MATERIALS AND METHODS

Human iPSC cell line

Human iPSCs cell line was selected based on their differentiation ability evaluated in our previous study (*43*). The generation of this human iPSC was approved by The University of Pennsylvania Human Subjects Research Institutional Review Board.

Nuclei isolation

The nuclei isolation for snRNA-seq and snATAC-seq was adapted from the 10X Genomics nuclei demonstration protocol. Fresh mouse liver tissue or frozen human liver tissue was dounced in 10ml of lysis dilution buffer (10mM Tris-HCl, PH7.5; 10mM NaCl; 3mM MgCl2; 1% BSA; cOmplete Protease Inhibitor (Roche, Cat# 11873580001) was added freshly before use) and passed through 200μm strainer. RNase inhibitors SUPERase•In RNase Inhibitor (Thermo Fisher Scientific, Cat# AM2694) and RNasin Ribonuclease Inhibitor (Promega) were added as 1:1000 freshly before use for snRNA-seq. Liver lysate was centrifuged at 500g for 5min at 4°C and the pellet was resuspended in 10ml of lysis dilution buffer. 1ml of lysis buffer (10mM Tris-HCl, PH7.5; 10mM NaCl; 3mM MgCl2; 1% BSA; 0.1% Tween; 0.1% NP40; 0.01% Digitonin; cOmplete Protease Inhibitor was added freshly before use) was added dropwise to the nuclei resuspension solution, incubated on ice for 5min and passed through 40μm strainer. The nuclei were washed with lysis dilution buffer 2 more times at 500g for 5min at 4°C and resuspended in lysis dilution buffer. Nuclei were further purified using Nuclei Isolation Kit (Sigma Cat# NUC201) through 1.8M sucrose cushion according to the manufacturer's instructions. After centrifuging for 45 minutes at 30,000 X g (13,000 rpm) at 4 °C, the nuclei pellet at the lower phase was resuspended in 1X nuclei buffer provided in Chromium Next GEM Single Cell ATAC kit or in 1% BSA in PBS for snRNA-seq. Nuclei were passed through a 40μm strainer before loading to the 10X Genomics Chromium controller. snATAC-seq and snRNA-seq libraries were prepared according to the

manufacturer's instructions of 10XGenomics (Cat#100128, Cat#100176). To maximize the recovery of *EPHB2* positive cells in human samples, targeted sequencing was performed using 98 probes detecting *EPHB2* which were pre-designed by 10XGenomics. The targeted sequencing libraries were prepared using xGen Hybridization and Wash Kit (IDT, Cat#1080577).

Processing and analyses of snRNA-seq data

The raw fastq files of snRNA-seq were mapped to *Mus musculus* assembly GRCm38 (mm10) or *Homo sapiens* genome assembly GRCh38 (hg38) using CellRanger (3.0.2). For the human snRNA-seq, the fastq files of targeted libraries were directly combined with the input libraries. Unique molecular identifiers (UMI) count matrices were imported into Seurat (3.2.0) (*58*) to generate Seurat object to each experiment. Before merging, each dataset was filtered with the parameters nFeature_RNA > 500 & nFeature_RNA < 8000 & percent.mt < 2 for mice and nFeature_RNA > 500 & nFeature_RNA < 8000 & percent.mt < 10 for humans. Different Seurat objects were integrated using Harmony (*59*) to correct batch effect. Unsupervised clustering was performed based on the first 40 principal components and visualized in UMAP plot. The cluster identities were annotated based on prior knowledge. The doublets with mixed cell features were manually removed. Pseudotime analysis of hepatocytes (mouse: all hepatocytes; human: randomly selected 2000 hepatocytes from each hepatocyte cluster) was performed with Monocle2 (*19*) with default parameters. The genes determined the bifurcation of hepatocyte trajectories were defined with q value < 1e-20. Gene ontology (GO) analysis was performed with Metascape (http://metascape.org). Cell–cell communication was analyzed with CellChat (*21*). To analyze correlation between *Ephb2* expression and the expression of inflammatory genes in hepatocytes, the expression value was transformed using MAGIC due to the sparsity of single-cell dataset (*60*).

Processing and analyses of snATAC-seq data

The raw fastq files of mouse livers were mapped to *Mus musculus* genome assembly GRCm38 (mm10) using Cellranger-atac (1.2.0). The fragment count matrix was imported into Signac (1.8.0). The nuclei with TSS enrichment score < 2 & pct_reads_in_peaks < 40 & nucleosome_signal > 4 & blacklist_ratio > 0.025 were removed. The chromatin assays were integrated using FindIntegrationAnchors with dimension 2:30. Unsupervised clustering was visualized in UMAP plot. The cluster identities were assigned based on transferred label from snRNA-seq in our study. ChromVAR motif Z score matrix was generated by ChromVAR and JASPAR 2018 core motifs (Homo sapiens) were used for peak annotation. ChromVAR motif Z score matrix was used as input for STREAM pseudotime analysis for identifying key TFs (*61*). The dimension reduction was performed using STREAM with parameters n_pc=15, n_components=5 and n_neighbors=30. The TFs in NAFL and NASH branch were prioritized by intersecting STREAM output with up-regulated genes in hepatocyte nuclei isolated from mice with mild and advanced NASH respectively in a published dataset (GSE162876). Log2FC > 0.5 and padj < 0.1 were used as cutoff for defining up-regulated genes.

Integration of snRNA-seq and snATAC-seq

snRNA-seq and snATAC-seq were integrated using Signac. The peak to gene links were defined by the correlation between peak accessibility (snATAC-seq) and gene expression (transferred snRNA-seq information) in the matched cells. Signaling pathway enrichment analysis for candidate TFs regulating EphB2 was performed with R package clusterProfiler (*62*) using all human genes as background and filtered by signaling transduction pathway annotation.

Bulk RNA-seq

For whole-liver RNA-seq, liver tissues were harvested from mice 1 week after injection with AAV-TBG-GFP or AAV-TBG-EphB2. 3 biological replicates were included in each experimental group. Tissue was homogenized in Trizol using TissueLyser and total RNA was extracted with RNeasy

Mini Kit (QIAGEN). Total RNA (>=1μg) was submitted to Novogene (https://en.novogene.com/) for library preparation and sequencing.

Processing and analyses of Bulk RNA-seq data

The raw fastq files were mapped to *Mus musculus* assembly GRCm38 (mm10) using STAR (2.5.4) (*63*) with --quantMode GeneCounts to generate gene count matrix. The matrix was imported into R package DEseq2 (*64*) to quantify gene expression. Genes with padj<0.1 and Log2(FC) > 0.5 or < -0.5 compared to Hep-GFP group were defined as significantly changed genes in Hep-EphB2 group. Transcript-specific quantification of GSE119340 and GSE130970 were analyzed by Salmon (*65*).

Sequencing

snRNA-seq, snATAC-seq and bulk RNA-seq libraries were sequenced on Illumina NextSeq5000 or NovaSeq. 400 million paired-end reads were obtained for each sample using the sequencing parameters recommended in 10X Genomics guide. EPHB2 targeted human snRNA-seq libraries were sequenced at the depth of 4000 read pairs per cell. 20 million reads were obtained with a paired-end 150bp module for each bulk RNA-seq library.

Generation of adeno-associated virus (AAV) and adenovirus

AAV-TBG-GFP plasmid were obtained from Viral Vector Core at University of Pennsylvania. Mouse *Ephb2* (ENSMUST00000059287.13) open reading frame (ORF) and human *EPHB2* (ENST00000400191.7) ORF were obtained from GenScript (U724TFF240 and U774PFH120) respectively. Flag-epitope was tagged on the C-terminus of EphB2. The non-targeting sequence in AAV-TBG-CasRX-U6-sgNC was acquired from previous publication (*66*) (**table S1**) and gRNAs against *Ephb2* were designed using cas13design (https://cas13design.nygenome.org/). Three combinations of tandem gRNAs against *Ephb2* were tested in HEK293 cells. The

combination resulted in the most efficient knockdown in vitro was selected to assemble AAV-TBG-CasRX-U6-sgEphB2 (**table S1**) by using Gibson Assembly (NEB). AAV-TBG-EphB2 plasmid DNA was purified using GenElute Endotoxin-free Plasmid Maxiprep Kit (Cat #NA0410) and submitted to Gene Vector Core at Baylor College of Medicine for AAV production. AAV was intravenously injected with 1.5 $X10^{11}$ genome copies (GC) per mouse.

Myc ORF is acquired from pD40-His/V5-c-Myc (Addgene #45597) and NICD1 ORF is from 3XFlagNICD1 (Addgene #20183). These two ORFs were cloned into AAV-TBG vector and AAV plasmid DNA was purified using GenElute Endotoxin-free Plasmid Maxiprep Kit (Cat #NA0410). Next, the AAV-TBG-GFP or AAV-TBG-Myc or AAV-NICD1 was co-transfected with packaging plasmid pAdDeltaF6 and RepCap Plasmids AAV2/8 (1:2:1) into AAVpro 293T cells at 80-90% confluency (gift from Hongjun Song's Lab at University of Pennsylvania) using Lipod293 In Vitro DNA Transfection Reagent (SignaGen Laboratories SL100668). The medium was replaced with fresh complete culture medium after 24 hours and the cells and medium were harvested 72 hours after transfection. The AAV was extracted using chloroform as previously described (*67*). The titer was quantified by qPCR for the inverted terminal repeat (ITR) region on the AAV vector according to the tutorial of Addgene(https://www.addgene.org/protocols/aav-titration-qpcr-using-sybr-green-technology).

Adenovirus Ad5-CMV-GFP was obtained from Vector Development Lab at Baylor College of Medicine. To produce adenovirus, human *EPHB2* (ENST00000400191.7) ORF was Flag-tagged on the C-terminus and cloned into pShuttle-IRES-hrGFP-1 vector (Agilent) and submitted to Vector Biolabs for adenovirus plasmid construction and virus production.

Glucose tolerance test

Mice after 8-week of diet feeding were fasted for 14 hours and then injected with glucose 1.5g/kg (body weight) intraperitoneally. Glucose levels were monitored at 0min, 15min, 30min, 60min, 90 min and 120min using glucometer (Relion).

Measurement of ALT in mouse plasma

Mice were fasted for 5 hours before plasma collection. Blood was collected in Microvette coated with heparin (SARSTEDT) by retro-orbital bleeding. Plasma was prepared by centrifugation at 12,000g for 2min and subjected to IDEXX chemistry analyzer at Rodent Metabolic Phenotyping Core of University of Pennsylvania.

Measurement of hepatic triglyceride

Liver pieces were homogenized in cell lysis buffer (50mM Tris-HCl, 140mM NaCl, 0.1% Triton X-100 with freshly added protease inhibitor) by TissueLyser. Triglycerides in homogenized liver lysate were measured using Stanbio Triglyceride LiquiColor (Cat# 2100-430).

Tissue processing, histology evaluation, immunofluorescent staining and RNA *in situ* hybridization

For H&E staining, picrosirius red staining and immunofluorescence staining, mouse livers were fixed in 4% PFA overnight and dehydrated in a serial ethanol, xylene and embedded in paraffin. 5μm-thick sections were stained with H&E and picrosirius red/fast green (https://pharm.ucsf.edu/xinchen/protocols/picrosirius-red) for fibrosis stage evaluation. Images for histology were acquired using OLYMPUS BX43 microscope. Sirus red positive area was quantified by a macro function in Image J to apply consistent thresholding across all images. The NAS and fibrosis stage were evaluated as previously described (*68-70*). The evaluation was performed blindly by the comparative pathology core from School of Veterinary Medicine at University of Pennsylvania.

For the immunofluorescence staining and RNAscope of the liver tissue, paraffin sections were used. Antibodies and reagents used for immunofluorescence staining were as follows: anti-GFP (Abcam Cat# ab290, RRID:AB_303395; Abcam Cat# ab6673, RRID:AB_305643, 1:1000), anti-FLAG (Millipore Sigma Cat #F1804, RRID:AB_262044,1:500), anti-EphB2(Cell Signaling Technology, Cat# 83029, RRID:AB_2800007,1:500), anti-HNF4a (Thermo Fisher Scientific, Cat# MA1199, RRID:AB_2633309,1:500), anti-CD45 (Cell Signaling Technology, Cat# 70257, RRID:AB_2799780,1:200), anti-CD11b(Abcam, Cat# ab133357, RRID:AB_2650514), TUNEL(In Situ Cell Death Detection Kit, TMR red, Millipore Sigma Cat#12156792910), anti-Ifit3(Proteintech,Cat#15201-1, RRID:AB_2248738,1:500), anti-CXCR3 (LSBio, Cat# LS-B10183), ImmPRESS Detection kits (Vector laboratories Cat#MP7401,MP7402,MP7405), TSA Fluorescein reagent pack (Akoya biosciences Cat# SAT701001), TSA plus(Akoya biosciences, Cat# NEL742001KT). RNAscope probes used for the staining were as follows: *Ly6d* (ACD, Cat# 532071), *Cxcl9*(ACD, Cat# 489341), *Hnf4a* (ACD, Cat# 497651-C2), *Ephb2*(ACD, Cat#447611), *Grip1*(ACD, Cat#574411), *EPHB2* (ACD, Cat# 406251), *Aldh3a2* (ACD, Cat# 861951), *Gls2* (ACD, Cat# 449281), HNF4A (ACD, Cat# 442921-C2). Images for immunofluorescence staining and RNAscope were acquired using Leica TCS SP8 Confocal microscope at Cell and Developmental Biology Microscopy Core at University of Pennsylvania. RNAscope images and cell counting were blindly analyzed with CellProfiler using a published script with modification and the same parameters were automatically applied to all images (*71*). For human samples, the cells stained with background signals (overlapping signals appeared both in green and red channels) which cannot be removed by CellProfiler were manually removed.

Differentiation of human iPSC into hepatocyte like cells (HLCs)

iPSCs were maintained in mTeSR medium at 37 °C with 5% CO2. The differentiation protocol of iPSCs into HLCs has been previously described (*43*). The culture plates were coated with GELTREX matrix (Thermo Fisher Scientific, Cat# A1413302). To start differentiation, STEMdiff Definitive Endoderm Kit (STEMCELL Technologies, Cat# 05110) was used for 4 days when iPSC reached 30-40% confluency according to the manufacture's instruction. Next, RPMI-B27 (made with 2% B-27 Supplement) supplemented with BMP-4 (Peprotech, Cat# 120-05ET, 20 ng/mL) and FGF basic (Peprotech, Cat# 100-18B,10 ng/mL) was applied to the cells for 5 days in hypoxia condition (5% oxygen/5% CO2), and followed by RPMI-B27 (made with 2% B-27 Supplement, Thermo Fisher Scientific Cat#17504044) supplemented with HGF (Peprotech, Cat# 100-39H, 20 ng/mL) for another 5 days in hypoxia condition (5% oxygen/5% CO2). Finally, cells were transferred back to normoxia condition and treated with HCM Hepatocyte Culture Medium BulletKit (Lonza, Cat# CC-3198) without EGF for further maturation for 5 days. The successful HLCs differentiation was verified by immunofluorescence staining of hepatocyte markers using antibodies against HNF4α (Thermo Fisher Scientific, Cat# MA1199) and E-cadherin (BD Biosciences Cat# 610181).

Adenovirus transduction and Ephrin ligand treatment of HLCs

Adenovirus Ad-CMV-GFP and Ad-CMV-hEphB2 were transduced to HLCs with GeneJammer (Agilent) following the manufacture's instruction. 10,000 adenovirus particles per cell were applied. After 24 hours, the medium containing adenovirus was removed and kept in starving medium (HBM supplemented with BSA, Transferrin, Ascorbic Acid and GA-1000 from HCM Hepatocyte Culture Medium BulletKit, Lonza) for 4 hours. Next, the cells were treated with starving medium containing 2 µg/ml of the pre-clustered control human IgG-Fc (R&D Systems, Cat #110-HG-100, RRID:AB_276244) or recombinant Human EphrinA5 Fc (R&D Systems, Cat #374-EA-200) for 4 hours. RNA was extracted from the HLCs after the ligand treatment. To cluster the ligand, Fc proteins were incubated with 0.2 µg anti-human IgG (Jackson Immuno Research, 109-005-098,

RRID:AB_2337541) per µg of Fc protein in HBM for 30min at a final concentration of 10 µg/ml (*72, 73*).

RNA extraction and quantitative-PCR

For gene expression analysis, liver tissue was homogenized in Trizol using TissueLyser and cells were lysed in RLT buffer provided in RNeasy Mini Kit. RNA was extracted using RNeasy mini kit (QIAGEN) according to the manufacturer's instructions. cDNA was synthesized using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Quantitative real-time PCR(q-PCR) was conducted using Power SYBR Green PCR Master Mix (Applied Biosystems). The relative expression levels were normalized to the internal controls (*Rplp0* for mouse samples and *HPRT1* for human samples). The sequence of the primers used for qPCR was provided in **table S1**.

Luciferase assay

The 2 genomic regions detected by snATAC-seq at *Ephb2* locus that positively associated with *Ephb2* transcription in mNASH-Hep1/2 were chr4:136687766-136688266 and chr4:136735850-136736350 and names as Peak1 and Peak2 respectively. These 2 regions were amplified with Q5 High-Fidelity DNA Polymerases (New England Biolabs) using genomic DNA extracted from C57BL/6J mouse livers as template. The amplified regions were cloned into XhoI and KpnI restriction enzyme sites of pGL3-promoter vector (Promega, Cat#1761) to generate luciferase reporter plasmids. Luciferase assay was performed using murine hepatocyte cell line AML12 (ATCC, Cat# CRL2254) in 24-well plate format. 1*10^5 cells/per well were seeded the day before transfection. Next day, 3 µg of Myc-expression plasmid pD40-His/V5-c-Myc (Addgene Cat#45597) or NICD1-expression plasmid 3XFlagNICD1 (Addgene Cat#20183) were co-transfected with 300ng of luciferase reporter plasmids (pGL3-promoter only plasmid, or pGL3-promoter plasmids containing Peak1 or Peak 2 region), and 30ng of renilla luciferase plasmid using lipofectamine

3000(Thermo Fisher Scientific). 48 hours after transfection, luciferase activity was measured with Synergy HT plate reader (Biotek) using Dual-Luciferase Kit (Promega).

Flow cytometry: Mice were euthanized by CO_2 inhalation and perfused with 10 mL cold PBS through the right ventricle, after which the liver was removed and transferred to cold DMEM. The liver was minced with scissors and digested in DMEM + 1 mg / mL collagenase type IV (ThermoFisher 17104019) + 0.1 mg / mL DNase I (Sigma 10104159001) for 30 min at 37°C with end-over-end rotation. Digested livers were passed through a 70 μM strainer. Liver digest was centrifuged at 20g for 5 min at 4°C to pellet hepatocytes. Supernatant (containing leukocytes) was transferred to a clean tube and centrifuged at 450 rcf for 5 min at 4°C to pellet leukocytes. To remove debris, the cell pellet was resuspended in 30% isotonic Percoll (GE Healthcare GE17-0891-01) and centrifuged at 800g for 30 min at 4°C with low brake. Immune cells were collected from the cell pellet. Red blood cells were lysed with ACK lysing buffer (Quality Biological 118-156-721EA) for 5 min at room temperature. Fc receptors were blocked with anti-mouse CD16/32 (1:50, BD 553142) prior to staining with fluorochrome-conjugated monoclonal antibodies. Antibodies (**table S2**) were diluted in PBS at 1:200 and used to stain cell suspensions for 30 min on ice in the dark. Dead cells were labeled using Live/Dead Blue (1:2500, ThermoFisher L23105). Cells were washed with FACs buffer (PBS + 2% FBS + 10 mM HEPES + 1 mM EDTA) and resuspended in the same prior to flow cytometric analysis. Compensation was performed with cells (Live/Dead) or One Comp eBeads (ThermoFisher 01-1111-42). Cells were analyzed using a Cytek Aurora running SpectroFlo software (version 2.2, Cytek). Data were analyzed using FlowJo software (version 10.8.1) and Graphpad Prism 9 (GraphPad Software).

Western blot

Mouse liver tissue was lysed in RIPA buffer (150mM NaCl, 50mM Tris-HCl, 1%NP40, 0.1%SDS and 0.5% Na-deoxycholate) using TissueLyser. Protein lysates were resolved in 10% SDS-PAGE gel, transferred to polyvinylidene difluoride membranes, and blotted with primary antibodies. The membrane was detected by secondary antibody conjugated to HRP. The antibodies used for western blot were as follows: anti-EphB2 (Cell signaling Technology Cat#83029, RRID:AB_2800007), anti-Flag (Cell signaling Technology Cat#8146, RRID:AB_10950495), anti-vinculin-HRP (Cell signaling Technology, Cat#18799, RRID:AB_2714181), anti-rabbit-HRP (Cell signaling Technology, Cat#7074, RRID:AB_2099233); anti-mouse-HRP (Cell signaling Technology, Cat#7076, RRID:AB_330924).

Fig.S1

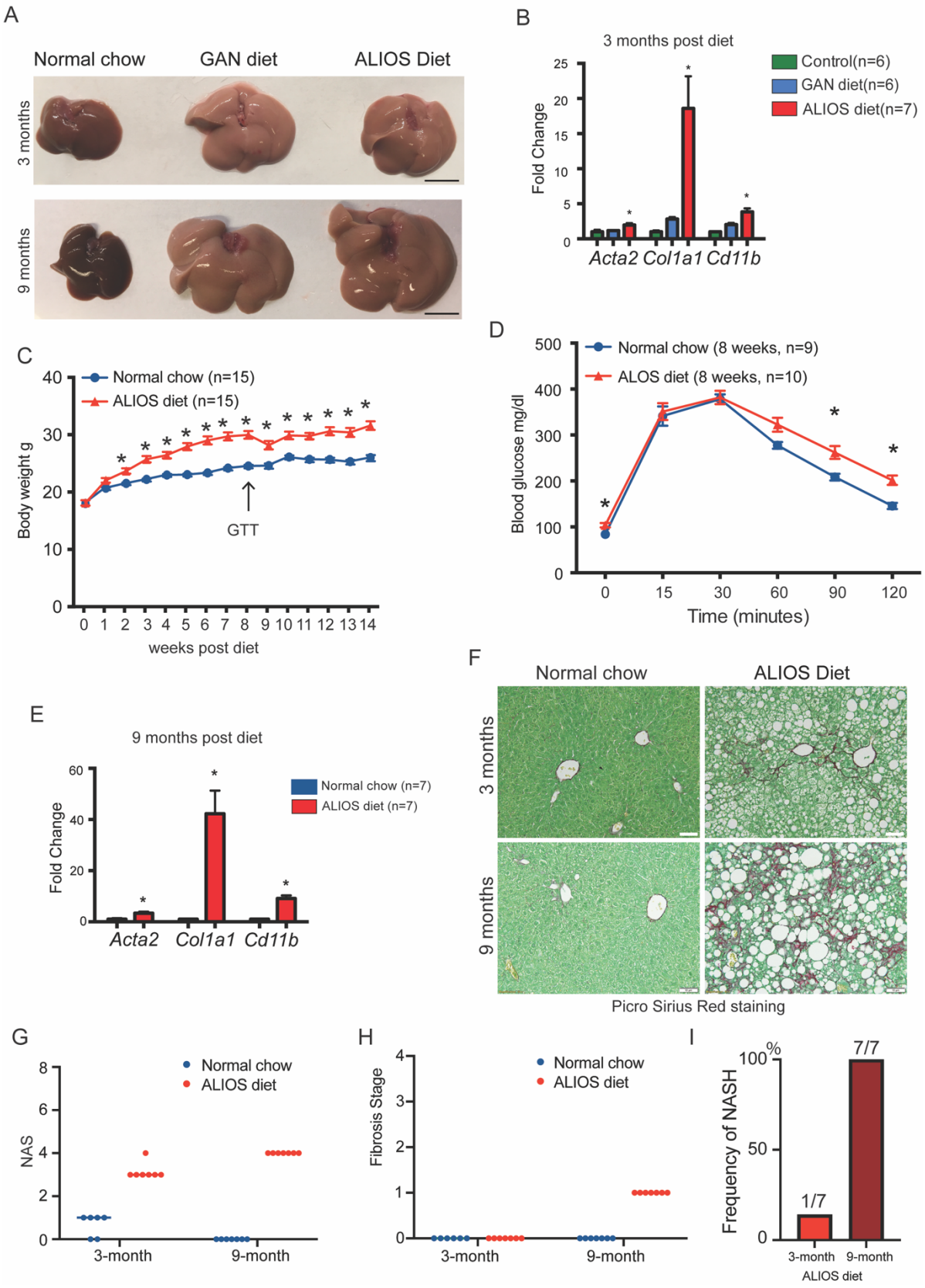
A
Normal chow
GAN diet
ALIOS Diet
3 months
9 months
B
3 months post diet
Control(n=6)
GAN diet(n=6)
ALIOS diet(n=7)
Fold Change
Acta2 Col1a1 Cd11b
C
Normal chow (n=15)
ALIOS diet (n=15)
Body weight g
GTT
weeks post diet
D
Normal chow (8 weeks, n=9)
ALOS diet (8 weeks, n=10)
Blood glucose mg/dl
Time (minutes)
E
9 months post diet
Normal chow (n=7)
ALIOS diet (n=7)
Fold Change
Acta2 Col1a1 Cd11b
F
Normal chow
ALIOS Diet
3 months
9 months
Picro Sirius Red staining
G
Normal chow
ALIOS diet
NAS
3-month
9-month
H
Normal chow
ALIOS diet
Fibrosis Stage
3-month
9-month
I
Frequency of NASH
100%
7/7
1/7
3-month
9-month
ALIOS diet

Figure S1. NASH promoting diet induces progressive NASH in mice and mimics metabolic changes in human patients.

(A) Representative liver gross morphology of the mice fed normal chow, GAN diet and ALIOS diet for 3 mo and 9 mo respectively (n=15 in each group, scale bar=1cm).

(B) mRNA of the fibrotic genes *Acta2* and *Col1a1* and inflammatory genes *Cd11b* and *Cd45* quantified by qPCR were elevated in the liver fed NASH promoting diet. Data are expressed as mean ± SEM, n=6-7 in each group, *p<0.05 Mann-Whitney U test.

(C) ALIOS diet induced body weight gain since the 2nd week of the diet treatment Data are expressed as mean ± SEM, n=15 in each group, *p<0.05 Mann-Whitney U test.

(D) Mice fed 8-week ALIOS diet developed glucose intolerance. Data are expressed as mean ± SEM, n=9-10 in each group,*p<0.05 Mann-Whitney U test.

(E) mRNA of *Acta2*, *Col1a1* and *Cd11b* quantified by qPCR were elevated in the liver fed ALISO diet for 9 mo. Data are expressed as mean ± SEM, n=7 in each group, *p<0.05 Mann-Whitney U test.

(F) Picro sirius red and fast green staining in mouse liver revealed mild fibrosis and severe fibrosis (red) after 3-mo and 9-mo ALIOS diet feeding respectively (n=15 in each group, scale bar=10μm).

(G) NAFLD activity score (NAS) of the liver fed normal chow or ALIOS diet for 3 mo or 9 mo (n=6-7 in each group).

(H) Fibrosis stage of the liver fed normal chow or ALIOS diet for 3 mo or 9 mo (n=6-7 in each group).

(I) The frequency of the mice fed ALIOS diet was diagnosed with NASH. Number of mice indicated above the bar graph.

Fig.S2

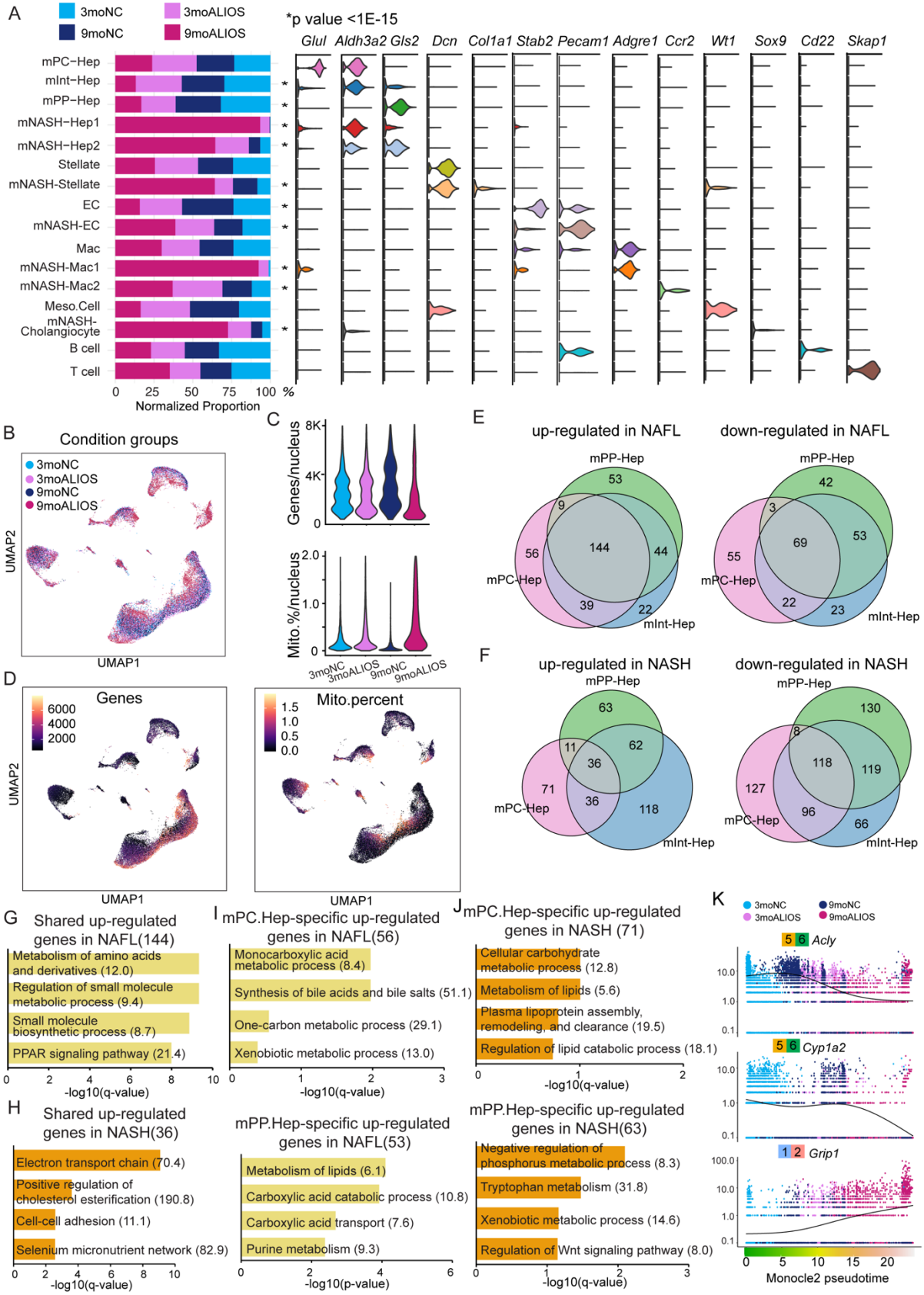
A
3moNC
3moALIOS
9moNC
9moALIOS
*p value <1E-15
mPC−Hep
mInt−Hep
mPP−Hep
mNASH−Hep1
mNASH−Hep2
Stellate
mNASH-Stellate
EC
mNASH-EC
Mac
mNASH-Mac1
mNASH-Mac2
Meso.Cell
mNASH-Cholangiocyte
B cell
T cell
0 25 50 75 100 %
Normalized Proportion
Glul Aldh3a2 Gls2 Dcn Col1a1 Stab2 Pecam1 Adgre1 Ccr2 Wt1 Sox9 Cd22 Skap1
B
Condition groups
3moNC
3moALIOS
9moNC
9moALIOS
UMAP2
UMAP1
C
Genes/nucleus
8K
4K
0
Mito.%/nucleus
2.0
1.0
0
3moNC 3moALIOS 9moNC 9moALIOS
D
Genes
6000
4000
2000
Mito.percent
1.5
1.0
0.5
0.0
UMAP2
UMAP1
UMAP1
E
up-regulated in NAFL
mPP-Hep
53
9
144
44
56
39
22
mPC-Hep
mInt-Hep
down-regulated in NAFL
mPP-Hep
42
3
69
53
55
22
23
mPC-Hep
mInt-Hep
F
up-regulated in NASH
mPP-Hep
63
11
62
36
71
36
118
mPC-Hep
mInt-Hep
down-regulated in NASH
mPP-Hep
130
8
118
119
127
96
66
mPC-Hep
mInt-Hep
G
Shared up-regulated genes in NAFL(144)
Metabolism of amino acids and derivatives (12.0)
Regulation of small molecule metabolic process (9.4)
Small molecule biosynthetic process (8.7)
PPAR signaling pathway (21.4)
0 2 4 6 8 10
-log10(q-value)
I
mPC.Hep-specific up-regulated genes in NAFL(56)
Monocarboxylic acid metabolic process (8.4)
Synthesis of bile acids and bile salts (51.1)
One-carbon metabolic process (29.1)
Xenobiotic metabolic process (13.0)
0 1 2 3
-log10(q-value)
J
mPC.Hep-specific up-regulated genes in NASH (71)
Cellular carbohydrate metabolic process (12.8)
Metabolism of lipids (5.6)
Plasma lipoprotein assembly, remodeling, and clearance (19.5)
Regulation of lipid catabolic process (18.1)
0 1 2
-log10(q-value)
K
3moNC
3moALIOS
9moNC
9moALIOS
5 6 Acly
10.0
1.0
0.1
5 6 Cyp1a2
10.0
1.0
0.1
1 2 Grip1
100.0
10.0
1.0
0.1
0 5 10 15 20
Monocle2 pseudotime
H
Shared up-regulated genes in NASH(36)
Electron transport chain (70.4)
Positive regulation of cholesterol esterification (190.8)
Cell-cell adhesion (11.1)
Selenium micronutrient network (82.9)
0 2 4 6 8 10
-log10(q-value)
mPP.Hep-specific up-regulated genes in NAFL(53)
Metabolism of lipids (6.1)
Carboxylic acid catabolic process (10.8)
Carboxylic acid transport (7.6)
Purine metabolism (9.3)
0 2 4 6
-log10(p-value)
mPP.Hep-specific up-regulated genes in NASH(63)
Negative regulation of phosphorus metabolic process (8.3)
Tryptophan metabolism (31.8)
Xenobiotic metabolic process (14.6)
Regulation of Wnt signaling pathway (8.0)
0 1 2 3
-log10(q-value)

Figure S2. Cell clustering of snRNA-seq on mouse livers and transcriptomic changes in hepatocytes across different zones.

(A) Left: The composition of individual clusters identified in snRNA-seq. Right: violin plot of transcript abundance for signature genes of each cluster colored according to Figure 1B.

(B) UMAP plot displayed cell clustering based on the experimental group origin.

(C-D) Quality control metrics of four snRNA-seq experimental groups. Violin plot showed the overall number of genes and percent of mitochondria genes recovery per nucleus and UMAP plot showed number of genes and percent of mitochondria genes recovery per nucleus by clusters.

(E-F) Venn diagram shown differentially expressed genes in hepatocytes across different liver zones in NAFL or NASH (abs(Log2FC) >0.5 and $p<0.05$).

(G-H) GO analysis of shared up-regulated genes in hepatocytes (Odds ratio indicated after each term).

(I-J) GO analysis of zone-specific up-regulated genes in hepatocytes (Odds ratio indicated after each term).

(K) Example genes in gene module 1/2 and 5/6 that determined hepatocyte trajectory bifurcation (the gene module color was consistent with Figure 1J-I and the Monocle2 pseudotime color was consistent with Figure 1F).

Fig.S3

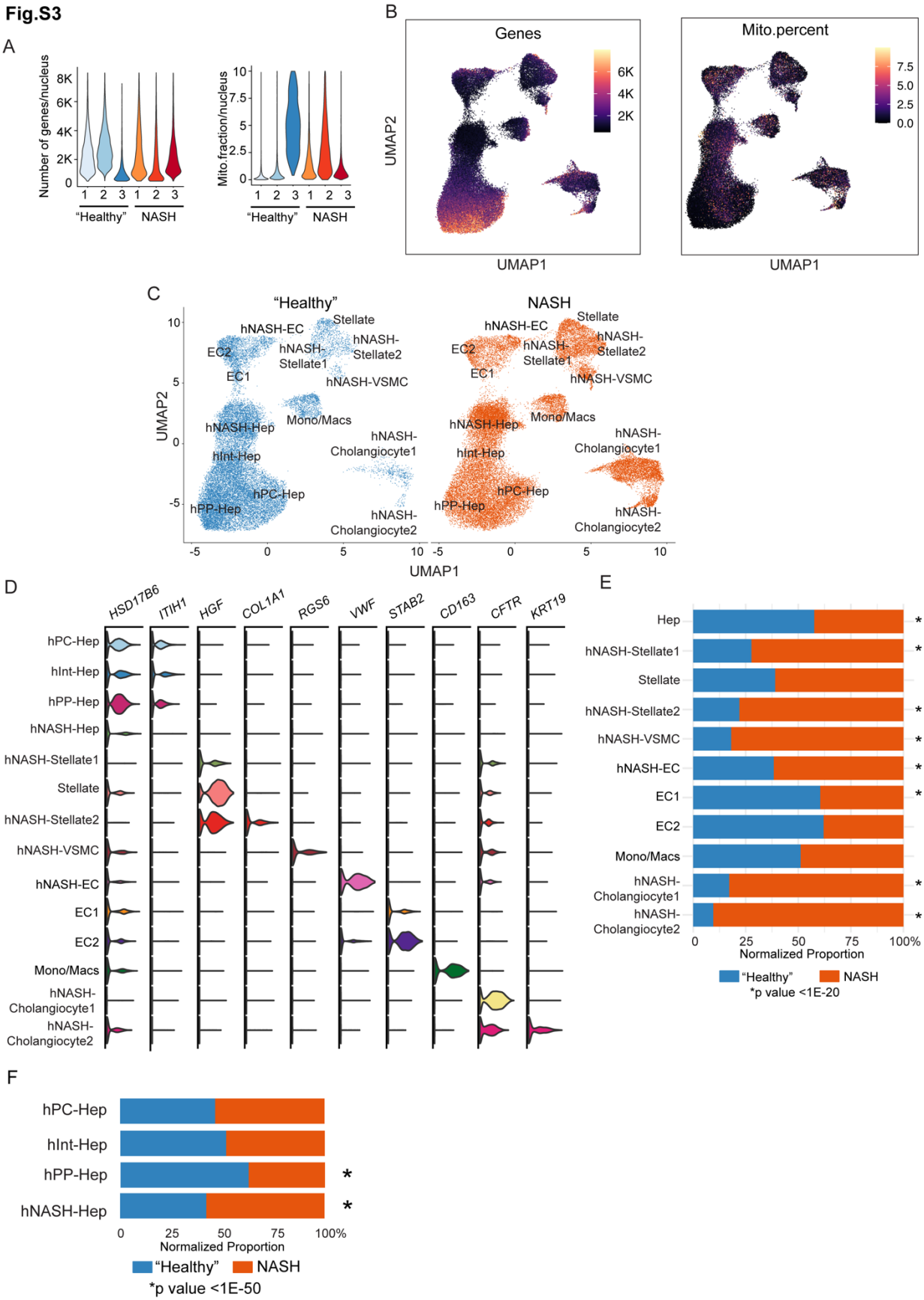
A
Number of genes/nucleus
8K
6K
4K
2K
0
1 2 3 1 2 3
"Healthy" NASH
Mito.fraction/nucleus
10
7.5
5
2.5
0
B
Genes
6K
4K
2K
UMAP2
UMAP1
Mito.percent
7.5
5.0
2.5
0.0
C
"Healthy"
NASH
hNASH-EC
Stellate
hNASH-Stellate2
EC2
hNASH-Stellate1
EC1
hNASH-VSMC
Mono/Macs
hNASH-Hep
hNASH-Cholangiocyte1
hInt-Hep
hPC-Hep
hPP-Hep
hNASH-Cholangiocyte2
UMAP2
UMAP1
D
HSD17B6
ITIH1
HGF
COL1A1
RGS6
VWF
STAB2
CD163
CFTR
KRT19
hPC-Hep
hInt-Hep
hPP-Hep
hNASH-Hep
hNASH-Stellate1
Stellate
hNASH-Stellate2
hNASH-VSMC
hNASH-EC
EC1
EC2
Mono/Macs
hNASH-Cholangiocyte1
hNASH-Cholangiocyte2
E
Hep
hNASH-Stellate1
Stellate
hNASH-Stellate2
hNASH-VSMC
hNASH-EC
EC1
EC2
Mono/Macs
hNASH-Cholangiocyte1
hNASH-Cholangiocyte2
0 25 50 75 100%
Normalized Proportion
"Healthy" NASH
*p value <1E-20
F
hPC-Hep
hInt-Hep
hPP-Hep
hNASH-Hep
0 25 50 75 100%
Normalized Proportion
"Healthy" NASH
*p value <1E-50

Figure S3. Cell clustering and cluster composition of snRNA-seq on human livers.

(A-B) Quality control metrics of snRNA-seq across different samples by violin plot and UMAP plot.

(C) UMAP plot visualization of clusters based on group conditions.

(D) Violin plot of expression level for signature genes of each cluster.

(E) Relative proportion of the group conditions in each cluster.

(F) Relative proportion of the group conditions in hepatocyte clusters.

Fig.S4

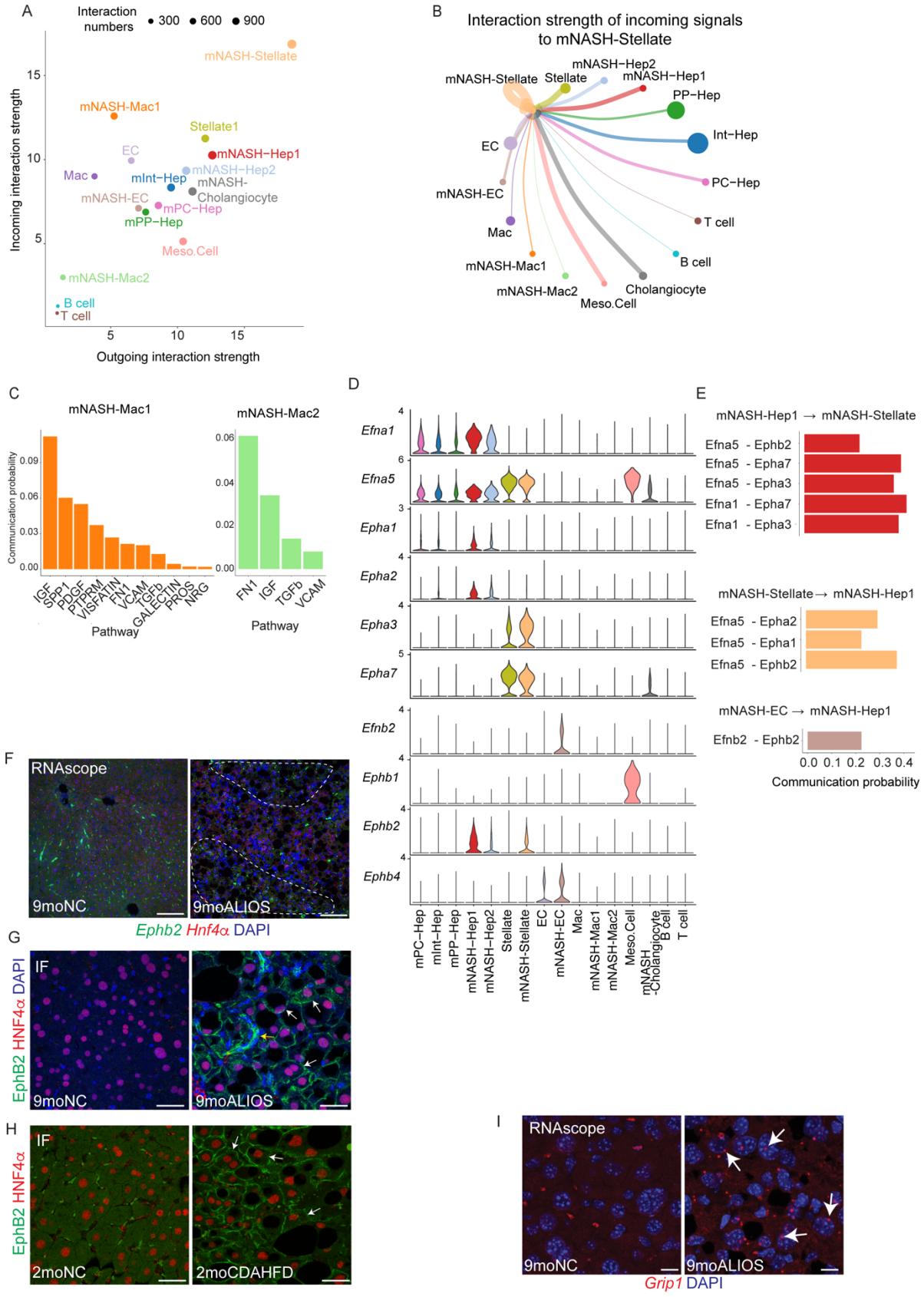
A
Interaction numbers
300
600
900
mNASH-Stellate
mNASH-Mac1
Stellate1
EC
mNASH−Hep1
Mac
mInt−Hep
mNASH−Hep2
mNASH-Cholangiocyte
mNASH-EC
mPC−Hep
mPP−Hep
Meso.Cell
mNASH-Mac2
B cell
T cell
Incoming interaction strength
Outgoing interaction strength
B
Interaction strength of incoming signals to mNASH-Stellate
mNASH-Stellate
Stellate
mNASH−Hep2
mNASH−Hep1
PP−Hep
Int−Hep
EC
PC−Hep
mNASH-EC
T cell
Mac
B cell
mNASH-Mac1
mNASH-Mac2
Cholangiocyte
Meso.Cell
C
mNASH-Mac1
mNASH-Mac2
Communication probability
IGF
SPP1
PDGF
PTPRM
VISFATIN
FN1
VCAM
TGFb
GALECTIN
PROS
NRG
Pathway
D
Efna1
Efna5
Epha1
Epha2
Epha3
Epha7
Efnb2
Ephb1
Ephb2
Ephb4
mPC−Hep
mInt−Hep
mPP−Hep
mNASH−Hep1
mNASH−Hep2
Stellate
mNASH-Stellate
EC
mNASH-EC
Mac
mNASH-Mac1
mNASH-Mac2
Meso.Cell
mNASH-Cholangiocyte
B cell
T cell
E
mNASH-Hep1 → mNASH-Stellate
Efna5 - Ephb2
Efna5 - Epha7
Efna5 - Epha3
Efna1 - Epha7
Efna1 - Epha3
mNASH-Stellate→ mNASH-Hep1
Efna5 - Epha2
Efna5 - Epha1
Efna5 - Ephb2
mNASH-EC → mNASH-Hep1
Efnb2 - Ephb2
Communication probability
F
RNAscope
9moNC
9moALIOS
Ephb2 Hnf4α DAPI
G
IF
EphB2 HNF4α DAPI
9moNC
9moALIOS
H
IF
EphB2 HNF4α
2moNC
2moCDAHFD
I
RNAscope
9moNC
9moALIOS
Grip1 DAPI

Figure S4. The intercellular communication in mouse livers.

(A) Intercellular communication strength computed by CellChat was visualized on a scatter plot. The color code was consistent with cell cluster colors in Figure 1B.

(B) Circle plots shown the interaction strength of incoming signals to mNASH-Stellate. The size of the dot corresponded to the cell number in each cluster. The color code was consistent with cell cluster colors in Figure 1B. The width of the line connecting any two dots indicated the interaction strength between the clusters.

(C) Signaling sent from mNASH-Mac1 (left) and mNASH-Mac2 (right) to mNASH-Stellate.

(D) The expression pattern of all detectable Ephrin-Eph family members. The color code was consistent with cell cluster colors in Figure 1B.

(E) The communication probability of Eph-Ephrin signaling pathway between mNASH-Hep1 and mNASH-Stellate/mNASH-EC.

(F) *Ephb2* was detected by RNAscope in a subset of hepatocytes (dotted circle, scale bar=100μm, n=3 in each group).

(G-H) Increased protein abundance of EphB2 in hepatocytes from 9moALIOS liver as well as 2moCDAHFD liver detected by immunofluorescent staining (white arrows pointing to the EphB2 positive hepatocytes which co-stained with Hnf4α; yellow arrows pointing to the EphB2 positive non-hepatocytes, scale bar=25μm, n=3 in each group).

(I) *Grip1* expression in hepatocytes visualized by RNA-scope was detected in 9moALIOS but not in 9moNC (scale bar=10μm, n=3 in each group).

Fig.S5

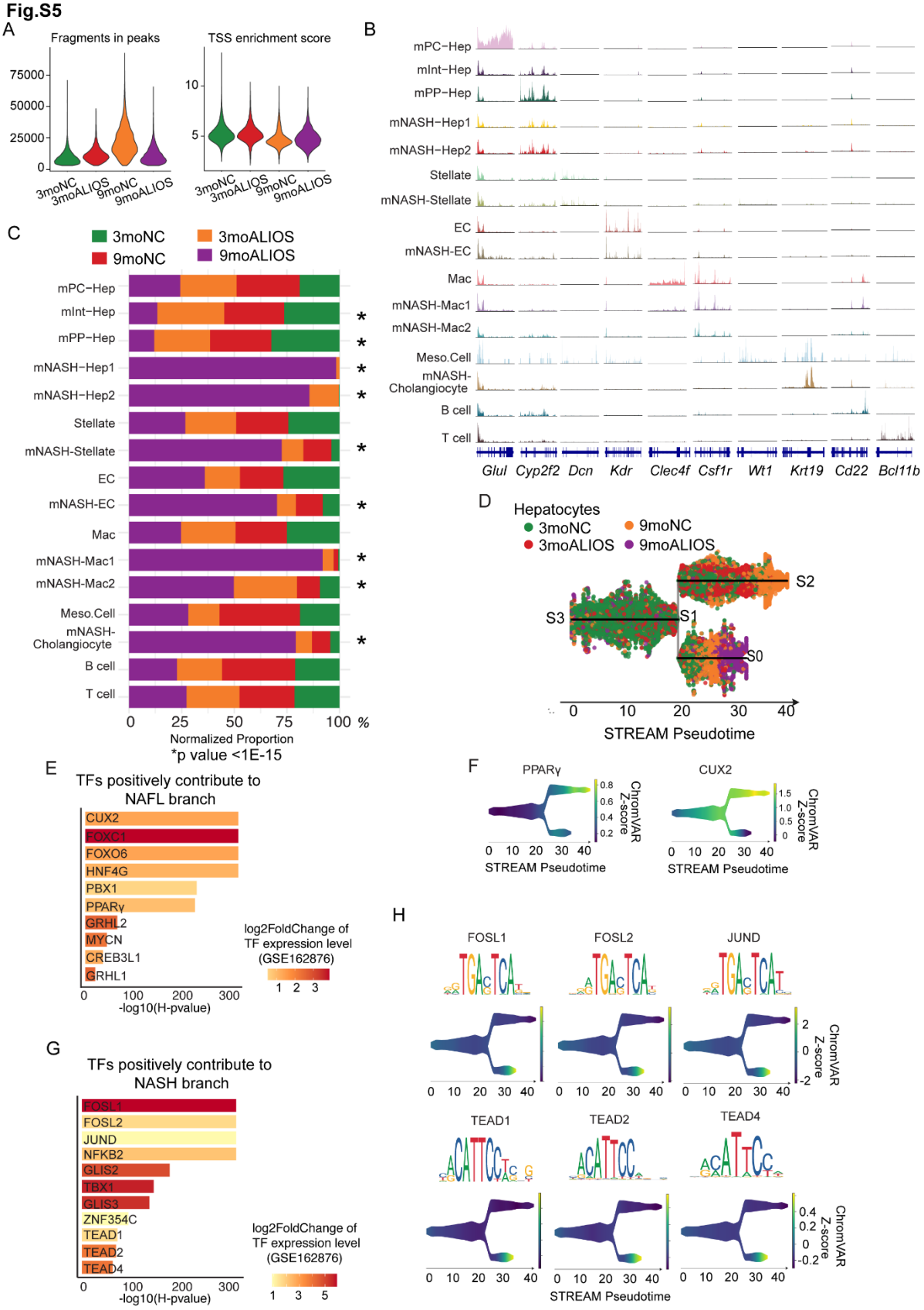
A
Fragments in peaks
TSS enrichment score
3moNC
3moALIOS
9moNC
9moALIOS
B
mPC−Hep
mInt−Hep
mPP−Hep
mNASH−Hep1
mNASH−Hep2
Stellate
mNASH-Stellate
EC
mNASH-EC
Mac
mNASH-Mac1
mNASH-Mac2
Meso.Cell
mNASH-Cholangiocyte
B cell
T cell
Glul Cyp2f2 Dcn Kdr Clec4f Csf1r Wt1 Krt19 Cd22 Bcl11b
C
Normalized Proportion
*p value <1E-15
D
Hepatocytes
S0
S1
S2
S3
STREAM Pseudotime
E
TFs positively contribute to NAFL branch
CUX2
FOXC1
FOXO6
HNF4G
PBX1
PPARγ
GRHL2
MYCN
CREB3L1
GRHL1
-log10(H-pvalue)
log2FoldChange of TF expression level (GSE162876)
F
PPARγ
CUX2
ChromVAR Z-score
STREAM Pseudotime
G
TFs positively contribute to NASH branch
FOSL1
FOSL2
JUND
NFKB2
GLIS2
TBX1
GLIS3
ZNF354C
TEAD1
TEAD2
TEAD4
-log10(H-pvalue)
log2FoldChange of TF expression level (GSE162876)
H
FOSL1
FOSL2
JUND
TEAD1
TEAD2
TEAD4
ChromVAR Z-score
STREAM Pseudotime

Figure S5. Characterization of snATAC-seq on mouse livers

(A) Violin plot shown comparable quality across different snATAC-seq mouse livers.

(B) Opening chromatin profiling of signature genes of all snATAC-seq clusters.

(C) Relative proportion of four experimental groups in each cell cluster of snATAC-seq (Chi-squared test, *p<10E-15).

(D) STREAM pseudotime plot of hepatocytes visualized at single cell level colored by experimental group.

(E) The transcription factors (TFs) were predicted to positively contribute to NAFL branch as well as their expression was up-regulated in FPC-mild condition from GSE162876.

(F) Selected examples of TFs that positively contribute to NAFL branch visualized in STREAM plot.

(G) The TFs were predicted to positively contribute to NASH branch as well as their expression is up-regulated in FPC-advanced condition from GSE162876.

(H) FOS/JUN family members that positively contributed to NASH branch were visualized in STREAM plot.

Fig.S6

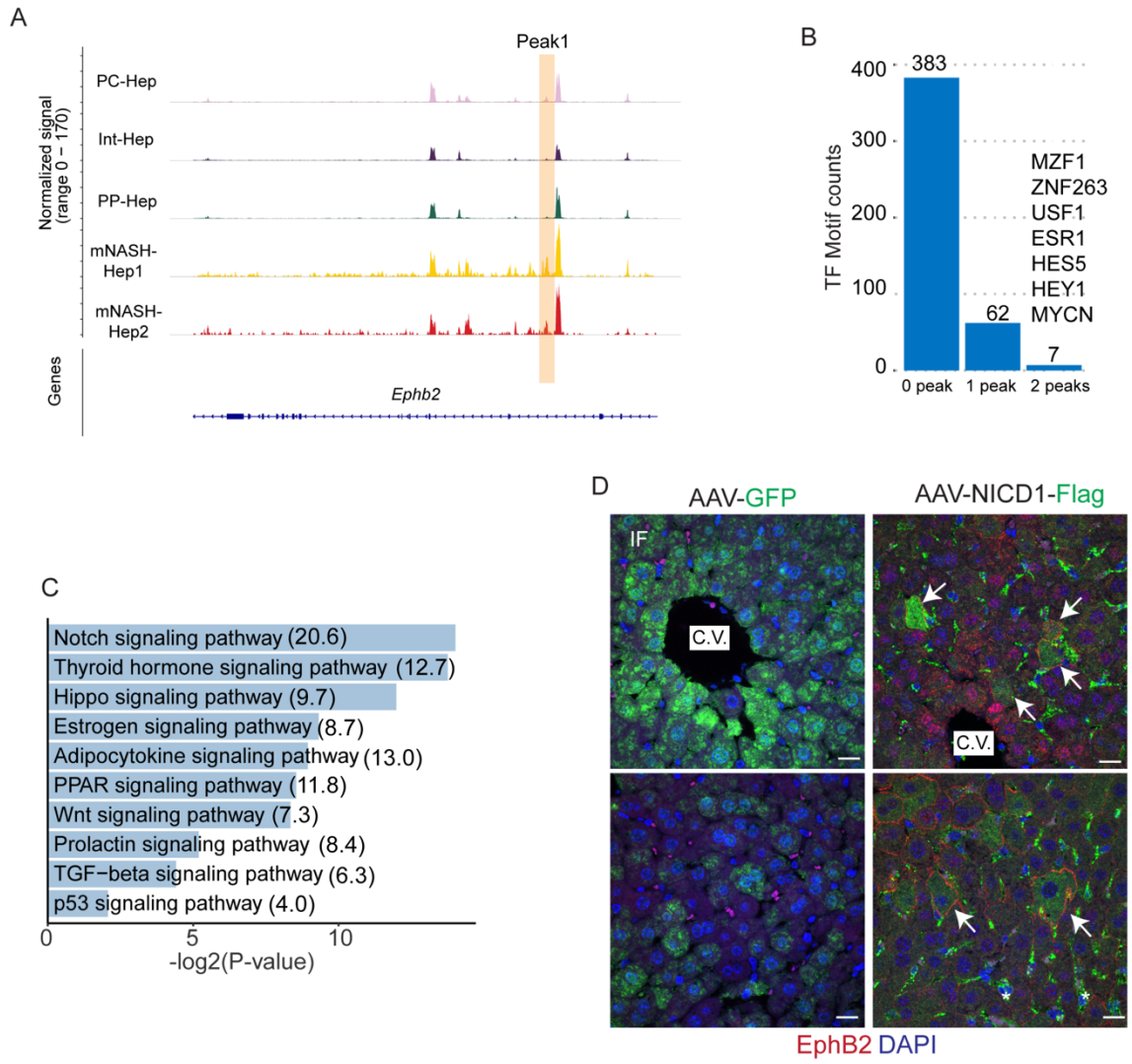

Figure S6. The candidate upstream signaling pathways of EphB2.

(A) Peak1 detected by snATAC-seq at *Ephb2* locus visualized in higher resolution.

(B) The distribution of TF binding motif occurrences in the 2 cis-elements at *Ephb2* locus identified by snATAC-seq.

(C) Enriched signaling pathway of all 69 TFs potentially regulating *Ephb2* transcription (Odds ratio indicated after each term).

(D) Elevated EphB2 (red) protein was detected in NICD-activated hepatocytes (green, right panels) but not in control hepatocytes stained with GFP (green, left panels) in pericentral zones (white arrows in upper panels) and intermediated zones (white arrows in bottom panels) (C.V.:

central vein, scale bar=12.5μm, stars marked non-specific staining due to the Flag antibody was produced in mouse, n=3 in each group).

Fig.S7

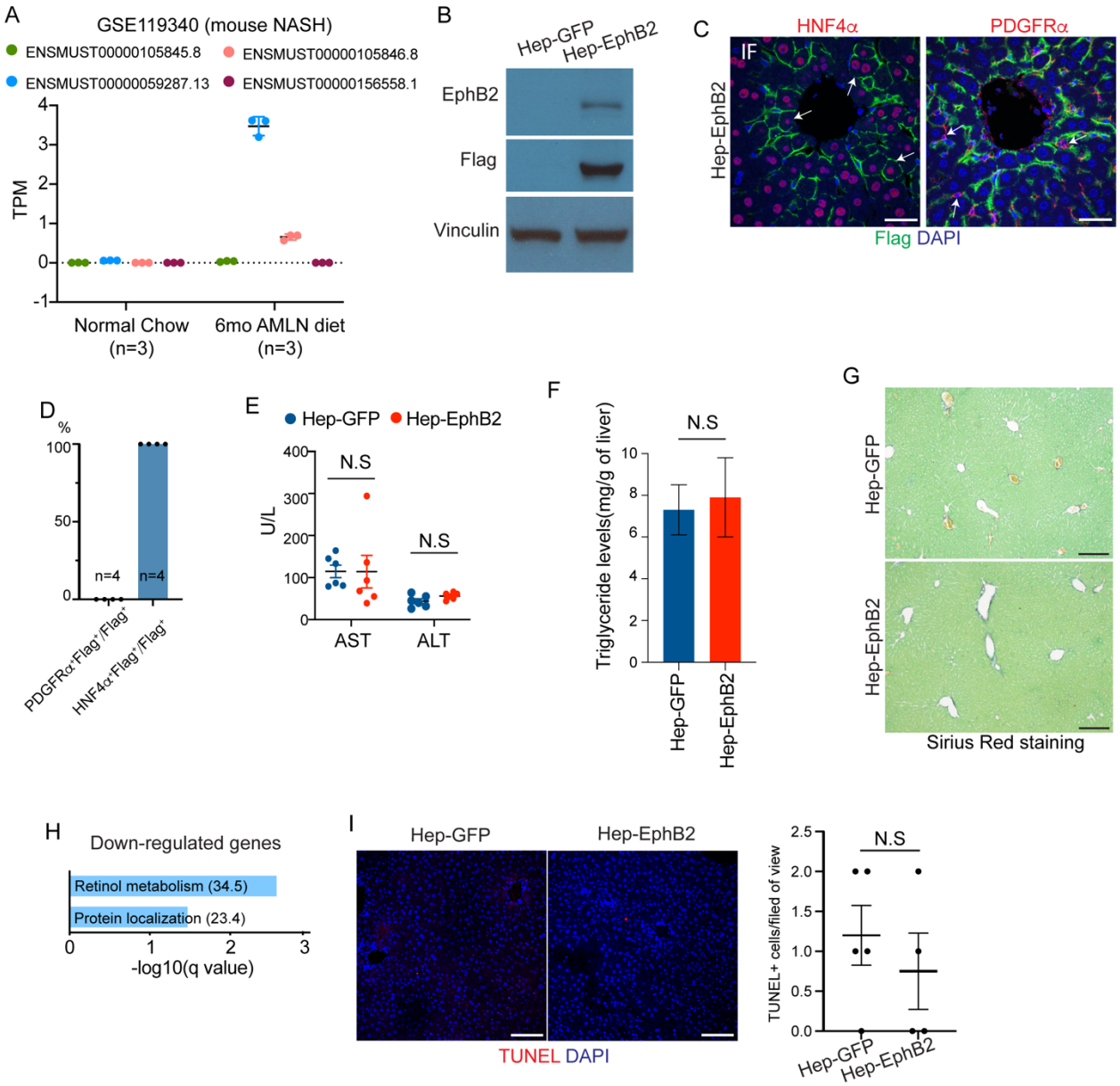

Figure S7. The characterization of EphB2 overexpression mouse model.

(A) Transcript-specific analysis of GSE119340 revealed ENSMUST00000059287.13 was the most responsive isoform of *Ephb2* to NASH (n=3 in each group).

(B) Western blot demonstrated an efficient activation of Flag epitope-tagged EphB2 in the livers from mice injected with AAV-EphB2-Flag using EphB2 or Flag antibodies (n=3 in each group).

(C-D) Force expression of Flag-tagged EphB2 in mouse livers (green) localized in hepatocytes (HNF4α positive cells, red on left panel) not in stellate cells (PDGFRα positive cells, red on right panel) (scale bar=25μm, n=4 in each group).

(E) AST and ALT concentrations were not different between Hep-GFP and Hep-EphB2 (n=6 in each group).

(F) Hepatic triglyceride concentrations were not different between Hep-GFP and Hep-EphB2 (n=3-4 in each group).

(G) No significant fibrosis was detected in Hep-EphB2 (n=3-4 in each group, scale bar=100μm).

(H) GO analysis of down-regulated genes of whole-liver RNA-seq in Hep-EphB2 (Odds ratio indicated after each term).

(I) Equivalent numbers of apoptotic cells stained by TUNEL were observed between Hep-EphB2 and Hep-GFP groups. Data are expressed as mean ± SEM, n=4-5 in each group, scale bar=100μm.

Fig.S8

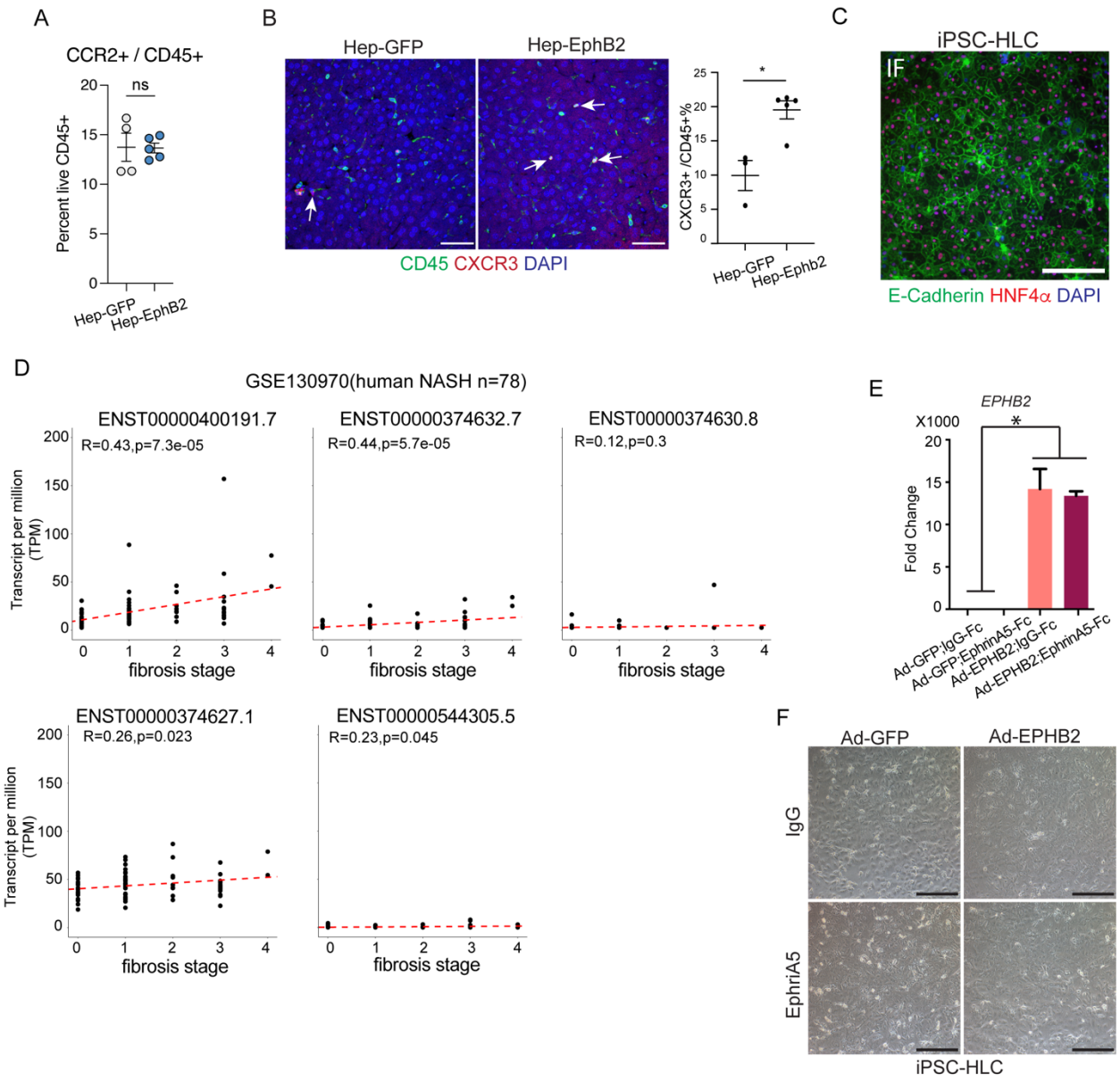

Figure S8. Characterization of EphB2 overexpression in hiPSC-HLC and mouse.

(A) The percentage of CCR2 cells were not different between Hep-GFP and Hep-EphB2 (n=4 in each group).

(B) IF demonstrated higher percentage of CXCR3 in Hep-EphB2. Data are expressed as mean ± SEM, *p<0.05, Mann-Whitney U test, n=3-5 in each group, scale bar=50μm.

(C) iPSC-HLC expressing hepatocyte markers HNF4α (red) and E-Cadherin (green),scale bar=100μm.

(D) Transcript-specific analysis of GSE130970 revealed ENST00000400191.7 was the most abundant isoform of *Ephb2* that significantly positively correlated with NASH fibrosis stage.

(E) Ad-EphB2 efficiently induced *EPHB2* expression. Data are expressed as mean ± SEM, n=3 in each group, *p<0.05, Mann-Whitney U test, n=3 in each group.

(F) Cell morphology of iPSC-HLC treated with adenovirus and ligands, n=3 in each group, scale bar=100μm.

Fig.S9

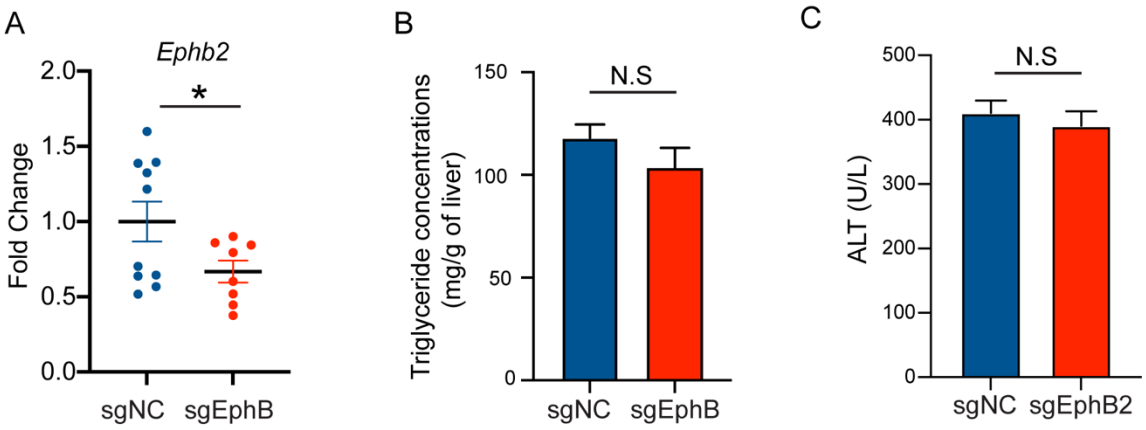

Figure S9. Characterization of EphB2 knockdown mouse model.

(A) The knockdown efficiency of *Ephb*2 detected by qPCR in sgEphB2 group. Data are expressed as mean ± SEM, *p<0.05, n=8-10 in each group.

(B) Hepatic triglyceride concentrations were not different between sgNC and sgEphB2, n=8-10 in each group.

(C) ALT concentrations were not different between sgNC and sgEphB2, n=8-10 in each group.

Fig.S10

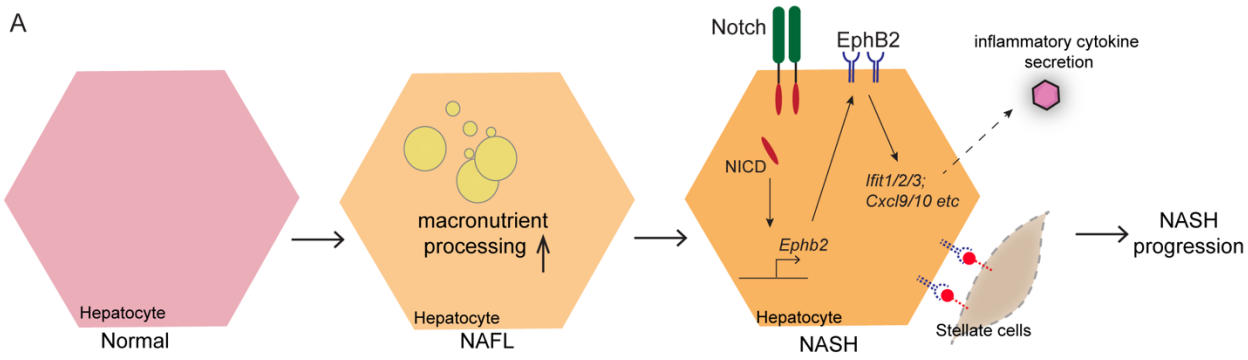

Figure S10. Proposed model for the function of EphB2 in hepatocytes during NASH progression.

(A) At the stage of NAFL, hepatocytes adaptively enhance macronutrient processing with little expression of EphB2. At the stage of NASH, hepatocytes acquire cell migration signatures and are marked with expression of EphB2. The transcription of *Ephb2* is induced by Notch signaling and drives cell-autonomous interferon pathway activation and inflammatory response, which contributes to fibrosis advancing. Moreover, EphB2-expressing hepatocytes possibly participates in intercellular crosstalk with NPCs such as stellate cells.

Table S1 Oligo sequence for qPCR and cloning

Arbp-Forward:	5'-TCCAGGCTTTGGGCATCA-3'
Arbp-Reverse:	5'-CTTTATCAGCTGCACATCACTCAGA-3'
Acta2-Forward:	5'-CAGCCATCTTTCATTGGGATGGAG-3'
Acta2-Reverse:	5'-AATGCCTGGGTACATGGTGG-3'
Cd11b-Forward:	5'-CTCACGTATCCGTGCCTTCTT-3'
Cd11b-Reverse:	5'-GTCCACGCAGTCCGGTAAA-3'
Cd45-Forward:	5'-ACTGAGCACAACAGAGAATGCC-3'
Cd45-Reverse:	5'-AGCGTGGATAACACACCTGGA-3'
Col1a1-Forward:	5'-GAACTGGACTGTCCCAACCC-3'
Col1a1-Reverse:	5'-TTGGGTCCCTCGACTCCTAC-3'
HPRT-Forward:	5'-TGACACTGGCAAAACAATGCA-3'
HPRT-Reverse:	5'-GGTCCTTTTCACCAGCAAGCT-3'
IFIT1-Forward:	5'-CCACAAGACAGAATAGCCAGAT-3'
IFIT1-Reverse:	5'-ATCTCAATTGCTCCAGACTATCC-3'
IFIT2-Forward:	5'-GCAACCATGAGTGAGAACAATAAG-3'
IFIT2-Reverse:	5'-TCCATCAAGTTCCAGGTGAAAT-3'
IFIT3-Forward:	5'-GTACAACTTGTTGGCCTACATAAA-3'
IFIT3-Reverse:	5'-GCATGTTCTTGCTGGATTAACT-3'
CCL2-Forward:	5'-TCATAGCAGCCACCTTCATTC-3'
CCL2-Reverse:	5'-CTCTGCACTGAGATCTTCCTATTG-3'
CXCL9-Forward:	5'-CATCATCTTGCTGGTTCTGATTG-3'
CXCL9-Reverse:	5'-GATTGTAGGTGGATAGTCCCTTG-3'
sgNC-Forward	5'-AAACGGTAATGCCTGGCTTGTCGACGCATAGTCT-3'
sgNC-Reverse	5'-CTTGAGACTATGCGTCGACAAGCCAGGCATTACC-3'
sgEphb2-Forward	5'-AAACGACTCAAAGACATTGCACACCTGCAAGTAAACCCCTACCAACTGGTCGGGGTTTGAAACATGACGATGCCATAGCTCCACAC-3'
sgEphb2-Reverse	5'-CTTGGTGTGGAGCTATGGCATCGTCATGTTTCAAACCCCGACCAGTTGGTAGGGGTTTACTTGCAGGTGTGCAATGTCTTTGAGTC-3'

Table S2. Antibodies used for flow cytometry

Antibody	Clone	Fluorochrome	Source	Identifier	RRID
Anti-mouse CD45	30 F11	BUV395	BD	564279	AB_2651134
Anti-mouse I-A/I-E	M5/114.15.2	Pacific Blue	BioLegend	107619	AB_493528
Anti-mouse Ly6C	HK1.4	BV510	BioLegend	128033	AB_2562351
Anti-mouse Ly6G	1A8	BV650	BioLegend	127641	AB_2565881
Anti-mouse CD11b	M1/70	BV785	BioLegend	101243	AB_2561373
Anti-mouse Tim4	RMT4-54	PerCP-Cy5.5	BioLegend	130019	AB_2876458
Anti-mouse CXCR3	S18001A	PE	BioLegend	155903	AB_2783131
Anti-mouse CD19	1D3	PE-CF594	BD	562291	AB_11154223
Anti-mouse CD8	53-6.7	PE-Cy5	BioLegend	100710	AB_312749
Anti-mouse CD4	RM4-5	AF 700	eBioscience	56-0042-80	AB_494002
Anti-mouse Siglec F	E50-2440	APC-Cy7	BD	565527	AB_2732831