MASLD-MASH subgroups (Figure 5A, B and Supplemental Table S7, http://links.lww.com/HC9/A597). Nine genes were identified in the Normal-MASLD/MASH

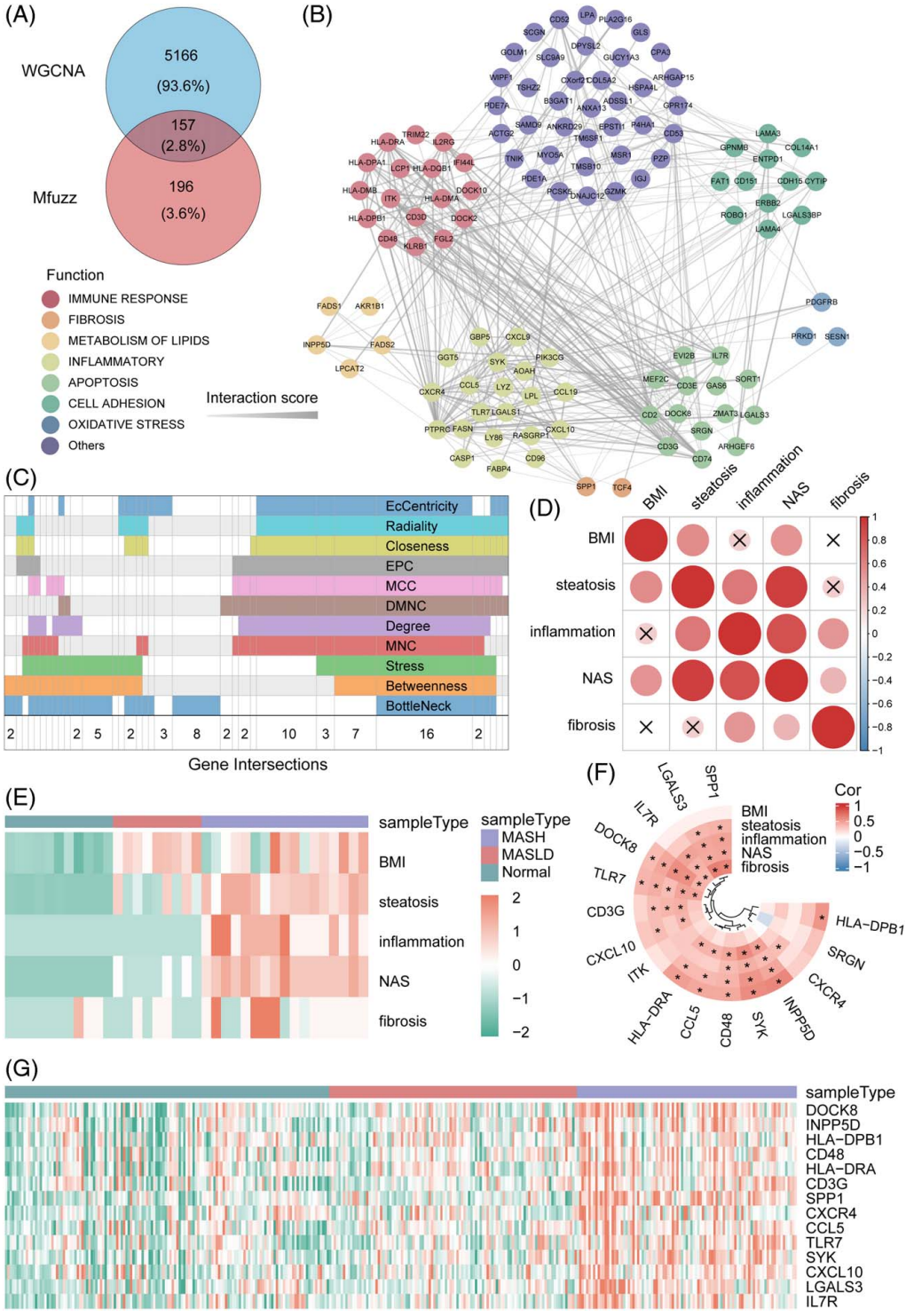


FIGURE 4 A Venn diagram was conducted to obtain the overlapping genes between the Mfuzz-differentially expressed genes and the genes in the crucial WGCNA modules (A). The PPI network among 113 genes and 11 algorithms to screen hub genes (B, C). Correlation analysis among clinical features (D). ×, Pearson correlation *p*-value > 0.05. Quantitative levels of clinical features indices in the samples (E). The visualization of the Pearson correlation between hub gene and clinical features (F). *Pearson correlation *p*-value ≤ 0.05. Expression levels of 14 hub genes with clinical significance in the samples (G). Abbreviations: CXCR4, C-X-C motif chemokine receptor 4; DMNC, Density of Maximum Neighborhood Component; EPC, edge-percolated component; MCC, maximal clique centrality; NAS, NAFLD activity score; PPI, protein-protein interaction; WGCNA, weighted gene coexpression network analysis.

group, including CXCL10, INPP5D, SPP1, dedicator of cytokinesis 8 (DOCK8), SYK, LGALS3, HLA class II histocompatibility antigen , DR alpha chain (HLA-DRA), C-X-C motif chemokine receptor 4 (CXCR4), and CD48. On the other hand, in the MASLD-MASH group, only 3 genes were identified, including CXCR4, TLR7, and DOCK8.

Notably, CXCR4 and DOCK8 were common features in both groups.

Based on the 2 Boruta algorithm–selected features, we employed 3 machine learning methods to develop ensemble learning models. To create effective base learners, we performed 10-fold cross-validation for hyperparameter tuning. Furthermore, we conducted ROC analysis to validate the diagnostic efficiency of MASLD and MASH progression models (Supplemental Table S8, http://links.lww.com/HC9/A598). As shown in Figure 5C, an MASLD/MASH diagnosis model was created based on the CXCR4 and DOCK8 signature, with an AUC of over 0.9, which indicated that the model had excellent diagnostic accuracy. As shown in Figure 5D, a MASH diagnosis model was created, with an AUC of over 0.8, showing great diagnostic accuracy.

Subsequently, we validated the diagnostic ability of the models using external validation data sets. As shown in Figure 5E, F, the AUC of the MASLD/MASH diagnosis model was 0.785, whereas that of the MASH diagnosis model was 0.733. We further verified the diagnostic effectiveness of CXCR4 and DOCK8 in early (F0–F2) to mid-late (F3–F4) MASH fibrosis and MASH-HCC by ROC analysis. As shown in Figure 5G, H, the AUC of the 2-gene diagnostic signature in early to mid-late MASH fibrosis reached 0.6, and the AUC in MASH-HCC reached 0.826. Overall, the CXCR4 and DOCK8 diagnostic signatures demonstrated strong predictive ability in MASLD progression.

Development and evaluation of nomograms based on a 2-gene diagnostic signature for MASLD progression

To improve the accuracy of diagnosing different stages of MASLD, we constructed nomograms based on the 2-gene diagnostic signature of MASLD/MASH and MASH (Figure 6A, D). Furthermore, we used calibration plots of the nomograms to demonstrate close agreement between the predicted and actual diagnoses (Figure 6B, E). To evaluate the performance of the prediction models, decision curve analysis was employed to calculate the net benefit without additional clinical information such as body mass index or NAFLD activity score. As shown in Figure 6C, F, the 2-gene diagnostic signature of MASLD/MASH provided a superior net benefit when the decision threshold was between 0 and 1. This finding suggests that the net benefit of the prediction model was superior to that of all treatments and no treatment. The threshold probability for MASH diagnosis ranged from 0 to 0.66, and the net benefit of the predictive model was negative when the threshold probability exceeded 0.66. Hence, the predictive model can be employed for MASH diagnosis when the threshold probability is between 0 and 0.66. These results imply that CXCR4 and DOCK8 diagnostic signatures hold promise as valuable tools for clinical practice.

Single-cell analysis reveals the mechanism of T-cell involvement in the progression of MASLD

We employed the Scissor algorithm to identify cell types associated with MASLD progression (Scissor-positive) (Figure 7A). Notably, besides hepatocytes, cholangiocytes, and HSCs, the majority of T cells were also connected to the progression of MASLD (Figure 7B). Furthermore, CXCR4 and DOCK8 exhibited high expression in T cells (Figure 7C). Based on AUCell scores for CXCR4 and DOCK8, we observed an enrichment of T cells with high scores within cells linked to MASLD deterioration (60%) (Figure 7D, E). Additionally, gene set variation analysis analysis unveiled that T cells with high scores shared akin functionalities with Scissor-positive T cells, encompassing interferon alpha and gamma responses and DNA repair. Conversely, T cells with low scores displayed similarities to Scissor-negative T cells, being enriched in functions such as inflammatory responses and epithelial-mesenchymal transition (Figure 7F). These findings imply that CXCR4-regulated and DOCK8-regulated T cells might participate in MASLD progression through distinct functional mechanisms.

In total, 2265 T cells were clustered into four distinct subtypes: T-C1, T-C2, T-C3, and T-C4 (Figure 8A). The identification and characterization of T-cell subtypes were conducted using the expression of functional markers and marker genes (Figure 8B). Specifically, T-C1 exhibited an increased expression of genes associated with naïve T cells, T-C2 displayed high expression levels of Chemokine genes, T-C3 showed heightened expression of proliferation markers, and T-C4 presented elevated expression of cytotoxicity-related genes. Notably, high-scored T cells were enriched in T-C1 (87.4%), T-C2 (71.2%), and T-C4 (60.2%), whereas low-scored T cells were abundant in T-C3 (54.2%) (Figure 8C). Furthermore, the proportion of T-C3 increased from Normal to MASH, while the proportions of T-C1, T-C2, and T-C4 decreased, with a reduction in the abundance of high-scored T cells (Figure 8D). These findings indicate the specific expression patterns of CXCR4 and DOCK8 in T cells across different stages of MASLD, suggesting the potential involvement of CXCR4 and DOCK8 in MASLD progression by regulating T-cell subpopulations in distinct mechanisms.

As HSCs play a crucial role in liver fibrosis, we examined the molecular interactions between T cells and HSCs at various stages of MASLD progression to elucidate the underlying molecular mechanism of T cells in MASLD (Figure 8E–G). Using iTALK, we constructed a cell communication network based on

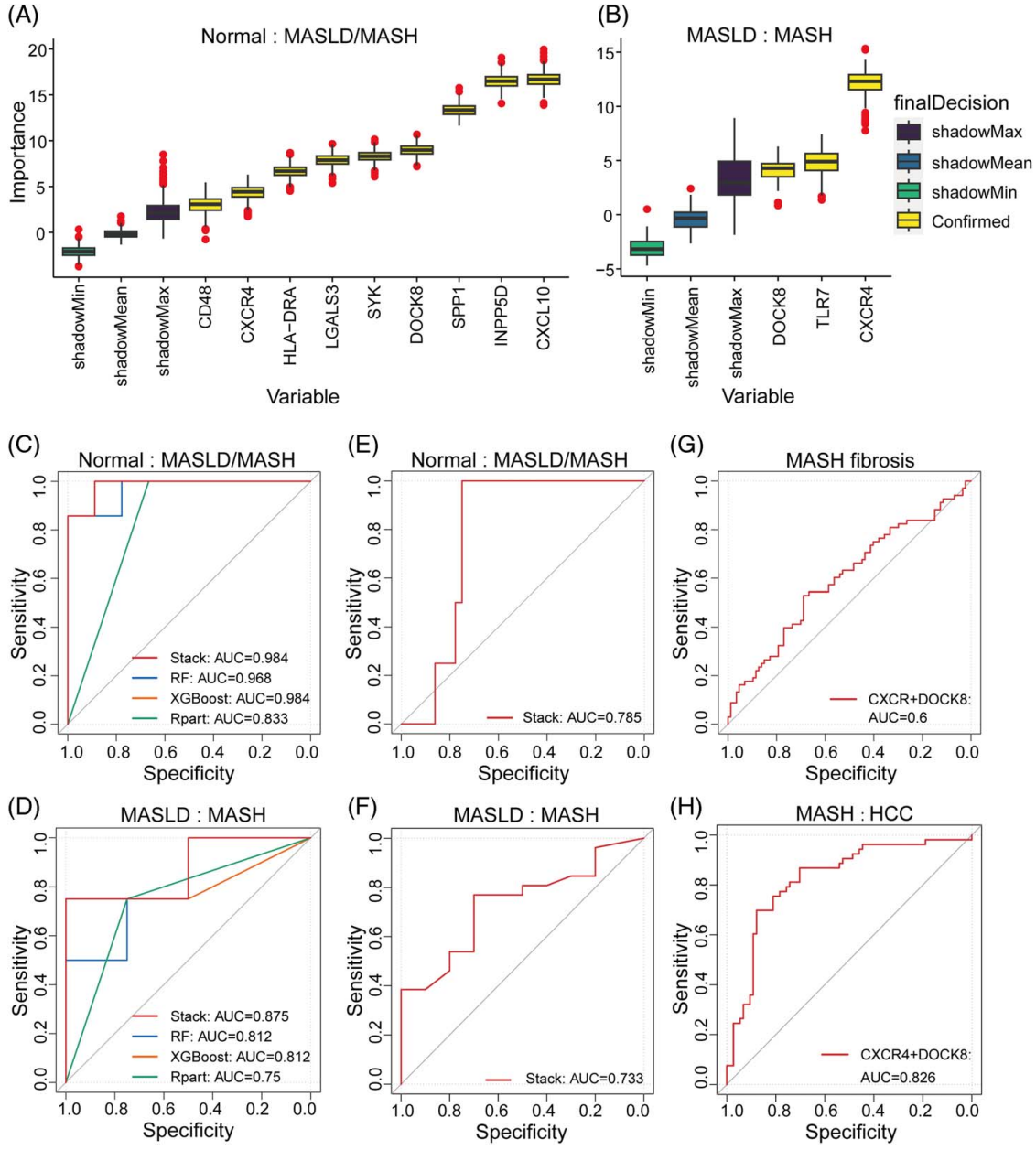


FIGURE 5 Feature selection in two groups based on the Boruta algorithm (A, B). Yellow indicates confirmed features, while other colors indicate unimportant shadow attributes. The ROC analysis of the ensemble model in the testing dataset (C, D) and verification data set (E, F) was shown. The ROC of diagnostic signature in MASH fibrosis (G) and MASH-HCC (H) cohorts. Abbreviations: CXCR4, C-X-C motif chemokine receptor 4; DOCK8, dedicator of cytokinesis 8; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease; RF, Random Forest.

potential receptor–ligand interactions. From Normal to MASH, the interactions between HSCs and high-scored T cells (T-C1 and T-C2) were intensified, while interactions with low-scored T cells (T-C3) were weakened. In normal samples, the influence of HSCs on T-C1 and T-C2 through HGF-CD44 and FN1-CD44 interactions suggests a potential role of early-stage HSCs in promoting T-cell activation and differentiation. However, interactions between HSCs and T-C4 were amplified in NAFL samples, revealing emerging interactions involving FGA-ITGAV, FGA-ITGB1, FGB-ITGAV, and FGB-ITGB1, which indicates that following activation of HSCs during liver injury, T cells might mediate the migration and differentiation of HSCs through integrins. In MASH samples, interactions between HSCs and T-C1, as well as T-C2, were augmented, impacting both T-cell subgroups through THBS1-ITGA4, THBS1-ITGA6, FN1-CD44, FN1-ITGA4, and FN1-ITGA6. Due to the stimulation of inflammatory responses, T cells underwent substantial proliferation. The above results suggest that T cells possess the potential to engage in interactions with

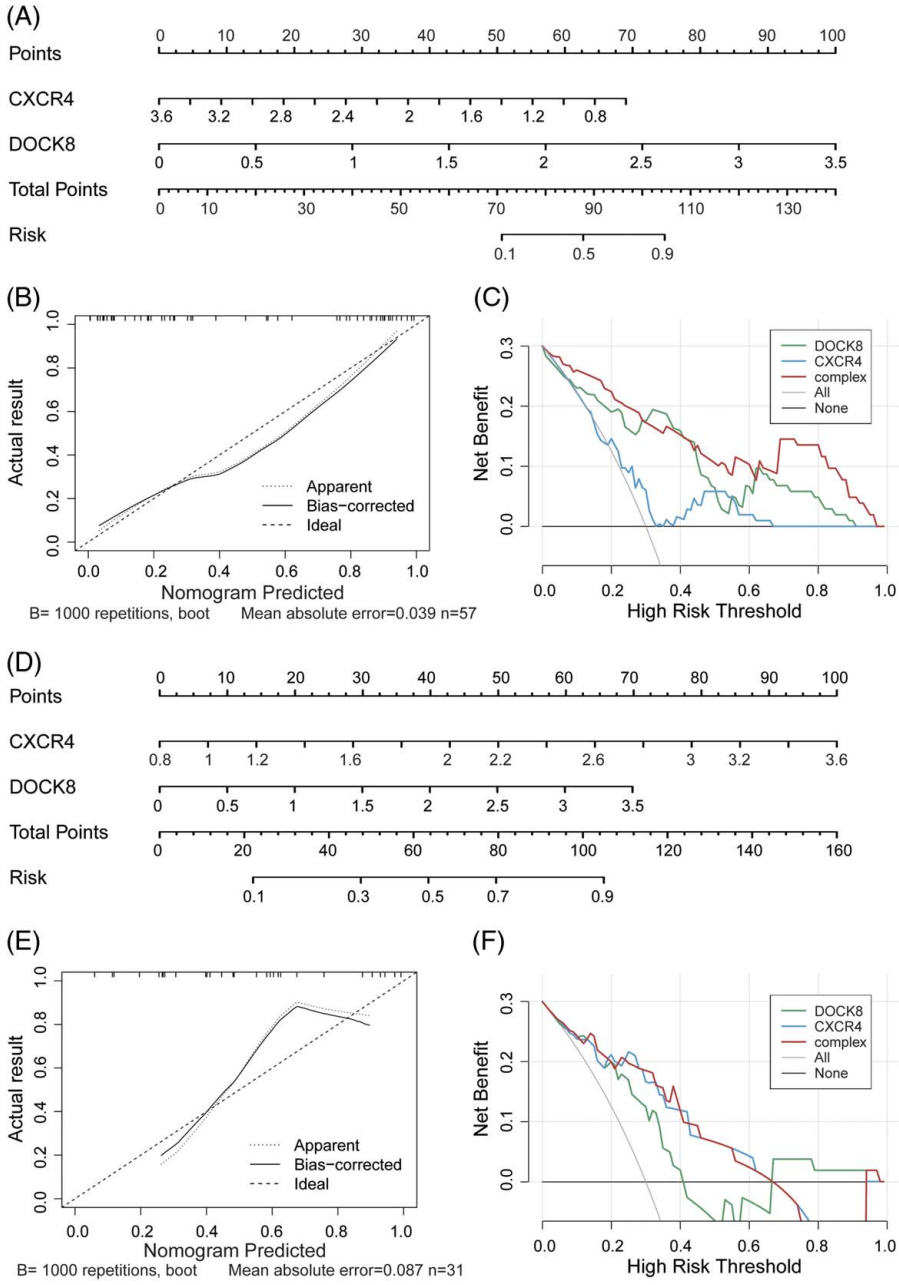


FIGURE 6 Nomograms were developed to diagnose Normal-MASLD/MASH (A) and MASLD-MASH (D) cohorts, and their predictive performance was assessed using calibration curves (B, E) and decision curve analysis (C, F). Abbreviations: CXCR4, C-X-C motif chemokine receptor 4; DOCK8, dedicator of cytokinesis 8; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease.

HSCs through specific cell adhesion molecules, thus participating in the progression of MASLD.

DISCUSSION

MASLD is a complex and multifactorial disease that is rapidly increasing in incidence and prevalence worldwide, and is considered the leading cause of chronic liver disease. MASLD is characterized by hepatic steatosis, which is often considered a milder condition that may not always exhibit symptoms. However, some patients with MASLD may progress to MASH over time. MASH is a more severe form of MASLD characterized by inflammation, hepatocellular injury, and fat accumulation in the liver and is associated with an increased risk of developing cirrhosis, liver failure, and liver cancer.[37] The diagnosis of MASLD and MASH typically involves a thorough medical history, physical examination, blood tests, and imaging studies; in some cases, a liver biopsy may be necessary. However, the invasiveness of the "gold standard" liver biopsy limits its widespread use.[38] Furthermore, existing imaging studies (such as ultrasound and elastography), when

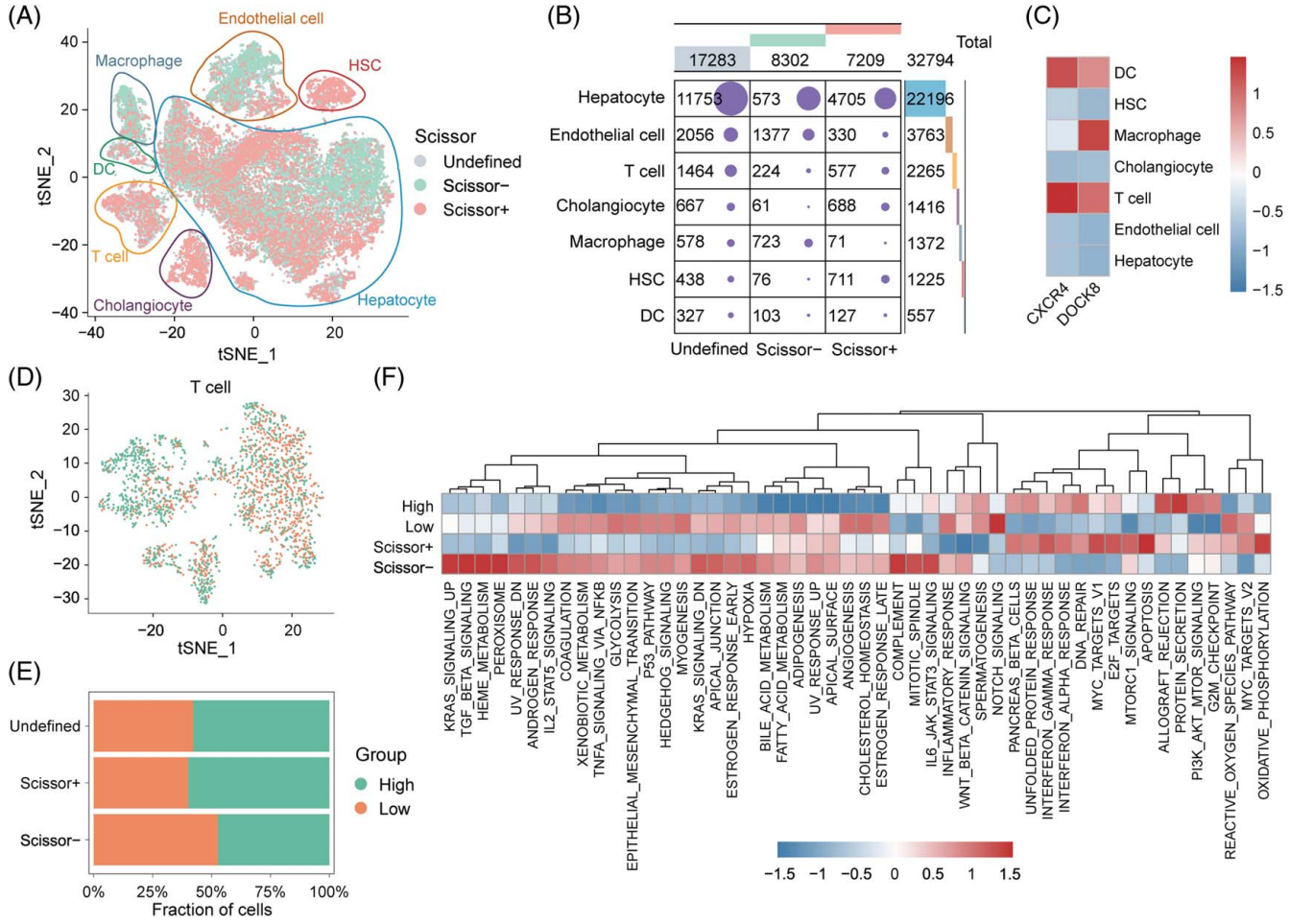


FIGURE 7 Cell types associated with MASLD progression identified by Scissor are illustrated by the t-SNE plot (A). The number of MASLD progression–associated cells in the major cell types (B). Heatmap showing normalized average expression of CXCR4 and DOCK8 in each cell type (C). t-SNE plot of high-scored and low-scored T cells (D). Proportions of high-scored and low-scored T cells in the MASLD progression–associated T cell (E). Hallmark pathways assessed by gene set variation analysis in high-scored, low-scored, and MASLD progression–associated T cells (F). Abbreviations: CXCR4, C-X-C motif chemokine receptor 4; DC, dendritic cell; DOCK8, dedicator of cytokinesis 8; MASLD, metabolic dysfunction–associated steatotic liver disease; t-SNE, t-Distributed Stochastic Neighbor Embedding.

combined with other indicators, provide only rough estimates of the degree of MASLD, advanced liver steatosis, and cirrhosis. In contrast, effective non-invasive assessment methods to identify patients with MASLD and MASH are still lacking.

This study presents a novel approach for diagnosing and monitoring the progression of MASLD, based on the analysis of immune cell infiltration in liver biopsies. Specifically, this study aimed to identify the diagnostic signature of MASLD progression using immune cell infiltration analysis and integrated a new set of immune cells to enhance the analysis. The results of the immune infiltration analysis were consistent with those of previous studies. As MASLD progresses (ie, from healthy to MASLD to MASH), there is a significant increase in the infiltration of various immune cells, including resting memory CD4 T cells, T cells, central memory CD4 T cells, naive CD4 T cells, CD8 T cells, regulatory T cell, central memory CD8 T cells, mast cells, activated CD8 T cells, effector memory CD4 T cells, resting natural killer cells, and monocytic lineage. This observation suggests that immune cells, especially T cells, are crucial for MASLD progression.

T cells play a distinct role in providing a potent adaptive immune response to potential threats to the liver. These cells participate in a range of processes including viral clearance, inflammation regulation, tissue repair, and maintenance of liver homeostasis.[39] In liver disease, T cells are involved in both the development and resolution of liver inflammation.[40] Moreover, some studies have suggested that CD4+ T cells also contribute to the promotion of liver inflammation and fibrosis through the production of proinflammatory cytokines such as IL-17 and IFN-γ.[41] In addition, T cells may promote chronic inflammation observed in MASLD by activating HSCs and promoting fibrosis, and they may also play a role in regulating liver lipid metabolism.[42] Furthermore, CD8+ T cells have been shown to promote MASLD progression by inducing hepatocyte death and fibrosis.[43]

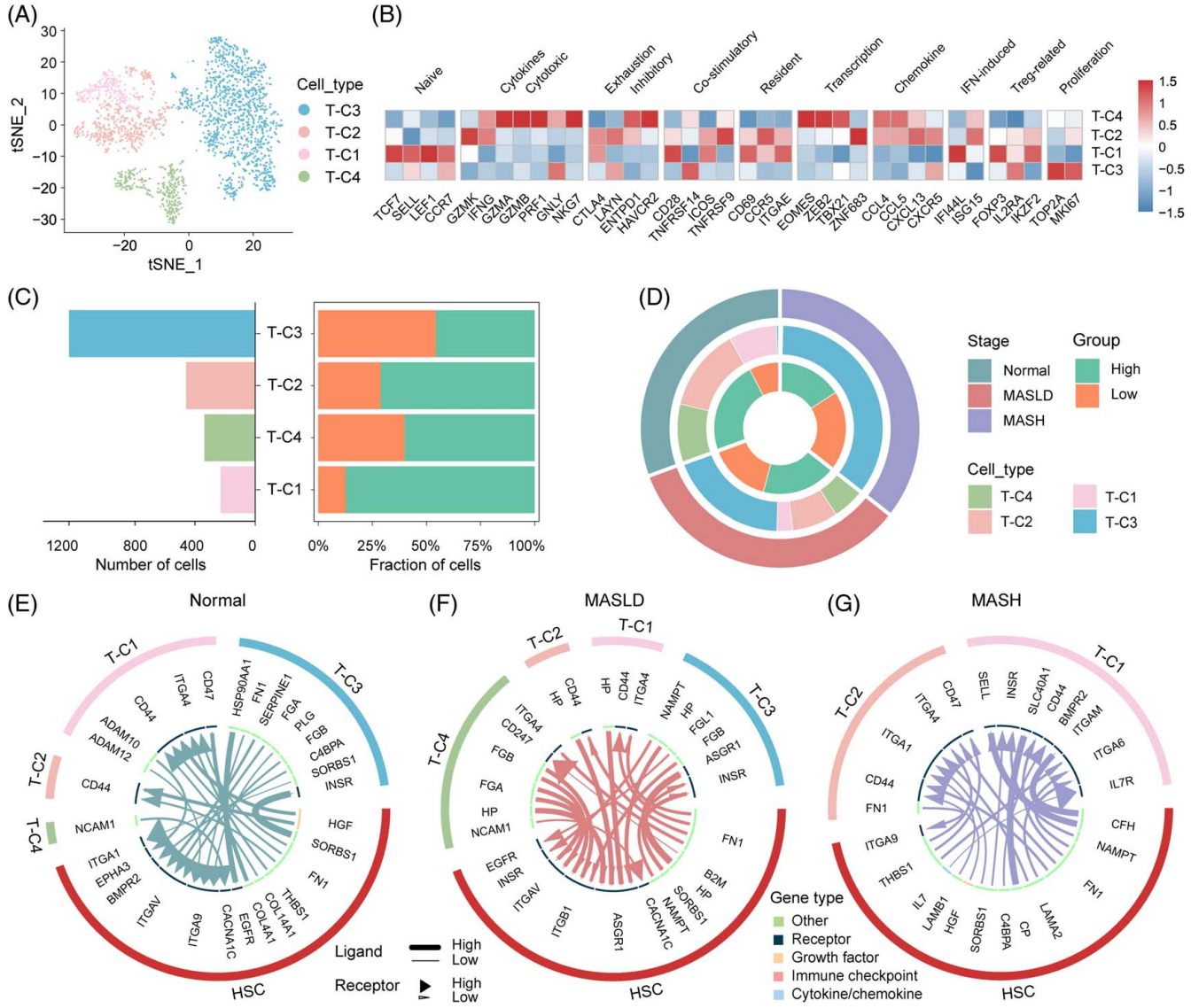


FIGURE 8 T-SNE plot of 2265 T cells (A). Heatmap showing the normalized average expression of functional markers in T-cell subtypes (B). Proportions of high-scored and low-scored T cells in each T-cell subtype (C). Proportions of T-cell subtypes in Normal, MASLD, and MASH samples (D). Ligand-receptor interaction in Normal (E), MASLD (F), and MASH (G). Abbreviations: CFH, complement factor H; CP, ceruloplasmin; EGFR, epidermal growth factor receptor; FGB, fibrinogen beta chain; FGA, fibrinogen alpha chain; HP, haptoglobin; ITGAV, integrin alpha-V; MASH, metabolic dysfunction–associated steatohepatitis; MASLD, metabolic dysfunction–associated steatotic liver disease; PLG, plasminogen; t-SNE, t-Distributed Stochastic Neighbor Embedding.

The present study successfully identified a pivotal diagnostic signature consisting of 2 genes, DOCK8 and CXCR4, which exhibit significant associations with T cells and are intricately linked to the progression of MASLD. Furthermore, the diagnostic potential of these genes in predicting MASLD progression was validated across diverse cohorts, including Normal-MASLD/MASH, MASLD-MASH, MASH fibrosis, and MASH-HCC, underscoring their critical relevance in the progression of MASLD. Moreover, confirmation at the single-cell level has further underscored the significance of these genes in MASLD progression, exerting an influence on T cells through various mechanisms encompassing epithelial-mesenchymal transition, inflammatory responses, and metabolism-related pathways.

CXCR4 encodes C-X-C chemokine receptor type 4, a chemokine receptor protein. It is located on chromosome 2 in humans and is widely expressed in various tissues, including the liver. CXCR4 plays a crucial role in immune responses and cell migration by binding to its ligand, stromal-derived factor 1.[44] In addition to immune cell recruitment, CXCR4 is involved in many functions and plays a pivotal role in the pathogenesis of liver disease.[45] The aberrant expression pattern of CXCR4 is related to the migration and movement of liver-specific cells in liver disease through cross talk with other signaling pathways.[46] In addition, studies

have reported that overexpression of CXCR4 in the liver in chronic liver injury decreases the expression of hepatocyte-specific genes and promotes liver injury.[47]

DOCK8 encodes the DOCK8 protein, which plays a crucial role in regulating various cellular processes, including cell adhesion, migration, and the immune response.[48] It is reported that DOCK8 plays a pivotal role in the survival and function of T cells.[49] Additionally, mutations in DOCK8 have been associated with DOCK8 deficiency syndrome, also known as autosomal recessive hyper-IgE syndrome, a rare immunodeficiency disorder.[50]

This study had several limitations that warrant discussion. First, the analysis solely hinges on transcriptomics, without the integration of multiomics data, thereby limiting the holistic approach to data interpretation. Second, the study falls short of addressing the pressing issue of noninvasive MASLD diagnosis, as further clinical experiments on blood samples are warranted. Finally, the functional interplay between immune cells and other hepatic cell types, including hepatocytes and endothelial cells, has not been evaluated, potentially undermining the nuanced understanding of the mechanistic underpinnings of MASLD progression.

In summary, this study employed a comprehensive analysis of bulk, scRNA-Seq and clinical information to investigate immune cell infiltration associated with MASLD progression. Through bioinformatics analysis, a 2-gene diagnostic signature relevant to MASLD progression was identified, including the CXCR4 and DOCK8 genes, which exhibited significant correlations with clinical characteristics. Moreover, this 2-gene diagnostic signature effectively differentiated the Normal-MASLD/MASH, MASLD-MASH, MASH fibrosis, and MASH-HCC cohorts. Finally, at the single-cell level, the crucial roles of diagnostic genes and immune cells in MASLD progression were elucidated, suggesting that CXCR4 and DOCK8 may facilitate MASLD progression by influencing signaling pathways and molecular interactions in T cells. Consequently, this study provides a fresh perspective on the diagnosis and underlying immune-related molecular mechanisms of MASLD by identifying immune cells and diagnostic gene signature associated with MASLD progression.

DATA AVAILABILITY STATEMENT

The data sets for this study can be found in the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/). Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

Lidao Bao, Yuheng Ma, and Yingqi Xu: conceptualization; Jing Bai, Lun Zhu, and Wanqi Mi: methodology; Lun Zhu, Zhengzheng Gao, and Wanqi Mi: investigation; Lin Song, Jing Bai, and Zhengzheng Gao: writing—original draft; Lun Zhu and Wanlu Sheng: writing-review and editing; Yuheng Ma and Yingqi Xu: funding acquisition; Zhengzheng Gao and Minyue Ouyang: resources.

FUNDING INFORMATION

This work was supported by the National Natural Science Foundation of China [61903105, 62172131, 62102123, 82160794], the Heilongjiang Postdoctoral Science Foundation [LBH-Q19045], the Heilongjiang Provincial Natural Science Foundation [YQ2021C026], the Key Technology Project Foundation of Inner Mongolia [2021GG0176], and Major projects of Inner Mongolia Natural Science Foundation [No.2023ZD15].

CONFLICTS OF INTEREST

The authors have no conflicts to report.

ORCID

Yingqi Xu https://orcid.org/0000–0001–6879–639X

REFERENCES

1. Nassir F. NAFLD: Mechanisms, treatments, and biomarkers. Biomolecules. 2022;12:824.
2. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. A multi-society Delphi consensus statement on new fatty liver disease nomenclature. Hepatology. 2023. Publish Ahead of Print. doi:10.1097/HEP.0000000000000520.
3. Powell EE, Wong VW, Rinella M. Non-alcoholic fatty liver disease. Lancet. 2021;397:2212–4.
4. Wang H, Mehal W, Nagy LE, Rotman Y. Immunological mechanisms and therapeutic targets of fatty liver diseases. Cell Mol Immunol. 2021;18:73–91.
5. Huby T, Gautier EL. Immune cell-mediated features of non-alcoholic steatohepatitis. Nat Rev Immunol. 2021;22:429–43.
6. Haas JT, Vonghia L, Mogilenko DA, Verrijken A, Molendi-Coste O, Fleury S, et al. Transcriptional network analysis implicates altered hepatic immune function in NASH development and resolution. Nat Metab. 2019;1:604–14.
7. Kim KH, Cheng N, Lau LF. Cellular communication network factor 1-stimulated liver macrophage efferocytosis drives hepatic stellate cell activation and liver fibrosis. Hepatol Commun. 2022;6:2798–811.
8. Hammoutene A, Rautou PE. Role of liver sinusoidal endothelial cells in non-alcoholic fatty liver disease. J Hepatol. 2019;70:1278–91.
9. Deng CJ, Lo TH, Chan KY, Li X, Wu MY, Xiang Z, et al. Role of B lymphocytes in the pathogenesis of NAFLD: A 2022 update. Int J Mol Sci. 2022;23:12376.
10. Merad M, Sathe P, Helft J, Miller J, Mortha A. The dendritic cell lineage: Ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. Annu Rev Immunol. 2013;31:563–604.
11. Kozumi K, Kodama T, Murai H, Sakane S, Govaere O, Cockell S, et al. Transcriptomics identify thrombospondin-2 as a biomarker for NASH and advanced liver fibrosis. Hepatology. 2021;74:2452–66.
12. Lake AD, Novak P, Fisher CD, Jackson JP, Hardwick RN, Billheimer DD, et al. Analysis of global and absorption,

distribution, metabolism, and elimination gene expression in the progressive stages of human nonalcoholic fatty liver disease. Drug Metab Dispos. 2011;39:1954–60.
13. Starmann J, Fälth M, Spindelböck W, Lanz KL, Lackner C, Zatloukal K, et al. Gene expression profiling unravels cancer-related hepatic molecular signatures in steatohepatitis but not in steatosis. PLoS One. 2012;7:e46584.
14. Ahrens M, Ammerpohl O, von Schönfels W, Kolarova J, Bens S, Itzel T, et al. DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. Cell Metab. 2013;18: 296–302.
15. Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schönfels W, Ahrens M, et al. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci USA. 2014;111:15538–43.
16. Xanthakos SA, Jenkins TM, Kleiner DE, Boyce TW, Mourya R, Karns R, et al. High prevalence of nonalcoholic fatty liver disease in adolescents undergoing bariatric surgery. Gastroenterology. 2015;149:623–34 e8.
17. Arendt BM, Comelli EM, Ma DWL, Lou W, Teterina A, Kim T, 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: 1565–78.
18. Pinyol R, Torrecilla S, Wang H, Montironi C, Piqué-Gili M, Torres-Martin M, et al. Molecular characterisation of hepatocellular carcinoma in patients with non-alcoholic steatohepatitis. J Hepatol. 2021;75:865–78.
19. Suppli MP, Rigbolt KTG, Veidal SS, Heebøll S, Eriksen PL, Demant M, et al. Hepatic transcriptome signatures in patients with varying degrees of nonalcoholic fatty liver disease compared with healthy normal-weight individuals. Am J Physiol Gastrointest Liver Physiol. 2019;316:G462–72.
20. Hoang SA, Oseini A, Feaver RE, Cole BK, Asgharpour A, Vincent R, et al. Gene expression predicts histological severity and reveals distinct molecular profiles of nonalcoholic fatty liver disease. Sci Rep. 2019;9:12541.
21. Govaere O, Cockell S, Tiniakos D, Queen R, Younes R, Vacca M, et al. Transcriptomic profiling across the nonalcoholic fatty liver disease spectrum reveals gene signatures for steatohepatitis and fibrosis. Sci Transl Med. 2020;12:eaba4448.
22. Filliol A, Saito Y, Nair A, Dapito DH, Yu LX, Ravichandra A, et al. Opposing roles of hepatic stellate cell subpopulations in hepatocarcinogenesis. Nature. 2022;610:356–65.
23. Alvarez M, Benhammou JN, Darci-Maher N, French SW, Han SB, Sinsheimer JS, et al. Human liver single nucleus and single cell RNA sequencing identify a hepatocellular carcinoma-associated cell-type affecting survival. Genome Med. 2022; 14:50.
24. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013;39:782–95.
25. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell. 2015;160:48–61.
26. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218.
27. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016; 352:189–96.
28. Aran D, Hu Z, Butte AJ. xCell: Digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18:220.
29. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep. 2017;18: 248–62.
30. Nirmal AJ, Regan T, Shih BB, Hume DA, Sims AH, Freeman TC. Immune cell gene signatures for profiling the microenvironment of solid tumors. Cancer Immunol Res. 2018;6: 1388–400.
31. Miao YR, Zhang Q, Lei Q, Luo M, Xie GY, Wang H, et al. ImmuCellAI: A unique method for comprehensive T-cell subsets abundance prediction and its application in cancer immunotherapy. Adv Sci (Weinh). 2020;7:1902880.
32. Steen CB, Liu CL, Alizadeh AA, Newman AM. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. Methods Mol Biol. 2020;2117:135–57.
33. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1): D638–46.
34. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. Cell. 2019;177:1888–1902 e21.
35. Hu C, Li T, Xu Y, Zhang X, Li F, Bai J, et al. CellMarker 2.0: An updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. Nucleic Acids Res. 2023;51(D1):D870–6.
36. Sun D, Guan X, Moran AE, Wu LY, Qian DZ, Schedin P, et al. Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. Nat Biotechnol. 2022;40: 527–38.
37. Schuppan D, Surabattula R, Wang XY. Determinants of fibrosis progression and regression in NASH. J Hepatol. 2018; 68:238–50.
38. Cotter TG, Rinella M. Nonalcoholic fatty liver disease 2020: The state of the disease. Gastroenterology. 2020;158: 1851–64.
39. Wang Y, Zhang C. The roles of liver-resident lymphocytes in liver diseases. Front Immunol. 2019;10:1582.
40. Bhattacharjee J, Kirby M, Softic S, Miles L, Salazar-Gonzalez RM, Shivakumar P, et al. Hepatic natural killer T-cell and CD8+ T-cell signatures in mice with nonalcoholic steatohepatitis. Hepatol Commun. 2017;1:299–310.
41. Van Herck MA, Weyler J, Kwanten WJ, Dirinck EL, De Winter BY, Francque SM, et al. The differential roles of T cells in non-alcoholic fatty liver disease and obesity. Front Immunol. 2019; 10:82.
42. McVey JC, Green BL, Ruf B, McCallen JD, Wabitsch S, Subramanyam V, et al. NAFLD indirectly impairs antigen-specific CD8(+) T cell immunity against liver cancer in mice. iScience. 2022;25:103847.
43. Mao T, Yang R, Luo Y, He K. Crucial role of T cells in NAFLD-related disease: A review and prospect. Front Endocrinol (Lausanne). 2022;13:1051076.
44. Daniel SK, Seo YD, Pillarisetty VG. The CXCL12-CXCR4/CXCR7 axis as a mechanism of immune resistance in gastrointestinal malignancies. Semin Cancer Biol. 2020;65:176–88.
45. Wang S, Gao S, Li Y, Qian X, Luan J, Lv X. Emerging importance of chemokine receptor CXCR4 and its ligand in liver disease. Front Cell Dev Biol. 2021;9:716842.
46. Yang J, Zhang L, Jiang Z, et al. TCF12 promotes the tumorigenesis and metastasis of hepatocellular carcinoma via upregulation of CXCR4 expression. Theranostics. 2019;9:5810–27.
47. Aguilar-Bravo B, Arino S, Blaya D, Pose E, Martinez Garcia de la Torre RA, Latasa MU, et al. Hepatocyte dedifferentiation profiling in alcohol-related liver disease identifies CXCR4 as a driver of cell reprogramming. J Hepatol. 2023;79: 728–40.

48. Kearney CJ, Randall KL, Oliaro J. DOCK8 regulates signal transduction events to control immunity. Cell Mol Immunol. 2017; 14:406–11.
49. Kunimura K, Uruno T, Fukui Y. DOCK family proteins: Key players in immune surveillance mechanisms. Int Immunol. 2020; 32:5–15.
50. Gharehzadehshirazi A, Amini A, Rezaei N. Hyper IgE syndromes: A clinical approach. Clin Immunol. 2022;237:108988.

How to cite this article: Bai J, Zhu L, Mi W, Gao Z, Ouyang M, Sheng W, et al. Multiscale integrative analyses unveil immune-related diagnostic signature for the progression of MASLD. Hepatol Commun. 2023;7:e0298. https://doi.org/10.1097/HC9.0000000000000298