






## ORIGINAL ARTICLE

OPEN

# Paired snRNA-seq and scRNA-seq analysis of MASLD patients to identify early-stage markers for disease progression

Suebin Park<sup>1,2</sup>  | Su-Hyeon Lee<sup>1</sup>  | Se-eun Han<sup>1,3</sup>  | Beom Kyung Kim<sup>4</sup>  |  
 Byungjin Hwang<sup>1,3</sup> 

<sup>1</sup>Department of Biomedical Sciences, Yonsei University College of Medicine, Seoul, Republic of Korea

<sup>2</sup>Department of Clinical Drug Discovery and Development, Yonsei University College of Medicine, Seoul, Republic of Korea

<sup>3</sup>Brain Korea 21 Project, Graduate School of Medical Science, Yonsei University College of Medicine, Seoul, Republic of Korea,

<sup>4</sup>Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

## Correspondence

Byungjin Hwang, Department of Biomedical Sciences, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea.  
 Email: [bjhwang113@yuhs.ac](mailto:bjhwang113@yuhs.ac)

Beom Kyung Kim, Department of Internal Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea.  
 Email: [beomkkim@yuhs.ac](mailto:beomkkim@yuhs.ac)

## Abstract

**Background and Aims:** Metabolic dysfunction–associated steatotic liver disease (MASLD) is a leading cause of chronic liver disease worldwide. Progression from simple metabolic dysfunction–associated steatotic liver (MASL) without necro-inflammation to metabolic dysfunction–associated steatohepatitis (MASH) triggers fibrosis, leading to liver-related morbidity and mortality. Early recognition of MASH is imperative to enable appropriate interventions aimed at preventing liver damage. Thus, this study aimed to elucidate molecular mechanisms driving MASLD progression and identify early-stage transcriptomic signatures by analyzing paired liver tissue and peripheral blood mononuclear cells (PBMCs).

**Methods:** We collected 16 paired liver and PBMC samples from 8 histologically confirmed patients with MASLD. Liver tissue was obtained by needle biopsy for single-nucleus RNA sequencing, and PBMCs underwent single-cell and bulk RNA sequencing. PBMC–liver interactions were examined to identify cross-tissue signaling, and machine learning was applied to derive transcriptomic signatures predictive of fibrosis stage.

**Results:** Hepatocyte transcriptomic profiling revealed distinct MASH-associated alterations, including downregulated fatty acid metabolism, upregulated immune activation pathways, and changes in tissue remodeling. PBMC analysis identified shifts in immune populations, with increased aTregs and chronic CD4+ T cell activation. Liver–PBMC interaction analysis highlighted

**Abbreviations:** aTreg, activated regulatory T cell; CD4 Tcm, CD4 central memory T cell; CD4 Tem, CD4 effector memory T cell; CD8 Tcm, CD8 central memory T cell; CD8 Tem, CD8 effector memory T cell; DEG, differentially expressed gene; ECM, extracellular matrix; ER, endoplasmic reticulum; GSEA, gene set enrichment analysis; ILC, innate lymphocyte cell; MASLD, metabolic dysfunction-associated steatotic liver disease; MASH, metabolic dysfunction-associated steatohepatitis; MASL, metabolic dysfunction–associated steatotic liver; MetALD, MASLD with increased alcohol intake; NK, natural killer; OXPHOS, oxidative phosphorylation; PBMC, peripheral blood mononuclear cell; PCA, principal component analysis; PtdIns, phosphatidylinositol; rTreg, resting regulatory T cell; RPCA, reciprocal PCA; scRNA-seq, single-cell RNA sequencing; snRNA-seq, single-nucleus RNA sequencing; SLD, steatotic liver disease; SNP, single nucleotide polymorphism; UMAP, uniform manifold approximation and projection.

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](http://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 © 2025 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Association for the Study of Liver Diseases.

enhanced HSC–natural killer cell signaling in MASH, linking immune responses to fibrosis progression. Machine learning identified liver-derived and PBMC-derived transcriptomic signatures that robustly distinguished mild (F0–F2) from advanced (F3–F4) fibrosis (AUC=0.93), suggesting their potential for early diagnostic stratification.

**Conclusions:** Significant molecular and immune alterations occur in disease progression of MASLD to MASH, reflecting both localized hepatic changes and systemic immune dysregulation. The identified transcriptomic signatures provide a promising tool for fibrosis prediction and monitoring, underscoring the need to target early disease mechanisms for improved diagnosis and therapeutic strategies.

**Keywords:** bulk RNA sequencing, early-stage MASLD, liver, PBMC, single cell/single nuclei

## INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) is a primary contributor to liver-related morbidity and mortality that affects ~30% of the global population.<sup>[1–3]</sup> Its incidence is rising parallel to the increasing prevalence of metabolic syndrome and obesity. The risk of serious pathological states such as cardiovascular diseases, liver complications, and extrahepatic and hepatic malignancies is substantially elevated by NAFLD. Recently, a new classification of steatotic liver disease (SLD) has emerged, taking into consideration the interplay between metabolic disease and alcohol intake, suggesting new terminologies: metabolic dysfunction–associated steatotic liver disease (MASLD), MASLD with increased alcohol intake (MetALD), and alcohol-associated liver disease.<sup>[4–6]</sup>

According to the presence of hepatic necro-inflammation, MASLD can be classified into 2 spectrums: (1) simple metabolic dysfunction–associated steatotic liver (MASL), which involves only accumulation of fat in the liver without hepatic necro-inflammation, and (2) metabolic dysfunction–associated steatohepatitis (MASH) is a more severe form of MASLD accompanying hepatic necro-inflammation in addition to fat accumulation with varying degrees of fibrosis.<sup>[7]</sup> Without appropriate pharmacological intervention or lifestyle modification,<sup>[8]</sup> hepatic necro-inflammation and resultant liver fibrosis progress, eventually leading to liver cirrhosis and/or HCC.<sup>[9]</sup>

Bulk RNA sequencing and microarray analyses have identified numerous molecular markers associated with MASLD progression.<sup>[10–12]</sup> While these approaches have provided valuable insights into disease pathophysiology, they lack the resolution to capture cellular heterogeneity within liver tissue. A

prevailing hypothesis suggests that chronic hepatic injury induced by intrahepatic fat accumulation disrupts the balance of metabolic and immune responses, promoting disease progression from simple MASLD to MASH.<sup>[13,14]</sup> However, the precise molecular mechanisms driving this transition remain poorly understood.

Recent advancements in single-cell RNA sequencing (scRNA-seq) have revolutionized biological research by enabling the characterization of intercellular heterogeneity and the identification of disease-specific cellular subtypes. Prior studies have primarily focused on nonparenchymal cells, such as HSCs, which play a central role in fibrosis development during MASLD progression.<sup>[15–17]</sup> While these studies have provided crucial insights into advanced disease stages, early-stage MASLD still remains underexplored, particularly regarding molecular changes in peripheral blood mononuclear cells (PBMCs). As PBMCs reflect systemic immune responses, they hold promise as a non-invasive biomarker source. However, studies investigating PBMC-associated molecular alterations linked to MASLD progression are scarce. Despite paired liver–PBMC analyses being established in viral hepatitis<sup>[18,19]</sup> research, such integrated approaches remain unexplored in MASLD. Furthermore, obtaining paired samples from MASLD patients presents significant challenges due to the asymptomatic nature of early disease and limited clinical indications for biopsy.<sup>[20]</sup> This gap underscores the need for integrative approaches that combine tissue-specific and systemic data to identify early molecular markers of MASLD progression.

To address these limitations, we conducted paired single-nucleus RNA sequencing (snRNA-seq) of liver tissue obtained via needle biopsy and scRNA-seq of

PBMCs from early-stage MASLD patients (F0–F2), a critical yet challenging stage for sample acquisition. By simultaneously profiling liver tissue and blood samples from matched patients, we aimed to identify molecular markers associated with the transition from simple MASLD to MASH. Our approach integrates bulk RNA sequencing with single-cell technologies to refine cellular resolution and uncover liver–PBMC cellular communication networks driving disease progression. By validating our findings in a large public cohort, applying machine learning to identify MASH-specific transcriptomic signatures, and utilizing cost-effective sample pooling, we establish a scalable framework for earlier MASH diagnosis and improved disease monitoring.

## METHODS

### Sample collection

This study adhered to the ethical principles of the Declaration of Helsinki and Istanbul and was approved by the Institutional Review Board (IRB) of Severance Hospital (IRB number: 4-2018-0537). Informed consent was obtained from all patients sequenced in this study. For blood samples, 5–10 mL fresh peripheral blood was collected in EDTA anticoagulant tubes. Liver tissue samples were obtained through the liver needle biopsy procedure. A 5.5–11 mg fragment of tissue was preserved for snRNA-seq by snap freezing in liquid nitrogen.

### PBMC isolation

PBMCs were isolated from fresh peripheral blood collected in EDTA tubes using HISTOPAQUE-1077 (Sigma-Aldrich) according to the manufacturer's instructions. Blood was layered onto HISTOPAQUE-1077, centrifuged, and the immune cell layer at the plasma–density gradient interface was collected. Cells were washed twice with PBS containing 2% FBS, resuspended and aliquoted at  $\sim 5 \times 10^6$  cells/mL, cryopreserved in CELLBANKER 2, and stored in liquid nitrogen. These PBMCs were used for both bulk and single-cell RNA sequencing.

### Nuclei isolation and preparation

Nuclei were isolated from snap-frozen tissue using the 10x Genomics Chromium Nuclei Isolation Kit with RNase Inhibitor (PN-1000494), following the manufacturer's protocol. Briefly, frozen samples were transferred to pre-chilled Sample Dissociation Tubes, homogenized in Lysis Buffer, and filtered through a Nuclei Isolation

Column. After debris removal and washing steps, nuclei were resuspended in Wash and Resuspension Buffer. The detailed protocol is available in the Chromium Nuclei Isolation Reagent Kits User Guide (CG000505 Rev A). Following isolation, nuclei were counted, normalized to the desired concentration, pooled, and processed using the 10x Genomics single-cell 3' v3.1 assay, followed by library preparation and Illumina sequencing.

### Library construction and sequencing

Library construction was performed using the 10x Chromium Next GEM Single-cell 3' Reagent Kits v3.1 (10x Genomics) according to the manufacturer's protocol (CG000315 Rev E). Purified libraries were analyzed by an Illumina NovaSeq. 6000 sequencing platform with 150 bp paired-end reads.

### Sequencing data processing

Bulk PBMC RNA-seq BAM files were processed for variant calling using GATK (v4.4.0.0) to genotype individual patients. Variant calling was also performed on pooled scRNA-seq (PBMC) and snRNA-seq (liver) samples to enable demultiplexing by patient using SoupCell. Raw scRNA-seq and snRNA-seq data were aligned to the GRCh38 reference genome and processed with CellBender<sup>[21]</sup> (v.0.3.0) to remove ambient RNA contamination. DoubletFinder<sup>[22]</sup> (v.2.0.3) was applied to detect and exclude potential doublets. Seurat<sup>[23]</sup> (v.5.0.0) was used for quality control and downstream analysis. Cells filtering criteria included expression of 300–6000 genes and mitochondrial UMI fractions below 10% for PBMCs and 20% for liver nuclei. Highly variable genes (2000 for PBMCs, 3000 for liver) were selected for principal component analysis (PCA), and the top 30 components were used for clustering. Batch effects were corrected using the reciprocal PCA (RPCA) method within Seurat.

### Machine learning analysis

To predict liver fibrosis stages (mild: F0–F2; advanced: F3–F4), we trained a LightGBM model using candidate gene features. The model was evaluated on the GSE135251 bulk liver dataset (n = 206) using 4-fold stratified cross-validation, with each fold comprising 154–155 training and 51–52 testing samples. Within each fold, data were standardized, and top-ranked features were selected based on feature importance. Model performance was assessed using the AUROC.

## RESULTS

### Single-cell and single-nucleus transcriptomic profiling of PBMCs and liver tissues

To investigate molecular heterogeneity in MASLD, we collected 16 paired PBMCs and liver tissue samples from 8 patients, including 4 with simple MASL and 4 with definite MASH. All cases were histologically confirmed, and detailed clinical and histological characteristics are presented in Supplemental Table S1, <http://links.lww.com/HC9/C137>. PBMCs were used for both single-cell and bulk RNA sequencing, while liver tissue, obtained via needle biopsy, was used for snRNA-seq. We specifically employed snRNA-seq for liver tissues because this method provides excellent utility for profiling single hepatocytes from frozen tissues, offering superior preservation of nuclear RNA integrity compared with scRNA-seq performed on snap-frozen samples.<sup>[24]</sup> We pooled PBMC samples from all 8 patients for scRNA-seq and liver samples from the same patients for snRNA-seq. Variant calling was performed on bulk PBMC RNA-seq data to identify patient-specific single-nucleotide polymorphisms (SNPs), which were then applied in Soupcorell<sup>[25]</sup> to demultiplex and accurately assign each cell and nucleus to its respective patient (Figures 1A, B). Quality control filtering retained only singlet-assigned cells for all further downstream analyses, with doublets and unassigned cells excluded (Supplemental Figure S1A, <http://links.lww.com/HC9/C138>).

To estimate the power and accuracy of donor assignment, we generated a genotype correlation heatmap (Supplemental Figure S1B, <http://links.lww.com/HC9/C138>) using SNP variants shared between PBMC and liver samples for each individual. The Pearson correlation coefficients between matched PBMC and liver samples for each donor were clearly higher than the correlations between unmatched donor pairs.

From a total of 36,588 captured single cells from PBMCs, 24,492 passed quality control and were profiled, revealing nine distinct cell lineages (Figure 1C). While most identified clusters contained cells from both conditions, distinct MASH-specific and MASL-specific clusters were also observed (Figure 1D). Each cluster was annotated based on the expression of known lineage markers (Figure 1E and Supplemental Figure S1C, <http://links.lww.com/HC9/C138>). From liver tissues, a total of 35,154 nuclei were captured, with 18,652 profiled after quality control, revealing 10 distinct clusters (Figure 1F). While nuclei from other liver cell types included both MASL and MASH conditions, hepatocytes exhibited greater heterogeneity, with distinct clustering patterns observed between the 2 conditions (Figure 1G). The liver cell types were

annotated based on known canonical markers (Figure 1H and Supplemental Figure S1D, <http://links.lww.com/HC9/C138>).

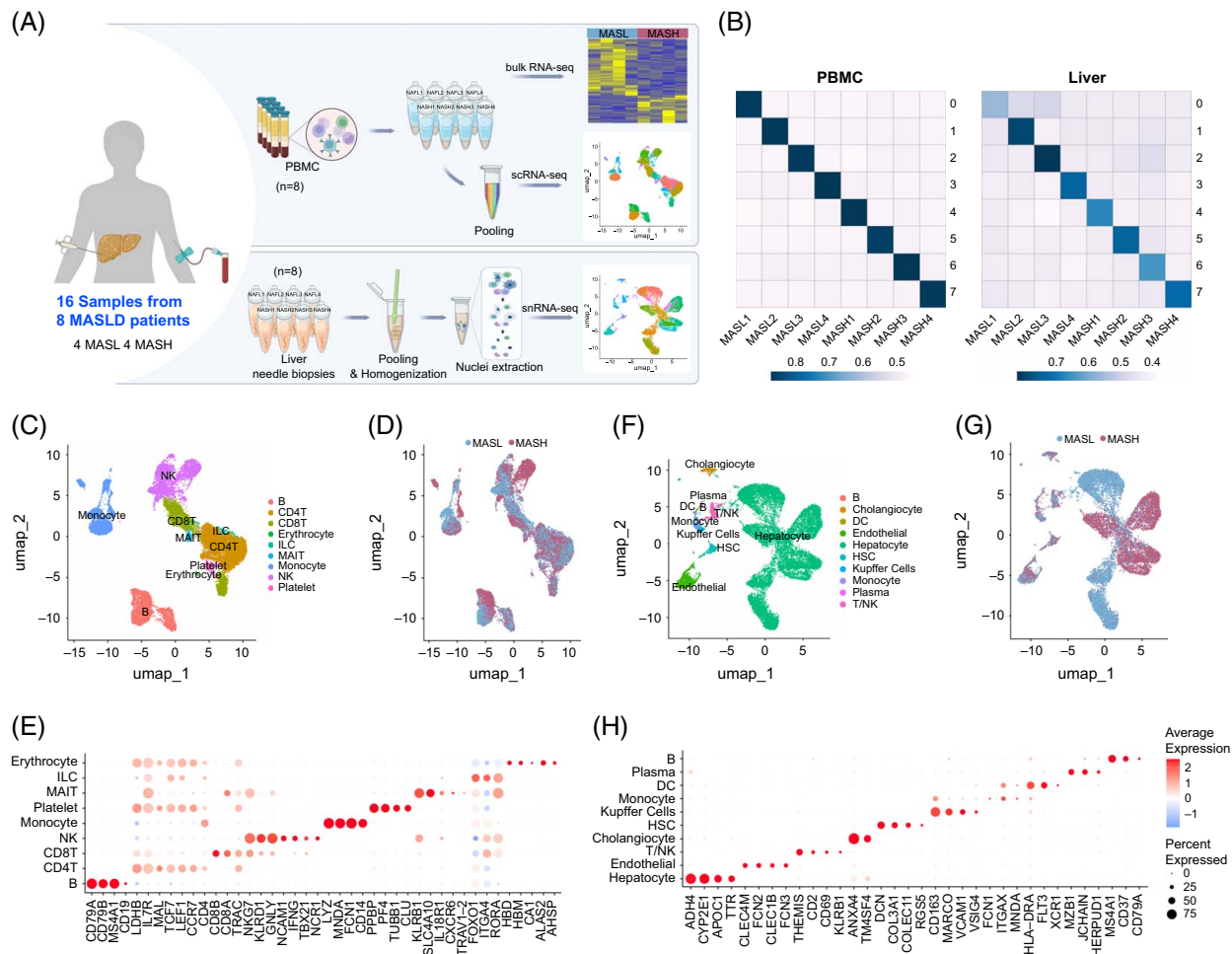
### Characterization of hepatocyte subtypes and functional pathways in MASLD progression

Further analysis of 16,098 hepatocytes revealed 6 distinct populations (Figure 2A). Hepatocyte clusters 2, 4, and 5 predominantly comprised MASL cells, while clusters 1 and 3 exhibited a higher proportion of MASH cells. Cluster 6 included a similar proportion of both MASL and MASH hepatocytes (Figures 2B, C). The hep1 cluster was characterized by the expression of RGN, AKR1D1, and ADAMTS17, which showed enrichment in oxidative phosphorylation and N-glycan biosynthesis pathways (Figure 2D and Supplemental Table S2, <http://links.lww.com/HC9/C137>). Hep2 showed high ALPK2, GRM8, and HYDIN expression, with ALPK2 polymorphisms associated with female MASH,<sup>[26]</sup> and enrichment of immune response-related pathways. Hep3, the cluster with the highest proportion of MASH cells, expressed IGF1, ZNF385D, and P4HA1. IGF1 expression has been associated with fibrosis stages, rising in stages 1 and 2 and decreasing by stage 3,<sup>[27]</sup> aligning with our findings of high IGF1 expression in early-stage MASH samples (fibrosis stages 1B and 2). Hep4 showed expression of CTNNA3, PDK4, and LRRTM3, while hep5 expressed PNPLA3, FDPS, and FDFT, with both clusters enriched for ERBB signaling pathways. Hep6, a cluster with similar proportions of MASL and MASH cells, expressed RBMS3, ZEB2, and MAML2, indicating activated Wnt/ $\beta$ -catenin and Notch signaling pathways.

### Hepatocyte transcriptomic changes during MASLD progression

Following our characterization of hepatocyte subtypes, we performed pseudobulk analysis to identify transcriptomic differences between MASL and MASH hepatocytes. This approach identified comprehensive transcriptomic differences between the 2 conditions, providing broader insights into MASL-specific and MASH-specific signatures. We identified 519 differentially expressed genes between the MASL and MASH groups, including 78 upregulated genes in MASH and 441 upregulated genes in MASL, showing condition-specific molecular signatures (Figure 2E and Supplemental Table S3, <http://links.lww.com/HC9/C137>). Gene set enrichment analysis (GSEA) showed increased immune and inflammatory activation in MASH hepatocytes, with enrichment in T cell, natural killer (NK) cell, chemokine, and cytokine signaling, indicating



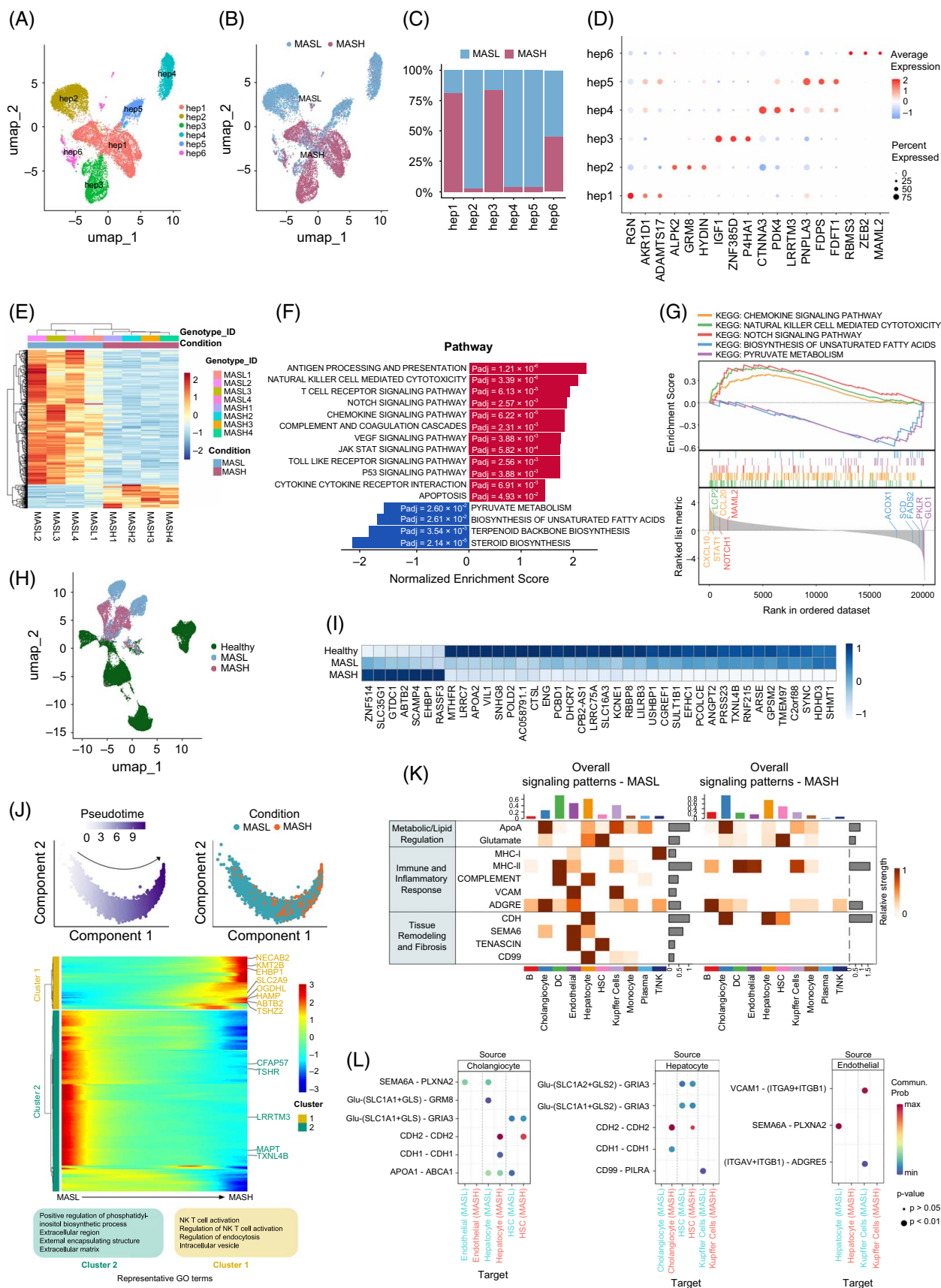


**FIGURE 1** Transcriptional landscape of PBMCs and liver nuclei in MASLD. (A) An experimental scheme illustrating the study design, including sample collection and sequencing methods for PBMC and liver tissues. Figure created with Biorender.com. (B) Heatmap of Pearson correlation coefficients between identified clusters (y-axis) and individual patient samples (x-axis), showing cluster-patient associations from Souporecell demultiplexing. Higher values indicate stronger representation of cells from specific individuals in each cluster. (C) UMAP visualization of 24,492 PBMCs from MASL (n=4) and MASH (n=4). (D) UMAP projection of PBMCs split by condition (MASL vs. MASH). (E) Dot plot showing the expression of PBMC cluster-defining genes. (F) UMAP visualization of 18,652 nuclei from liver tissue (MASL n=4, MASH n=4). (G) UMAP visualization of liver cells, split by condition (MASL vs. MASH). (H) Dot plot of known lineage markers for each liver cluster. Abbreviations: MASH, metabolic dysfunction–associated steatohepatitis; MASL, metabolic dysfunction–associated steatotic liver; MASLD, metabolic dysfunction–associated steatotic liver disease; PBMC, peripheral blood mononuclear cell; scRNA-seq, single-cell RNA sequencing; snRNA-seq, single-nucleus RNA sequencing; UMAP, uniform manifold approximation and projection.

heightened immune cell recruitment and response (eg, CXCL10, CCL20, LCP2). In addition, upregulation of JAK–STAT (STAT1) and Notch (NOTCH1, MAML2) pathways suggests their involvement in inflammation and fibrosis,<sup>[28]</sup> promoting HSC activation, epithelial–mesenchymal transition, and ductular reactions.<sup>[29]</sup> Meanwhile, downregulation of pyruvate metabolism and fatty acid biosynthesis pathways in MASH reflects disrupted hepatocyte metabolism, contributing to cellular stress and injury (Figures 2F, G).

To identify progressive transcriptomic signatures using healthy tissue as a baseline, we integrated healthy hepatocyte data and examined sequential gene expression changes across healthy, MASL, and MASH conditions using the healthy liver snRNA-seq dataset (GSE185477; Figure 2H).<sup>[24]</sup> By assessing

differentially expressed genes throughout disease progression (Figure 2I), we found genes progressively increased from healthy to MASH, associated with transcriptional regulation and potential roles in tumor progression (ZNF514), cell growth and apoptosis (ABTB2, RASSF3), glycosylation and metabolism (SLC35G1, GTDC1), and cellular trafficking and signaling (SCAMP4, EHBP1). In contrast, genes that gradually decreased in expression were linked to folate metabolism (MTHFR, SHMT1), lipid metabolism (APOA2, DHCR7), and DNA repair (POLD2, RBBP8), suggesting roles in metabolic dysregulation and cellular stress. This approach highlights genes with potential as robust biomarkers, reflecting key regulatory shifts in hepatocyte transcriptomics during early MASLD progression.



**FIGURE 2** Characterization and interactions of hepatocytes in human MASLD liver. (A) Clustering of 16,098 hepatocytes from MASL (n=4) and MASH (n=4) human livers. (B) UMAP projection of hepatocytes group by condition (MASL vs. MASH). (C) Proportions of hepatocyte

subpopulations in MASL and MASH livers. (D) Dot plot of hepatocyte cluster marker genes. (E) Heatmap of significant (adjusted  $p$ -value  $< 0.05$ ) DEGs from pseudobulk analysis (MASL vs. MASH), color-coded by genotype ID and condition. (F) Bar plot of significant (adjusted  $p$ -value  $< 0.05$ ) KEGG pathways from GSEA using pseudobulk DEGs (MASL vs. MASH). Red bars indicate pathways enriched in MASH, while blue bars indicate pathways enriched in MASL. (G) GSEA enrichment plot of selected KEGG pathways. Leading-edge genes for each pathway are displayed. (H) UMAP projection of integrated hepatocyte data from Healthy ( $n=4$ , GSE185477), MASL ( $n=4$ ), and MASH ( $n=4$ ) human livers. (I) Heatmap of genes showing progressive changes across Healthy, MASL, and MASH conditions. (J) Trajectory visualization of hepatocytes along pseudotime, with cells colored by pseudotime value (top left) and condition (MASL and MASH, top right). Pseudotemporal heatmap (middle) is grouped by hierarchical clustering ( $k=2$ ), with exemplar genes labeled, and representative GO terms associated with each cluster (bottom). (K) Heatmap showing the contribution of signaling pathways in MASL (left) and MASH (right). The right bar plot shows the total signaling strength of a signaling pathway by summarizing all cell groups displayed in the heatmap. (L) Dot plot of significant ligand–receptor interactions predicted by CellChat. Abbreviations: DEGs, differentially expressed genes; GO, Gene Ontology; GSEA, gene set enrichment analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; MASH, metabolic dysfunction–associated steatohepatitis; MASL, metabolic dysfunction–associated steatotic liver; MASLD, metabolic dysfunction–associated steatotic liver disease; UMAP, uniform manifold approximation and projection.

To better understand the dynamic transcriptional changes within hepatocytes during MASLD progression, we employed trajectory analysis and pseudotime ordering (Monocle2)<sup>[30]</sup> to capture the continuous spectrum of hepatocyte states transitioning from MASL to MASH phenotypes (Figure 2J and Supplemental Figure S2, <http://links.lww.com/HC9/C139>). Specifically, we aimed to determine whether the DEGs identified in our pseudobulk analysis exhibit progressive expression changes along the MASLD trajectory. The pseudotime analysis revealed a well-aligned transition, with early pseudotime cells predominantly corresponding to MASL and late pseudotime cells aligning with MASH. In early cells, pathways related to phosphatidylinositol biosynthetic processes and extracellular matrix (ECM) organization were activated, whereas late cells showed predominant activation of pathways associated with NK/T cell activation and regulation of endocytosis (Supplemental Table S4, <http://links.lww.com/HC9/C137>). Beyond parenchymal cells, we also characterized liver-resident immune cell populations. Kupffer cells accounted for the largest fraction in both conditions, while CD8<sup>+</sup> T cells and NK cells were the next most represented populations in MASL and MASH, respectively (Supplemental Figure S3, <http://links.lww.com/HC9/C140>).

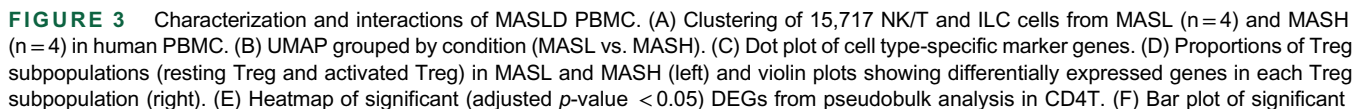
## Enhanced fibrosis and inflammation signaling during MASLD progression

Next, we identified ligand–receptor signaling pathways in MASLD using CellChat,<sup>[31]</sup> categorized into metabolic and lipid regulation (ApoA, Glutamate), immune and inflammatory response (MHC-I, MHC-II, Complement, VCAM, ADGRE), and tissue remodeling and fibrosis (CDH, SEMA6, TENASCIN, CD99) (Figure 2K). The ApoA1–ABCA1 signaling pathway facilitates cholesterol efflux via ABCA1 and HDL synthesis through ApoA1.<sup>[32]</sup> This pathway was enriched in cholangiocyte-to-hepatocyte interactions within MASL and may contribute to lipid homeostasis by mitigating fat accumulation during

early disease stages. Glutamate signaling involved cholangiocytes, hepatocytes, and HSCs and exhibited 2 distinct regulatory modes: inhibitory GRM8 (mGluR8) signaling in MASL and excitatory GRIA3 (AMPA) signaling in MASH. In MASL, GRM8 signaling potentially maintained metabolic balance while suppressing pro-inflammatory cytokine release.<sup>[33]</sup> In contrast, GRIA3 signaling in MASH correlated with lipid accumulation and oxidative stress.<sup>[34]</sup> Complement signaling, along with MHC-I and MHC-II activation, was observed in MASL, indicating an early-phase immune response. VCAM and ADGRE signaling between endothelial cells and Kupffer cells suggested innate immune cell infiltration during this stage. MHC-II activation became more prominent in MASH, reflecting a shift toward adaptive immune responses as the disease progresses. CDH signaling, which regulates cell adhesion and migration through cadherin interactions, was active in hepatocytes, cholangiocytes, and HSCs during MASH, consistent with enhanced fibrosis and tissue remodeling. SEMA6A–PlxnA2 interactions were enriched in MASL and were associated with cell migration and architectural regulation. These findings reveal distinct ligand–receptor signaling patterns across MASLD stages, reflecting shifts in metabolic regulation, immune responses, and tissue remodeling processes during disease progression (Figure 2L and Supplemental Figure S4, <http://links.lww.com/HC9/C141>).

## Immune activation and stress response in CD4<sup>+</sup> T cells during progression to MASH

In total, 15,717 cells were clustered into 10 groups, including CD4 T naive, CD4 central memory T (CD4 Tcm), CD4 effector memory T (CD4 Tem), CD8 T naive, CD8 central memory T (CD8 Tcm), CD8 effector memory T (CD8 Tem), innate lymphoid cells (ILC), NK cells, activated regulatory T cells (aTreg), and resting regulatory T cells (rTreg) (Figure 3A), with both MASL and MASH cells present in each cluster (Figure 3B). Each cluster was annotated using canonical marker





(adjusted  $p$ -value  $< 0.05$ ) pathways from GSEA using CD4T pseudobulk DEGs (MASH vs. MASL). Red bars indicate pathways enriched in MASH, while blue bars indicate pathways enriched in MASL. (G) Violin plot of HAVCR2 expression in CD8 Tem cells (left) and LAG3 expression in NK cells (right) comparing MASL and MASH conditions. (H) Dot plot of significant pathways (adjusted  $p$ -value  $< 0.05$ ,  $|\text{NES}| > 1$ ) from GSEA using PBMC bulk RNA-seq data. Dot size represents the adjusted  $p$ -value, and color intensity indicates the NES, with NES  $> 1$  indicating pathways enriched in MASH and NES  $< -1$  indicating pathways enriched in MASL. (I) Heatmap showing the contribution of signaling pathways in MASL (left) and MASH (right). Color intensity indicates the relative signaling strength. (J) Significant ligand–receptor interactions predicted by CellChat. Abbreviations: DEGs, differentially expressed genes; GSEA, gene set enrichment analysis; ILC, innate lymphocyte cell; MASH, metabolic dysfunction–associated steatohepatitis; MASL, metabolic dysfunction–associated steatotic liver; MASLD, metabolic dysfunction–associated steatotic liver disease; NES, normalized enrichment score; NK, natural killer; PBMC, peripheral blood mononuclear cell; UMAP, uniform manifold approximation and projection.

genes to confirm cell type identities (Figure 3C). Cell type proportion comparison analysis revealed that in MASH, CD4 Tcm, aTreg, and NK cells were augmented, whereas CD4 T naive cells and ILC were depleted, reflecting shifts in immune cell balance during disease progression (Supplemental Figure S5, <http://links.lww.com/HC9/C142>).

Regulatory T cells (Tregs) are essential for immune tolerance and inflammation control and can be divided into functionally distinct subsets.<sup>[35]</sup> However, their roles in MASLD PBMCs remain poorly understood. Here, we identified 2 major Treg subsets—resting Tregs (rTregs) and activated Tregs (aTregs)—and compared their proportions between MASL and MASH conditions (Figure 3D). The analysis revealed an increase in aTreg proportion in MASH compared with MASL. Gene expression analysis showed that aTregs exhibited elevated expression of CTLA4, CCR6, ITGB1, and IL32, which are associated with immune suppression and activation, whereas rTregs expressed higher levels of CCR7, BACH2, and SELL, indicating their role in maintaining stability and migratory capacity. These findings suggest that aTregs exist in a more activated state within the inflammatory environment of MASH, reflecting their dual role in MASLD pathogenesis, where Tregs typically exert protective effects against inflammation while potentially promoting fibrosis through TGF- $\beta$  secretion. The increased proportion of aTregs in MASH PBMCs observed in our study supports the emerging evidence that Tregs may contribute to fibrosis progression in MASLD.

In the pseudobulk analysis of CD4+ T cells, MASH showed increased expression of genes linked to immune activation (HLA-DPA1), oxidative stress (DDIT4, JUNB), mitochondrial function (NR4A2, MTFP1), and endoplasmic reticulum (ER) stress (HERPUD1, TIPARP). In contrast, MASL showed increased expression of genes associated with metabolic regulation (LINC00861, AK5), cell signaling (ARL17A/B, PTPRM), and energy homeostasis (AK5) (Figure 3E and Supplemental Table S5, <http://links.lww.com/HC9/C137>). In the pathway analysis of CD4+ T cells, upregulation of pathways related to immune activation and cellular stress was observed, such as antigen processing and presentation, unfolded protein

response, apoptosis, and IL-6–JAK–STAT3 signaling for MASH. Metabolic pathways, including Myc targets v1, mTORC1 signaling, and adipogenesis, were also enriched in MASH, suggesting metabolic dysregulation. In contrast, pathways related to oxygen transport (hemoglobin binding), lipid homeostasis (lipoprotein clearance), and immune cell adhesion (leukocyte tethering) were downregulated (Figure 3F).

In summary, these results highlight the critical involvement of CD4+ T cells in the inflammatory and stress response in MASH, with their activation and stress-related pathways potentially driving immune dysregulation and metabolic disturbances that contribute to disease progression.

## Immune exhaustion and metabolic reprogramming in MASH PBMCs

Next, we analyzed exhaustion marker expression across all cell types. Notably, CD8+ Tem and NK cells showed distinct patterns (Figure 3G). In MASL, HAVCR2 (TIM-3) was higher in CD8+ Tem cells, while LAG3 was lower in NK cells. In MASH, HAVCR2 (TIM-3) was lower in CD8+ Tem cells, and LAG3 was higher in NK cells, reflecting differential exhaustion marker expression between the conditions.

To explore whether these immune alterations extend to global transcriptional changes, we analyzed PBMC bulk RNA sequencing data, revealing significant MASH-associated pathways (Figure 3H and Supplemental Table S6, <http://links.lww.com/HC9/C137>). Pathways related to immune activation, inflammation, cell cycle regulation, stress response, and cell death were upregulated in MASH, whereas pathways linked to energy metabolism (glycolysis) and heme/oxygen transport (heme metabolism, oxygen binding, hemoglobin binding) were downregulated. Interestingly, unlike the reduced hepatic OXPHOS capacity typically observed in MASH due to mitochondrial dysfunction,<sup>[36]</sup> PBMCs showed increased OXPHOS pathways alongside decreased glycolysis. This metabolic adaptation is characteristic of chronic immune activation, where sustained stimulation drives T cells toward exhaustion or alternative metabolic programming. While activated T cells

predominantly rely on glycolysis for rapid energy production and robust immune responses, chronically stimulated or memory T cells shift their metabolic preference toward OXPHOS.<sup>[37,38]</sup>

### Distinct PBMC interaction patterns reflect immune response differences in early-stage MASLD

Cell–cell interaction analysis revealed distinct immune signaling patterns in early-stage (F0–F2) MASL and MASH, reflecting differences in immune responses during disease progression. Key signaling pathways fall into 4 major categories: immune cell migration and adhesion (ICAM, SELPLG, PECAM), inflammatory response and cytokine signaling (MIF, IL16, CysLTs, PARs), T cell activation and immune regulation (CD6, CD160, ICOS), and metabolic and tissue remodeling (CHOLESTEROL, THBS, COLLAGEN) (Figure 3I). In MASH, interactions promoting immune cell migration and activation were enhanced. Notably, increased ICAM1/2–ITGAL/ITGB2 signaling facilitated leukocyte–endothelial adhesion, mainly in monocytes and NK cells. MIF–(CD74+CXCR4/CD44) signaling mediated immune activation and trafficking, enriched in NK, T, and B cells. THBS1–receptor interactions drove immune infiltration and inflammatory cytokine production, especially in monocytes. ICOS–ICOSL/CD28 signaling between B and T cells was present but less prominent than other MASH-related pathways (Figure 3J). In contrast, MASL exhibited distinct patterns, with early immune activation involving CD160-mediated NK and T cell regulation and SELPLG–SELL adhesion. Metabolic and tissue remodeling signals such as CHOLESTEROL–LIPA–RORA, linked to lipid metabolism and inflammation, and COL6A2–CD44, associated with ECM reorganization, were also observed (Supplemental Figure S6, <http://links.lww.com/HC9/C143>). These pathways indicate an initial immune and metabolic response, but were less intense than in MASH.

These results suggest PBMC immune responses evolve from early activation in MASL to heightened immune activity and metabolic reprogramming in MASH, marking a shift from transient immune engagement to chronic inflammation. Importantly, such changes appear before advanced fibrosis (F3–F4), indicating immune signaling alterations are closely tied to early MASH progression.

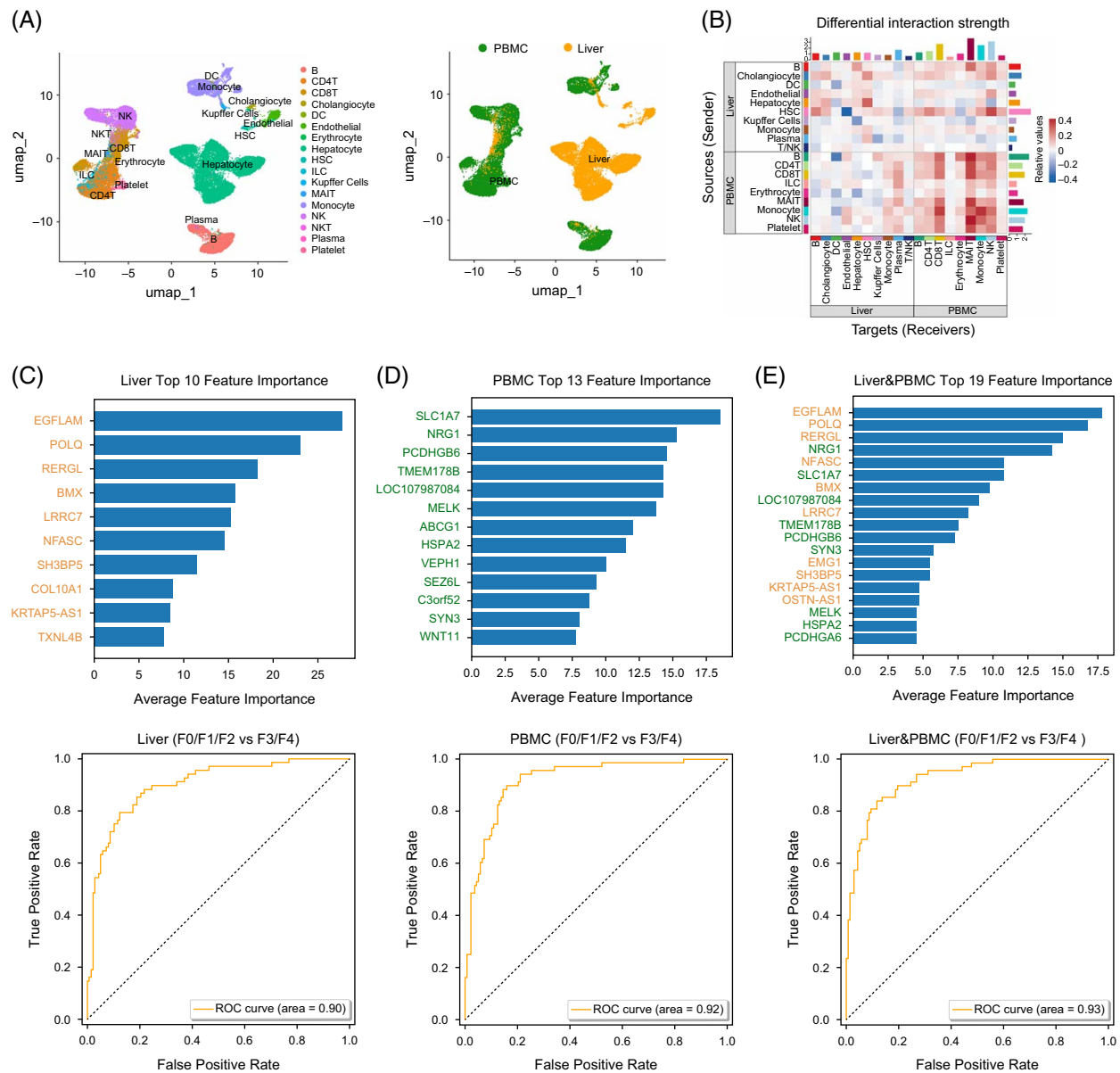
### Validation using public bulk RNA-seq datasets

To validate our findings, we analyzed public bulk RNA-seq datasets: liver (GSE130970<sup>[39]</sup>) and PBMC

(GSE267032<sup>[40]</sup>) datasets grouped into MASL-like and MASH-like categories. CIBERSORTx digital cytometry using custom signature matrices showed correlations between bulk deconvolution and single-cell mean cell fractions. Liver compositions showed correlations of  $r=0.90$ – $0.91$ , while PBMC populations displayed variable concordance (MASL:  $r=0.82$ , MASH:  $r=0.55$ ), with higher liver correlations likely influenced by hepatocyte-dominant composition and limited immune cell capture during nuclei isolation (Supplemental Figure S7A, <http://links.lww.com/HC9/C144>, and S8A, <http://links.lww.com/HC9/C145>). GSVA analysis showed substantial concordance. For hepatocyte pathways, 87.5% (14/16) were directionally consistent, with 6 reaching statistical significance: apoptosis, chemokine signaling, p53 signaling, T cell receptor signaling, Toll-like receptor signaling, and VEGF signaling (Supplemental Figure S7B, <http://links.lww.com/HC9/C144>). For PBMC pathways, 78.3% (18/23) were consistent, with 5 significant: E2F targets, IL2–STAT5 signaling, mTORC1 signaling, MYC targets V1, and unfolded protein response (Supplemental Figure S8B, <http://links.lww.com/HC9/C145>).

### PBMC–liver tissue interactions reveal immune–liver communication in MASLD

To investigate PBMC–liver interactions in MASLD, we leveraged our paired samples to integrate liver and PBMC datasets, providing a more comprehensive view of the cellular landscape. This integration captured diverse cell types, including hepatocytes, nonparenchymal cells, and various immune populations (Figure 4A). Using this integrated dataset, we performed cell–cell interaction analysis to examine differential interaction strength between liver and PBMC cell types in MASL and MASH. A key observation was the increased predicted interaction between liver HSCs and PBMC NK cells in MASH compared with MASL, suggesting a significant shift in PBMC–liver communication that may contribute to fibrosis progression (Figure 4B and Supplemental Figure S9, <http://links.lww.com/HC9/C146>). In MASH, the predicted upregulation of LAMA2–CD44 and LAMB1–CD44 signaling, which mediates ECM–immune cell interactions, strengthens PBMC–liver interactions between PBMC NK cells and liver HSCs. Given HSCs' role in fibrosis and NK cells' ability to modulate HSC activation, this CD44-mediated interaction may facilitate the recruitment and retention of PBMC-derived NK cells in fibrotic liver tissue, enhancing local fibrogenic signaling. This hypothesis is supported by previous findings showing CD44 expression on HSCs and its functional role in ECM remodeling.<sup>[41,42]</sup> Together, these findings suggest that CD44 signaling may promote HSC activation and ECM remodeling, potentially accelerating fibrosis progression



**FIGURE 4** PBMC–liver interaction and AUROC. (A) Integrated UMAP plots of 43,144 cells from liver ( $n=8$ ) and PBMC ( $n=8$ ) datasets, displaying clustering by cell type (left) and tissue origin (right). (B) Heatmap showing differential interaction strength in the cell–cell communication network between MASL and MASH. Red indicates increased signaling, and blue indicates decreased signaling in MASH compared with MASL. (C–E) Bar plots showing feature importance (top) and AUROC for fibrosis stage classification (F0–F2 vs. F3–F4) (bottom). (C) 10 liver-derived genes, AUC = 0.90, (D) 13 PBMC-derived genes, AUC = 0.92, and (E) 21 combined genes from liver and PBMC datasets, AUC = 0.93. Abbreviations: MASH, metabolic dysfunction–associated steatohepatitis; MASL, metabolic dysfunction–associated steatotic liver; PBMC, peripheral blood mononuclear cell; UMAP, uniform manifold approximation and projection.

and contributing to a fibrotic microenvironment that amplifies immune responses in MASH.

## Early transcriptional changes as indicators of fibrosis progression

To validate our previously identified candidate genes as early markers of disease progression in a large public cohort, we established a robust classifier for fibrosis stages (F0–F2 vs. F3–F4). We first identified 695

candidate genes (310 liver-derived and 386 PBMC-derived) from differential expression analysis between MASL (F0, F1A) and MASH (F1A, F1B, F2). These candidates were further refined using a public liver bulk RNA-seq dataset (GSE135251) and feature selection with LightGBM (LGBM). The dataset was randomly split into 75% training and 25% testing sets (154–155 training and 51–52 testing samples per fold), with 4-fold stratified cross-validation to ensure robustness. After standardizing gene expression, we trained an initial LGBM model with all candidate genes and ranked

them by feature importance. The top 10 liver-derived genes were selected based on feature importance, and their classification performance was evaluated using the ROC curve. The model achieved an AUC of 0.90, indicating strong classification performance (Figure 4C). Similarly, feature selection on the PBMC-derived candidate genes identified 13 key genes with the highest predictive value. This model slightly outperformed the liver-derived model, with an AUC of 0.92 (Figure 4D). We then performed feature selection across the full set of 695 candidate genes (310 liver-derived, 386 PBMC-derived), identifying a refined set of 19 key genes (10 liver-derived, 9 PBMC-derived). Most overlapped with top-ranking genes from previous analyses, but 3 additional genes (PCDHGA6, OSTN-AS1, EMG1) emerged from this combined approach. This final model achieved the highest AUC of 0.93, outperforming models using liver or PBMC genes alone (Figure 4E). This suggests that the newly identified genes may play a potential role in fibrosis progression by capturing interactions between liver-specific signals and systemic immune responses reflected in PBMCs. To further evaluate the generalizability of our classifier, we validated the model using an independent external liver bulk RNA-seq dataset (GSE130970). The liver, PBMC, and combined gene models achieved AUCs of 0.85, 0.84, and 0.87, respectively, indicating good predictive performance even in an independent cohort (Supplemental Figure S10, <http://links.lww.com/HC9/C147>).

Notably, the fibrosis classifier, built from F0–F2 stage patients, demonstrated strong predictive power even when distinguishing between F0–F2 and F3–F4 stages. This finding implies that molecular changes occurring in early-stage MASLD may not only reflect the current fibrosis stage but also serve as early indicators of future fibrosis progression.

## DISCUSSION

Since some novel pharmacological approaches, in addition to lifestyle modification, have become available to manage MASLD nowadays, recognition of MASLD disease progression earlier from MASL to MASH might be quite important to prevent liver-related morbidity and mortality. This study provides new insights into the early disease progression of MASLD from MASL to MASH by integrating paired snRNA-seq of liver tissue and scRNA-seq of PBMCs. While most previous research has focused on immune and nonparenchymal cells within liver tissue, our approach of analyzing both liver and PBMCs offers a more comprehensive view of the disease. By identifying key transcriptomic changes that distinguish MASL from MASH, we highlight the early molecular shifts that drive disease progression. This work emphasizes the importance of studying early-

stage MASLD and reveals how both localized liver changes and systemic immune alterations contribute to the transition from MASL to MASH.

Our hepatocyte analysis revealed fundamental molecular reprogramming during MASLD progression. First, immune activation pathways, including TLR, JAK–STAT, and NOTCH signaling, were significantly upregulated in MASH compared with MASL, consistent with previous reports. This immune activation was further supported by our pseudotime analysis, which demonstrated a striking shift in hepatocyte function as the disease progressed. Early MASL hepatocytes exhibited enrichment of phosphatidylinositol biosynthesis and ECM remodeling pathways, reflecting structural and metabolic adaptations to lipid accumulation. As cells transitioned to the MASH state, we observed a marked increase in NK/T cell activation and endocytosis regulation pathways, suggesting hepatocytes actively participate in immune modulation during disease progression.

The increased endocytosis pathways in MASH hepatocytes are particularly noteworthy, as they likely facilitate recognition of danger signals through pattern recognition receptors like TLR4, which has been implicated in MASH pathogenesis.<sup>[43]</sup> Concurrently, we observed a gradual reduction in phosphatidylinositol (PtdIns) biosynthesis as cells progressed along the pseudotime trajectory toward the MASH phenotype, which may contribute to disease progression through multiple mechanisms. Recent studies have shown that disruption of cellular PtdIns supply induces steatosis similar to inositol deficiency.<sup>[41,44]</sup> As PtdIns is a major component of ER membranes, insufficient PtdIns can activate ER stress response pathways and promote MASLD development. In addition, we observed a progressive decrease in fatty acid metabolism in MASH, supporting the role of lipotoxicity and mitochondrial dysfunction in MASH pathogenesis.

These findings collectively indicate a fundamental shift in hepatocyte function during MASLD progression—from maintaining tissue homeostasis and extracellular environment remodeling to adopting a defensive posture characterized by immune activation and altered endocytic processing. The increased NK/T cell activation pathways in MASH hepatocytes are particularly interesting, as recent studies have shown complex interactions between hepatocytes and NKT cells during MASLD progression.<sup>[45,46]</sup> Hepatocytes can present lipid antigens to NKT cells via CD1d, which is upregulated in MASH livers. This interaction can lead to NKT cell activation, which in turn can affect hepatocytes through various mechanisms, including TNF- $\alpha$  secretion that may either promote hepatocyte death or regeneration depending on the context. In addition, activated NKT cells can enhance lipid uptake by hepatocytes through secretion of TNFSF14 (LIGHT), potentially exacerbating steatosis.<sup>[46]</sup> This metabolic



reprioritization likely represents hepatocytes' response to increasing inflammatory signals and cellular stress. Our results emphasize that hepatocytes are not merely passive targets in MASLD but active participants driving both immune activation and disease pathophysiology, even in early disease stages.

PBMC analysis revealed important systemic immune changes in MASH, with immune regulation mechanisms engaged even in early-stage disease. We identified a shift in CD4<sup>+</sup> T cell gene expression toward immune activation and oxidative stress, corroborating previous studies linking these features to MASLD progression. Notably, we observed an increased proportion of activated Tregs in MASH PBMCs, aligning with recent studies showing Tregs' complex role in MASLD, where they may exert both anti-inflammatory effects through IL-10 secretion and pro-fibrotic effects through TGF- $\beta$  secretion.<sup>[47]</sup> Mouse models have demonstrated that Treg depletion can alleviate steatosis and hepatocyte injury at early disease stages, suggesting their potential contribution to disease progression.<sup>[48]</sup> Our findings of heightened immune activation in early MASH support the broader understanding of MASLD progression, where both innate and adaptive immune responses are initially activated at early disease stages. However, studies of advanced disease suggest that immune activity may decline over time, transitioning to functionally less active phenotypes in later stages.<sup>[49]</sup> These results characterize MASLD as a systemic disorder with coordinated immune responses across multiple compartments, with immune dysregulation present even in early disease stages (F0–F2), before fibrosis develops. Importantly, this systemic nature may explain why our PBMC-derived gene signatures demonstrated strong predictive performance in independent liver bulk RNA-seq datasets—reflecting hepatic–systemic immune crosstalk, where intrahepatic immune and fibrotic processes are partially associated with peripheral blood immune transcriptomes. In addition, the highly vascularized liver tissue may capture immune cells during biopsy sampling, especially in inflamed or fibrotic states, further supporting the predictive capacity of PBMC-derived markers in liver datasets.

This study is limited by its small sample size (16 paired liver–PBMC samples from 8 patients), which restricts its ability to capture the full heterogeneity of MASLD. Future studies should validate these findings in larger, multicenter cohorts, with both in vivo and in vitro experiments to further confirm and elucidate the predicted cell–cell interactions and their functional roles. While our findings demonstrated strong predictive performance in an independent validation dataset (AUC = 0.93), additional validation with histological and serum protein-level analyses is needed to confirm the clinical relevance of these molecular markers. Incorporating these approaches will strengthen the biological implications of our study and improve the translation of these findings to clinical practice.

## AUTHOR CONTRIBUTIONS

Suebin Park: sample processing, bulk RNA-seq sample preparation, scRNA-seq/snRNA-seq library construction, data analysis, and writing; Su-Hyeon Lee: sample processing, bulk RNA-seq sample preparation, and scRNA-seq/snRNA-seq library construction; Se-eun Han: sample processing; Beom Kyung Kim: writing, funding acquisition, and supervision; Byungjin Hwang: conceptualization, funding acquisition, supervision, and writing.

## FUNDING INFORMATION

This work was in part supported by the National Research Foundation of Korea (NRF) grant funded by Korean Government (MSIT) (RS-2021-NR064716; RS-2023-00276271; RS-2024-00455210) and by a new faculty research seed money grant of Yonsei University College of Medicine for 2023 (6-2023-0213).

## ACKNOWLEDGMENTS

The authors thank the patients who contributed through biopsies and blood sampling for this study.

## CONFLICTS OF INTEREST

The authors have no conflicts to report.

## ORCID

Suebin Park  <https://orcid.org/0009-0004-7319-9935>

Su-Hyeon Lee  <https://orcid.org/0000-0002-2510-5621>

Se-eun Han  <https://orcid.org/0009-0002-4933-3909>

Beom Kyung Kim  <https://orcid.org/0000-0002-5363-2496>

Byungjin Hwang  <https://orcid.org/0000-0002-7221-3641>

## REFERENCES

1. Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): A systematic review. *Hepatology*. 2023;77:1335–47.
2. Kim G-A, Moon JH, Kim W. Critical appraisal of metabolic dysfunction-associated steatotic liver disease: Implication of Janus-faced modernity. *Clin Mol Hepatol*. 2023;29:831.
3. Han SK, Baik SK, Kim MY. Non-alcoholic fatty liver disease: Definition and subtypes. *Clin Mol Hepatol*. 2022;29(suppl):S5.
4. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology*. 2023;78:1966–86.
5. Yan M, Man S, Ma L, Guo L, Huang L, Gao W. Immunological mechanisms in steatotic liver diseases: An overview and clinical perspectives. *Clin Mol Hepatol*. 2024;30:620.
6. Yoon EL, Jun DW. Waiting for the changes after the adoption of steatotic liver disease. *Clin Mol Hepatol*. 2023;29:844.
7. Chon YE, Jin Y-J, An J, Kim HY, Choi M, Jun DW, et al. Optimal cut-offs of vibration-controlled transient elastography and magnetic resonance elastography in diagnosing advanced liver fibrosis in patients with nonalcoholic fatty liver disease: A systematic review and meta-analysis. *Clin Mol Hepatol*. 2024;30(suppl):S117.

8. Sohn W, Lee Y-S, Kim SS, Kim JH, Jin YJ, Kim GA, et al. KASL clinical practice guidelines for the management of metabolic dysfunction-associated steatotic liver disease 2025. *Clin Mol Hepatol*. 2025;31(suppl):S1.
9. Kim D, Manikat R, Wijarnpreecha K, Cholankeril G, Ahmed A. Burden of mortality from hepatocellular carcinoma and biliary tract cancers by race and ethnicity and sex in US, 2018–2023. *Clin Mol Hepatol*. 2024;30:756.
10. 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.
11. 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;149:623–634.e8.
12. Kim D, Shah M, Kim JH, Kim J, Baek YH, Jeong JS, et al. Integrative transcriptomic and genomic analyses unveil the IFI16 variants and expression as MASLD progression markers. *Hepatology*. 2025;81:962–75.
13. Tacke F, Weiskirchen R. Non-alcoholic fatty liver disease (NAFLD)/non-alcoholic steatohepatitis (NASH)-related liver fibrosis: Mechanisms, treatment and prevention. *Ann Transl Med*. 2021;9:729.
14. Ramadori P, Kam S, Heikenwalder M. T cells: Friends and foes in NASH pathogenesis and hepatocarcinogenesis. *Hepatology*. 2022;75:1038–49.
15. Ramachandran P, Dobie R, Wilson-Kanamori JR, Dora EF, Henderson BEP, Luu NT, et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature*. 2019;575:512–8.
16. Fred RG, Steen Pedersen J, Thompson JJ, Lee J, Timshel PN, Stender S, et al. Single-cell transcriptome and cell type-specific molecular pathways of human non-alcoholic steatohepatitis. *Sci Rep*. 2022;12:13484.
17. Su Q, Kim SY, Adewale F, Zhou Y, Aldler C, Ni M, et al. Single-cell RNA transcriptome landscape of hepatocytes and non-parenchymal cells in healthy and NAFLD mouse liver. *Iscience*. 2021;24:103233.
18. Zhang C, Li J, Cheng Y, Meng F, Song JW, Fan X, et al. Single-cell RNA sequencing reveals intrahepatic and peripheral immune characteristics related to disease phases in HBV-infected patients. *Gut*. 2023;72:153–67.
19. Genshaft AS, Subudhi S, Keo A, Sanchez Vasquez JD, Conceição-Neto N, Mahamed D, et al. Single-cell RNA sequencing of liver fine-needle aspirates captures immune diversity in the blood and liver in chronic hepatitis B patients. *Hepatology*. 2023;78:1525–41.
20. Boel F, Akimov V, Teuchler M, Terkelsen MK, Wernberg CW, Larsen FT, et al. Deep proteome profiling of metabolic dysfunction-associated steatotic liver disease. *Commun Med*. 2025;5:56.
21. Fleming SJ, Chaffin MD, Arduini A, Akkad AD, Banks E, Marioni JC, et al. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat Methods*. 2023;20:1323–35.
22. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst*. 2019;8:329–337.e4.
23. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. 2024;42:293–304.
24. Andrews TS, Atif J, Liu JC, Perciani CT, Ma XZ, Thoeni C, et al. Single-cell, single-nucleus, and spatial RNA sequencing of the human liver identifies cholangiocyte and mesenchymal heterogeneity. *Hepatol Commun*. 2022;6:821–40.
25. Heaton H, Talman AM, Knights A, Imaz M, Gaffney DJ, Durbin R, et al. Souporecell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat Meth*. 2020;17:615–20.
26. McIntosh AT, Wei R, Ahn J, Aouizerat BE, Kassaye SG, Augenbraun MH, et al. A genomic variant of ALPK2 is associated with increased liver fibrosis risk in HIV/HCV coinfecting women. *PLoS One*. 2021;16:e0247277.
27. Stanley TL, Fourman LT, Zheng I, McClure CM, Feldpausch MN, Torriani M, et al. Relationship of IGF-1 and IGF-binding proteins to disease severity and glycemia in nonalcoholic fatty liver disease. *J Clin Endocrinol Metab*. 2021;106:e520–33.
28. Zhu C, Kim K, Wang X, Bartolome A, Salomao M, Dongiovanni P, et al. Hepatocyte Notch activation induces liver fibrosis in nonalcoholic steatohepatitis. *Sci Transl Med*. 2018;10:eaat0344.
29. Gadd VL, Skoien R, Powell EE, Fagan KJ, Winterford C, Horsfall L, et al. The portal inflammatory infiltrate and ductular reaction in human nonalcoholic fatty liver disease. *Hepatology*. 2014;59:1393–405.
30. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Meth*. 2017;14:979–82.
31. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell–cell communication using CellChat. *Nat Commun*. 2021;12:1088.
32. Segrest JP, Tang C, Song HD, Jones MK, Davidson WS, Aller SG, et al. ABCA1 is an extracellular phospholipid translocase. *Nat Commun*. 2022;13:4812.
33. Xie Y, Chen L, Chen J, Luo Y, Peng Z, Zhang H, et al. Metabotropic glutamate receptor 8 suppresses M1 polarization in microglia by alleviating endoplasmic reticulum stress and mitochondrial dysfunction. *J Integr Neurosci*. 2024;23:26.
34. Shi N, Sun K, Tang H, Mao J. The impact and role of identified long noncoding RNAs in nonalcoholic fatty liver disease: A narrative review. *J Clin Lab Anal*. 2023;37:e24943.
35. Kim J-R, Chae J-N, Kim S-H, Ha J-S. Subpopulations of regulatory T cells in rheumatoid arthritis, systemic lupus erythematosus, and Behcet's disease. *J Korean Med Sci*. 2012;27:1009.
36. Shum M, Ngo J, Shirihai OS, Liesa M. Mitochondrial oxidative function in NAFLD: Friend or foe? *Mol Metab*. 2021;50:101134.
37. Cao J, Liao S, Zeng F, Liao Q, Luo G, Zhou Y. Effects of altered glycolysis levels on CD8+ T cell activation and function. *Cell Death Dis*. 2023;14:407.
38. Bettonville M, d'Aria S, Weatherly K, Porporato PE, Zhang J, Bousbata S, et al. Long-term antigen exposure irreversibly modifies metabolic requirements for T cell function. *Elife*. 2018;7:e30938.
39. 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.
40. Steixner-Kumar AA, Santacruz D, Geiger T, Rust W, Böttner D, Krenkel O, et al. Single-cell landscape of peripheral immune cells in MASLD/MASH. *Hepatol Commun*. 2025;9:e0643.
41. Alb JG, Cortese JD, Phillips SE, Albin RL, Nagy TR, Hamilton BA, et al. Mice lacking phosphatidylinositol transfer protein- $\alpha$  exhibit spinocerebellar degeneration, intestinal and hepatic steatosis, and hypoglycemia. *J Biol Chem*. 2003;278:33501–18.
42. Osawa Y, Kawai H, Tsunoda T, Komatsu H, Okawara M, Tsutsui Y, et al. Cluster of differentiation 44 promotes liver fibrosis and serves as a biomarker in congestive hepatopathy. *Hepatol Commun*. 2021;5:1437–47.
43. Sharifnia T, Antoun J, Verriere TGC, Suarez G, Wattacheril J, Wilson KT, et al. Hepatic TLR4 signaling in obese NAFLD. *Am J Physiol Gastrointest Liver Physiol*. 2015;309:G270–8.
44. Alb JG, Phillips SE, Wilfley LR, Philpot BD, Bankaitis VA. The pathologies associated with functional titration of phosphatidylinositol transfer protein  $\alpha$  activity in mice. *J Lipid Res*. 2007;48:1857–72.

45. Syn WK, Oo YH, Pereira TA, Karaca GF, Jung Y, Omenetti A, et al. Accumulation of natural killer T cells in progressive nonalcoholic fatty liver disease. *Hepatology*. 2010;51: 1998–2007.
46. Gu X, Chu Q, Ma X, Wang J, Chen C, Guan J, et al. New insights into iNKT cells and their roles in liver diseases. *Front Immunol*. 2022;13:1035950.
47. Wu K, Qian Q, Zhou J, Sun D, Duan Y, Zhu X, et al. Regulatory T cells (Tregs) in liver fibrosis. *Cell Death Discov*. 2023;9:53.
48. Dywicky J, Buitrago-Molina LE, Noyan F, Davalos-Misslitz AC, Hupa-Breier KL, Lieber M, et al. The detrimental role of regulatory T cells in nonalcoholic steatohepatitis. *Hepatol Commun*. 2022;6:320–33.
49. Waller KJ, Saihi H, Li W, Brindley JH, De Jong A, Syn W, et al. Single-cell phenotypes of peripheral blood immune cells in early and late stages of non-alcoholic fatty liver disease. *Clin Mol Hepatol*. 2022;29:417.

**How to cite this article:** Park S, Lee S, Han S, Kim BK, Hwang B. Paired snRNA-seq and scrRNA-seq analysis of MASLD patients to identify early-stage markers for disease progression. *Hepatol Commun*. 2025;9:e0820. <https://doi.org/10.1097/HC9.0000000000000820>