Resolving the fibrotic niche of human liver cirrhosis at single-cell level

https://doi.org/10.1038/s41586-019-1631-3

Received: 4 September 2018

Accepted: 4 September 2019

Published online: 9 October 2019

P. Ramachandran[1]*, R. Dobie[1], J. R. Wilson-Kanamori[1], E. F. Dora[1], B. E. P. Henderson[1], N. T. Luu[2,3], J. R. Portman[1], K. P. Matchett[1], M. Brice[1], J. A. Marwick[1,4], R. S. Taylor[1], M. Efremova[5], R. Vento-Tormo[5], N. O. Carragher[4], T. J. Kendall[1,6], J. A. Fallowfield[1], E. M. Harrison[7], D. J. Mole[1,7], S. J. Wigmore[1,7], P. N. Newsome[2,3], C. J. Weston[2,3], J. P. Iredale[8], F. Tacke[9], J. W. Pollard[10,11], C. P. Ponting[12], J. C. Marioni[5,13,14], S. A. Teichmann[5,13,15] & N. C. Henderson[1]*

Liver cirrhosis is a major cause of death worldwide and is characterized by extensive fibrosis. There are currently no effective antifibrotic therapies available. To obtain a better understanding of the cellular and molecular mechanisms involved in disease pathogenesis and enable the discovery of therapeutic targets, here we profile the transcriptomes of more than 100,000 single human cells, yielding molecular definitions for non-parenchymal cell types that are found in healthy and cirrhotic human liver. We identify a scar-associated TREM2$^+$CD9$^+$ subpopulation of macrophages, which expands in liver fibrosis, differentiates from circulating monocytes and is pro-fibrogenic. We also define ACKR1$^+$ and PLVAP$^+$ endothelial cells that expand in cirrhosis, are topographically restricted to the fibrotic niche and enhance the transmigration of leucocytes. Multi-lineage modelling of ligand and receptor interactions between the scar-associated macrophages, endothelial cells and PDGFRα$^+$ collagen-producing mesenchymal cells reveals intra-scar activity of several pro-fibrogenic pathways including TNFRSF12A, PDGFR and NOTCH signalling. Our work dissects unanticipated aspects of the cellular and molecular basis of human organ fibrosis at a single-cell level, and provides a conceptual framework for the discovery of rational therapeutic targets in liver cirrhosis.

Recent estimates suggest that 844 million people worldwide have chronic liver disease, with two million deaths per year and a rising incidence[1]. Iterative liver injury secondary to any cause leads to progressive fibrosis and ultimately results in liver cirrhosis. Notably, the degree of liver fibrosis predicts adverse patient outcomes[2]. Hence, effective antifibrotic therapies for patients with chronic liver disease are urgently required[3,4].

Liver fibrosis involves a complex interplay between multiple non-parenchymal cell (NPC) lineages including immune, endothelial and mesenchymal cells spatially located within areas of scarring, termed the fibrotic niche. Despite progress in our understanding of liver fibrogenesis accrued using rodent models, there remains a considerable 'translational gap' between putative targets and effective patient therapies[3,4]. This is in part due to limited definition of the functional heterogeneity and interactome of cell lineages that contribute to the fibrotic niche of human liver cirrhosis, which is imperfectly recapitulated by rodent models[3].

Single-cell RNA sequencing (scRNA-seq) is delivering a step change in our understanding of disease pathogenesis, allowing the interrogation of individual cell populations at unprecedented resolution[5]. Here, we studied the mechanisms that regulate human liver fibrosis using scRNA-seq.

Single-cell atlas of human liver NPCs

Hepatic NPCs were isolated from healthy and cirrhotic human livers spanning a range of aetiologies of cirrhosis (Fig. 1a, Extended Data Fig. 1a). Leucocytes (CD45$^+$) or other NPC (CD45$^-$) fractions (Extended Data Fig. 1b) were sorted by flow cytometry before scRNA-seq analysis. To discriminate between liver-resident and circulating leucocytes, we also performed scRNA-seq on CD45$^+$CD66b$^-$ peripheral blood mononuclear cells (PBMCs) (Extended Data Fig. 1c, g–i). The combined tissue and PBMC dataset was partitioned into clusters (Extended Data Fig. 1d) and annotated using signatures of known lineage markers (Extended

[1]University of Edinburgh Centre for Inflammation Research, The Queen's Medical Research Institute, Edinburgh BioQuarter, Edinburgh, UK. [2]NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK. [3]Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK. [4]Cancer Research UK Edinburgh Centre, MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Edinburgh, UK. [5]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK. [6]Division of Pathology, University of Edinburgh, Edinburgh, UK. [7]Clinical Surgery, University of Edinburgh, Royal Infirmary of Edinburgh, Edinburgh, UK. [8]Office of the Vice Chancellor, Beacon House and National Institute for Health Research, Biomedical Research Centre, Bristol, UK. [9]Department of Hepatology and Gastroenterology, Charité University Medical Center, Berlin, Germany. [10]MRC Centre for Reproductive Health, The Queen's Medical Research Institute, University of Edinburgh, Edinburgh, UK. [11]Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, New York, NY, USA. [12]MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Edinburgh, UK. [13]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge, UK. [14]Cancer Research UK Cambridge Institute, Li Ka Shing Centre, University of Cambridge, Cambridge, UK. [15]Theory of Condensed Matter Group, The Cavendish Laboratory, University of Cambridge, Cambridge, UK. *e-mail: prakash.ramachandran@ed.ac.uk; neil.henderson@ed.ac.uk

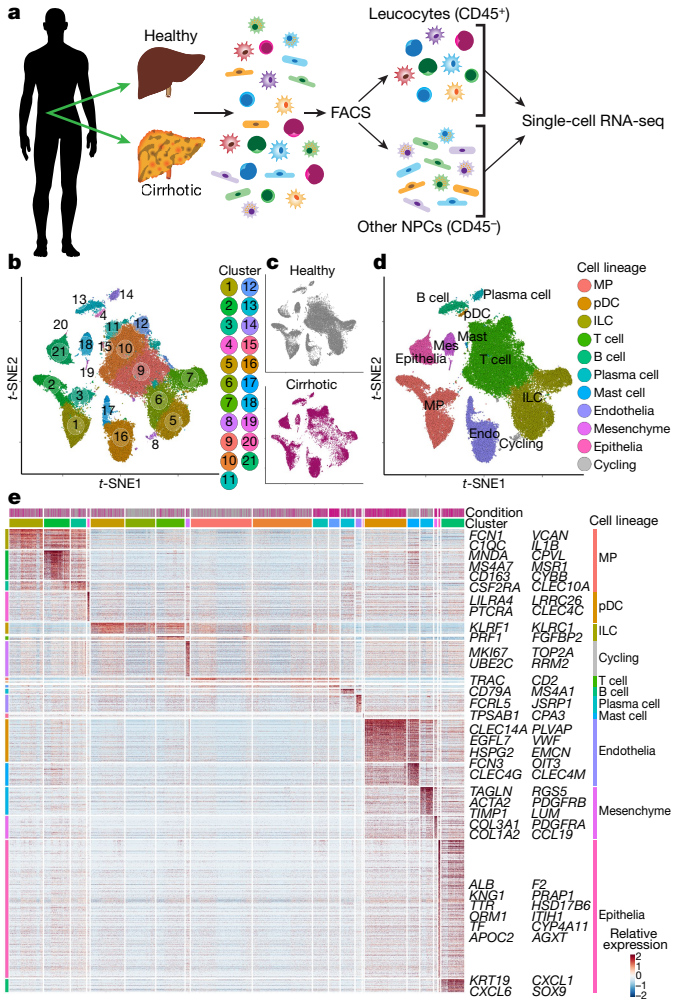


Fig. 1 | Single-cell atlas of human liver NPCs. a, Overview, illustrating the isolation, FACS sorting and scRNA-seq analysis of leucocytes (CD45$^+$) and other NPC fractions (CD45$^-$). **b**, Clustering of 66,135 cells from healthy ($n = 5$) and cirrhotic ($n = 5$) human livers. **c**, Annotation by injury condition. **d**, Cell lineage inferred from expression of marker gene signatures. ILC, innate lymphoid cell; MP, mononuclear phagocyte; pDC, plasmacytoid dendritic cell. **e**, Heat map of cluster marker genes (top, colour-coded by cluster and condition), with cell lineage of exemplar genes labelled (right). Columns denote cells; rows denote genes.

Data Fig. 1d, e, Supplementary Table 1). To generate an atlas of liver-resident cells, contaminating circulating cells were removed from the liver tissue datasets, by excluding cells from the tissue samples which mapped transcriptionally to blood-derived clusters 1 and 13 (Extended Data Fig. 1d). Liver-resident cells expressed higher levels of tissue-residency markers such as CXCR4 compared with PBMCs (Extended Data Fig. 1f).

Re-clustering the 66,135 liver-resident cells from 10 livers ($n = 5$ healthy and $n = 5$ cirrhotic) revealed 21 populations (Fig. 1b), each containing cells from both healthy and cirrhotic livers (Fig. 1c, Extended Data Fig. 2), across 10 cell lineages (Fig. 1d, Extended Data Fig. 2a, b). Subpopulation markers were identified across all clusters and lineages (Fig. 1e, Supplementary Tables 3, 4). Quality control metrics were highly reproducible between individual samples and conditions (Extended Data Fig. 2c–f, Supplementary Table 2). Expression of collagens type I and III, the main fibrillar collagens within the fibrotic niche, was restricted to cells of the mesenchymal lineage (Fig. 1e).

We proceeded to annotate all human liver NPC lineages (below, Supplementary Notes 1–3, Extended Data Fig. 3), and provide an open-access gene browser (http://www.livercellatlas.mvm.ed.ac.uk) that allows assessment of NPC gene expression between healthy and cirrhotic livers.

Distinct macrophages inhabit the fibrotic niche

Previous studies in rodents have highlighted macrophage subpopulations that orchestrate both the progression and regression of liver fibrosis[6–8]. Clustering of mononuclear phagocytes (MPs) identified ten clusters; annotated as scar-associated macrophages (SAMacs), Kupffer cells (KCs), tissue monocytes (TMs), conventional dendritic cells (cDCs) and cycling (proliferating) cells (Fig. 2a, Extended Data Fig. 4a, Supplementary Note 2). Clusters MP(4) and MP(5)—named SAMac(1) and SAMac(2), respectively—were expanded in cirrhotic livers (Fig. 2b), as confirmed by quantification of the MP cell composition of each liver individually (Fig. 2c).

Clusters MP(6) and MP(7) were enriched in the expression of *CD163*, *MARCO* and *TIMD4* (Extended Data Fig. 4b); tissue staining confirmed these as KCs (resident liver macrophages), facilitating the annotation of these clusters as KC(1) and KC(2), respectively (Extended Data Fig. 4c). A lack of *TIMD4* expression distinguished cluster KC(2) from KC(1) (Extended Data Fig. 4b); cell counting demonstrated TIMD4$^+$ cell numbers to be equivalent between healthy and cirrhotic livers, but showed a loss of MARCO$^+$ cells, consistent with a selective reduction in MARCO$^+$TIMD4$^-$ KC(2) in liver fibrosis (Fig. 2c, Extended Data Fig. 4d, e).

Scar-associated clusters SAMac(1) and SAMac(2) expressed the unique markers *TREM2* and *CD9* (Fig. 2d, e). These macrophages displayed a hybrid phenotype, with features of both TMs and KCs (Fig. 2d, e), analogous to monocyte-derived macrophages in mouse liver injury models[7]. Flow cytometry confirmed expansion of TREM2$^+$CD9$^+$ macrophages in human fibrotic livers (Fig. 2f, Extended Data Fig. 4f). Conditioned medium from SAMacs after fluorescence-activated cell sorting (FACS) promoted fibrillar collagen expression by primary human hepatic stellate cells (HSCs) (Fig. 2g), indicating that SAMacs have a pro-fibrogenic phenotype. Tissue staining demonstrated the presence of TREM2$^+$CD9$^+$MNDA$^+$ SAMacs topographically localized in collagen-positive scar regions (Fig. 2h, Extended Data Fig. 4g–i), and significantly expanded in cirrhotic livers (Extended Data Fig. 4j, k). Cell counting of stained cirrhotic livers morphologically segmented into regions of fibrotic septae and parenchymal nodules, confirmed SAMac accumulation within the fibrotic niche (Extended Data Fig. 4l).

Local proliferation has a significant role in macrophage expansion at sites of fibrosis in rodent models[7,9]. Cycling MP cells (Fig. 2a) subclustered into subpopulations of conventional dendritic cells (cDC1 and cDC2), KCs and SAMacs (Extended Data Fig. 4m, Supplementary Table 8). Cycling SAMacs expanded in cirrhosis (Extended Data Fig. 4m), which highlights the potential role of macrophage proliferation in promoting SAMac accumulation in the fibrotic niche.

Pro-fibrogenic phenotype of SAMacs

To delineate the functional profile of SAMacs, we visualized co-ordinately expressed gene groups across the MP subpopulations using self-organizing maps (Extended Data Fig. 5a). We identified six optimally differentiating metagene signatures, denoted as A–F (Extended Data Fig. 5a, Supplementary Table 9). Signatures A and B defined SAMacs and were enriched for ontology terms relevant to tissue fibrosis (Extended Data Fig. 5b). These SAMac-defining signatures included genes such as *TREM2*, *IL1B*, *SPP1*, *LGALS3*, *CCR2* and *TNFSF12*, some of which are known to regulate the function of scar-producing myofibroblasts in fibrotic liver diseases[10–13]. The remaining MP subpopulations were defined by signature C (KCs), signatures D, E (TMs) and signature F (cDC1); ontology terms matched known functions for the associated cell type (Extended Data Fig. 5b, Supplementary Table 9).

In mice, under homeostatic conditions, embryologically derived self-renewing tissue-resident KCs predominate[14–16]. However, after

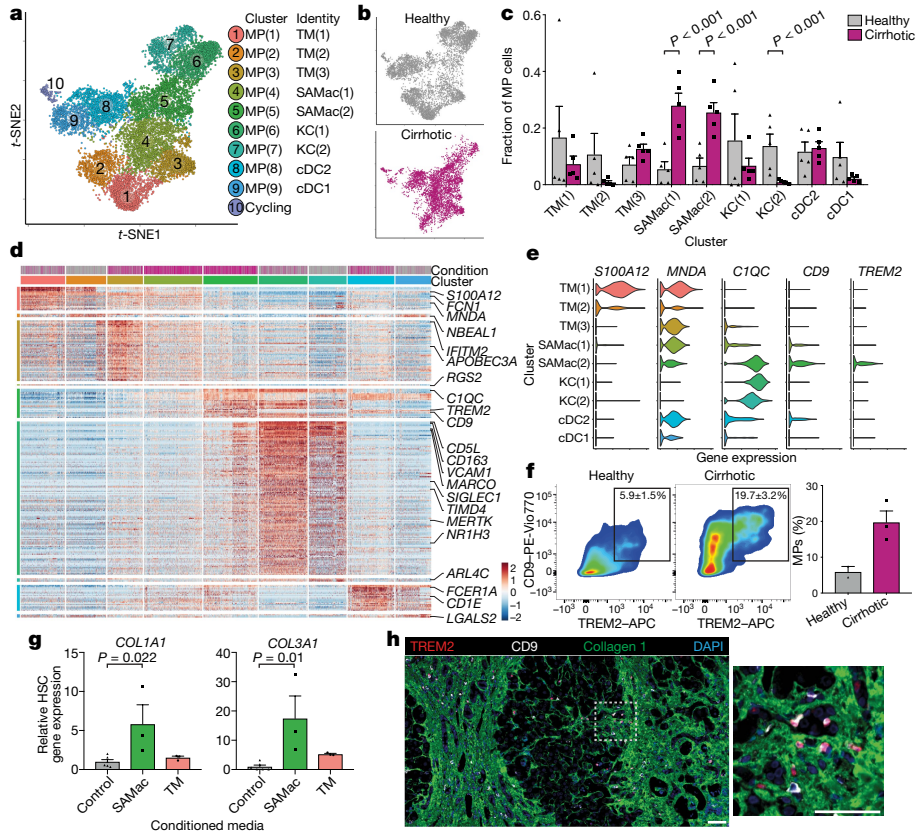


Fig. 2 | Identifying SAMac subpopulations. a, Clustering of 10,737 MPs from healthy ($n = 5$) and cirrhotic ($n = 5$) human livers. **b**, Annotation by injury condition. **c**, Fractions of MP subpopulations in healthy ($n = 5$) and cirrhotic ($n = 5$) livers. **d**, Heat map of MP cluster marker genes (top, colour-coded by cluster and condition), with exemplar genes labelled (right). Columns denote cells; rows denote genes. **e**, Scaled gene expression of SAMac and TM cluster markers across MP cells from healthy ($n = 5$) and cirrhotic ($n = 5$) livers. **f**, Flow cytometry analysis of TREM2$^+$CD9$^+$ MP fraction in healthy ($n = 2$) and cirrhotic ($n = 3$) liver. **g**, Primary human HSCs were treated with conditioned medium from SAMacs ($n = 3$) or TMs ($n = 3$), and indicated genes were analysed by quantitative PCR (qPCR). Expression is shown relative to mean expression of untreated control HSCs ($n = 6$). **h**, Representative immunofluorescence images ($n \geq 3$) of TREM2 (red), CD9 (white), collagen 1 (green) and DAPI (blue) in cirrhotic liver. All scale bars, 50 μm. Data are mean ± s.e.m. *P* values determined by Wald test (**c**) or Kruskal–Wallis and Dunn tests (**g**).

injury, macrophages derived from circulating monocytes accumulate in the liver and regulate fibrosis[7,8]. The ontogeny of human hepatic macrophage subpopulations is unknown. TREM2$^+$CD9$^+$ SAMacs demonstrated a monocyte-like morphology (Fig. 2h, Extended Data Fig. 4g–i) and a distinct topographical distribution from KCs (Extended Data Fig. 4l). To assess the origin of SAMacs, we performed in silico trajectory analysis on a combined dataset of peripheral blood monocytes and liver-resident MPs. We visualized the transcriptional profile of these cells (Fig. 3a, Extended Data Fig. 5c), mapped them along a pseudotemporal trajectory and interrogated their directionality via spliced and unspliced mRNA ratios (RNA velocity[17]). These analyses suggested a differentiation trajectory from peripheral blood monocytes into either SAMacs or cDCs, with no differentiation from KCs to SAMacs, and no progression from SAMacs to KCs (Fig. 3a, Extended Data Fig. 5c). Additional RNA velocity analyses[17] showed downregulation (negative velocity) of the monocyte gene *MNDA* in SAMacs, upregulation (positive velocity) of the SAMac marker gene *CD9* in TMs, and a lack of KC gene *TIMD4* velocity in SAMacs (Extended Data Fig. 5d). Furthermore, assessment of the probabilities of cells in this dataset transitioning into SAMacs indicated a higher likelihood of TMs than KCs differentiating into SAMacs (Fig. 3b). Overall, these data suggest that SAMacs are monocyte-derived, and represent a terminally differentiated cell state within the fibrotic niche.

To characterize the SAMac phenotype further, we identified differentially expressed genes along monocyte differentiation trajectories. We defined three gene co-expression modules, with module 1 representing upregulated genes during blood monocyte-to-SAMac differentiation (Fig. 3c). Module 1 contained multiple pro-fibrogenic genes including *SPP1*, *LGALS3*, *CCL2*, *CXCL8*, *PDGFB* and *VEGFA*[10–13], and displayed ontology terms that are consistent with the promotion of tissue fibrosis and angiogenesis (Fig. 3c, d, Supplementary Table 10). Module 2 contained genes that were downregulated during the differentiation of monocytes to SAMacs (Fig. 3c, Extended Data Fig. 5e), whereas module 3 encompassed a group of upregulated genes during the differentiation from monocytes to cDCs (Fig. 3c, Extended Data Fig. 5f, Supplementary Table 10). SAMacs isolated from cirrhotic human livers (Fig. 2f, Extended Data Fig. 4f) demonstrated enhanced protein secretion of several of the mediators identified by transcriptional analysis (Extended Data Fig. 5g) and promoted fibrillar collagen expression by primary human HSCs (Fig. 2g), which confirms that SAMacs have a pro-fibrogenic phenotype.

To enable cross-species comparison, we performed scRNA-seq on liver MP cells isolated from control mice or mice treated with chronic carbon tetrachloride (CCl_4)—a mouse model of liver fibrosis[7]. MP cells from fibrotic livers were isolated 24 h after the final CCl_4 injection, a time of active fibrogenesis[7]. Five MP cell clusters were defined (Extended Data Fig. 6a–d, Supplementary Table 11), and injury-specific cluster mMP(2) was differentiated by high expression of *Cd9*, *Trem2*, *Spp1* and *Lgals3* (Extended Data Fig. 6a–d). We confirmed expansion of this CD9$^+$ mSAMac population in liver fibrosis (Extended Data Fig. 6e, f) and co-culture of mSAMacs with quiescent primary mouse HSCs promoted fibrillar collagen expression in HSCs (Extended Data Fig. 6g). Canonical

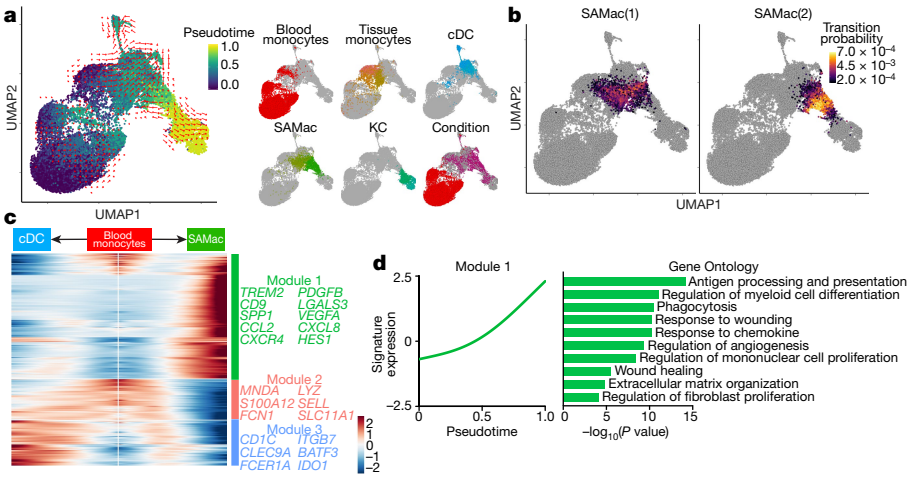


Fig. 3 | Pro-fibrogenic phenotype of SAMacs. a, Uniform manifold approximation and projection (UMAP) visualization of 23,075 cells from liver-resident MPs (healthy, $n = 5$; cirrhotic, $n = 5$) and blood monocytes (PBMCs, $n = 5$), annotating monocle pseudotemporal dynamics (purple to yellow). RNA velocity field (red arrows) visualized using Gaussian smoothing on regular grid. Right, annotation of MP subpopulations and injury condition. **b**, Transition probabilities per SAMac subpopulation, indicating for each cell the likelihood of transition into either SAMac(1) or SAMac(2), calculated using RNA velocity (yellow, high; purple, low; grey, below threshold of 2×10^{-4}). **c**, Heat map with spline curves fitted to genes differentially expressed across blood monocyte-to-SAMac (right arrow) and blood monocyte-to-cDC (left arrow) pseudotemporal trajectories, grouped by hierarchical clustering ($k = 3$). Gene co-expression modules (colour) and exemplar genes from each module are labelled (right). **d**, Spline curve fitted to averaged expression of all genes in module 1 along the monocyte-to-SAMac pseudotemporal trajectory (left), with selected enrichment of Gene Ontology terms (right). *P* values determined by Fisher's exact test.

correlation analysis between human and mouse MP datasets[18] demonstrated that human and mouse SAMacs clustered together (Extended Data Fig. 6h, i) and that this cluster was enriched for SAMac markers *CD9*, *TREM2* and *SPP1* (Extended Data Fig. 6j), confirming that mouse SAMacs represent a corollary population to human SAMacs.

To identify potential transcriptional regulators of human SAMacs, we defined sets of genes co-expressed with known transcription factors (regulons) along the tissue monocyte-to-macrophage pseudotemporal trajectory and in KCs (Extended Data Fig. 5g, h, Supplementary Table 12). This identified regulons and corresponding transcription factors associated with distinct macrophage phenotypes, highlighting *HES1* and *EGR2* activity in SAMacs.

To determine whether SAMacs expand in earlier-stage human liver disease, we analysed cohorts of patients with non-alcoholic fatty liver disease (NAFLD). Application of differential gene expression signatures of human SAMacs, KCs and TMs to a deconvolution algorithm[19] enabled the assessment of hepatic monocyte–macrophage composition in whole liver microarray data across the spectrum of early-stage NAFLD[20] (Extended Data Fig. 7a). This demonstrated expansion of SAMacs in patients with non-alcoholic steatohepatitis (NASH) (Extended Data Fig. 7a, b), an increased frequency of SAMacs with worsening histological NAFLD activity score (NAS) and fibrosis score (Extended Data Fig. 7c), but no association with other patient demographics (Extended Data Fig. 7d). In a separate NAFLD biopsy cohort, the expansion of SAMacs increased with NAFLD activity (Extended Data Fig. 7e) and positively correlated with the degree of fibrosis across the full severity spectrum of NAFLD-induced liver fibrosis (Extended Data Fig. 7f).

In summary, these data demonstrate that TREM2$^+$CD9$^+$ SAMacs derive from the recruitment and differentiation of circulating monocytes, are conserved across species, display a pro-fibrogenic phenotype and expand early in the course of liver disease progression.

Endothelial subpopulations inhabit the fibrotic niche

In rodent models, hepatic endothelial cells are known to regulate fibrogenesis. Clustering of human liver endothelial cells identified seven subpopulations (Fig. 4a). Classical endothelial cell markers did not discriminate between the seven clusters, although Endo(1) was distinct in lacking *CD34* expression (Extended Data Fig. 8a). To annotate endothelial subpopulations fully (Supplementary Note 3, Extended Data Fig. 8k), we identified differentially expressed markers (Fig. 4c, Supplementary Table 13), determined functional expression profiles (Extended Data Fig. 8g, Supplementary Table 14), performed analysis

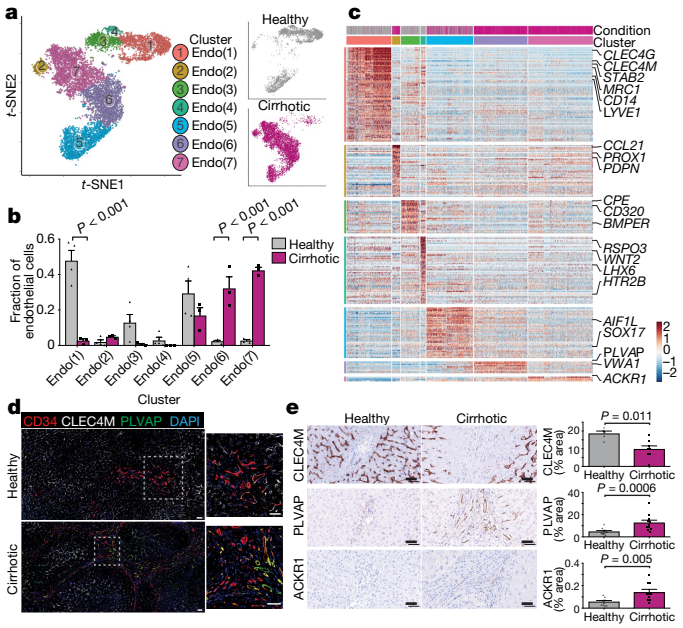


Fig. 4 | Identifying scar-associated endothelial subpopulations. a, Clustering 8,020 endothelial cells from healthy ($n = 4$) and cirrhotic ($n = 3$) human livers, annotating injury condition (right). **b**, Fractions of endothelial subpopulations in healthy ($n = 4$) and cirrhotic ($n = 3$) livers. **c**, Heat map of endothelial cluster marker genes (colour-coded by cluster and condition), with exemplar genes labelled (right). Columns denote cells; rows denote genes. **d**, Representative immunofluorescence images ($n \geq 3$) of CD34 (red), CLEC4M (white), PLVAP (green) and DAPI (blue) in healthy and cirrhotic human liver. **e**, Digital pixel quantification of CLEC4M staining in healthy ($n = 5$) and cirrhotic ($n = 8$) liver, PLVAP staining in healthy ($n = 11$) and cirrhotic ($n = 11$) liver, and ACKR1 staining in healthy ($n = 10$) and cirrhotic ($n = 10$) liver. All scale bars, 50 µm. Data are mean ± s.e.m. *P* values determined by Wald test (**b**) or two-tailed Mann–Whitney test (**e**).

of transcription factor regulons (Extended Data Fig. 8h, Supplementary Table 15) and assessed spatial distribution via tissue staining (Fig. 4d, Extended Data Fig. 8j).

Disease-specific endothelial cells Endo(6) and Endo(7) (CD34$^+$PLVAP$^+$VWA1$^+$ and CD34$^+$PLVAP$^+$ACKR1$^+$, respectively; Fig. 4a–c, Extended Data Fig. 8b) expanded in cirrhotic liver tissue (Fig. 4e) and were restricted to the fibrotic niche (Fig. 4d, e, Extended Data Fig. 8c), allowing annotation as scar-associated endothelia SAEndo(1) and SAEndo(2), respectively. By contrast, CD34$^-$CLEC4M$^+$ Endo(1) (annotated as liver sinusoidal endothelial cells), were reduced in cirrhotic livers (Fig. 4b, e). Metagene signature analysis demonstrated that Endo(6) (SAEndo(1)) cells expressed pro-fibrogenic genes including *PDGFD*, *PDGFB*, *LOX* and *LOXL2*; associated ontology terms included extracellular matrix organization (signature A; Extended Data Fig. 8g). Endo(7) (SAEndo(2)) cells displayed an immunomodulatory phenotype (signature B; Extended Data Fig. 8g). The most discriminatory marker for this cluster, *ACKR1*, has a role in regulating leucocyte recruitment[21]. We confirmed increased expression of PLVAP, CD34 and ACKR1 on endothelial cells isolated from cirrhotic livers (Extended Data Fig. 8d). Flow-based adhesion assays[22] demonstrated that cirrhotic endothelial cells display enhanced leucocyte transmigration (Extended Data Fig. 8e), which was attenuated by ACKR1 knockdown (Extended Data Fig. 8f).

PDGFRA expression defines SAMes cells

Clustering of human liver mesenchymal cells identified four populations (Fig. 5a, b, Extended Data Fig. 9a, Supplementary Table 16). Cluster Mes(1), distinguished by *MYH11* expression (Fig. 5b, Extended Data Fig. 9a), was identified as vascular smooth muscle cells (VSMCs) (Fig. 5c). Mes(4) demonstrated expression of mesothelial markers (Fig. 5b, Extended Data Fig. 9a). Cluster Mes(2) expressed high levels of *RGS5* (Fig. 5b, Extended Data Fig. 9a), and RGS5 staining identified this population as HSCs (Fig. 5c). RGS5$^+$ cells were absent from the fibrotic niche (Fig. 5c). Cluster Mes(3) (distinguished by *PDGFRA* expression) expressed high levels of fibrillar collagens and pro-fibrogenic genes (Fig. 5b, d, Extended Data Fig. 9a). PDGFRα$^+$ cells expanded in cirrhotic livers (Fig. 5a, e, f) and were mapped to the fibrotic niche (Fig. 5f), enabling annotation as scar-associated mesenchymal (SAMes) cells.

To study SAMes cell heterogeneity, further clustering (Extended Data Fig. 9b) identified two populations of SAMes cells (Extended Data Fig. 9c,d, Supplementary Table 17). *OSR1* expression distinguished cluster SAMesB (Extended Data Fig. 9c), and labelled a subpopulation of periportal cells in healthy liver and scar-associated cells in the fibrotic niche (Extended Data Fig. 9e, f). Cluster SAMesB also expressed other known portal fibroblast markers[23] (Extended Data Fig. 9g).

In rodent liver fibrosis models, HSCs differentiate into scar-producing myofibroblasts[24–26]. Pseudotemporal ordering and RNA velocity analyses demonstrated a trajectory from human HSCs to SAMes cells (Extended Data Fig. 9h). Assessment of gene co-expression modules along the HSC-to-SAMes differentiation continuum indicated upregulation of fibrogenic genes including *COL1A1*, *COL1A2*,*COL3A1* and *TIMP1* and downregulation of genes including *RGS5*, *IGFBP5*, *ADAMTS1* and *GEM*, which are known to be downregulated in mouse HSC in response to liver injury[25] (Extended Data Fig. 9i).

The multi-lineage interactome in the fibrotic niche

Having defined the populations of scar-associated macrophages, endothelial and mesenchymal cells, we confirmed the close topographical association of these cells within the fibrotic niche (Extended Data Fig. 10a, b), and used CellPhoneDB[27] to perform an unbiased ligand–receptor interaction analysis between these populations.

Numerous statistically significant paracrine and autocrine interactions were detected between ligands and cognate receptors expressed

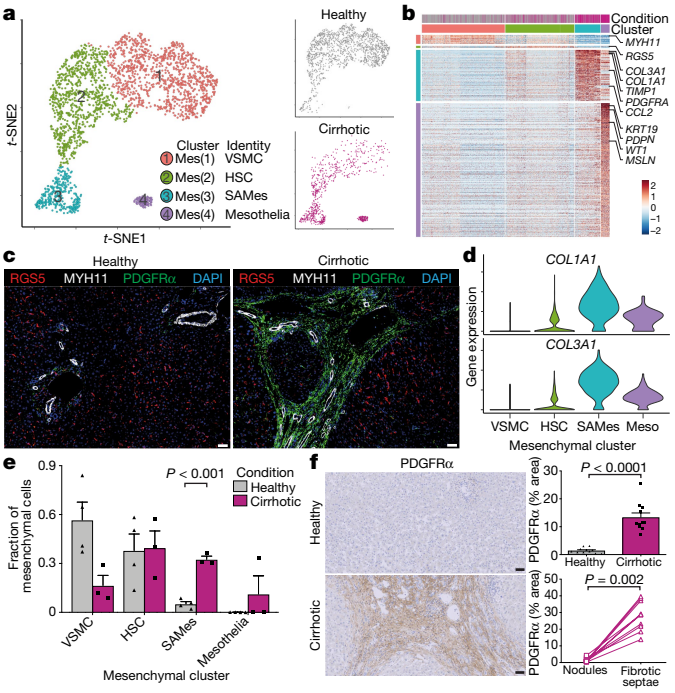


Fig. 5 | Identifying a SAMes subpopulation. a, Clustering of 2,318 mesenchymal cells from healthy ($n = 4$) and cirrhotic ($n = 3$) human livers, annotating injury condition (right). **b**, Heat map of mesenchymal cluster marker genes (top, colour-coded by cluster and condition), with exemplar genes labelled (right). Columns denote cells; rows denote genes. **c**, Representative immunofluorescence images ($n \geq 3$) of RGS5 (red), MYH11 (white), PDGFRα (green) and DAPI (blue) in healthy and cirrhotic human liver. **d**, Scaled gene expression of fibrillar collagens across mesenchymal cells from healthy ($n = 4$) and cirrhotic ($n = 3$) livers. **e**, Fraction of mesenchymal subpopulations in healthy ($n = 4$) and cirrhotic ($n = 3$) livers. **f**, PDGFRα immunohistochemistry (left) and digital pixel quantification (right) in healthy ($n = 11$) and cirrhotic ($n = 11$) livers (top) and in fibrotic septae and parenchymal nodules in cirrhotic livers ($n = 11$; bottom). All scale bars, 50 µm. Data are mean ± s.e.m. *P* values determined by Wald test (**e**), two-tailed Mann–Whitney test (**f**, top), or two-tailed Wilcoxon test (**f**, bottom).

by SAMac, SAEndo and SAMes cells (Supplementary Table 18, Extended Data Fig. 10f–m). To interrogate how scar-associated NPCs regulate fibrosis and to identify tractable therapeutic targets, we focused functional analyses on interactions with SAMes (Fig. 6a, e, Extended Data Fig. 10d). In keeping with our data demonstrating that SAMacs promote fibrillar collagen expression in HSCs (Fig. 2g), SAMacs expressed epidermal growth factor receptor (EGFR) ligands that are known to regulate mesenchymal cell activation[28] (Fig. 6a). In addition, SAMacs expressed the mesenchymal cell mitogens *TNFSF12* and *PDGFB*, signalling to cognate receptors *TNFRSF12A* and *PDGFRA* on SAMes (Fig. 6a). We confirmed localization of these ligand–receptor pairs within the fibrotic niche (Fig. 6b). Both TNFSF12 and PDGF-BB induced proliferation of primary human HSCs, which was inhibited by blockade of TNFRSF12A and PDGFRA, respectively (Fig. 6c, d). Conditioned medium from primary human SAMacs promoted primary human HSC proliferation ex vivo (Extended Data Fig. 10c), demonstrating a functional role for SAMacs in regulating SAMes cell expansion.

SAEndo cells expressed high levels of Notch ligands *JAG1*, *JAG2* and *DLL4* interacting with Notch receptor *NOTCH3* on SAMes cells (Fig. 6e). NOTCH3 was identified on PDGFRα$^+$ SAMes cells within the fibrotic niche (Fig. 6f), and primary endothelial cells from cirrhotic human liver demonstrated increased expression of JAG1 (Fig. 6g). Co-culture of primary human HSCs and endothelial cells from cirrhotic livers promoted fibrillar collagen production by HSCs, which was inhibited by addition

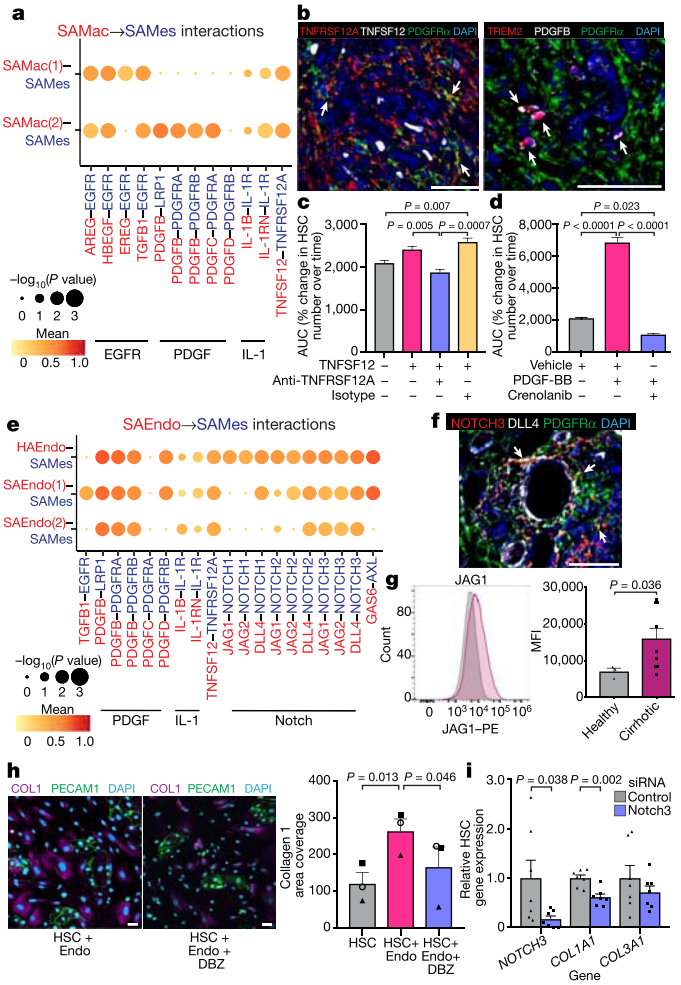


Fig. 6 | Multi-lineage interactions in the fibrotic niche. a, Dot plot of ligand–receptor interactions between SAMac ($n = 10$ human livers) and SAMes ($n = 7$ human livers) subpopulations. Ligand (red) and cognate receptor (blue) are shown on the *x* axis; cell populations that express ligand (red) and receptor (blue) are shown on the *y* axis. Circle size denotes *P* value (permutation test); colour (red, high; yellow, low) denotes average ligand and receptor expression levels in interacting subpopulations. **b**, Left, representative immunofluorescence image ($n \geq 3$) of TNFRSF12A (red), TNFSF12 (white), PDGFRα (green) and DAPI (blue); arrows denote TNFRSF12A$^+$PDGFRα$^+$ cells. Right, representative immunofluorescence image ($n \geq 3$) of TREM2 (red), PDGFB (white), PDGFRα (green) and DAPI (blue); arrows denote TREM2$^+$PDGFB$^+$ cells. Scale bars, 50 µm. **c**, **d**, HSC proliferation assay. The area under the curve (AUC) of the percentage change in HSC number over time (hours) is shown on the *y* axis. $n = 3$ (all conditions in **c** and **d**). **e**, Dot plot of ligand–receptor interactions between SAEndo ($n = 7$ human livers) and SAMes ($n = 7$ human livers) subpopulations as in **a**. **f**, Representative immunofluorescence image ($n \geq 3$) of NOTCH3 (red), DLL4 (white), PDGFRα (green) and DAPI (blue) in the fibrotic niche; arrows denote NOTCH3$^+$PDGFRα$^+$ cells. Scale bar, 50 µm. **g**, Flow cytometry analysis of JAG1 in endothelial cells from healthy ($n = 3$) or cirrhotic ($n = 9$) liver. Left, representative histogram; right, mean fluorescence intensity (MFI). **h**, Co-culture of primary human HSCs and endothelial cells from cirrhotic livers, with or without the Notch inhibitor dibenzazepine (DBZ). Left, representative immunofluorescence images ($n = 3$) of collagen 1 (COL1; magenta), PECAM1 (green) and DAPI (blue). Scale bars, 50 µm. Right, digital pixel analysis of the collagen 1 area ($n = 3$). **i**, Gene knockdown in HSCs using control ($n = 7$) or *NOTCH3* ($n = 7$) siRNA. Indicated genes were analysed by qPCR, with expression relative to mean expression of control siRNA-treated HSCs. Data are mean ± s.e.m. *P* values determined by one-way ANOVA and Tukey test (**c**, **d**), two-tailed Mann–Whitney test (**g**, **i**) or repeated-measures one-way ANOVA and Tukey test (**h**).

of the Notch-signalling inhibitor dibenzazepine (Fig. 6h). Furthermore, knockdown of *NOTCH3* expression in primary human HSCs resulted in reduced fibrillar collagen expression (Fig. 6i).

In summary, our unbiased dissection of the key ligand–receptor interactions between scar-associated NPCs highlights TNFRSF12A, PDGFRA and Notch signalling as important regulators of mesenchymal cell function within the human liver fibrotic niche.

Discussion

Here, using scRNA-seq and spatial mapping, we resolve the fibrotic niche of human liver cirrhosis, identifying pathogenic subpopulations of TREM2$^+$CD9$^+$ macrophages, ACKR1$^+$ and PLVAP$^+$ endothelial cells and PDGFRα$^+$ collagen-producing myofibroblasts. We dissect a complex, pro-fibrotic interactome between multiple scar-associated cell lineages and identify highly relevant intra-scar pathways that are potentially druggable. In this era of precision medicine, this unbiased multi-lineage approach should inform the design of highly targeted combination therapies that will very likely be necessary to achieve effective antifibrotic potency[3,4].

Application of our novel scar-associated cell markers could also potentially inform molecular pathology-based patient stratification, which is fundamental to the prosecution of successful antifibrotic clinical trials. Our work illustrates the power of single-cell transcriptomics to decode the cellular and molecular basis of human organ fibrosis, providing a conceptual framework for the discovery of relevant therapeutic targets to treat patients with a broad range of fibrotic diseases.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at [https://doi.org/10.1038/s41586-019-1631-3].

1. Marcellin, P. & Kutala, B. K. Liver diseases: a major, neglected global public health problem requiring urgent actions and large-scale screening. *Liver Int.* **38** (Suppl. 1), 2–6 (2018).
2. Angulo, P. et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* **149**, 389–97.e10 (2015).
3. Ramachandran, P. & Henderson, N. C. Antifibrotics in chronic liver disease: tractable targets and translational challenges. *Lancet Gastroenterol. Hepatol.* **1**, 328–340 (2016).
4. Friedman, S. L., Neuschwander-Tetri, B. A., Rinella, M. & Sanyal, A. J. Mechanisms of NAFLD development and therapeutic strategies. *Nat. Med.* **24**, 908–922 (2018).
5. Stubbington, M. J. T., Rozenblatt-Rosen, O., Regev, A. & Teichmann, S. A. Single-cell transcriptomics to explore the immune system in health and disease. *Science* **358**, 58–63 (2017).
6. Duffield, J. S. et al. Selective depletion of macrophages reveals distinct, opposing roles during liver injury and repair. *J. Clin. Invest.* **115**, 56–65 (2005).
7. Ramachandran, P. et al. Differential Ly-6C expression identifies the recruited macrophage phenotype, which orchestrates the regression of murine liver fibrosis. *Proc. Natl Acad. Sci. USA* **109**, E3186–E3195 (2012).
8. Karlmark, K. R. et al. Hepatic recruitment of the inflammatory Gr1$^+$ monocyte subset upon liver injury promotes hepatic fibrosis. *Hepatology* **50**, 261–274 (2009).
9. Minutti, C. M. et al. Local amplifiers of IL-4Rα-mediated macrophage activation promote repair in lung and liver. *Science* **356**, 1076–1080 (2017).
10. Pradere, J.-P. et al. Hepatic macrophages but not dendritic cells contribute to liver fibrosis by promoting the survival of activated hepatic stellate cells in mice. *Hepatology* **58**, 1461–1473 (2013).
11. Henderson, N. C. et al. Galectin-3 regulates myofibroblast activation and hepatic fibrosis. *Proc. Natl Acad. Sci. USA* **103**, 5060–5065 (2006).
12. Seki, E. et al. CCR2 promotes hepatic fibrosis in mice. *Hepatology* **50**, 185–197 (2009).
13. Syn, W. K. et al. Osteopontin is induced by hedgehog pathway activation and promotes fibrosis progression in nonalcoholic steatohepatitis. *Hepatology* **53**, 106–115 (2011).
14. Scott, C. L. et al. Bone marrow-derived monocytes give rise to self-renewing and fully differentiated Kupffer cells. *Nat. Commun.* **7**, 10321 (2016).
15. Gomez Perdiguero, E. et al. Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* **518**, 547–551 (2015).
16. Mass, E. et al. Specification of tissue-resident macrophages during organogenesis. *Science* **353**, aaf4238 (2016).
17. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

18. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
19. Schelker, M. et al. Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat. Commun.* **8**, 2032 (2017).
20. Ahrens, M. et al. DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab.* **18**, 296–302 (2013).
21. Pruenster, M. et al. The Duffy antigen receptor for chemokines transports chemokines and supports their promigratory activity. *Nat. Immunol.* **10**, 101–108 (2009).
22. Shetty, S., Weston, C. J., Adams, D. H. & Lalor, P. F. A flow adhesion assay to study leucocyte recruitment to human hepatic sinusoidal endothelium under conditions of shear stress. *J. Vis. Exp.* **85**, 51330 (2014).
23. Iwaisako, K. et al. Origin of myofibroblasts in the fibrotic liver in mice. *Proc. Natl Acad. Sci. USA* **111**, E3297–E3305 (2014).
24. Henderson, N. C. et al. Targeting of α_v integrin identifies a core molecular pathway that regulates fibrosis in several organs. *Nat. Med.* **19**, 1617–1624 (2013).
25. De Minicis, S. et al. Gene expression profiles during hepatic stellate cell activation in culture and *in vivo*. *Gastroenterology* **132**, 1937–1946 (2007).
26. Mederacke, I., Dapito, D. H., Affò, S., Uchinami, H. & Schwabe, R. F. High-yield and high-purity isolation of hepatic stellate cells from normal and fibrotic mouse livers. *Nat. Protocols* **10**, 305–315 (2015).
27. Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).
28. Minutti, C. M. et al. A macrophage-pericyte axis directs tissue restoration via amphiregulin-induced transforming growth factor beta activation. *Immunity* **50**, 645–654.e6 (2019).

METHODS

Study subjects

Local approval for procuring human liver tissue and blood samples for scRNA-seq, flow cytometry and histological analysis was obtained from the NRS BioResource and Tissue Governance Unit (study number SR574), following review at the East of Scotland Research Ethics Service (reference 15/ES/0094). All subjects provided written informed consent. Healthy background non-lesional liver tissue was obtained intraoperatively from patients undergoing surgical liver resection for solitary colorectal metastasis at the Hepatobiliary and Pancreatic Unit, Department of Clinical Surgery, Royal Infirmary of Edinburgh. Patients with a known history of chronic liver disease, abnormal liver function tests or those who had received systemic chemotherapy within the last four months were excluded from this cohort. Cirrhotic liver tissue was obtained intraoperatively from patients undergoing orthotopic liver transplantation at the Scottish Liver Transplant Unit, Royal Infirmary of Edinburgh. Blood from patients with a confirmed diagnosis of liver cirrhosis were obtained from patients attending the Scottish Liver Transplant Unit, Royal Infirmary of Edinburgh. Patients with liver cirrhosis due to viral hepatitis were excluded from the study. Patient demographics are summarized in Extended Data Fig. 1a. Isolation of primary hepatic macrophage subpopulations and endothelial cells from healthy and cirrhotic livers for cell culture and analysis of secreted mediators was performed at the University of Birmingham, UK. Local ethical approval was obtained (reference 06/Q2708/11, 06/Q2702/61) and all patients provided written, informed consent. Liver tissue was acquired from explanted diseased livers from patients undergoing orthotopic liver transplantation, resected liver specimens or donor livers rejected for transplant at the Queen Elizabeth Hospital, Birmingham. For histological assessment of NAFLD biopsies, anonymized unstained formalin-fixed paraffin-embedded liver biopsy sections encompassing the complete NAFLD spectrum were provided by the Lothian NRS Human Annotated Bioresource under authority from the East of Scotland Research Ethics Service REC 1, reference 15/ES/0094.

Human tissue processing

Importantly, to minimize artefacts[29], we developed a rapid tissue processing pipeline, obtaining fresh non-ischaemic liver tissue taken by wedge biopsy before the interruption of the hepatic vascular inflow during liver surgery or transplantation, and immediately processing this for FACS. This enabled a workflow time of under three hours from patient to single-cell droplet encapsulation.

For human liver scRNA-seq and flow cytometry analyses, a wedge biopsy of non-ischaemic fresh liver tissue (2–3 g) was obtained by the operating surgeon. This was immediately placed in HBSS (Gibco) on ice. The tissue was then transported directly to the laboratory and dissociation routinely commenced within 20 min of the liver biopsy. To enable paired histological assessment, a segment of each liver specimen was also fixed in 4% neutral-buffered formalin for 24 h followed by paraffin-embedding. Additional liver samples, obtained via the same method, were fixed in an identical manner and used for further histological analysis. For human macrophage cell sorting and endothelial cell isolation, liver tissue (40 g) was used from cirrhotic patients undergoing orthotopic liver transplantation or control samples from donor liver or liver resection specimens.

Mice

Adult male C57BL/6JCrl mice aged 8–10 weeks were purchased from Charles River. Mice were housed under specific pathogen-free conditions at the University of Edinburgh. All experimental protocols were approved by the University of Edinburgh Animal Welfare and Ethics Board in accordance with UK Home Office regulations. Liver fibrosis was induced with 4 weeks (nine injections) of twice-weekly intraperitoneal CCl_4 at a dose of 0.4 µl g^{-1} body weight, diluted 1:3 in olive oil as previously described[7]. Mice were randomly assigned to receive CCl_4 or to serve as healthy controls. No sample size calculation or blinding was performed. Liver tissue was obtained 24 h after the final CCl_4 injection, a time of active fibrogenesis[7]. Comparison was made to age-matched uninjured mice.

Preparation of single-cell suspensions

For human liver scRNA-seq, liver tissue was minced with scissors and digested in 5 mg ml^{-1} pronase (Sigma-Aldrich, P5147-5G), 2.93 mg ml^{-1} collagenase B (Roche, 11088815001) and 0.019 mg ml^{-1} DNase (Roche, 10104159001) at 37 °C for 30 min with agitation (200–250 r.p.m.), then strained through a 120-µm nybolt mesh along with PEB buffer (PBS, 0.1% BSA and 2mM EDTA) including DNase (0.019 mg ml^{-1}). Thereafter, all processing was done at 4 °C. The cell suspension was centrifuged at 400*g* for 7 min, supernatant removed, cell pellet resuspended in PEB buffer and DNase added (0.019 mg ml^{-1}), followed by additional centrifugation (400*g*, 7 min). Red blood cell lysis was performed (BioLegend, 420301), followed by centrifugation (400*g*, 7 min), resuspension in PEB buffer and straining through a 35-µm filter. Following another centrifugation at 400*g* for 7 min, cells were blocked in 10% human serum (Sigma-Aldrich, H4522) for 10 min at 4 °C before antibody staining.

For human liver macrophage flow cytometry analysis and cell sorting, and for mouse liver macrophage flow cytometry, cell sorting and scRNA-seq, single-cell suspensions were prepared as previously described, with minor modifications[30]. In brief, liver tissue was minced and digested in an enzyme cocktail 0.625 mg ml^{-1} collagenase D (Roche, 11088882001), 0.85 mg ml^{-1} collagenase V (Sigma-Aldrich, C9263-1G), 1 mg ml^{-1} dispase (Gibco, Invitrogen, 17105-041) and 30 U ml^{-1} DNase (Roche, 10104159001) in RPMI-1640 at 37 °C for 20 min (mouse) or 45 min (human) with agitation (200–250 r.p.m.), before being passed through a 100-µm filter. After lysis of red blood cells (BioLegend, 420301), cells were washed in PEB buffer and passed through a 35-µm filter. Before the addition of antibodies, cells from human samples were blocked in 10% human serum (Sigma-Aldrich, H4522) and mouse samples were blocked in anti-mouse CD16/32 antibody (1:100; BioLegend, 101302) and 10% normal mouse serum (Sigma, M5905) for 10 min at 4 °C.

For human PBMC scRNA-seq, 4.9-ml peripheral venous blood samples were collected in EDTA-coated tubes (Sarstedt, S-Monovette 4.9ml K3E) and placed on ice. Blood samples were transferred into a 50-ml Falcon tube. After lysis of red blood cells (BioLegend, 420301), blood samples were then centrifuged at 500*g* for 5 min and supernatant was removed. Pelleted samples were then resuspended in staining buffer (PBS plus 2% BSA; Sigma-Aldrich) and centrifugation was repeated. Samples were then blocked in 10% human serum (Sigma-Aldrich, H4522) in staining buffer on ice for 30 min. Cells were then resuspended in staining buffer and passed through a 35-µm filter before antibody staining.

Flow cytometry and cell sorting

Incubation with primary antibodies was performed for 20 min at 4 °C. All antibodies, conjugates, lot numbers and dilutions used in this study are presented in Supplementary Table 19. After antibody staining, cells were washed with PEB buffer. For human macrophage flow cytometry analysis and cell sorting, cells were then incubated with streptavidin-BV711 for 20 min at 4 °C (BioLegend 405241; 1:200). For human and mouse cell sorting (FACS) and mouse flow cytometry analysis, cell viability staining (DAPI; 1:1,000) was then performed, immediately before acquiring the samples.

Human cell sorting for scRNA-seq was performed on a BD Influx (Becton Dickinson). Viable single $CD45^+$ (leucocytes) or $CD45^-$ (other non-parenchymal cells) cells were sorted from human liver tissue (Extended Data Fig. 1b) and viable $CD45^+$ $CD66b^-$ (PBMC) cells were sorted from peripheral blood (Extended Data Fig. 1c) and processed for droplet-based scRNA-seq.

To generate conditioned medium from cirrhotic liver macrophage subpopulations, cells were sorted on a BD FACSAria Fusion (Becton Dickinson). Sorted SAMacs (viable $CD45^+Lin^-HLA\text{-}DR^+CD14^+CD16^+$ $CD163^-TREM2^+CD9^+$), TMs (viable $CD45^+Lin^-$ $HLA\text{-}DR^+CD14^+CD16^+$

$CD163^-TREM2^-CD9^-$) and KCs (viable $CD45^+Lin^-HLA\text{-}DR^+CD14^+CD16^+CD163^+CD9^-$) were plated in 12-well plates (Corning, 3513) in DMEM (Gibco, 41965039) containing 2% fetal bovine serum (FBS; Gibco, 10500056) at 1×10^6 cells per ml for 24 h at 37°C 5% CO_2. Control wells contained medium alone. Conditioned medium was collected, centrifuged at 400*g* for 10 min, and supernatant was stored at −80 °C.

For human macrophage flow cytometry analysis, after surface antibody staining, cells were stained with Zombie NIR fixable viability dye (BioLegend, 423105) according to the manufacturer's instructions. Cells were washed in PEB then fixed in Intracellular (IC) Fixation Buffer (Thermo Fisher, 00-8222-49) for 20 min at 4 °C. Fixed samples were stored in PEB at 4 °C until acquisition. Flow cytometry acquisition was performed on a six-laser Fortessa flow cytometer (Becton Dickinson). The gating strategy is shown in Extended Data Fig. 4f and Fig. 2f.

Mouse macrophage cell sorting for scRNA-seq and co-culture experiments was performed on a BD FACSAriaII (Becton Dickinson). For scRNA-seq, viable $CD45^+Lin(CD3, NK1.1, Ly6G, CD19)^-$ cells were sorted from healthy ($n = 3$) and CCl_4-treated ($n = 3$) mice and processed for droplet-based scRNA-seq. For transwell co-culture, viable $CD45^+Lin^-CD11b^+F4/80^+TIMD4^-CD9^+$ (SAMacs) or $CD9^-$ (TMs) cells were sorted from CCl_4-treated mice (Extended Data Fig. 6e). Flow cytometry analysis on macrophages from healthy and CCl_4-treated mice was also performed on a BD FACSAriaII (Becton Dickinson), using the same gating strategy (Extended Data Fig. 6e). All flow cytometry data were analysed using FlowJo software (Treestar).

Luminex assay

Detection of CCL2, galectin-3, IL-1β, CXCL8 and osteopontin (SPP1) and CD163 proteins in conditioned medium from human liver macrophage subpopulations was performed using a custom human luminex assay (R&D Systems), according to the manufacturer's protocol. Data were acquired using a Bio-Plex 200 (Bio-Rad) and are presented as MFI for each analyte.

Cell culture

Primary human HSCs (ScienCell, 5300) were cultured in stellate cell medium (SteCM, ScienCell, 5301) on poly-L-lysine (Sigma, P4832)-coated T75 tissue culture flasks, according to the supplier's protocol. All experiments were performed using cells between passages 3 and 5. For assessment of fibrillar collagen gene expression, HSCs were plated at 75,000 cells per well in 24-well plates (Costar, 3524) in HSC medium consisting of DMEM (Gibco, 21969-035) with 20 μM HEPES (Sigma, H3375), 2 mM L-glutamine (Gibco, 25030-024), 1% penicillin streptomycin (Gibco, 15140-122) and 2% FBS (Gibco, 10270). HSCs were serum-starved overnight (in HSC medium without FBS), washed with PBS, then 250 μl of conditioned medium from primary human macrophage subpopulations was added for 24 h. HSCs were obtained for RNA.

Human HSC proliferation

For proliferation assays, after serum starvation, HSCs were obtained using TrypLE Express (Gibco, 12604013), re-suspended in HSC medium at 2.5×10^4 per ml with Incucyte NucLight Rapid Red (Essen Biosciences, 4717) at a dilution of 1:500 and seeded into 384-well plates (Greiner Bio-One, 781090) at 25 μl per well. HSCs were then treated with (1) control medium; (2) PDGF-BB (10 ng ml^{-1}; Peprotech, 100-14B) or TNFSF12 (100 ng ml^{-1}; Peprotech, 310-06-5) with or without the PDGFRα inhibitor crenolanib[31] (1 μM; Cayman Chemicals, CAY1873), anti-TNFRSF12A (2 μg ml^{-1}; Life Technologies, 16-9018-82, clone ITEM-4), mouse IgG2b kappa isotype control antibody (2 μg ml^{-1}; Life Technologies, 16-4732-82, clone eBMG2b) or vehicle control as indicated; (3) conditioned medium from human hepatic macrophage subpopulations as indicated. The final volume was 50 μl for all conditions. Cells were then incubated in an Incucyte ZOOM live cell analysis system (Essen Biosciences) humidified at 37 °C with 5% CO_2 with imaging every 3 h using the 10× optic for either 87 h (recombinant cytokines/inhibitors) or 39 h (macrophage-conditioned medium). Analysis was performed with the Incucyte proprietary analysis software (v.2018A) by using machine learning to distinguish the individual nuclei (stained red by the NucLight Rapid Red dye) and perform nuclear counts of the images at each 3 h time point over the period of culture. Data are expressed as the AUC for percentage change in nuclear number from baseline versus time (hours), calculated in GraphPad Prism.

Gene knockdown in human HSCs

Knockdown of NOTCH3 in human HSCs was performed using siRNA. HSCs were plated at 75,000 cells per well in a 12-well plate (Costar, 3513) followed by serum starvation overnight (in HSC medium without FBS). siRNA duplexes with Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher, 13778075) were prepared in OptiMEM (Thermo Fisher, 31985070) according to the manufacturer's recommendations, and used at a concentration of 50 nM. Cells were exposed to the duplex for 48 h, in HSC medium containing 2% FBS. Cells were collected for RNA and quantitative PCR with reverse transcription (RT–qPCR). Knockdown efficiency was assessed by *NOTCH3* RT–qPCR. The best siRNA for knockdown was determined empirically using the FlexiTube GeneSolution kit (Qiagen, GS4854). HSCs treated with control siRNA (Qiagen, 1027280) and siRNA for *NOTCH3* (Qiagen, Hs_NOTCH3_3, SI00009513; knockdown 83%) were then assessed for fibrillar collagen gene expression.

Mouse HSC activation

Primary mouse HSCs were isolated from healthy mice as previously described[26]. In brief, after cannulation of the inferior vena cava, the portal vein was cut to allow retrograde step-wise perfusion with pronase (Sigma, P5147) and collagenase D (Roche, 11088882001) -containing solutions, before ex vivo digestion in a solution containing pronase, collagenase D and DNase (Roche, 10104159001). HSCs were isolated from the digest solution by Histodenz (Sigma, D2158-100G) gradient centrifugation. HSCs were plated at a density of 400,000 cells per well in a 24-well plate (Costar, 3524) in HSC medium containing 10% FBS. After overnight culture, cells were washed with PBS and cultured in HSC medium containing 2% FBS. For macrophage co-culture, transwell inserts (0.4-μm polyester membrane; Costar, 3470) were then placed above adherent HSCs. FACS-sorted $CD9^+$ mouse SAMacs or $CD9^-$ mouse TMs from CCl_4-treated mice were resuspended in HSC medium containing 2% FBS at 400,000 cells per ml and 200,000 cells were added to the top of the transwell insert. Co-culture proceeded for 48 h and HSCs were collected for RNA. Quiescent HSCs (collected at the start of co-culture) were used as a control population.

Isolation of human liver endothelial cells

Human liver endothelial cells were isolated from cirrhotic explant livers and non-fibrotic control donor liver as previously described[32]. Endothelial cells were cultured on plasticware coated with rat-tail collagen (Sigma, C3867) in complete endothelial medium consisting of endothelial basal media (Thermo Fisher, 11111044) containing 10% heat-inactivated human serum (tcsBiosciences, CS100-500), 100 U penicillin, 100 μg ml^{-1} streptomycin, 2 mM glutamine (Sigma, G6784), VEGF (10 ng ml^{-1}; Peprotech, 100-20) and 10 ng ml^{-1} HGF (10 ng ml^{-1}; Peprotech, 100-39). Expression of PLVAP, CD34, ACKR1 and JAG1 was assessed using flow cytometry.

Flow-based adhesion assays

Flow-based adhesion assays were performed as previously described[22,32]. In brief, endothelial cells from healthy and cirrhotic liver were seeded onto a rat-tail collagen-coated Ibidi slide $VI^{0.4}$ (Ibidi, 80606) at a density to give a monolayer and incubated overnight. Peripheral blood was collected from healthy donors in EDTA-coated tubes. PBMCs were isolated using a lympholyte density gradient (Cedarlane Laboratories) then washed in PBS containing 1 mM Ca^{2+}, 0.5 mM Mg^{2+} and 0.15% bovine serum albumin (BSA). Monocytes were enriched from PBMCs using a pan-monocyte isolation kit (Miltenyi Biotech, 130-096-537) according

to the manufacturer's protocol. For flow-based adhesion assays, cells were resuspended at 10^6 cells per millilitre in endothelial basal media (Thermo Fisher, 11111044) containing 0.15% BSA, then perfused over the endothelial cell monolayer for 5 min at 0.28 ml min^{-1}. Non-adherent cells were washed off during 5 min perfusion of 0.15% BSA human basal endothelial medium and 10 random non-overlapping images were randomly recorded from each channel. Total adherent (bright-phase; expressed as cell number per mm^2 per million cells perfused) and transmigrating cells (dark-phase; expressed as percentage total adherent cells) on an endothelial cell monolayer from each patient were counted and quantified as previously described[22].

Gene knockdown in endothelial cells

Knockdown of *ACKR1* and *PLVAP* gene expression in human cirrhotic endothelial cells was performed using siRNA as previously described[32]. In brief, siRNA duplexes for *PLVAP*, *ACKR1* or negative control (Qiagen, 1027280) with Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher, 13778075) were prepared in OptiMEM (Thermo Fisher, 31985070) according to the manufacturer's recommendations, and used at a concentration of 25 nM. Cells were exposed to the duplex for 4 h at 37 °C, after which time the medium was replaced with endothelial basal medium containing 10% heat-inactivated human serum for 24 h. The medium was then replaced with complete endothelial medium and incubated at 37 °C with 5% CO_2 for a further 24 h. Knockdown efficacy was assessed by flow cytometry and the MFI (Extended Data Fig. 8f). The best siRNA for knockdown was determined empirically using the FlexiTube GeneSolution kit (Qiagen, GS83483 (*PLVAP*) and GS2532 (*ACKR1*)). For flow-based adhesion assays, siRNAs against *PLVAP* (Qiagen, Hs_PLVAP_1, SI00687547; knockdown 50.6%), *ACKR1* (Qiagen, Hs_Fy_5, SI02627667; knockdown 37.7%) or control siRNA were selected. Then, 90,000 endothelial cells from cirrhotic patients ($n = 6$) were seeded into channels of a rat-tail collagen-coated Ibidi slide VI$^{0.4}$ and gene knockdown was performed, followed by flow-based adhesion assay as described above.

Co-culture of endothelial cells and HSCs

HSCs (15,000 cells) were seeded onto an Ibidi slide VI$^{0.4}$ with and without primary human endothelial cells (15,000 cells) from individual patients with cirrhosis ($n = 3$) in complete endothelial medium. After 2 h, all growth factor supplements were removed and cells were cultured for a further 72 h in endothelial basal medium containing 10% heat-inactivated human serum with or without the Notch signalling inhibitor dibenzazepine (10 μM; Bio-Techne, 4489/10) or vehicle (DMSO) control. Cells were fixed in 4% paraformaldehyde (PFA) for 30 min, permeabilized with 0.3% Triton X-100 in PBS for 5 min and blocked with 10% goat serum in PBS for 30 min followed by primary antibody incubation (mouse anti-PECAM1 and rabbit anti-collagen 1; see Supplementary Table 19) for 1 h. Cells were washed in 0.1% Triton X-100 in PBS followed by addition of fluorescently conjugated secondary antibodies (1:500 dilution) for 1 h. Cells were mounted with Pro-long Gold anti-fade DAPI, images were taken on the Confocal Microscope Zeiss LSM780, and the collagen 1 staining area was quantified using IMARIS.

RNA extraction and RT–qPCR

RNA was isolated from HSCs using the RNeasy Plus Micro Kit (Qiagen, 74034) and cDNA synthesis performed using the QuantiTect Reverse Transcription Kit (Qiagen, 205313) according to the manufacturer's protocol. Reactions were performed in triplicate in 384-well plate format and were assembled using the QIAgility automated pipetting system (Qiagen). RT–qPCR for human HSCs was performed using PowerUp SYBR Green Master Mix (Thermo Fisher, A25777) with the following primers (all Qiagen): *GAPDH* (QT00079247), *COL1A1* (QT00037793), *COL3A1* (QT00058233) and *NOTCH3* (QT00003374). RT–qPCR for mouse HSCs was performed using TaqMan Fast Advanced Master Mix (Thermo Fisher, 4444557) with the following primers: *Gapdh* (Thermo Fisher, Mm99999915_g1) and *Col3a1* (Thermo Fisher, Mm00802300_m1). Samples were amplified on an ABI 7900HT FAST PCR system (Applied Biosystems, Thermo Fisher Scientific). Data were analysed using Thermo Fisher Connect cloud qPCR analysis software (Thermo Fisher Scientific). The $2^{-\Delta\Delta Ct}$ quantification method, using *GAPDH* for normalization, was used to estimate the amount of target mRNA in samples, and expression calculated relative to average mRNA expression levels from control samples.

Immunohistochemistry, immunofluorescence and single-molecule FISH

Formalin-fixed paraffin-embedded human liver tissue was cut into 4-μm sections, dewaxed, rehydrated, then incubated in 4% neutral-buffered formalin for 20 min. After heat-mediated antigen retrieval in pH 6 sodium citrate (microwave; 15 min), slides were washed in PBS and incubated in 4% hydrogen peroxide for 10 min. Slides were then washed in PBS, blocked using protein block (GeneTex, GTX30963) for 1 h at room temperature before incubation with primary antibodies for 1 h at room temperature. A full list of primary antibodies and conditions is shown in Supplementary Table 19. Slides were washed in PBS plus 0.1% Tween 20 (PBST; Sigma-Aldrich, P1379) then incubated with ImmPress HRP Polymer Detection Reagents (depending on species of primary; rabbit, MP-7401; mouse, MP-6402-15; goat, MP-7405; all Vector Laboratories) for 30 min at room temperature. Slides were washed in PBS followed by detection. For DAB staining, sections were incubated with DAB (DAKO, K3468) for 5 min and washed in PBS before a haematoxylin (Vector Laboratories, H3404) counterstain. For multiplex immunofluorescence staining, following the incubation with ImmPress and PBS wash, initial staining was detected using Cy3, Cy5, or fluorescein tyramide (Perkin-Elmer, NEL741B001KT) at a 1:1,000 dilution. Slides were then washed in PBST followed by further heat treatment with pH 6 sodium citrate (15 min), washes in PBS, protein block, incubation with the second primary antibody (incubated overnight at 4 °C), ImmPress Polymer and tyramide as before. This sequence was repeated for the third primary antibody (incubated at room temperature for 1 h) and a DAPI-containing mountant was then applied (Thermo Fisher Scientific, P36931).

For AMEC staining (only CLEC4M immunohistochemistry), all washes were carried out with TBST (dH_2O, 200 mM Tris, 1.5 M NaCl, 1% Tween-20 (all Sigma-Aldrich) pH 7.5) and peroxidase blocking was carried out for 30 min in 0.6% hydrogen peroxide in methanol. Sections were incubated with AMEC (Vector Laboratories, SK-4285) for 20 min and washed in TBST before a haematoxylin (Vector Laboratories, SK-4285) counterstain.

For combined single-molecule fluorescent in situ hybridization (smFISH) and immunofluorescence, detection of TREM2 was performed using the RNAscope 2.5 LS Reagent Kit BrownAssay (Advanced Cell Diagnostics) in accordance with the manufacturer's instructions. In brief, 5-μm tissue sections were dewaxed, incubated with endogenous enzyme block, boiled in pre-treatment buffer and treated with protease, followed by target probe hybridization using the RNAscope LS 2.5 Hs-TREM2 (420498, Advanced Cell Diagnostics) probe. Target RNA was then detected with Cy3 tyramide (Perkin-Elmer, NEL744B001KT) at a 1:1,000 dilution. The sections were then processed through a pH 6 sodium citrate heat-mediated antigen retrieval, hydrogen peroxidase treatment and protein block (all as for multiplex immunofluorescence staining as above). MNDA antibody was applied overnight at 4 °C, completed using a secondary ImmPress HRP Anti-Rabbit Peroxidase IgG (Vector Laboratories, MP7401), visualized using a fluorescein tyramide (Perkin-Elmer, NEL741B001KT) at a 1:1,000 dilution and stained with DAPI.

Bright-field and fluorescently stained sections were imaged using the slide scanner AxioScan.Z1 (Zeiss) at 20× magnification (40× magnification for smFISH). Images were processed and scale bars added using Zen Blue (Zeiss) and Fiji software[33].

Cell counting and image analysis

Automated cell counting was performed using QuPath software[34]. In brief, DAB-stained whole tissue section slide-scanned images (CZI files)

were imported into QuPath. Cell counts were carried out using the positive cell detection tool, detecting haematoxylin-stained nuclei and then thresholding for positively stained DAB cells, generating DAB-positive cell counts per mm^2 tissue. Identical settings and thresholds were applied to all slides for a given stain and experiment. For cell counts of fibrotic septae versus parenchymal nodules, the QuPath segmentation tool was used to segment the DAB-stained whole tissue section into fibrotic septae or non-fibrotic parenchymal nodule regions using tissue morphological characteristics (Fig. 2j). Positive cell detection was then applied to the fibrotic and non-fibrotic regions in turn, providing DAB-positive cell counts per mm^2 in fibrotic septae and non-fibrotic parenchymal nodules for each tissue section.

Digital morphometric pixel analysis was performed using the Trainable Weka Segmentation (TWS) plugin[35] in Fiji software[33]. In brief, each stained whole tissue section slide-scanned image was converted into multiple TIFF files in Zen Blue software (Zeiss). TIFF files were imported into Fiji and TWS plugin trained to produce a classifier which segments images into areas of positive staining, tissue background and white space[35]. The same trained classifier was then applied to all TIFF images from every tissue section for a particular stain, providing a percentage area of positive staining for each tissue section. For digital morphometric quantification of positive staining of fibrotic septae versus parenchymal nodules, TIFF images were segmented into fibrotic septae or non-fibrotic parenchymal nodule regions using tissue morphological characteristics, followed by analysis using the TWS plugin in Fiji software.

Histological assessment of NASH sections

Sections stained with haematoxylin and eosin or picrosirius red were whole-slide imaged using a NanoZoomer imager (Hamamatsu Photonics). Images of stained sections were independently scored by a consultant liver transplant histopathologist (T.J.K.) at the national liver transplant centre with experience in trial scoring by applying the ordinal NAFLD activity score[36]. For observer-independent quantification of the area of positive picrosirius red staining, images were split using ndpisplit[37] into tiles of ×5 magnification before the application of a classifier that had been trained by the liver histopathologist using the machine learning WEKA plugin in FIJI[33,35], as previously described[38]. All analysis was undertaken blinded to all other data.

Droplet-based scRNA-seq

Single cells were processed through the Chromium Single Cell Platform using the Chromium Single Cell 3′ Library and Gel Bead Kit v2 (10X Genomics, PN-120237) and the Chromium Single Cell A Chip Kit (10X Genomics, PN-120236) as per the manufacturer's protocol. In brief, single cells were sorted into PBS plus 0.1% BSA, washed twice and counted using a Bio-Rad TC20. Then, 10,800 cells were added to each lane of the 10X chip. The cells were partitioned into Gel Beads in Emulsion in the Chromium instrument, in which cell lysis and bar-coded reverse transcription of RNA occurred, followed by amplification, fragmentation and 5′ adaptor and sample index attachment. Libraries were sequenced on an Illumina HiSeq 4000.

Computational analysis

In total, we analysed 67,494 human cells from healthy ($n = 5$) and cirrhotic ($n = 5$) livers, 30,741 PBMCs from patients with cirrhosis ($n = 4$) and compared our data with a publicly available reference dataset of 8,381 PBMCs from a healthy donor (https://support.10xgenomics.com/single-cell-gene-expression/datasets/2.1.0/pbmc8k).

Pre-processing scRNA-seq data

We aligned to the GRCh38 and mm10 (Ensembl 84) reference genomes as appropriate for the input dataset, and estimated cell-containing partitions and associated unique molecular identifiers (UMIs), using the Cell Ranger v.2.1.0 Single-Cell Software Suite from 10X Genomics. Genes expressed in fewer than three cells in a sample were excluded, as were cells that expressed fewer than 300 genes or mitochondrial gene content >30% of the total UMI count. We normalized by dividing the UMI count per gene by the total UMI count in the corresponding cell and log-transforming. Variation in UMI counts between cells was regressed according to a negative binomial model, before scaling and centring the resulting value by subtracting the mean expression of each gene and dividing by its standard deviation (E_n), then calculating $\ln(10^4 \times E_n + 1)$.

Dimensionality reduction, clustering and differential expression analysis

We performed unsupervised clustering and differential gene expression analyses in the Seurat R package v.2.3.0[39]. In particular, we used shared nearest neighbour graph-based clustering, in which the graph was constructed using from 2 to 11 principal components as determined by dataset variability shown in principal component analysis (PCA); the resolution parameter to determine the resulting number of clusters was also tuned accordingly. To assess cluster similarity we used the 'BuildClusterTree' function from Seurat.

In total, we present scRNA-seq data from ten human liver samples (named healthy 1–5 and cirrhotic 1–5), five human blood samples ($n = 4$ cirrhotic named blood 1–4 and $n = 1$ healthy named PBMC8K; pbmc8k dataset sourced from single-cell gene expression datasets hosted by 10X Genomics), and two mouse liver samples ($n = 3$ uninjured and $n = 3$ fibrotic). For seven human liver samples (healthy 1–4 and cirrhotic 1–3), we performed scRNA-seq on both leucocytes ($CD45^+$) and other non-parenchymal cells ($CD45^-$); for the remaining three human livers (healthy 5, cirrhotic 4–5) we performed scRNA-seq on leucocytes only (Extended Data Fig. 2e, f).

Initially, we combined all human scRNA-seq datasets (liver and blood) and performed clustering analysis with the aim of isolating a population of liver-resident cells, by identifying contaminating circulatory cells within datasets generated from liver digests and removing them from downstream analysis. Specifically, we removed from our liver datasets cells that fell into clusters 1 and 13 of the initial dataset in Extended Data Fig. 1d.

Using further clustering followed by signature analysis, we interrogated this post-processed liver-resident cell dataset for robust cell lineages. These lineages were isolated into individual datasets, and the process was iterated to identify robust lineage subpopulations. At each stage of this process we removed clusters expressing more than one unique lineage signature in more than 25% of their cells from the dataset as probable doublets. This resulted in removal of 1,351 cells. Where the cell proliferation signature identified distinct cycling subpopulations, we re-clustered these again to ascertain the identity of their constituent cells.

The mouse scRNA-seq datasets were combined, clustered and interrogated for cell lineages in a similar manner to their human counterparts.

All heat maps, *t*-distributed stochastic neighbour embedding (*t*-SNE) and UMAP visualizations, violin plots and dot plots were produced using Seurat functions in conjunction with the ggplot2, pheatmap and grid R packages. *t*-SNE and UMAP visualizations were constructed using the same number of principal components as the associated clustering, with perplexity ranging from 30 to 300 according to the number of cells in the dataset or lineage. We conducted differential gene expression analysis in Seurat using the standard AUC classifier to assess significance. We retained only those genes with a log-transformed fold change of at least 0.25 and expression in at least 25% of cells in the cluster under comparison.

Defining cell lineage signatures

For each cell, we obtained a signature score across a curated list of known marker genes per cell lineage in the liver (Supplementary Table 2). This signature score was defined as the geometric mean of the expression of the associated signature genes in that cell. Lineage signature scores were scaled from 0 to 1 across the dataset, and the score for each cell

with a signature less than a given threshold (the mean of said signature score across the entire dataset) was set to 0.

Batch effect and quality control

To investigate agreement between samples, we extracted the average expression profile for a given cell lineage in each sample, and calculated the Pearson correlation coefficients between all possible pairwise comparisons of samples per lineage[40].

Imputing dropout in T cell and ILC clusters

To impute dropout of low-abundance transcripts in our T cell and ILC clusters so that we might associate them with known subpopulations, we downsampled to 7,380 cells from 36,900 and applied the scImpute R package v.0.0.8[41], using as input both our previous annotation labels and *k*-means spectral clustering ($k = 5$), but otherwise default parameters.

Analysing functional phenotypes of scar-associated cells

For further analysis of function we adopted the self-organizing maps approach as implemented in the SCRAT R package v.1.0.0[42]. For each lineage of interest, we constructed a self-organizing map in SCRAT using default input parameters and according to its clusters. We defined the signatures expressed in a cell by applying a threshold criterion ($e^{thresh} = 0.95 \times e^{max}$) selecting the highest-expressed metagenes in each cell, and identified for further analysis those metagene signatures defining at least 30% of cells in at least one cluster within the lineage. We smoothed these self-organizing maps using the 'disaggregate' function from the raster R package for visualization purposes, and scaled radar plots to maximum proportional expression of the signature. Gene Ontology enrichment analysis on the genes in these spots was performed using PANTHER 13.1 (pantherdb.org).

Inferring injury dynamics and transcriptional regulation

To generate cellular trajectories (pseudotemporal dynamics) we used the monocle R package v.2.6.1[43]. We ordered cells in a semi-supervised manner on the basis of their Seurat clustering, scaled the resulting pseudotime values from 0 to 1, and mapped them onto either the *t*-SNE or UMAP visualizations generated by Seurat or diffusion maps as implemented in the scater R package v.1.4.0[44] using the top 500 variable genes as input. We removed mitochondrial and ribosomal genes from the gene set for the purposes of trajectory analysis. Differentially expressed genes along this trajectory were identified using generalized linear models via the 'differentialGeneTest' function in monocle.

When determining significance for differential gene expression along the trajectory, we set a *q*-value threshold of 1×10^{-20}. We clustered these genes using hierarchical clustering in pheatmap, cutting the tree at $k = 3$ to obtain gene modules with correlated gene expression across pseudotime. Cubic smoothing spline curves were fitted to scaled gene expression along this trajectory using the smooth.spline command from the stats R package, and Gene Ontology enrichment analysis again performed using PANTHER 13.1.

We verified the trajectory and its directionality using the velocyto R package v.0.6.0[17], estimating cell velocities from their spliced and unspliced mRNA content. We generated annotated spliced and unspliced reads from the 10X BAM files via the 'dropEst' pipeline, before calculating gene-relative velocity using *k*NN pooling with $k = 25$, determining slope gamma with the entire range of cellular expression, and fitting gene offsets using spanning reads. Aggregate velocity fields (using Gaussian smoothing on a regular grid) and transition probabilities per lineage subpopulations were visualized on *t*-SNE, UMAP, or diffusion map visualizations as previously generated. Gene-specific phase portraits were plotted by calculating spliced and unspliced mRNA levels against steady-state inferred by a linear model; levels of unspliced mRNA above and below this steady-state indicate increasing and decreasing expression of said gene, respectively. Similarly, we plotted the unspliced count signal residual per gene, based on the estimated gamma fit, with positive and negative residuals indicating expected upregulation and downregulation, respectively.

For transcription factor analysis, we obtained a list of all genes identified as acting as transcription factors in humans from AnimalTFDB[45]. To analyse transcription factor regulons further, we adopted the SCENIC v.0.1.7 workflow in R[46], using default parameters and the normalized data matrices from Seurat as input. For visualization, we mapped the regulon activity (AUC) scores thus generated to the pseudotemporal trajectories from monocle and the clustering subpopulations from Seurat.

Analysing inter-lineage interactions within the fibrotic niche

For comprehensive systematic analysis of inter-lineage interactions within the fibrotic niche, we used CellPhoneDB[27]. CellPhoneDB is a manually curated repository of ligands, receptors and their interactions, integrated with a statistical framework for inferring cell–cell communication networks from single-cell transcriptomic data. In brief, we derived potential ligand–receptor interactions on the basis of the expression of a receptor by one lineage subpopulation and a ligand by another; as input to this algorithm, we used cells from the fibrotic niche as well as liver sinusoidal endothelial cells and KCs as controls, and we considered only ligands and receptors expressed in greater than 5% of the cells in any given subpopulation. Subpopulation-specific interactions were identified as follows: (1) randomly permuting the cluster labels of all cells 1,000 times and determining the mean of the average receptor expression of a subpopulation and the average ligand expression of the interacting subpopulation, thus generating a null distribution for each ligand–receptor pair in each pairwise comparison between subpopulations; (2) calculating the proportion of these means that were 'as or more extreme' than the actual mean, thus obtaining a *P* value for the likelihood of subpopulation specificity for a given ligand–receptor pair; (3) prioritizing interactions that displayed specificity to subpopulations interacting within the fibrotic niche.

Canonical correlation analysis

To compare human and mouse populations of monocytic phagocytes, we used canonical correlation analysis as implemented in Seurat[18]. We map the genes in the human dataset to their mouse orthologues using biomaRt, discarding any genes for which no orthologues can be found. We then calculate the shared low-dimensional subspace on the union of genes that are variably expressed in both datasets ($n = 159$), and align using six canonical components as determined by evaluating the biweight midcorrelation. Results are visualized by *t*-SNE analysis as previously described.

Deconvolution of whole liver microarray data

To assess the macrophage composition of early-stage NAFLD, we performed deconvolution analysis on publicly available microarray data from annotated liver biopsy specimens taken across the NAFLD disease spectrum (GEO accession GSE48452)[20]. Tissue MP cells from our human scRNA-seq data were manually clustered into the main annotated MP populations. Signature gene expression profiles of SAMacs, TMs and KCs were used to deconvolve the monocyte–macrophage composition of liver biopsy samples from GSE48452 using Cibersort[47], as previously described[19]. The monocyte–macrophage composition of each biopsy sample was then compared to the associated histological and demographic features, available from the GEO database.

Statistics and reproducibility

To assess whether our identified subpopulations were significantly overexpressed in injury, we posited the proportion of injured cells in each cluster as a random count variable using a Poisson process, as previously described[40]. We modelled the rate of detection using the total number of cells in the lineage profiled in a given sample as an offset, with the condition of each sample (healthy versus cirrhotic) provided as a

covariate factor. The model was fitted using the R command 'glm' from the stats package. The *P* value for the significance of the proportion of injured cells was assessed using a Wald test on the regression coefficient.

Remaining statistical analyses were performed using GraphPad Prism. Comparison of changes between two groups was performed using a Mann–Whitney test (unpaired; two-tailed) or a Wilcoxon matched-pairs signed rank test (paired; two-tailed). Comparison of changes between multiple groups was performed using a Kruskal–Wallis and Dunn, one-way ANOVA and Tukey or repeated measures one-way ANOVA and Tukey tests. Correlations were preformed using Pearson correlation and best-fit line plotted using linear regression. $P < 0.05$ was considered statistically significant. All immunofluorescence stains were repeated in a minimum of three patients and representative images are displayed.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Our expression data are freely available for user-friendly interactive browsing online at http://www.livercellatlas.mvm.ed.ac.uk. CellPhoneDB is available at www.CellPhoneDB.org. All raw sequencing data have been deposited in the Gene Expression Omnibus (GEO) under accession GSE136103.

Code availability

R scripts enabling the main steps of the analysis are available from the corresponding authors on reasonable request.

29. Searle, B. C., Gittelman, R. M., Manor, O. & Akey, J. M. Detecting sources of transcriptional heterogeneity in large-scale RNA-seq data sets. *Genetics* **204**, 1391–1396 (2016).
30. Bain, C. C. et al. Long-lived self-renewing bone marrow-derived macrophages displace embryo-derived cells to inhabit adult serous cavities. *Nat. Commun.* **7**, 11852 (2016).
31. Heinrich, M. C. et al. Crenolanib inhibits the drug-resistant PDGFRA D842V mutation associated with imatinib-resistant gastrointestinal stromal tumors. *Clin. Cancer Res.* **18**, 4375–4384 (2012).
32. Patten, D. A. et al. SCARF-1 promotes adhesion of $CD4^+$ T cells to human hepatic sinusoidal endothelium under conditions of shear stress. *Sci. Rep.* **7**, 17600 (2017).
33. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
34. Bankhead, P. et al. QuPath: open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
35. Arganda-Carreras, I. et al. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424–2426 (2017).
36. Kleiner, D. E. et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).
37. Deroulers, C. et al. Analyzing huge pathology images with open source software. *Diagn. Pathol.* **8**, 92 (2013).
38. Kendall, T. J. et al. Hepatic elastin content is predictive of adverse outcome in advanced fibrotic liver disease. *Histopathology* **73**, 90–100 (2018).
39. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
40. Haber, A. L. et al. A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
41. Li, W. V. & Li, J. J. An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nat. Commun.* **9**, 997 (2018).
42. Camp, J. G. et al. Multilineage communication regulates human liver bud development from pluripotency. *Nature* **546**, 533–538 (2017).
43. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
44. McCarthy, D. J., Campbell, K. R., Lun, A. T. L. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).
45. Zhang, H. M. et al. AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. *Nucleic Acids Res.* **43**, D76–D81 (2015).
46. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
47. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).

Acknowledgements This work was supported by an MRC Clinician Scientist Fellowship (MR/N008340/1) to P.R., a Wellcome Trust Senior Research Fellowship in Clinical Science (ref. 103749) to N.C.H., an AbbVie Future Therapeutics and Technologies Division grant to N.C.H., a Guts UK–Children's Liver Disease Foundation grant (ref. R43927) to N.C.H., a Tenovus Scotland grant (ref. E18/05) to R.D. and N.C.H. and British Heart Foundation grants (RM/17/3/33381; RE/18/5/34216) to N.C.H. R.V.-T. was funded by EMBO and Human Frontiers long-term fellowships. C.J.W. was funded by a BBSRC New Investigator Award (BB/N018869/1). P.N.N., C.J.W. and N.T.L. are funded by the NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. This paper presents independent research supported by the NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. J.P.I. is funded by the NIHR Bristol Biomedical Research Centre, University Hospitals Bristol Foundation Trust and the University of Bristol. C.P.P. was funded by the UK Medical Research Council, MC_UU_00007/15. This work was also supported by Wellcome Sanger core funding (WT206194). We thank the patients who donated liver tissue and blood for this study. We thank J. Davidson, C. Ibbotson, J. Black and A. Baird of the Scottish Liver Transplant Unit and the research nurses of the Wellcome Trust Clinical Research Facility for assistance with consenting patients for this study. We thank the liver transplant coordinators and surgeons of the Scottish Liver Transplant Unit and the surgeons and staff of the Hepatobiliary Surgical Unit, Royal Infirmary of Edinburgh for assistance in procuring human liver samples. We thank S. Johnston, W. Ramsay and M. Pattison for technical assistance with FACS and flow cytometry. We thank J. Henderson for technical support and G. Muirhead for assistance with isolation of liver endothelial cells. This publication is part of the Human Cell Atlas (www.humancellatlas.org/publications).

Author contributions P.R. performed experimental design, tissue procurement, data generation, data analysis and interpretation, and manuscript preparation; R.D. performed experimental design, data generation and data analysis; E.F.D., K.P.M., B.E.P.H., M.B., J.A.M. and N.T.L. performed data generation and analysis; J.R.P. generated the interactive online browser; M.E. and R.V.-T. assisted with CellPhoneDB analyses and critically appraised the manuscript; T.J.K. performed pathological assessments and provided intellectual contribution; N.O.C., J.A.F. and P.N.N. provided intellectual contribution; C.J.W. performed tissue procurement, data generation, interpretation and intellectual contribution; J.R.W.-K. performed computational analysis with assistance from J.R.P. and R.S.T. and advice from C.P.P., J.C.M. and S.A.T.; J.R.W.-K. also helped with manuscript preparation, and C.P.P., J.C.M. and S.A.T. critically appraised the manuscript; E.M.H., D.J.M. and S.J.W. procured human liver tissue and critically appraised the manuscript. J.P.I., F.T. and J.W.P. provided intellectual contribution and critically appraised the manuscript; N.C.H. conceived the study, designed experiments, interpreted data and prepared the manuscript.

Competing interests The authors declare no competing interests.

Additional information

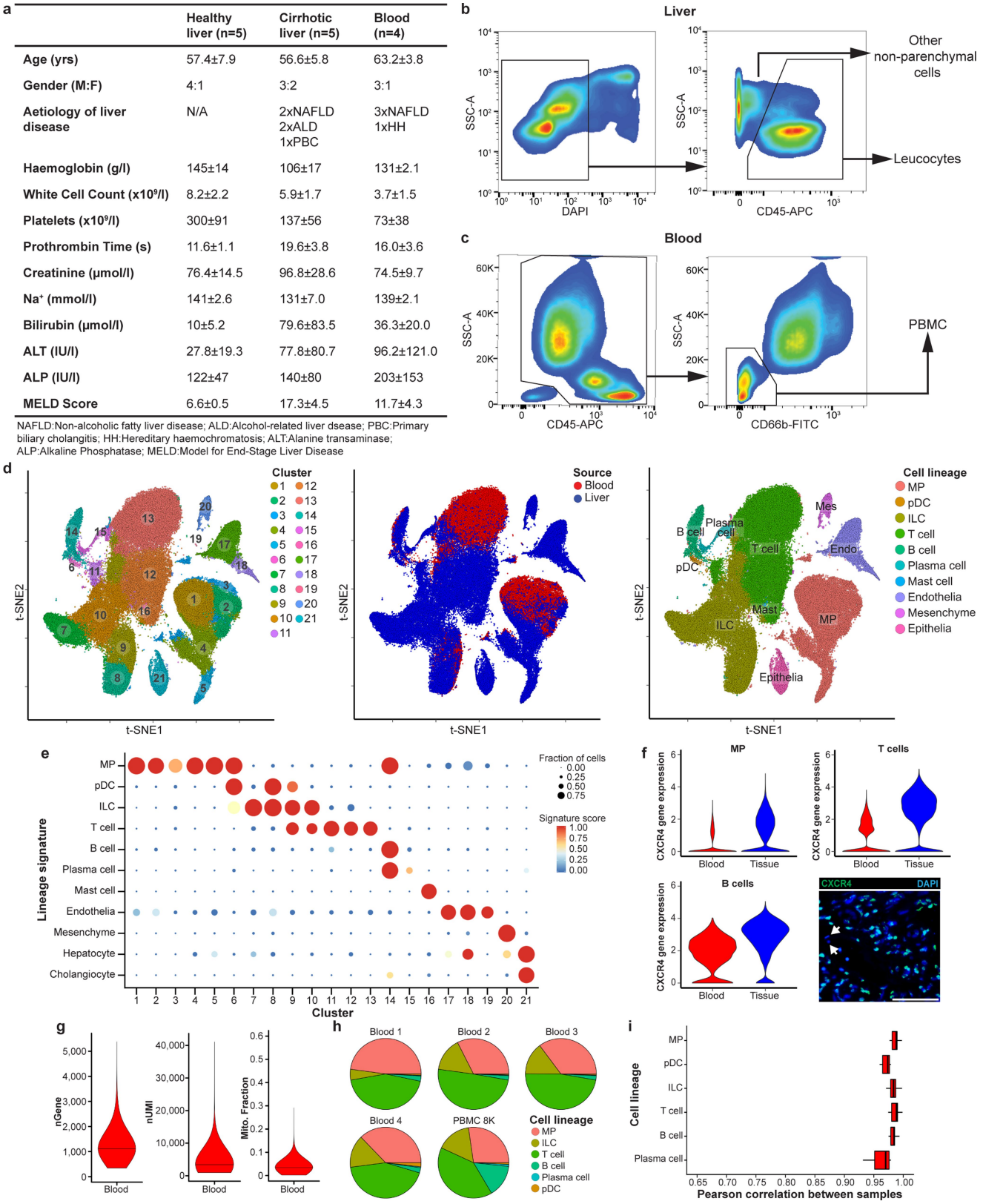
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1631-3.

Correspondence and requests for materials should be addressed to P.R. or N.C.H.

a

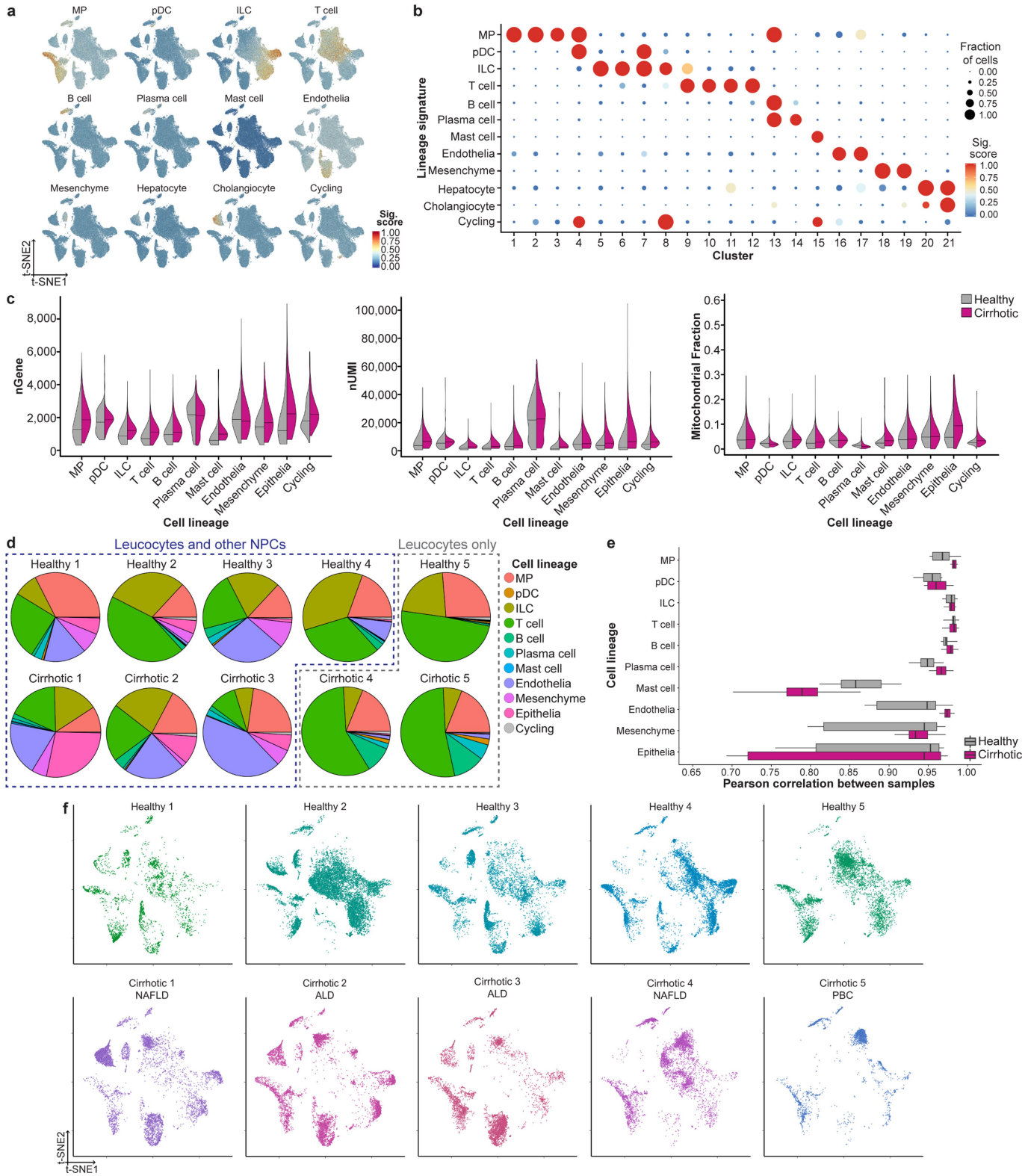
	Healthy liver (n=5)	Cirrhotic liver (n=5)	Blood (n=4)
Age (yrs)	57.4±7.9	56.6±5.8	63.2±3.8
Gender (M:F)	4:1	3:2	3:1
Aetiology of liver disease	N/A	2xNAFLD 2xALD 1xPBC	3xNAFLD 1xHH
Haemoglobin (g/l)	145±14	106±17	131±2.1
White Cell Count (x10⁹/l)	8.2±2.2	5.9±1.7	3.7±1.5
Platelets (x10⁹/l)	300±91	137±56	73±38
Prothrombin Time (s)	11.6±1.1	19.6±3.8	16.0±3.6
Creatinine (μmol/l)	76.4±14.5	96.8±28.6	74.5±9.7
Na⁺ (mmol/l)	141±2.6	131±7.0	139±2.1
Bilirubin (μmol/l)	10±5.2	79.6±83.5	36.3±20.0
ALT (IU/l)	27.8±19.3	77.8±80.7	96.2±121.0
ALP (IU/l)	122±47	140±80	203±153
MELD Score	6.6±0.5	17.3±4.5	11.7±4.3

NAFLD:Non-alcoholic fatty liver disease; ALD:Alcohol-related liver dsease; PBC:Primary biliary cholangitis; HH:Hereditary haemochromatosis; ALT:Alanine transaminase; ALP:Alkaline Phosphatase; MELD:Model for End-Stage Liver Disease

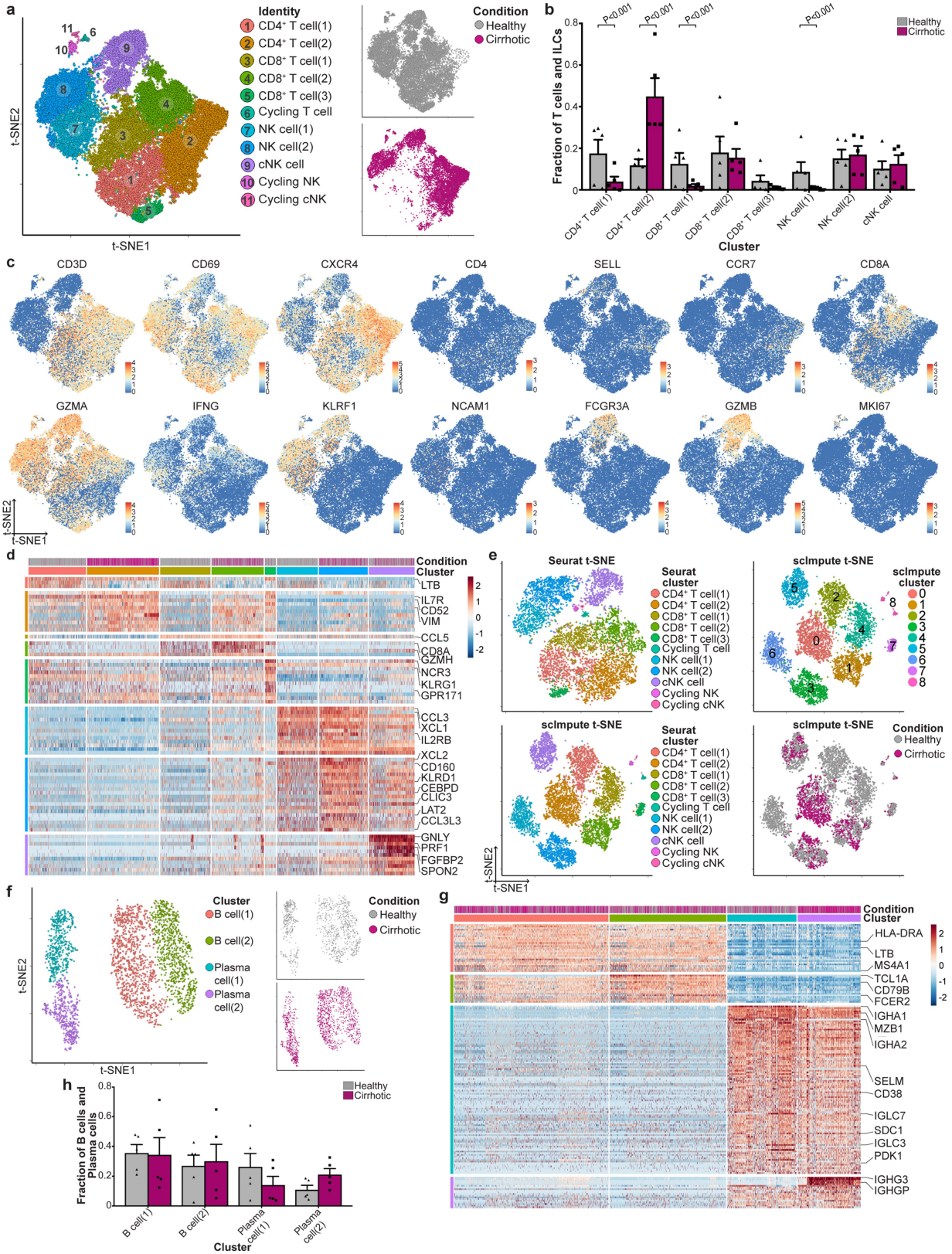


Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Strategy for isolation of human liver non-parenchymal cells. a, Patient demographics and clinical information. Data are mean ± s.e.m. **b**, Flow cytometry gating strategy for isolation of leucocytes ($CD45^+$) and other non-parenchymal cells ($CD45^-$) from human liver; representative plots from ten livers. **c**, Flow cytometry gating strategy for isolation of PBMCs; representative plots from four patients. **d**, Clustering 103,568 cells from healthy ($n = 5$) and cirrhotic ($n = 5$) livers, healthy PBMCs ($n = 1$) and cirrhotic PBMCs ($n = 4$) (left), annotating the source (PBMC versus liver; middle) and cell lineage inferred from known marker gene signatures (right). **e**, Dot plot annotating PBMC and liver clusters by lineage signatures. Circle size indicates cell fraction expressing signature greater than mean; colour indicates mean signature expression (red, high; blue, low). **f**, *CXCR4* gene expression in single cells derived from blood or liver tissue, divided by cell lineage. Bottom right, representative immunofluorescence image ($n \geq 3$) of CXCR4 (green) and DAPI (blue) in human liver; arrows denote $CXCR4^-$ cells in the lumen of a blood vessel. Scale bar, 50 µm. **g**, Violin plots showing the number of unique genes (nGene), number of total unique molecular identifiers (nUMI) and mitochondrial gene fraction expressed in five PBMC samples. Black lines denote the median. **h**, Pie charts showing the proportion of cell lineages per PBMC sample. **i**, Box and whisker plots showing the agreement in expression profiles across five PBMC samples. Pearson correlation coefficients between average expression profiles for cells in each lineage, across all pairs of samples. Black bars denote the median; box edges denote the twenty-fifth and seventy-fifth percentiles; whiskers denote the full range.

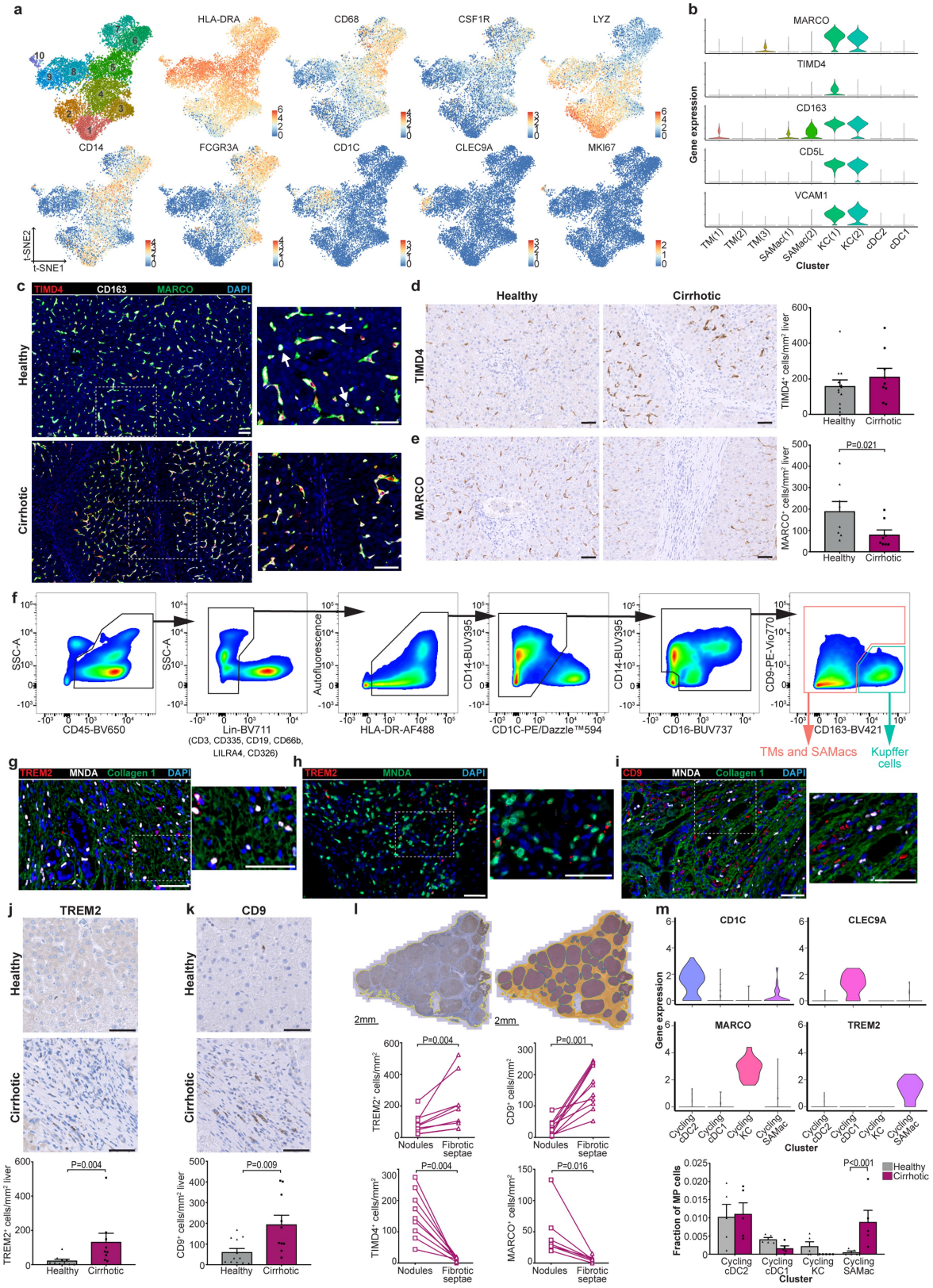


Extended Data Fig. 2 | Quality control and annotation of human liver-resident cells. a, Lineage signature expression across 66,135 liver-resident cells from healthy ($n = 5$) and cirrhotic ($n = 5$) human livers (red, high; blue, low). **b**, Dot plot annotating liver-resident cell clusters by lineage signature. Circle size indicates cell fraction expressing signature greater than mean; colour indicates mean signature expression (red, high; blue, low). **c**, Violin plots of the number of unique genes (left), number of total UMIs (middle) and mitochondrial gene fraction (right) across 66,135 liver-resident cells from healthy ($n = 5$) and cirrhotic ($n = 5$) livers. Black lines denote the median. **d**, Pie charts of the proportion of cell lineage per liver sample. **e**, Box and whisker plots of the agreement in expression profiles across healthy ($n = 5$) and cirrhotic ($n = 5$) liver samples, as in Extended Data Fig. 1i. **f**, *t*-SNE visualization of liver-resident cells per liver sample, with cirrhotic samples annotated by aetiology of underlying liver disease. ALD, alcohol-related liver disease; PBC, primary biliary cholangitis.



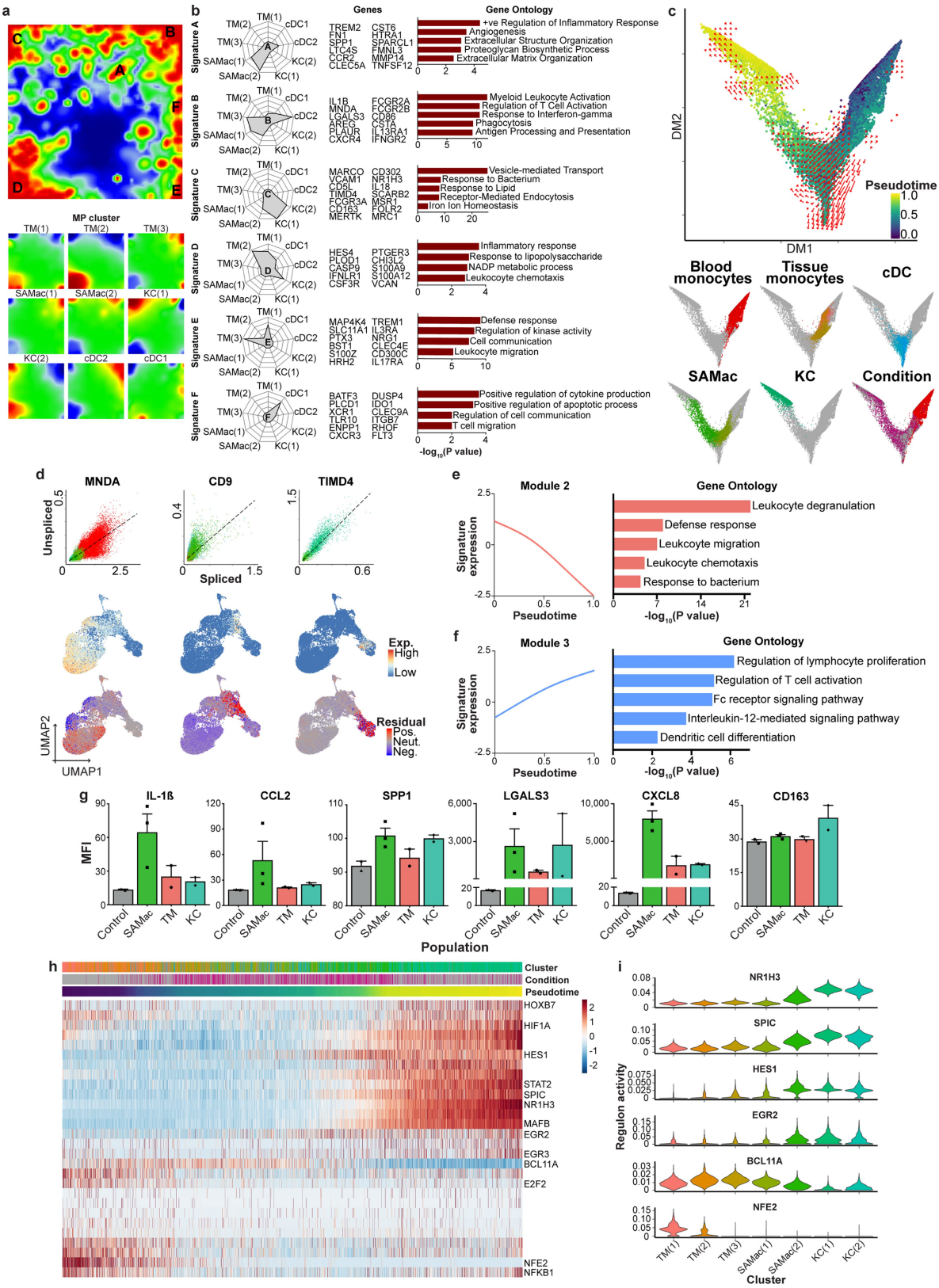
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Annotating human liver lymphoid cells. a, Clustering of 36,900 T cells and ILCs (left) from healthy ($n = 5$) and cirrhotic ($n = 5$) human livers, annotating the injury condition (right). NK, natural killer cell; cNK, cytotoxic NK cell. **b**, Fractions of T cell and ILC subpopulations in healthy ($n = 5$) and cirrhotic ($n = 5$) livers. **c**, Selected gene expression in 36,900 T cells and ILCs. **d**, Heat map of T cell and ILC cluster marker genes (colour-coded by cluster and condition), with exemplar genes labelled (right). Columns denote cells; rows denote genes. **e**, *t*-SNE visualizations of downsampled T cell and ILC dataset (7,380 cells from healthy ($n = 5$) and cirrhotic ($n = 5$) human livers) before and after imputation (scImpute); annotating data used for visualization and clustering, inferred lineage and injury condition. No additional heterogeneity was observed after imputation. **f**, Clustering 2,746 B cells and plasma cells (left) from healthy ($n = 5$) and cirrhotic ($n = 5$) human livers, annotating the injury condition (right). **g**, Heat map of B cell and plasma cell cluster marker genes (colour-coded by cluster and condition), with exemplar genes labelled (right). Columns denote cells; rows denote genes. **h**, Fractions of B cell and plasma cell subpopulations in healthy ($n = 5$) and cirrhotic ($n = 5$) livers. Data are mean ± s.e.m. *P* values determined by Wald test (**b**).



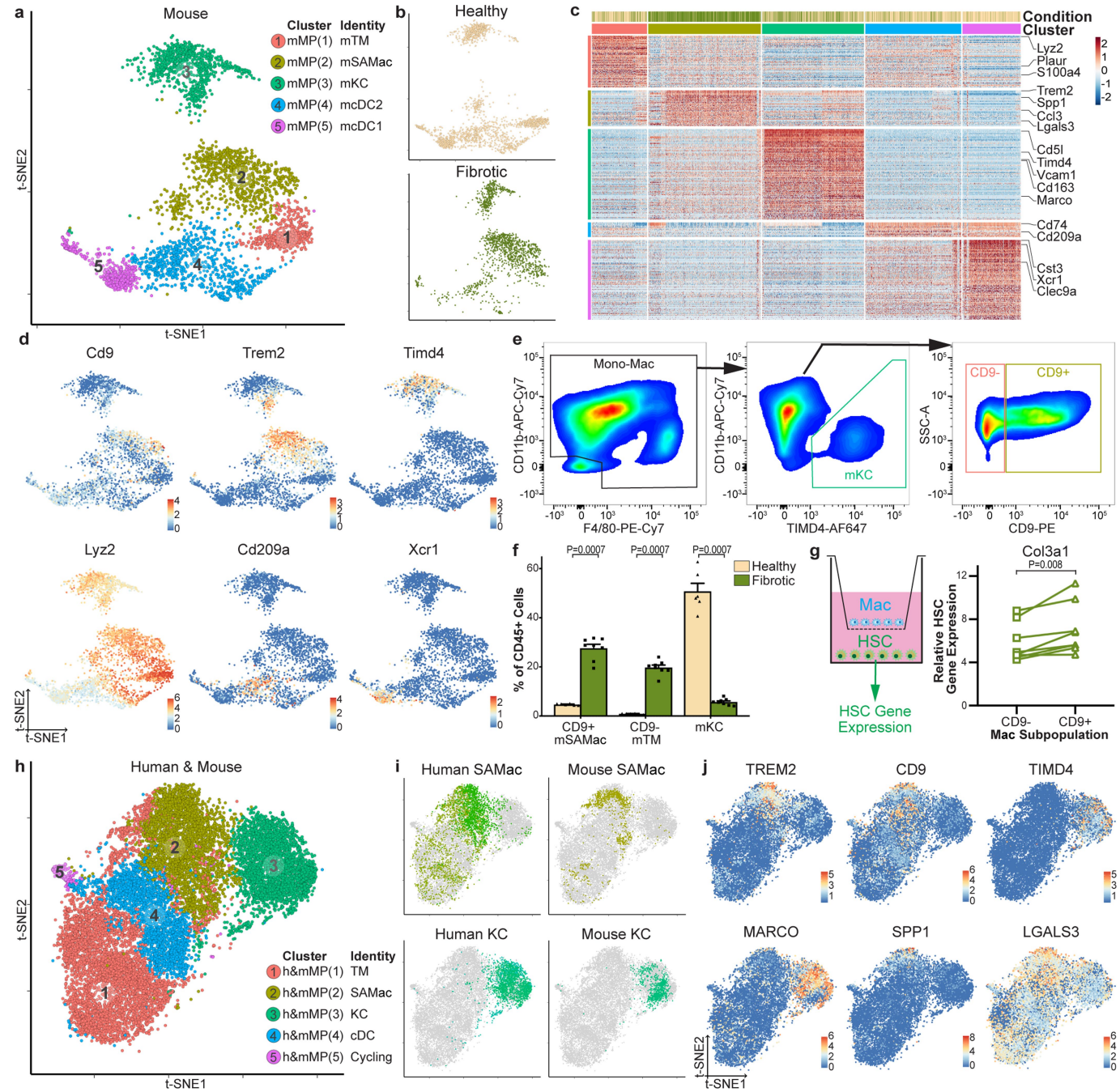
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Annotating human liver MPs. a, Clustering and selected genes expressed in 10,737 MPs from healthy ($n = 5$) and cirrhotic ($n = 5$) human livers. **b**, Scaled gene expression of KC cluster markers across MP cells from healthy ($n = 5$) and cirrhotic ($n = 5$) livers. **c**, Representative immunofluorescence images ($n \geq 3$) of TIMD4 (red), CD163 (white), MARCO (green) and DAPI (blue) in healthy and cirrhotic liver; arrows denote $CD163^{+}MARCO^{+}TIMD4^{-}$ cells.
d, Immunohistochemistry (left) and cell counts (right) of TIMD4 expression in healthy ($n = 12$) and cirrhotic ($n = 9$) human liver. **e**, Immunohistochemistry (left) and cell counts (right) of MARCO expression in healthy ($n = 8$) and cirrhotic ($n = 8$) liver. **f**, Flow cytometry gating strategy for identifying KCs, TMs and SAMacs in healthy ($n = 2$) and cirrhotic ($n = 3$) liver. SAMacs are detected as $TREM2^{+}CD9^{+}$ cells within the TM and SAMac gate (see Fig. 2f). **g**, Representative immunofluorescence images ($n \geq 3$) of TREM2 (red), MNDA (white), collagen 1 (green) and DAPI (blue) in cirrhotic liver. **h**, Representative images ($n = 2$) of TREM2 (smFISH; red), MNDA (immunofluorescence; green) and DAPI (blue) in cirrhotic liver. **i**, Representative immunofluorescence images ($n \geq 3$) of CD9 (red), MNDA (white), collagen 1 (green) and DAPI (blue) in cirrhotic liver. **j**, Immunohistochemistry (top) and cell counts (bottom) of TREM2 expression in healthy ($n = 10$) and cirrhotic ($n = 9$) liver. **k**, Immunohistochemistry (top) and cell counts (bottom) of CD9 expression in healthy ($n = 12$) and cirrhotic ($n = 10$) liver. **l**, Top, exemplar tissue segmentation of cirrhotic liver section into fibrotic septae (orange) and parenchymal nodules (purple). Bottom, cell counts based on immunohistochemistry analysis of TREM2 ($n = 9$), CD9 ($n = 11$), TIMD4 ($n = 9$) and MARCO ($n = 7$) in parenchymal nodules and fibrotic septae. **m**, Top, clustering and annotation of 208 cycling MP cells from healthy ($n = 5$) and cirrhotic ($n = 5$) livers, with scaled gene expression of MP subpopulation markers across four clusters of cycling MP cells. Bottom, fractions of cycling MP subpopulations in healthy ($n = 5$) and cirrhotic ($n = 5$) livers. All scale bars, 50 μm. Data are mean ± s.e.m. *P* values determined by two-tailed Mann–Whitney (**e**, **j**, **k**),

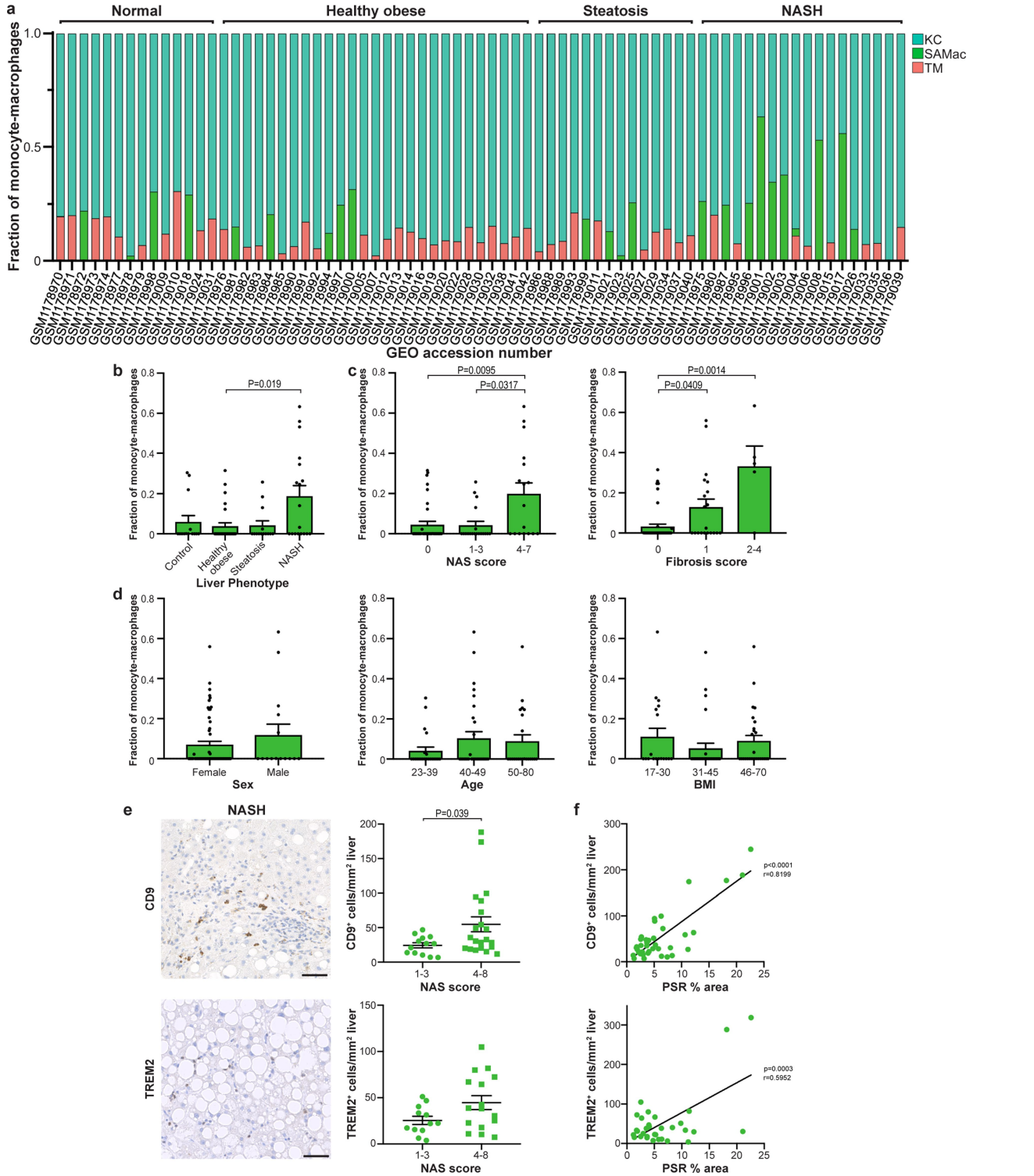


Extended Data Fig. 5 | See next page for caption.

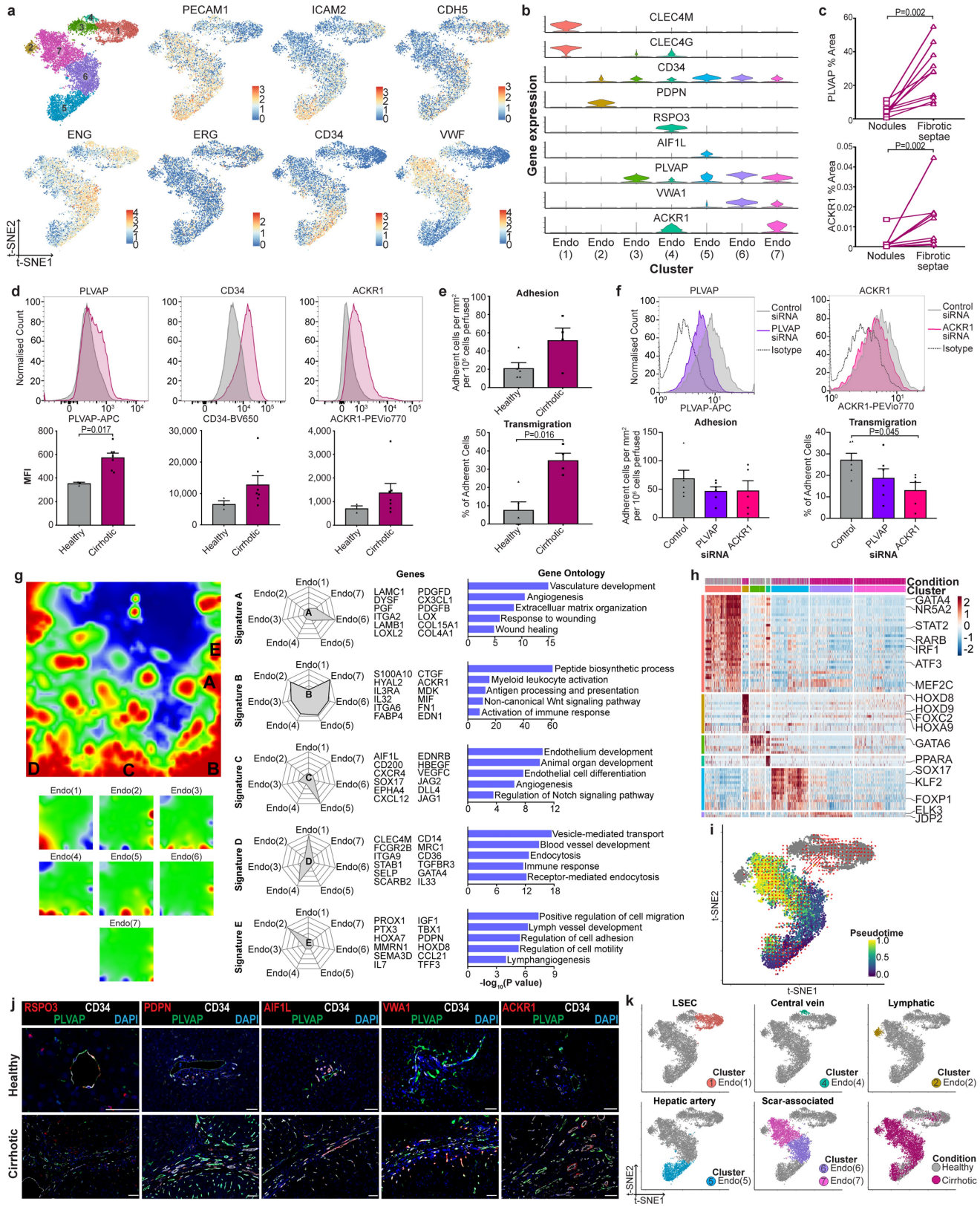
Extended Data Fig. 5 | Phenotypic characterization of mononuclear phagocytes in healthy and cirrhotic human livers. a, Top, self-organizing map (60 × 60 grid) of smoothed scaled metagene expression of 10,737 MPs from healthy ($n = 5$) and cirrhotic ($n = 5$) livers. In total, 20,952 genes, 3,600 metagenes and 44 signatures were identified. A–F denote metagene signatures overexpressed in one or more MP subpopulations. Bottom, smoothed mean metagene expression profile for each MP subpopulation. **b**, Radar plots (left), exemplar genes (middle) and selected GO enrichment (right) of metagene signatures A–F showing distribution of signature expression across MP subpopulations from 10,737 MP cells. **c**, Diffusion map (DM) visualization of blood monocytes and liver-resident MP lineages (23,075 cells from healthy ($n = 5$) and cirrhotic ($n = 5$) liver samples and PBMCs ($n = 5$)), annotating monocle pseudotemporal dynamics (purple to yellow). Top, RNA velocity field (red arrows) visualized using Gaussian smoothing on regular grid. Bottom, annotation of MPs by subpopulation and injury condition. **d**, Unspliced–spliced phase portraits (top); 23,075 cells coloured and visualized as in Fig. 3a; monocyte (*MNDA*), SAMac (*CD9*) and KC (*TIMD4*) marker genes. Cells plotted above or below the steady-state (black dashed line) indicate increasing or decreasing expression of gene, respectively. Spliced expression profile for stated genes (middle row; red, high, blue, low). Unspliced residuals for stated genes (bottom row), positive (red) indicating expected upregulation, negative (blue) indicating expected downregulation. *MNDA* displays negative velocity in SAMacs; *CD9* displays positive velocity in monocytes and SAMacs; *TIMD4* velocity is restricted to KCs. **e**, Cubic smoothing spline curve fitted to averaged expression of all genes in module 2 from the blood monocyte-to-SAMac pseudotemporal trajectory (see Fig. 3c), with selected GO enrichment (right). **f**, Cubic smoothing spline curve fitted to averaged expression of all genes in module 3 from the blood monocyte-to-cDC pseudotemporal trajectory (see Fig. 3c), with selected GO enrichment (right). **g**, Luminex assay showing quantification of levels of stated proteins in culture medium from FACS-isolated SAMacs ($n = 3$), TMs ($n = 2$) and KCs ($n = 2$). Control denotes medium alone ($n = 2$). Data are mean ± s.e.m. **h**, Heat map of transcription factor regulons across MP pseudotemporal trajectory and in KCs (colour-coded by MP cluster, condition and pseudotime), with selected regulons labelled (right). Columns denote cells; rows denote genes. **i**, Scaled regulon expression of selected regulons across MP clusters from healthy ($n = 5$) and cirrhotic ($n = 5$) livers. All *P* values determined by Fisher's exact test.



Extended Data Fig. 6 | Characterization of macrophages in mouse liver fibrosis. a, Clustering and annotating 3,250 mouse (m)MPs from healthy ($n = 3$) and fibrotic (4 weeks CCl_4 treatment; $n = 3$) livers. **b**, Annotating mouse MP cells by injury condition. **c**, Heat map of mouse MP cluster marker genes (top; colour-coded by cluster and condition), with exemplar genes labelled (right). Columns denote cells; rows denote genes. **d**, Selected genes expressed in 3,250 mouse MPs. **e**, Representative flow cytometry plots of the gating strategy ($n = 8$ from two independent experiments) for identifying mouse KCs, $CD9^-$ TMs and $CD9^+$ SAMacs in fibrotic mice. **f**, Quantifying mouse macrophage subpopulations by flow cytometry in healthy ($n = 6$) and fibrotic ($n = 8$) mouse livers from two independent experiments. The macrophage subpopulation (x axis) is shown as a percentage of total viable $CD45^+$ cells (y axis). Data are mean ± s.e.m. P values determined by two-tailed Mann–Whitney test. **g**, Co-culture of primary mouse HSCs from uninjured livers and either FACS-isolated $CD9^-$ mouse TMs or $CD9^+$ mouse SAMacs from fibrotic livers ($n = 8$ mice; two independent experiments). Right, qPCR of *Col3a1* expression in HSCs; expression relative to mean expression of quiescent HSC. P value determined by two-tailed Wilcoxon test. **h**, Clustering 3,250 mouse MPs and 10,737 human (h)MPs into five clusters using canonical correlation analysis. Annotation of cross-species clusters (identity). **i**, Annotation of human and mouse macrophage subpopulations from 3,250 mouse MPs and 10,737 human MPs. **j**, Selected genes expressed in 3,250 mouse MPs and 10,737 human MPs.

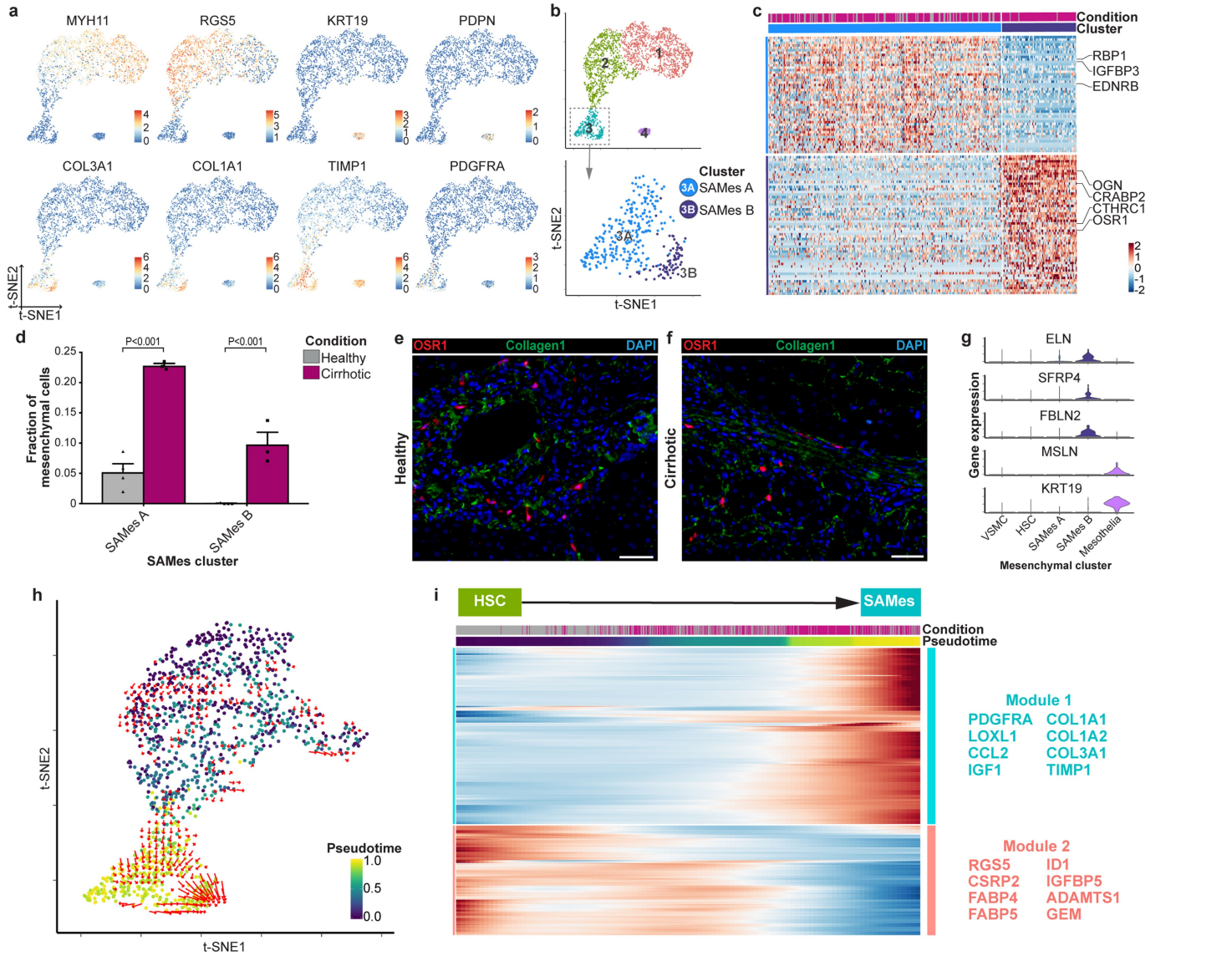


Extended Data Fig. 7 | SAMac expansion in human NASH. a–**d**, Deconvolution of publicly available whole liver microarray data ($n = 73$) assessed for frequency of SAMacs, KCs and TMs using the Cibersort algorithm. **a**, Macrophage composition. GEO accession numbers are shown on the *x* axis; the fraction of monocyte–macrophages is shown on the *y* axis. Liver phenotypes are annotated at the top. **b**, Frequency of SAMacs in control ($n = 14$), heathy obese ($n = 27$), steatosis ($n = 14$) and NASH ($n = 18$) livers. **c**, Left, frequency of SAMacs in patients with histological NAFLD activity scores (NAS) of 0 ($n = 37$), 1–3 ($n = 19$) and 4–7 ($n = 17$). Right, frequency of SAMacs in patients with histological fibrosis scores of 0 ($n = 46$), 1 ($n = 20$) and 2–4 ($n = 5$). **d**, Left, frequency of SAMacs in female ($n = 58$) and male ($n = 15$) patients. Middle, frequency of SAMacs in patients aged 23–39 ($n = 22$), 40–49 ($n = 29$) and 50–80 ($n = 22$). Right, frequency of SAMacs in patients with a body mass index (BMI) of 17–30 ($n = 18$), 31–45 ($n = 28$) and 46–70 ($n = 27$). **e**, Left, immunohistochemistry of CD9 and TREM2 expression in NAFLD liver biopsy sections. Scale bars, 50 µm. Right, cell counts of CD9 and TREM2 expression. CD9: NAS 1–3 ($n = 13$), NAS 4–8 ($n = 21$). TREM2: NAS 1–3 ($n = 12$), NAS 4–8 ($n = 16$). **f**, Correlation of cell counts with picrosirius red (PSR) digital morphometric pixel quantification in NAFLD liver biopsy tissue with CD9 staining (top; $n = 39$) or TREM2 staining (bottom; $n = 32$). Data are mean ± s.e.m. *P* values determined by Kruskal–Wallis and Dunn test (**b**, **c**), two-tailed Mann–Whitney test (**e**) or Pearson's correlation and linear regression (**f**).

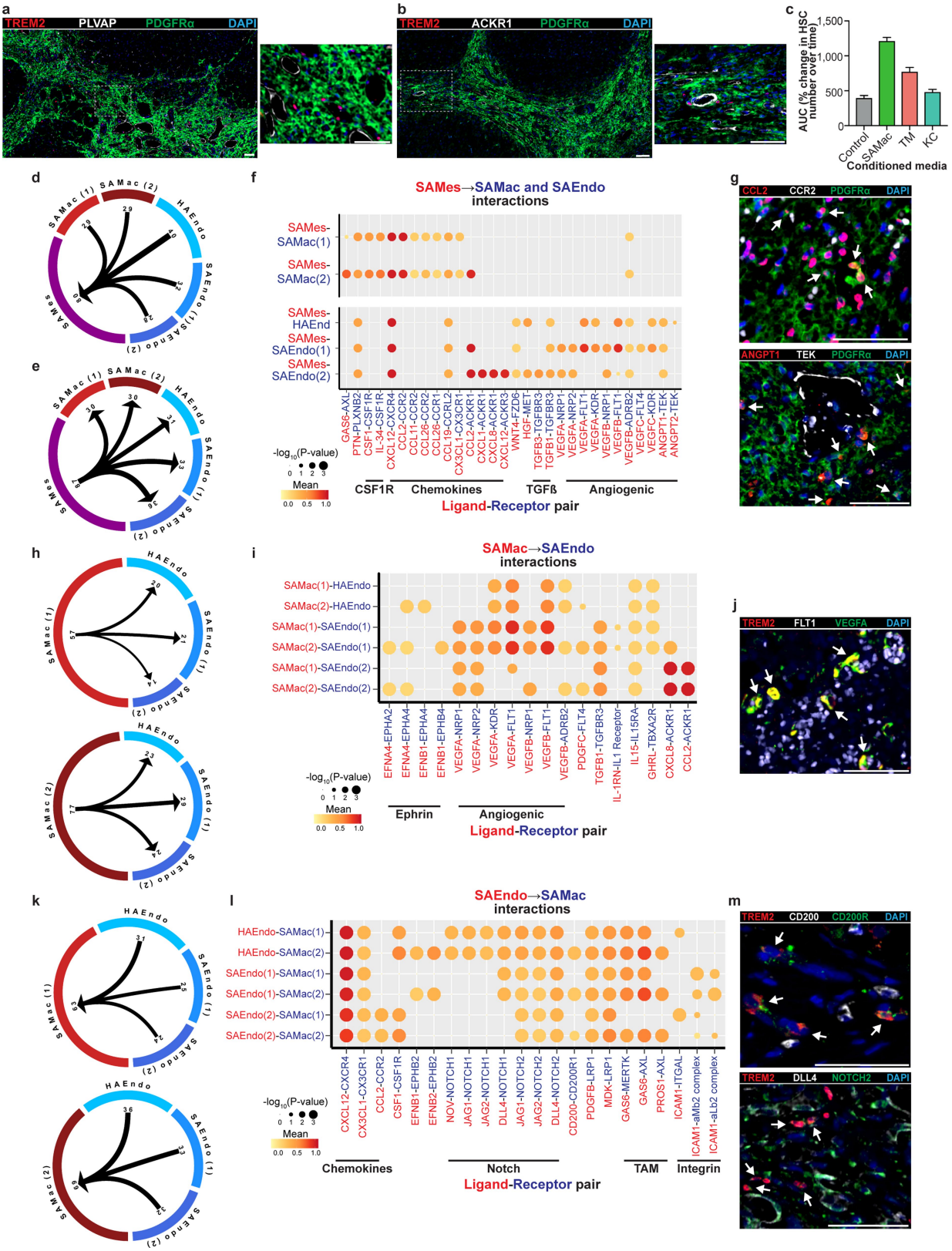


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Phenotypic characterization of endothelial cells in healthy and cirrhotic human livers. a, Clustering and selected genes expressed in 8,020 endothelial cells from healthy ($n = 4$) and cirrhotic ($n = 3$) human livers. **b**, Scaled gene expression of endothelial cluster markers across endothelial cells from healthy ($n = 4$) and cirrhotic ($n = 3$) livers. **c**, Top, digital pixel quantification of PLVAP immunostaining of cirrhotic liver sections ($n = 10$) in parenchymal nodules and fibrotic septae. Bottom, ACKR1 immunostaining of cirrhotic liver sections ($n = 10$) in parenchymal nodules and fibrotic septae. **d**, Flow cytometry analysis of PLVAP, CD34 and ACKR1 in endothelial cells from healthy ($n = 3$, grey) or cirrhotic ($n = 7$, red) livers. Top, representative histograms; bottom, MFI values. **e**, Flow-based adhesion assay. Peripheral blood monocytes assessed for adhesion to primary human liver endothelial cells (top) and the percentage of adherent monocytes that transmigrate (bottom); endothelial cells isolated from healthy ($n = 5$) or cirrhotic ($n = 4$) livers. **f**, Endothelial cell gene knockdown. Cirrhotic endothelial cells were treated with siRNA against *PLVAP* ($n = 6$) or *ACKR1* ($n = 5$) or with control siRNA ($n = 6$). Top, representative flow cytometry histograms for stated markers, with comparison to isotype control antibody. Bottom, flow-based adhesion assay, with PBMCs assessed for adhesion (bottom left) and the percentage of adherent cells that transmigrate (bottom right) after siRNA treatment of endothelial cells. **g**, Top left, self-organizing map (60 × 60 grid) of smoothed scaled metagene expression of endothelia lineage. In total, 21,237 genes, 3,600 metagenes and 45 signatures were identified. A–E denote metagene signatures overexpressed in one or more endothelial subpopulations. Bottom left, smoothed mean metagene expression profile for each endothelial subpopulation. Middle, radar plots of metagene signatures A–E showing distribution of signature expression across endothelial subpopulations, exemplar genes (middle) and Gene Ontology enrichment (right). **h**, Heat map of endothelial subpopulation transcription factor regulon expression (colour-coded by cluster and condition) across 8,020 endothelial cells from healthy ($n = 4$) and cirrhotic ($n = 3$) human livers. Exemplar regulons are labelled (right). Columns denote cells; rows denote regulons. **i**, *t*-SNE visualization of endothelial lineage (8,020 cells from healthy ($n = 4$) and cirrhotic ($n = 3$) livers), annotating monocle pseudotemporal dynamics (purple to yellow; grey indicates lack of inferred trajectory). RNA velocities (red arrows) visualized using Gaussian smoothing on regular grid. **j**, Representative immunofluorescence images ($n \geq 3$) of RSPO3, PDPN, AIF1L, VWA1 or ACKR1 (red), CD34 (white), PLVAP (green) and DAPI (blue) in healthy and cirrhotic liver. Scale bars, 50 µm. **k**, Annotation of 8,020 endothelial cells by subpopulation and injury condition. LSEC, liver sinusoidal endothelial cells. Data are mean ± s.e.m. *P* values determined by two-tailed Wilcoxon test (**c**), two-tailed Mann–Whitney test (**d**, **e**), Kruskal–Wallis and Dunn test (**f**), or Fisher's exact test (**g**).



Extended Data Fig. 9 | Characterization of mesenchymal cells in healthy and cirrhotic human livers. a, Selected genes expressed in 2,318 mesenchymal cells from healthy ($n = 4$) and cirrhotic ($n = 3$) human livers. **b**, Clustering 319 SAMes into two further subclusters. **c**, Heat map of SAMes subcluster marker genes (colour-coded by cluster and condition), with exemplar genes labelled (right). Columns denote cells; rows denote genes. **d**, Fractions of SAMes subpopulations in healthy ($n = 4$) and cirrhotic ($n = 3$) livers. **e**, **f**, Representative immunofluorescence images ($n \geq 3$) of OSR1 (red), collagen 1 (green) and DAPI (blue) in portal region of healthy liver (**e**) or fibrotic niche of cirrhotic liver (**f**). Scale bars, 50 µm. **g**, Scaled gene expression of selected genes across 2,318 mesenchymal cells from healthy ($n = 4$) and cirrhotic ($n = 3$) livers. **h**, *t*-SNE visualization of 1,178 HSCs and SAMes from healthy ($n = 4$) and cirrhotic ($n = 3$) livers annotated by monocle pseudotemporal dynamics (purple to yellow). RNA velocity field (red arrows) visualized using Gaussian smoothing on regular grid. **i**, Heat map of cubic smoothing spline curves fitted to genes differentially expressed across HSC-to-SAMes pseudotemporal trajectories, grouped by hierarchical clustering ($k = 2$); colour-coded by pseudotime and condition (top). Gene co-expression modules (colour) and exemplar genes are labelled (right). Data are mean ± s.e.m. *P* values determined by Wald test (**d**).



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Characterization of the cellular interactome in the fibrotic niche. a, **b**, Representative immunofluorescence images ($n \geq 3$) of fibrotic niche in cirrhotic liver. **a**, TREM2 (red), PLVAP (white), PDGFRα (green) and DAPI (blue). **b**, TREM2 (red), ACKR1 (white), PDGFRα (green) and DAPI (blue). **c**, Proliferation assay. Human HSCs were treated with conditioned medium from primary hepatic macrophage subpopulations SAMac ($n = 2$), TMs ($n = 2$), KCs ($n = 2$) or control medium ($n = 2$). The AUC of the percentage change in HSC number over time (hours) is shown on the *y* axis. Data are mean ± s.e.m. **d**, Circle plot showing potential interaction magnitude from ligands expressed by SAMacs and SAEndos to receptors expressed on SAMes. **e**, Circle plot showing potential interaction magnitude from ligands expressed by SAMes to receptors expressed on SAMacs and SAEndos. **f**, Dot plot of ligand–receptor interactions between SAMes ($n = 7$ human livers), SAMacs ($n = 10$ human livers) and SAEndos ($n = 7$ human livers). Ligand (red) and cognate receptor (blue) shown on the *x* axis; populations that express ligand (red) and receptor (blue) are shown on the *y* axis; circle size denotes *P* value (permutation test); colour (red, high; yellow, low) denotes average ligand and receptor expression levels in interacting subpopulations. **g**, Top, representative immunofluorescence image ($n \geq 3$) of CCL2 (red), CCR2 (white), PDGFRα (green) and DAPI (blue) in fibrotic niche in cirrhotic liver; arrows denote $CCL2^+PDGFR\alpha^+$ cells. Bottom, representative immunofluorescence image ($n \geq 3$) of ANGPT1 (red), TEK (white), PDGFRα (green) and DAPI (blue) in fibrotic niche in cirrhotic liver; arrows denote $ANGPT1^+PDGFR\alpha^+$ cells. **h**, Circle plot denotes potential interaction magnitude from ligands expressed by SAMacs to receptors expressed on SAEndos. **i**, Dot plot of ligand–receptor interactions between SAMacs ($n = 10$ human livers) and SAEndos ($n = 7$ human livers) as in **f**. **j**, Representative immunofluorescence image ($n \geq 3$) of TREM2 (red), FLT1 (white), VEGFA (green) and DAPI (blue) in fibrotic niche in cirrhotic liver; arrows denote $TREM2^+VEGFA^+$ cells. **k**, Circle plot of the potential interaction magnitude from ligands expressed by SAEndos to receptors expressed on SAMacs. **l**, Dot plot of ligand–receptor interactions between SAEndo ($n = 7$ human livers) and SAMacs ($n = 10$ human livers) as in **f**. **m**, Top, representative immunofluorescence image ($n \geq 3$) of TREM2 (red), CD200 (white), CD200R (green) and DAPI (blue) in fibrotic niche in cirrhotic liver; arrows denote $TREM2^+CD200R^+$ cells. Bottom, representative immunofluorescence image ($n \geq 3$) of TREM2 (red), DLL4 (white), NOTCH2 (green) and DAPI (blue) in fibrotic niche in cirrhotic liver; arrows denote $TREM2^+NOTCH2^+$ cells. All scale bars, 50 µm.

Corresponding author(s): Prakash Ramachandran
Neil C. Henderson

Last updated by author(s): Aug 20, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Initial processing of single-cell RNA-sequencing data was performed using the commercial CellRanger pipeline (10X Genomics, version 2.1.0, see Methods). Subsequent analysis was performed using the open-source R programming language (version 3.4.1). BD FACS Sortware software was used for cell sorting on BD Influx equipment. BD FACS Diva software was used for flow cytometry on BD LSR Fortessa equipment and for cell sorting on BD FACSAria Fusion and FACSAriaII. Fluorescent and brightfield microscopy images were acquired using Zen Blue software (Zeiss) on an Axioscan.Z1 instrument (Zeiss) or Confocal Microscope Zeiss LSM780. Luminex data was acquired on a Bio-Plex 200 (Bio-Rad). Cell Proliferation data was acquired on an Incucyte ZOOM live cell analysis system (Essen biosciences). RT-qPCR data was acquired on ABI 7900HT FAST PCR system (Applied Biosystems).
Data analysis	Immunoflurorescent images were processed and analysed using Zen Blue software (Zeiss) and Fiji software (ImageJ version 2.00). Cell proliferation data were analysed on the Incucyte proprietary analysis software (version 2018A). Immunohistochemistry images were analysed using QuPAth software (version 0.1.2) for automated cell counting and using Fiji software (ImageJ version 2.00) with Trainable Weka Segmentation plugin (see Methods). Co-culture immunocytochemistry data was analysed using Imaris x64 (version 8.1.2). Flow cytometry analysis was performed using FlowJo software (version 10.2). RT-qPCR data was analysed using ThermoFisher Connect cloud qPCR software (version 2019.1.8-Q1-19-build4). Statistical analysis was performed either in R (version 3.4.1) or using Graphpad Prism software (version 7.0a). Single-cell RNA-sequencing analysis was performed in R, based around the following packages: Seurat 2.3.0, scImpute 0.0.8, SCRAT 1.0.0, monocle 2.6.1, scater 1.4.0, velocyto 0.6.0, SCENIC 0.1.7 (see Methods). We also made use of the CellPhoneDB repository of ligands, receptors, and interactions. Deconvolution was performed using Cibersort. Gene Ontology enrichment analysis was performed using PANTHER 13.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Our expression data will be freely available for user-friendly interactive browsing online at www.livercellatlas.mvm.ed.ac.uk. CellPhoneDB is available at www.CellPhoneDB.org, along with lists of membrane proteins, ligands and receptors, and heteromeric complexes. All raw sequencing data have been deposited in the Gene Expression Omnibus (GEO Accession GSE136103). We make available as Supplementary Tables: lists of lineage-specific genes for signature analysis (Extended Data Figure 1e, 2b); lists of marker genes and regulons from clustering results (Figure 1e, 2d, 4c, 5b, Extended Data Figure 3d, e, g, 5h, 6c, 8h, 9c); lists of module / signature genes from trajectory and self-organising map analyses and corresponding lists of gene ontology terms from enrichment analysis (Figure 3c, d, Extended Data Figure 5a, b, e, f, 8g); lists of significant interactions in the fibrotic niche as identified using CellPhoneDB (Figure 6a, e, Extended Data Figure 10f, i, l).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In total, we present scRNA-seq data from ten human liver samples (n=5 healthy and n=5 cirrhotic), five human blood samples (n=4 cirrhotic and n=1 healthy named PBMC8K; pbmc8k dataset sourced from single-cell gene expression datasets hosted by 10X Genomics) and two mouse samples (n=3 healthy and n=3 fibrotic). No statistical methods were used to predetermine sample size. Patient number was selected to give a balanced representation of healthy and cirrhotic liver cells and to provide sufficient cells of each lineage to facilitate more detailed analysis. Histology, flow cytometry, luminex, RT-qPCR and cell proliferation analysis were performed on multiple independent biological replicates (n shown in figure legends).
Data exclusions	Described in detail in Methods. Exclusion criteria were determined following initial assessment and QC of the data. Low gene expression (fewer than 300 genes) or a high mitochondrial gene content (>30% of the total UMI count) are indicators of outlier low quality cells and were excluded. At each stage of the analysis we used signature analysis to identify and exclude potential doublet clusters.
Replication	All experimental findings reported here were successfully replicated across multiple biological samples (n reported in each figure legend). All immunofluorescence was performed on a minimum of 3 liver samples to identify representative images.
Randomization	One group of randomly selected healthy livers and another group of randomly selected cirrhotic livers were analysed in this study. All subsequent analyses were performed in randomly selected healthy or cirrhotic liver samples. For mouse experiments, age-matched littermate mice were randomly assigned to be healthy controls or receive carbon tetrachloride.
Blinding	Blinding to the origin of the tissue samples was not possible. All analyses were performed in an automated manner across conditions.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	All antibodies used in this work, clone, application, supplier and lot number are listed in Supplementary Table 19.
Validation	All antibodies used are commerically available and validated by the vendor for the assay and species used in this study. Specific validation information for each antibody is available on the vendors website. The specificity of each primary flow cytometry antibody was validated by staining directly against species-matched isotype and unstained controls. Validation of each primary antibody used for immunostaining was performed by comparison to species-matched isotype antibodies and unstained controls

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Male C57BL/6JCrl mice aged 8 to 10 weeks
Wild animals	Study did not involve wild animals
Field-collected samples	Study did not involve samples collected in the field
Ethics oversight	All experiments were performed in accordance with UK Home Office regulations.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Please see Extended Data Figure 1a for the clinical characteristics of patients used for single-cell RNA sequencing.
Recruitment	Patients were recruited as described in Methods. Healthy background non-lesional liver tissue was obtained intraoperatively from patients undergoing surgical liver resection for solitary colorectal metastasis at the Hepatobiliary and Pancreatic Unit, Department of Clinical Surgery, Royal Infirmary of Edinburgh. Patients with a known history of chronic liver disease, abnormal liver function tests or those who had received systemic chemotherapy within the last four months were excluded from this cohort. Cirrhotic liver tissue was obtained intraoperatively from patients undergoing orthotopic liver transplantation at the Scottish Liver Transplant Unit, Royal Infirmary of Edinburgh. Blood from patients with a confirmed diagnosis of liver cirrhosis were obtained from patients attending the Scottish Liver Transplant Unit, Royal Infirmary of Edinburgh. Patients with liver cirrhosis due to viral hepatitis were excluded from the study. For cell sorting of macrophages or isolation of human endothelial cells, liver tissue was acquired from explanted diseased livers from patients undergoing orthotopic liver transplantation, resected liver specimens or donor livers rejected for transplant at the Queen Elizabeth Hospital, Birmingham.
Ethics oversight	NRS BioResource and Tissue Governance Unit (Study Number SR574), following review at the East of Scotland Research Ethics Service (Reference 15/ES/0094) For University of Birmingham samples, separate local ethical approval was obtained (Reference 06/Q2708/11, 06/Q2702/61).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Please see Methods for detailed sample preparation protocol for FACS and flow cytometry
Instrument	BD Influx and BD FACSAriaII were used for cell sorting at University of Edinburgh. BD LSR Fortessa was used for flow cytometry analysis. BD FACSAria Fusion for cell sorting at the University of Birmingham
Software	BD FACS Sortware software was used for cell sorting on BD Influx equipment. BD FACS Diva software for cell sorting on BD FACSAriaII and BD FACSAria Fusion. BD FACS Diva software was used for flow cytometry on BD LSR Fortessa equipment. Flow

cytometry analysis was performed using FlowJo software (version 10.2).

Cell population abundance

Sort purity was routinely over 95% on post-sort checks

Gating strategy

Please see Methods and Extended Data Figures 1b,c, 4f and 6e for gating strategies. Initial gating for all experiments: Cells (FSC-A vs SSC-A), Singlets (FSC-A vs FSC-H (or FSC-A vs TPW for BD Influx)), Viable (SSC-A vs viability dye (See methods)). For human PBMC sort, CD45+ CD66b- cells were sorted. For human liver single-cell RNA-seq sorting, CD45+ cells (leukocytes) or CD45- cells (other NPCs) were sorted. For human liver macrophage flow cytometry quantification and cell sorting, tissue monocyte-macrophages were identified as CD45+, Lin- (CD3, CD335, CD19, CD66b, LILRA4, CD326), HLA-DR+, CD1C-, CD14+ and/or CD16+ cells. SAM were then identified as CD163- TREM2+ CD9+, KCs were identified as CD163+ CD9- and TMo were identified as CD163-. For mouse liver single-cell RNA-seq sorting, tissue mononuclear phagocytes identified as CD45+ Lin-(CD3, NK1.1, Ly6G, CD19) cells were sorted. For mouse liver macrophage cell sorting, CD45+ Lin- CD11b+ F4/80+ TIMD4- CD9+ or CD9- cells were sorted from CCl4-treated mice. For human liver endothelial cell flow cytometry, cultured endothelial cells were stained with antibodies to PLVAP, ACKR1, JAG1 and CD34. Gates and boundaries were defined by comparison to FMO and unstained samples.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.