Single-Cell Transcriptomics Uncovers Zonation of Function in the Mesenchyme during Liver Fibrosis

**Highlights**

- scRNA-seq reveals spatial zonation of hepatic stellate cells (HSCs)
- HSCs partition into topographically diametric lobule regions
- Functional zonation of HSCs during centrilobular injury-induced fibrosis is uncovered
- LPAR1 is a therapeutic target on pathological central vein-associated HSC

**Authors**

Ross Dobie, John R. Wilson-Kanamori, Beth E.P. Henderson, ..., Sarah A. Teichmann, John C. Marioni, Neil C. Henderson

**Correspondence**

neil.henderson@ed.ac.uk

**In Brief**

Dobie et al. use scRNA-seq to reveal spatial and functional zonation of hepatic stellate cells (HSCs) across the hepatic lobule, identifying central vein-associated HSCs as the dominant pathogenic collagen-producing cells during centrilobular injury-induced fibrosis. This illustrates the power of scRNA-seq to resolve the key collagen-producing cells driving liver fibrosis.
Single-Cell Transcriptomics Uncovers Zonation of Function in the Mesenchyme during Liver Fibrosis


1Centre for Inflammation Research, The Queen’s Medical Research Institute, Edinburgh BioQuarter, University of Edinburgh, Edinburgh EH16 4TJ, UK
2Karolinska Institutet (KI), Science for Life Laboratory, Tomtebodavägen 23, Solna 171 65, Sweden
3Gilead Sciences, Foster City, CA 94404, USA
4Clinical Surgery, University of Edinburgh, Royal Infirmary of Edinburgh, Edinburgh EH16 4SA, UK
5MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine at the University of Edinburgh, Edinburgh EH4 2XU, UK
6Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK
7European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, Cambridge CB10 1SD, UK
8Theory of Condensed Matter Group, The Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK
9Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, UK
10These authors contributed equally
11Lead Contact
*Correspondence: neil.henderson@ed.ac.uk

SUMMARY

Iterative liver injury results in progressive fibrosis disrupting hepatic architecture, regeneration potential, and liver function. Hepatic stellate cells (HSCs) are a major source of pathological matrix during fibrosis and are thought to be a functionally homogeneous population. Here, we use single-cell RNA sequencing to deconvolute the hepatic mesenchyme in healthy and fibrotic mouse liver, revealing spatial zonation of HSCs across the hepatic lobule. Furthermore, we show that HSCs partition into topographically diametric lobule regions, designated portal vein-associated HSCs (PaHSCs) and central vein-associated HSCs (CaHSCs). Importantly we uncover functional zonation, identifying CaHSCs as the dominant pathogenic collagen-producing cells in a mouse model of centrilobular fibrosis. Finally, we identify LPAR1 as a therapeutic target on collagen-producing CaHSCs, demonstrating that blockade of LPAR1 inhibits liver fibrosis in a rodent NASH model. Taken together, our work illustrates the power of single-cell transcriptomics to resolve the key collagen-producing cells driving liver fibrosis with high precision.

INTRODUCTION

Liver cirrhosis is a major global healthcare burden, with an estimated 844 million people suffering from chronic liver disease worldwide (Marcellin and Kutala, 2018). Mortality rates secondary to liver cirrhosis continue to increase, with no Food and Drug Administration (FDA)- or European Medicines Agency (EMA)-approved antifibrotic treatments currently available, and liver transplantation only accessible to a select few (Friedman et al., 2018; Koyama et al., 2016; Tapper and Parikh, 2018). An ideal antifibrotic therapy would specifically target the pathogenic collagen-producing cell population without perturbing homeostatic mesenchymal function. Therefore, increasing our understanding of the precise cellular and molecular mechanisms regulating liver fibrosis is fundamental to the rational design and development of effective, highly targeted anti-fibrotic therapies for patients with chronic liver disease (Ramachandran and Henderson, 2016; Trautwein et al., 2015).

Myofibroblasts are the key source of pathogenic extracellular matrix deposition during hepatic fibrogenesis and therefore have attracted considerable interest as a potential therapeutic target (Dobie and Henderson, 2016; Friedman, 2015; Hinz et al., 2012; Kisselova, 2017). Although different mesenchymal cell types have been proposed as the predominant source of myofibroblasts following liver injury (Iwaisako et al., 2014; Kisselova et al., 2006; Li et al., 2013; Mederacke et al., 2013), recent studies suggest that hepatic stellate cells (HSCs), first described by Kupffer in 1876 as vitamin A+ lipid droplet-containing cells that reside in the perisinusoidal space of the liver (Wake, 1971), are the dominant contributors to the myofibroblast pool independent of the etiology of liver fibrosis (Iwaisako et al., 2014; Mederacke et al., 2013). Furthermore, since the discovery 35 years ago that HSCs are major collagen-producing cells in the liver (Friedman et al., 1985; de Leeuw et al., 1984), these cells have been regarded as a functionally homogeneous population, with
Figure 1. Deconvolution of the Mouse Hepatic Mesenchyme Identifies Three Distinct Subpopulations in Liver Homeostasis

(A) Overview: representative immunofluorescence image depicts GFP reporting in the liver of healthy Pdgfrb-BAC-eGFP reporter mice. Scale bar, 100 μm; portal vein (*) as indicated. CV, central vein; PV, portal vein. GFP+ cells were processed for droplet- and plate-based scRNA-seq.

(B) Subpopulation

- Fibroblast (FB)
- Hepatic stellate cell (HSC)
- Vascular smooth muscle cell (VSMC)

(C) Mouse

CD34 = FB, Reelin = HSC, Calponin 1 = VSMC

(D) Human

MFAP4 = FB, RG5 = HSC, MYH11 = VSMC

(E) HSC

Gene Ontology

- Extracellular structure organization
- Extracellular matrix organization
- Connective tissue development
- Collagen fibril organization

(F) VSMC

Gene Ontology

- Regulated lipid localization
- Regulation of actin filament-based process
- Muscle structure development
- Muscle system process

(legend continued on next page)
the potential to transition to the activated, collagen-secreting myofibroblast phenotype thought to be equally distributed across all HSCs.

Single-cell RNA sequencing (scRNA-seq) is transforming our understanding of disease pathogenesis (Lee et al., 2017; Stubbington et al., 2017; Zepp et al., 2017). Here, we use scRNA-seq to resolve the hepatic mesenchyme in an unbiased manner at high resolution, analyzing the transcriptomes of over 30,000 hepatic mesenchymal cells. Our data: (1) deconvolve the hepatic mesenchyme in healthy and fibrotic mouse liver; (2) reveal spatial zonation of HSCs across the hepatic lobule; (3) generate gene signatures and markers that partition HSCs into two topographically diaphragmatically lobule regions, namely portal vein-associated HSCs (PaHSCs) and central vein-associated HSCs (CaHSCs); (4) importantly, uncover functional zonation of HSCs, identifying that CaHSCs, but not PaHSCs, are the dominant pathogenic collagen-producing cells in a mouse model of centrilobular liver injury; and (5) identify LPAR1 as a therapeutic target on collagen-producing HSCs and demonstrate that pharmacological antagonism of LPAR1 inhibits liver fibrosis. These studies allow us to further define and resolve the spatial, cellular, and molecular complexity present within the hepatic fibrotic niche. Our work highlights the power of scRNA-seq in identifying the key collagen-producing cells driving centrilobular liver fibrosis with high precision and therefore should serve as a framework for the high-resolution identification of the critical pathogenic cells and related therapeutic targets in a broad range of fibrotic diseases.

RESULTS

Deconvolution of the Mouse Hepatic Mesenchyme Identifies Three Distinct Subpopulations in Liver Homeostasis

We used a Pdgfrb-GFP knockin reporter mouse to label all mesenchymal cells in the mouse liver (Figure 1A). This reporter strain has previously been shown to label all mesenchymal cells (including HSCs) (Henderson et al., 2013). Here, we show that the Pdgfrb-GFP mouse labeled PDGFRβ+ cells in liver with high efficiency and specificity (Figure S1A). Two independent digestion protocols and gating strategies were used to isolate the different GFP+ mesenchymal cell populations (Figure S1B; STAR Methods). To initially characterize the hepatic mesenchyme at single-cell resolution, we used the 10X Chromium protocol to sequence 12,533 cells from mice (n = 3 digestion protocol 1; n = 3 digestion protocol 2) at a mean read depth of ~85 K reads per cell, which show negligible endothelial, epithelial, and leucocyte contamination (Figures 1A, S1C, and S1D).

We identified three subpopulations of mesenchymal (PDGFRβ+) cells (Figure 1B) with distinct sets of marker genes (Figure S2A; Table S1). Identifying highly specific marker genes (Figure S2B) and performing immunofluorescence co-staining (Figure 1C) validated the three mesenchymal subpopulations and delineated their topography. We found that CD34+ PDGFRβ+ cells reside primarily in the portal niche, adjacent to PanCK+ biliary epithelial cells, with rare cells also found around the central vein, possibly representing second layer cells (Figure S2C) (Bhunchet and Wake, 1992). Reelin+ PDGFRβ+ cells were found in the perisinusoidal space throughout the parenchyma (Figure S2D). Calponin 1+ PDGFRβ+ cells were located within both the hepatic artery and the portal vein walls (Figure S2E). Given the topographic distribution of these three mesenchymal subpopulations, we labeled them as fibroblasts (FBs), HSCs, and vascular smooth muscle cells (VSMCs), respectively (Figure 1D). We annotated the CD34+ mesenchymal subpopulation as FBs, which are known to be major mediators of matrix turnover in the portal niche (Wells, 2014). The Reelin+ subpopulation represents HSCs, located throughout the parenchyma, with functions including vitamin A storage and antigen presentation (Friedman, 2008; Winau et al., 2007). Finally, the topography of the Calponin 1+ subpopulation is consistent with a VSMC phenotype (Patel et al., 2016).

To determine whether similar mesenchymal subpopulations exist in healthy human liver, we performed immunofluorescence co-staining using genes identified as markers in the mouse dataset (Figure 1E). Akin to our findings in mouse liver, we identified three topographically distinct mesenchymal subpopulations. MFAP4+ PDGFRβ+ cells were confined to the portal niche, consistent with the CD34+ PDGFRβ+ subpopulation observed in the portal niche of mouse liver. RGS5+ PDGFRβ+ cells were found in locations throughout the hepatic parenchyma, consistent with HSCs. MYH11+ PDGFRβ+ cells were located around portal vein walls, consistent with VSMCs.

To assess the functional profile of the three mouse mesenchymal subpopulations, we generated self-organizing maps using the SCRAT R package (Camp et al., 2017) to visualize coordinately expressed gene groups across the transcriptomic landscape (Figure S2F). We identified three metagene signatures, denoted as A–C, that strongly define the subpopulations (Table S2). Signature A, enriched for gene ontology (GO) terms relating to extracellular structure organization, defined both FBs and VSMCs mesenchymal subpopulations. Signature B defined the HSCs subpopulation and was enriched for terms including retinoid metabolic process and antigen processing and presentation. Signature C defined
VSMCs exclusively and was enriched for terms such as actin filament-based processes (Figures 1F and S2F).

Using a single-cell approach also allowed us to interrogate "traditional" hepatic mesenchymal markers at high resolution. We found that certain historic HSC markers, such as Des and Vim, were expressed variably across all three mesenchymal subpopulations (Figure S2G). In keeping with recent findings (Mederacke et al., 2013), Gltap expression was negligible in our dataset. We confirmed Lrat and Rein as specific markers for HSCs within the hepatic mesenchyme (Lua et al., 2016; Mederacke et al., 2013), and Ngfr displayed a spectrum of expression across the HSC population. Pdgfra expression was confined to the FB and HSC subpopulations as opposed to Pdgfrb, which was pan-mesenchymal.

To reproduce our 10X Chromium-based findings, and to assess whether a plate-based full-length transcript approach would identify similar mesenchymal subpopulations, we also obtained scRNA-seq data of PDGFRα+ cells in liver using SmartSeq2 (SSeq2). We sequenced 905 cells isolated using both digestion protocols at a mean read depth of ~456 K reads per cell. Analysis of this SSeq2 dataset identified the same three mesenchymal subpopulations (Figure S3A; Table S1) with negligible non-mesenchymal cell contamination (Figure S3B). This alternative sequencing approach replicated our findings both in terms of the marker genes identified previously and the GO profiles generated using SCRAT (Figures S3C–S3G; Table S2).

Uncovering HSC Zonation across the Healthy Liver Lobule

The micro-architecture of the hepatic lobule displays highly ordered three-dimensional structural motifs consisting of a portal triad, hepatocytes arranged in linear cords between a sinusoidal capillary network, and a central vein and is highly conserved across species (Burke et al., 2009; Gebhardt, 1992; Kietzmann, 2017). Given the known zonation of hepatocytes (Halpern et al., 2017; Lamers et al., 1989) and endothelia (Halpern et al., 2018) across the liver lobule, and having observed variable patterns of gene expression in the HSC population (Figure S2G), we investigated the existence of similar zonation in HSCs. We used independent component analysis (ICA) to identify a set of highly variable genes in uninjured HSCs (Figure 2A). Thresholding on the gene weight loadings along this component, we extracted 81 genes consisting of two opposed signatures: 52 genes associated with and including Ngfr and Itgb3 and 29 genes associated with and including Adamtsl2 and Rspo3 (Figure 2A; Table S1). Supervised clustering using this signature allowed us to separate the HSCs into two further subpopulations.

To determine the topography of these subpopulations, we selected marker genes using the following criteria: (1) high gene weight loading on the independent component (IC) (Figure 2A), (2) mesenchyme specificity (Figure S4A), and (3) within the mesenchyme, greatest specificity to HSCs (Figure S4B). We identified Adamtsl2 and Ngfr as the best candidate markers. Using a combination of a highly sensitive modified in situ RNA hybridization procedure (RNAscope) and immunofluorescence staining, we confirmed Adamtsl2 and NGFR to be mesenchymal markers each labeling a subpopulation of HSCs (Figure S4A). To explore how this topography related to zonally distributed hepatocytes, we co-stained with peri-portal (E-cadherin) and pericentral (Cyp2e1) hepatocyte markers (Doi et al., 2007; Rocha et al., 2015). We found NGFRhi HSCs located in the same region of the liver lobule as portal vein-associated (E-cadherin+) hepatocytes, whereas Adamtsl2hi HSC was located in the same region of the liver lobule as central vein-associated hepatocytes (Cyp2e1+) (Figures 2B and S4C). This allowed us to annotate these two HSC subpopulations as NGFRhi portal vein-associated HSCs (PaHSCs) and Adamtsl2hi central vein-associated HSCs (CaHSCs) (Figure 2C).

To determine whether healthy human liver exhibits similar HSC zonation, we performed immunofluorescence co-staining using marker genes orthologous to those that delineated mouse HSC zonation—Ngfr and Adamtsl2—with the previously identified HSC marker RGS5 (Figure S4D). Akin to our findings in mouse liver, we observed zonal expression of NGFR and ADAMTS12 across the human liver lobule (Figure 2D).

Certain genes within our zonation signature, such as Itgb3 and Rspo3, have previously been identified as landmark genes used to zonally define endothelial cells across the liver lobule (Halpern et al., 2018). Spatial mapping of these populations using RNA-seq and immunofluorescence staining confirmed the presence of zonally distributed Itgb3 (Integrin β3) and Rspo3 HSCs within the parenchyma (Figure S5A). In line with previous studies, Integrin β3 HSCs were observed in the peri-portal region and Rspo3 HSCs in the peri-central region, suggesting possible spatial correlation between the endothelial and mesenchymal lineages (Figures S5A and S5B). Immunofluorescence staining of PaHSC marker NGFR and previously identified central-associated endothelial cell marker thrombomodulin (Halpern...
et al., 2018) further demonstrated zonation, with each marker defining a distinct region within the hepatic lobule (Figure S5C).

**HSC Populate the Fibrotic Niche in a Mouse Model of Centrilobular Fibrotic Liver Injury**

Chronic carbon tetrachloride (CCl4) administration is a broadly utilized, highly reproducible, and robust mouse model of centrilobular liver fibrosis that recapitulates many of the features of human fibrotic liver disease (Figures S6A and S6B). To investigate mesenchymal cell heterogeneity in fibrotic mouse liver, we sequenced 10,758 Pdgfrb–GFP⁺ reporter cells from murine liver following 6 weeks CCL4 administration using the 10X Chromium workflow and performed unsupervised clustering alongside our homeostatic hepatic mesenchyme dataset (Figures S6C and S6D).

We observed the same three mesenchymal subpopulations in both healthy and fibrotic livers (Figures 3A and 3B), with previously identified markers maintaining their specificity following chronic liver injury (Table S1). Expression of fibrillar collagens (Col1a1, Col1a2, and Col3a1) remained highest in FBs, however, a marked increase in expression of Col1a1 was observed in HSCs following induction of fibrotic injury (Figure 3C). Although overall marker gene expression profiles remained constant between uninjured and fibrotic HSCs compared to other mesenchymal populations, we also observed decreased expression of certain marker genes including Reln in HSC overexpressing fibrillar collagen (Figure S6E). Immunofluorescence staining confirmed diminished Reelin positivity within the fibrotic niche (Figure S6F). In contrast, staining for Lhx2 (a mesenchyme–specific marker of HSCs in both uninjured and 6 weeks CCL4 liver) confirmed an expansion of HSCs within the fibrotic niche (Figures 3D, S6F, and S6G). Previous studies have shown that HSCs are the major source of pathogenic collagen-producing cells following liver injury (Iwaisako et al., 2014; Mederacke et al., 2013); in accordance we did not identify FBs and VSMCs within the fibrotic niche, as evidenced by immunofluorescence staining for markers Mfap4 and Calponin 1 (Figure 3E). We observed minimal proliferation across all three populations (2.4% of mesenchymal cells in the dataset expressed proliferation marker Mki67, confirmed by ethynyldeoxyuridine (EdU) staining in CCL4-treated livers (2.6% of mesenchymal cells; Figure 3F).

**CaHSCs Are the Dominant Pathogenic Collagen-Producing Cells in a Mouse Model of Centrilobular Fibrotic Liver Injury**

Since the discovery 35 years ago that HSCs are major collagen-producing cells in the liver (Friedman et al., 1985; de Leeuw et al., 1984), they have been regarded as a functionally homogeneous population, with the potential to transition to the activated, collagen-secreting myofibroblast phenotype thought to be equally distributed across all HSCs. Having identified HSC zonation in the homeostatic liver, we therefore investigated the relative contributions of PaHSCs and CaHSCs to the fibrotic process. We performed supervised clustering on the combined healthy and fibrotic HSC populations based on their expression of the 81 zonation genes previously identified in homeostatic HSCs (Table S1) and found clear conservation of the zonation genes observed in homeostasis: 51/52 PaHSC-associated genes including Ngfr and Itgb3 and 26/29 CaHSC-associated genes including Adamts12 and Rspo3 continued to define the zonation profile of the combined HSCs cluster (Figure 4A). This again allowed delineation of HSCs into PaHSC and CaHSC subpopulations (Figure 4B).

To investigate whether our zonation signature correlated with pathogenic collagen-producing HSCs, we created a myofibroblast signature by thresholding HSCs on fibrillar collagen (Col1a1, Col1a2, and Col3a1) overexpression, a hallmark of HSC activation. This signature was then mapped onto our zonation profile to identify the contribution of PaHSCs and CaHSCs to the pathogenic myofibroblasts. In this way, we identify CaHSCs as the dominant pathogenic collagen-producing HSCs, representing 88% of myofibroblasts classified in this manner (Figure 4C). These findings were validated by spatial mapping of CaHSCs and PaHSCs in fibrotic liver using RNA-seq and immunofluorescence staining. Adamts12 CaHSCs were located throughout the fibrotic niche whereas Ngfr PaHSCs resided predominately in the parenchyma distal to the fibrotic region. (Figures 4D and S7A). A similar difference in topography was observed with other zonation markers Rspo3 and Integrin β3 (Figure S7B).

To further investigate the contribution of CaHSCs to pathogenic collagen production during fibrosis, we used self-organizing maps to identify metagene signatures enriched in the collagen-producing HSCs (Figure S7C; Table S2). Signature A...
was expressed across both PaHSCs and CaHSCs, with enriched GO terms such as retinoid metabolic process that also defined uninjured HSCs (Figure 1F). Signature B was highly enriched for terms related to collagen production, and all HSCs that expressed signature B were CaHSCs.

Having identified CaHSCs as the predominant pathogenic collagen-producing HSCs, we used the monocle R package to further investigate changes in gene expression within the CaHSCs. This highlighted the transition from a quiescent to a collagen-producing phenotype, with upregulation following injury of pro-fibrogenic genes including Col1a1, Col1a2, Col3a1, and Lox, with associated enriched GO terms such as extracellular structure organization and collagen fibril organization, and downregulation of uninjured HSC-related genes including Ecm1, Reln, Hgf, and Rgs5 (Figures 4E and 4F; Table S3).

CaHSCs Are the Dominant Pathogenic Collagen-Producing Cells following Acute Centrilobular Liver Injury
Acute CCl₄-induced liver injury is characterized by significant HSC proliferation (14.4% of Lhx2⁺ HSCs) and activation to a collagen-producing myofibroblast phenotype in the centrilobular region (Figures S8A–S8C). To further interrogate the dynamics of PaHSC and CaHSC differentiation into pathogenic collagen-producing cells, we sequenced 7,260 HSCs from uninjured and fibrotic (6 weeks CCl₄) livers on zonation signature separates them into distinct PaHSC and CaHSC subpopulations (Figures 5A and S8F). Both subpopulations contained cells expressing the known proliferation marker Mki67, however, only CaHSCs showed elevated levels of known genes associated with fibrogenesis (Figures 5B and 5C). Spatial mapping of these subpopulations using RNAscope combined with immunofluorescence staining confirmed the pathogenic collagen-producing cells in the fibrotic niche as overwhelmingly Adamtsl2⁺ CaHSCs (Figures 5D and S8G).

CaHSCs, but Not PaHSCs, Differentiate into Pathogenic Collagen-Producing Cells following Acute Centrilobular Liver Injury
We used the velocity R package (La Manno et al., 2018) to interrogate the HSC injury response by calculating cellular velocity from spliced and unspliced mRNA content. We found that the likelihood of transition between CaHSCs and PaHSCs was negligible, thus inferring absence of pseudotemporal dynamics between the two subpopulations (Figure S9A). Furthermore, fibrogenic genes such as Col1a1, Col3a1, and Acta2 display positive residuals (unspliced/spliced mRNA ratio) for CaHSCs but not for PaHSCs, reinforcing their potential for myofibroblast transition (Figure S9B). This demonstrated that HSC differentiation into pathogenic collagen-producing cells occurred in CaHSCs, but not in PaHSCs (Figure 6A). We thus used the monocle R package to independently define trajectories for CaHSCs and PaHSCs (Figure 6B). CaHSCs exhibited a branching trajectory (Figure 6B), with one branch transitioning from quiescence into a collagen-producing phenotype with upregulation of pro-fibrogenic genes including Col1a1, Col1a2, Col3a1, and Acta2, and the other branch displaying a primarily proliferative response (Figures 6B–6D; Table S3). PaHSCs displayed a proliferative response but did not transition to collagen-producing cells (Figures 6B, 6E, and 6F; Table S3). Immunofluorescence co-staining verified similar levels of proliferation between PaHSCs and CaHSCs following acute CCl₄-induced liver injury (Figure 6G).

Having identified CaHSCs as the pathogenic collagen-producing cells, we used the SCENIC R package (Albar et al., 2017) to provide mechanistic insight into the transcriptional regulation of HSC activation following both acute and chronic CCl₄-induced liver injury. SCENIC identifies sets of genes that co-express with known transcription factors and are differentially expressed along the HSC activation trajectories. We observed 50 such regulons in the acute activation trajectory and 29 in the chronic trajectory of which 18 were shared between both, including Egr2, Sox4, Plagl1, Rxra, Foxf1, and Kif7 (Mann and Smart, 2002; Vollmann et al., 2017) (Figures S9C and S9D; Table S3). As both fibroblasts and HSC-derived myofibroblasts are responsible for collagen deposition, we were keen to identify unique regulatory elements for the latter. We identified transcription factors Sox4 and Rxra as specific to HSCs following chronic CCl₄-induced liver injury (Figure S9E). Again focusing on potential regulatory target genes conserved between the two injury models, SCENIC uncovers 6 genes associated with Rxra and 5 genes associated with Sox4 (Figure S9F). Genes associated with the Sox4 regulon include Mdk and Hmcn1 that have previously been shown to...
Targeting of LPAR1 on Collagen-Producing HSCs Inhibits Liver Fibrosis

The Lpar1 gene encodes lysophosphatidic acid receptor 1 (LPAR1), a G protein-coupled receptor that binds the lipid signaling molecule lysophosphatidic acid (LPA). Previous studies have shown LPAR1 expression to be restricted to non-parenchymal cells and elevated in activated HSCs, with minimal expression reported in other hepatic lineages, including hepatocytes (Nakagawa et al., 2016; Simo et al., 2014). Having identified CaHSC as the predominant pathogenic collagen-producing cell during CCl4-induced centrilobular murine liver fibrosis, we identified that Lpar1 was expressed in CaHSC but not PaHSC following acute and chronic CCl4-induced liver injury (Figure 7A). Lpar1 was not expressed in hepatic leucocytes and endothelial cells following chronic CCl4-induced liver injury (Figure 7B).

To investigate whether LPAR1 is expressed on pathogenic collagen-producing cells in human liver cirrhosis, we interrogated the hepatic mesenchyme at single-cell resolution in a previously published dataset (Figure 7C) (Ramachandran et al., 2019). Unsupervised clustering identified three mesenchymal subpopulations, including a subpopulation which expands in cirrhosis and is defined by upregulation of fibrillar collagen expression, referred to as myofibroblasts (Figures 7C and 7D). Akin to our observations in murine liver injury, LPAR1 expression was restricted to the collagen-producing subpopulation (Figure 7D).
Previous studies have revealed a role for LPA signaling in HSC activation (Yanase et al., 2000), and LPAR1 antagonism has been shown to reduce fibrosis in a rodent model of hepatocellular carcinoma and mouse model of centrilobular fibrosis (Bollong et al., 2017; Nakagawa et al., 2016). As proof-of-concept of our scRNA-seq approach, we investigated the effects of LPAR1 antagonism on human HSC contractility and activation in vitro. LPA, a known ligand of LPAR1, is a major driver of actin polymerization and actomyosin contraction in HSCs (Yanase et al., 2000). LPAR1 antagonism inhibited LPA-induced HSC contractility, pro-fibrogenic connective tissue growth factor (CTGF) expression, F-actin polymerization, and phosphorylation of myosin light chain 2 (Figure S10).

The prevalence of non-alcoholic steatohepatitis (NASH) as a leading cause of chronic liver disease has reached epidemic proportions (Friedman et al., 2018). To assess whether LPAR1 expression increases during the evolution of human NASH-induced liver fibrosis, we performed bulk RNA-seq of human liver samples from a cohort of biopsy-confirmed NASH patients with a range of fibrosis stages (F1–F4). LPAR1 expression increased with fibrosis stage (Figure 7E) and correlated with fibrillar collagen expression (Figure 7F). To further assess whether LPAR1 antagonism inhibits liver fibrosis in vivo, we used a 12-week choline-deficient high-fat diet (CDHFD) rodent model of NASH. LPAR1 antagonism markedly reduced liver fibrosis as measured by digital morphometry of picrosirius red and αSMA (Figure 7G) and hydroxyproline assay (Figure 7H).

DISCUSSION

Despite significant advances in our understanding of the cellular and molecular mechanisms driving liver fibrosis over the past 40 years, there are still no FDA- or EMA-approved antifibrotic treatments currently available. Therefore, there remains a clear imperative to further resolve and understand the complex mechanisms that regulate the fibrotic niche, both in the liver and other organs.

scRNA-seq has facilitated the interrogation of mesenchymal heterogeneity at unprecedented resolution and has greatly advanced our understanding of mesenchymal cell biology and function in disease pathogenesis across various tissues (Croft et al., 2019; Guerrero-Juarez et al., 2019; Peyser et al., 2019; Xie et al., 2018). Investigating specific mesenchymal populations in normal and fibrotic liver has been hampered by the lack of reliable markers required to distinguish these subpopulations (Wells, 2014). In this study, we use a scRNA-seq approach to deconvolve for the first time the entire hepatic mesenchyme in healthy and fibrotic mouse liver, identifying three distinct populations of mesenchymal cells. We make this data freely available to browse at http://livermesenchyme.hendersonlab.mvm.ed.ac.uk, where it should serve as a useful reference resource for future studies of the hepatic mesenchyme.

Multiple mesenchymal cell types have been proposed as the major source of myofibroblasts following liver injury (Iwaisako et al., 2014; Kisseleva et al., 2006; Li et al., 2013; Mederacke et al., 2013), however, recent studies suggest that HSCs are the predominant contributors to the myofibroblast pool irrespectively of the cause of liver fibrosis (Iwaisako et al., 2014; Mederacke et al., 2013). Our data confirm HSCs as the dominant mesenchymal contributor to pathogenic collagen production in CCl4-induced centrilobular injury.

Zonation across the homeostatic liver lobule has recently been characterized in hepatocytes and endothelial cells using scRNA-seq approaches (Halpern et al., 2017, 2018), however, zonation of function in the context of a fibrotic injury response has not previously been documented in the liver. HSC zonation has previously been described in porcine liver (Wake and Sato, 1993), and a recent scRNA-seq study in mice concluded that HSCs from healthy liver are a transcriptionally homogeneous population (Krenkel et al., 2019). Furthermore, since the discovery 35 years ago that HSCs are major collagen-producing cells in the liver (Friedman et al., 1985; de Leeuw et al., 1984), their potential to transition to the activated, collagen-secreting myofibroblast phenotype has been thought to be equally distributed across the entire population. In this study, we used scRNA-seq to uncover heterogeneity within the mesenchyme, including zonation of HSCs across the hepatic lobule. We show that HSCs partition into two topographically distinct regions, designated portal vein-associated HSCs (PaHSCs) and central vein-associated HSCs (CaHSCs).

Importantly, we also uncover zonation of function in HSCs, with CaHSCs, but not PaHSCs, responsible for the vast majority of pathogenic fibrillar collagen-production in the CCl4 mouse model of centrilobular liver injury. The zonal activation of...
Figure 7. Targeting of LPAR1 on Collagen-Producing HSCs Inhibits Liver Fibrosis

(A) Violin plots: expression of fibrillar collagen genes (Col1a1, Col1a2, and Col3a1) and Lpar1 in PaHSCs versus CaHSCs following acute (72 h post single CCl4 injection; top) and chronic (6 weeks CCl4; bottom) liver injury.

(B) Violin plot: expression of Lpar1 in hepatic mesenchymal cells (Mes), endothelial cells (Endo), and leucocytes (Leuc) from chronic liver.

(legend continued on next page)
CaHSCs following induction of CCl₄ is likely to be secondary to the topographical location of injury, with CCl₄ causing necrosis of hepatocytes in the centrlobular region (Tanaka and Miyajima, 2016). It is possible that PaHSCs may represent the major collagen-producing HSCs in peri-portal injury models, however, the relative functional roles of PaHSCs and portal fibroblasts in the context of biliary injury requires further investigation (Wells, 2014).

Together, this study provides a high-resolution examination of the hepatic fibrotic niche, via a comprehensive analysis and partitioning of all the hepatic mesenchymal lineages, and investigation of their relative contributions to the fibrogenic process. This scRNA-seq approach has clear implications for the rational development of antifibrotic therapies; facilitating and informing specific targeting of pathogenic scar-forming cells without perturbing homeostatic mesenchymal function, which is of particular importance in patients with chronic liver disease who may already have very limited hepatic functional reserve.

In an era of precision medicine, where molecular profiling underpins the development of highly targeted therapies, we used scRNA-seq to resolve the healthy and fibrotic hepatic mesenchyme in high-definition. Our work illustrates the power of single-cell transcriptomics to identify the key collagen-producing cells driving liver fibrosis with high precision and should also serve as a framework for the high-resolution identification of pathogenic cells and related therapeutic targets in a broad range of fibrotic diseases.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **LEAD CONTACT AND MATERIALS AVAILABILITY**
- **EXPERIMENTAL MODELS AND SUBJECT DETAILS**
  - Mice
  - Rats
  - Human tissue
  - Fibrosis Models
  - Primary cells and cell lines
- **METHOD DETAILS**
  - Immunofluorescence Staining
  - RNAscope
  - EdU Click-iT Immunofluorescence staining
  - Picosirius Red Staining
  - Image Quantification
  - Hydroxyproline Assay
  - Immunocytochemistry
  - Contraction Assay
  - qPCR
  - Primary cell isolation
  - Cell sorting
  - Human liver tissue RNA-seq
  - Single-cell workflows
  - Pre-processing scRNA-seq data
  - Dimensionality reduction, clustering, and DE analysis
  - Defining cell expression signatures
  - Identifying and applying an HSC zonation signature
  - Analyzing functional phenotypes of mesenchymal cells
  - Inferring injury dynamics and transcriptional regulation

**QUANTIFICATION AND STATISTICAL ANALYSIS**

- **DATA AND CODE AVAILABILITY**
  - Additional Resources

**SUPPLEMENTAL INFORMATION**

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.10.024.

**ACKNOWLEDGMENTS**

We thank the Eukaryotic Single-Cell Genomics facility at SciLifeLab Stockholm for generating Smart-Seq2 single-cell sequencing libraries. We thank C. Betsholtz (Uppsala University and Karolinska Institutet) for providing PDGFRα-EGFP-BAC mice. We thank the patients who donated liver tissue and blood for this study. We thank J. Davidson, C. Ibbotson, 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 Zhi-Hua Cui, David Newstrow, and Shiva Zaboli (Gilead) for support with the rodent choline-deficient high-fat diet study and David Lopez (Gilead) for support with human FFPE RNA-seq data analysis. We thank J. Henderson for technical support and S. Johnston, W. Ramsay, and M. Patterson (QMRI Flow Cytometry and Cell Sorting Facility, University of Edinburgh) for technical assistance with fluorescence activated cell sorting (FACS). This work was supported by a Welcome Trust Senior Research Fellowship in Clinical Science (103749 to N.C.H.), a Guts UK-Children’s Liver Disease Foundation grant (R43927 to N.C.H.), and a Tenevus Scotland grant (E18/05 to R.D. and N.C.H.). P.R. was supported by an MRC Clinician Scientist Fellowship.


References


enzymic zonation: a reevaluation of the concept of the liver acinus. Hepatology 70, 72–76.


# STAR★METHODS

## KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-Adamts12</td>
<td>Abcam</td>
<td>Cat# ab97603; RRID:AB_10696110</td>
</tr>
<tr>
<td>Anti-Actin, alpha-Smooth Muscle</td>
<td>Biocare</td>
<td>Cat# CM001 [1A1]</td>
</tr>
<tr>
<td>Anti-CD31 Platelet Endothelial Cell Adhesion Molecule (PECAM)</td>
<td>BD Biosciences</td>
<td>Cat# 550274; RRID:AB_393571</td>
</tr>
<tr>
<td>Anti-CD34</td>
<td>Abcam</td>
<td>Cat# ab81289; RRID:AB_1640331</td>
</tr>
<tr>
<td>Anti-Cytokeratin 19</td>
<td>Abcam</td>
<td>Cat# ab220193</td>
</tr>
<tr>
<td>Anti-Type I Collagen</td>
<td>SouthernBiotech</td>
<td>Cat# 1310-01; RRID:AB_2753206</td>
</tr>
<tr>
<td>Anti-Type III Collagen</td>
<td>SouthernBiotech</td>
<td>Cat# 1330-01; RRID:AB_2794734</td>
</tr>
<tr>
<td>Anti-Cyp2e1</td>
<td>Atlas Antibodies</td>
<td>Cat# HPA009128; RRID:AB_1078613</td>
</tr>
<tr>
<td>Anti-Cd34</td>
<td>Abcam</td>
<td>Cat# ab46794; RRID:AB_2291941</td>
</tr>
<tr>
<td>Anti-E-Cadherin</td>
<td>Abcam</td>
<td>Cat# 610181; RRID:AB_397580</td>
</tr>
<tr>
<td>Anti-F4/80</td>
<td>Abcam</td>
<td>Cat# ab6640; RRID:AB_1140040</td>
</tr>
<tr>
<td>Anti-FGF</td>
<td>Abcam</td>
<td>Cat# ab13970; RRID:AB_300798</td>
</tr>
<tr>
<td>Anti-Glutamine Synthetase</td>
<td>Abcam</td>
<td>Cat# ab73593; RRID:AB_2247588</td>
</tr>
<tr>
<td>Anti-HNF-4alpha</td>
<td>Santa Cruz Biotechnology</td>
<td>Cat# sc-8987; RRID:AB_2116913</td>
</tr>
<tr>
<td>Anti-Lhx2/LH2</td>
<td>Abcam</td>
<td>Cat# ab184337</td>
</tr>
<tr>
<td>Anti-MFAP4</td>
<td>Abcam</td>
<td>Cat# ab80319; RRID:AB_1658848</td>
</tr>
<tr>
<td>Anti-MYH11</td>
<td>Atlas Antibodies</td>
<td>Cat# HPA015310; RRID:AB_1854261</td>
</tr>
<tr>
<td>Anti-p75 NGF Receptor</td>
<td>Abcam</td>
<td>Cat# ab52987; RRID:AB_881682</td>
</tr>
<tr>
<td>Anti-Cytokeratin, Wide Spectrum Screening</td>
<td>Agilent</td>
<td>Cat# Z0622; RRID:AB_2650434</td>
</tr>
<tr>
<td>Anti-PDGFR Receptor beta</td>
<td>Abcam</td>
<td>Cat# ab32570; RRID:AB_777165</td>
</tr>
<tr>
<td>Anti-Reelin</td>
<td>R and D Systems</td>
<td>Cat# AF3820; RRID:AB_2253745</td>
</tr>
<tr>
<td>Anti-RGS5</td>
<td>Abcam</td>
<td>Cat# ab196799</td>
</tr>
<tr>
<td>Anti-Thrombomodulin</td>
<td>Abcam</td>
<td>Cat# ab230010</td>
</tr>
<tr>
<td>PE/Cy7 anti-mouse CD45</td>
<td>BioLegend</td>
<td>Cat# 103114; RRID:AB_312979</td>
</tr>
<tr>
<td>AF647 anti-mouse CD102</td>
<td>BioLegend</td>
<td>Cat# 105612; RRID:AB_2122182</td>
</tr>
<tr>
<td><strong>Biological Samples</strong></td>
<td>Samsara Sciences</td>
<td>Cat #HLSC</td>
</tr>
<tr>
<td>Human Hepatic Stellate Cells</td>
<td>Samsara Sciences</td>
<td>Cat #HLSC</td>
</tr>
</tbody>
</table>

| Chemicals, Peptides, and Recombinant Proteins | | |
| Histodenz | Sigma-Aldrich | Cat#D2158 |
| AM095 | DSK Biopharma | N/A, custom order |
| BMS-986020 | MedChem Express | Cat#HY-100619 |

| Critical Commercial Assays | | |
| Click-IT Plus EdU Cell Proliferation Kit for Imaging, Alexa Fluor 647 dye | ThermoFisher Scientific | Cat# C10640 |
| RNAscope® Multiplex Fluorescent Reagent Kit v2 | Advanced Cell Diagnostics | Cat# 323100 |
| RNAscope® Mm-Adamts12-2-No-XHs | Advanced Cell Diagnostics | Cat# 465521 |
| RNAscope® Mm-Rspo3 | Advanced Cell Diagnostics | Cat# 402011 |
| RNAscope® Mm-Col1a1-C2 | Advanced Cell Diagnostics | Cat# 31937-C2 |
| Chromium™ Single Cell 3’ Library and Gel Bead Kit | 10X Genomics | Cat# PN-120237 |
| Chromium™ Single Cell A Chip Kit | 10X Genomics | Cat# PN-120236 |

(Continued on next page)
**LEAD CONTACT AND MATERIALS AVAILABILITY**

This study did not generate new unique reagents. Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Neil C. Henderson (Neil.Henderson@ed.ac.uk).

**EXPERIMENTAL MODELS AND SUBJECT DETAILS**

**Mice**

*Pdgfrb*-BAC-eGFP reporter mice (on a C57BL/6 background) were obtained from C. Betsholtz. For all experiments, the mice used were 10–16 week old males housed under pathogen–free conditions at the University of Edinburgh. All experiments were performed in accordance with the UK Home Office Regulations.

**Rats**

Wistar-Han rats were obtained from Charles River Laboratories (Kingston). Male 12-week-old rats were used for experiments. All rat experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals published by the National Research Council (National Academies Press, 2011) and the National Institutes of Health, Office of Laboratory Animal Welfare. Rat studies were run at Covance Laboratories Inc. Greenfield, Indiana, USA.

**Human tissue**

Local approval for procuring human liver tissue for immunofluorescence staining 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 male and female 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.

Formalin-fixed, paraffin-embedded (FFPE) liver samples used for tissue RNA-seq were obtained from commercial tissue vendors (Capital Biosciences, Tissue Solutions, and BioIVT); number of samples, n = 95; Gender (Female), n (%), 59 (62%); fibrosis stage, F1,F2,F3/F4 – 40,31,24; NAS score (mean ± SD), 5.2 ± 0.9; age (means ± SD, years), 52.4 ± 11.5.

**Fibrosis Models**

Carbon tetrachloride (CCl₄) liver injury was induced as described previously (Henderson et al., 2013). For acute CCl₄-induced liver injury, mice were injected i.p with 1µl/g body weight sterile CCl₄ in a 1:3 ratio with olive oil (0.25µl/g CCl₄) after overnight fast (with free access to water); livers were harvested 72 hours post injection. For chronic CCl₄-induced liver fibrosis mice were injected i.p
with 1 μl/g body weight CCl₄ in a 1:3 ratio with olive oil (0.25 μl/g CCl₄) twice weekly for 6 weeks; livers were harvested 48 hours post final injection. To assess proliferation in vivo, mice were injected i.p. with 5-ethyl-2'-deoxyuridine (EdU); 50 mg/kg; ThermoFisher Scientific, C10840) 3 hours prior to sacrifice.

Nonalcoholic steatohepatitis (NASH) was induced by feeding choline-deficient high-fat diet (CDHFD; Research Diets, Inc., A06071302) or standard chow (LabDiet, SCR4) to male Wistar-Han rats to induce nonalcoholic steatohepatitis (NASH). Rats were administered LPAR1 antagonist AM095 (Swaney et al., 2011) (30 mg/kg) (DSK Biopharma) or vehicle (25% (v/v) PEG 200, 74.625% (v/v) deionized water, 0.375% (w/v) Methyl Cellulose (A4M grade)) twice-daily (BiD) via oral gavage while being fed a CDHFD for 12 weeks. Rats were group-housed throughout the experiment and water and feed was provided ad libitum.

Primary cells and cell lines

Primary human hepatic stellate cells (hHSC) were isolated from viable male livers by density gradient centrifugation (Samsara Sciences; Cat, #HLS; Donor ID, HL150002SC). TWNT4 cells were obtained courtesy of Bryan Fuchs, PhD (Fuchs et al., 2014; Nakagawa et al., 2016). TWNT4 cells and hHSC were cultured in DMEM (4.5 mg/mL glucose, 110 mg/L sodium pyruvate, 4 mM L-glutamine) with 10% fetal bovine serum (FBS) supplemented with 100 units/mL penicillin and 100 mg/mL streptomycin (all from Mediatech, Manassus, VA). Cells were maintained at 37 °C in a humidified incubator with 5% CO₂ in air.

METHOD DETAILS

Immunofluorescence Staining

Mouse liver was briefly perfused through the inferior vena cava with PBS, then excised. For staining which included intrinsic GFP reporting, tissue was fixed/frozen. For all other staining, tissue was formalin-fixed paraffin-embedded.

Fixed/frozen sections

Liver was fixed in 4% paraformaldehyde for 2 hours at 4 °C then immersed in graded sucrose solutions, embedded in OCT and stored at −80 °C. 7 μm frozen sections were cut and left to air dry for 30 minutes, washed in PBS, then blocked using protein block (GeneTex, GTX30963) for 30 minutes. Sections were then incubated with antibodies listed in Table S4 overnight at 4 °C. Following a further PBS wash sections were incubated with fluorescently conjugated secondary antibodies depending on host species (Alexa Fluor 555 goat anti-rat; Alexa Fluor 555 goat anti-rabbit; Alexa Fluor 555 donkey anti-mouse; Alexa Fluor 488 goat anti-chicken (Life Technologies, A21434, Lot. 1722994; A21429, Lot.1937155; A31570, Lot.1850121; A11039; Lot.1869581, respectively). Co-stains were completed sequentially. Slides were washed further in PBS before DAPI-containing mountant was applied (ThermoFisher Scientific, P36931).

For PDGFRβ staining, before blocking, heat-mediated antigen retrieval in pH9 Tris-EDTA (microwave; 2 minutes) was performed. Slides were incubated in PBS, incubated in 3% hydrogen peroxide for 10 minutes and washed again in PBS before proceeding with the above protocol. Instead of incubating with a fluorescently conjugated secondary antibody sections were washed with PBS and then incubated with ImmPress HRP Polymer Detection Reagents (rabbit, MP-7401 Vector Laboratories) for 30 minutes and washed in PBS again. Staining was detected using Cy3 tyramide (Perkin-Elmer, NEL744B001KT) at 1:1000 dilution. Sections were imaged using a slide scanner (AxioScan.Z1, Zeiss) at 20X magnification. Images were processed using Zen Blue (Zeiss) and Fiji image software.

Formalin-fixed paraffin-embedded (FFPE) sections:

Liver was fixed in 4% neutral-buffered formalin for 24 hours followed by paraffin-embedding. 5 μm sections were cut, dewaxed, rehydrated, then incubated in 4% neutral-buffered formalin for 20 minutes. Following heat-mediated antigen retrieval in pH9 sodium citrate or pH9 Tris-EDTA (microwave; 15 minutes), slides were washed in PBS and incubated in 3% hydrogen peroxide for 10 minutes. Slides were then washed in PBS, blocked using protein block (GeneTex, GTX30963) for 1 hour at room temperature, and incubated with primary antibodies. A full list of primary antibodies and conditions are shown in Table S4. Slides were then washed in PBS/T (PBS plus 0.1% Tween 20; Sigma-Aldrich, P1379), 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 minutes at room temperature, and washed again with PBS/T. Staining was detected using either Cy3, Cy5, or Fluorescein tyramide (Perkin-Elmer, NEL744B001KT) at 1:1000 dilution. For multiplex stains slides were then washed in PBS/T followed by a further heat treatment with pH6 sodium citrate or pH9 Tris-EDTA (15 minutes), washed in PBS, incubated in 3% hydrogen peroxide for 10 minutes, washed in PBS, protein blocked, and finally incubated with the second primary antibody followed by the ImmPress Polymer and tyramide as before. When required this sequence was repeated for the third primary antibody. A DAPI-containing mountant was then applied (ThermoFisher Scientific, P36931). Sections were imaged using a slide scanner (AxioScan.Z1, Zeiss) at 20X magnification. Images were processed using Zen Blue (Zeiss) and Fiji image software.

αSMA (Biocare CM001 [1A1], Ms mAb, 0.12 μg/mL) IHC staining of rat liver was performed using the Biocare Intellipath autostainer utilizing Biocare Medical reagents. Following deparaffinization, slides were sequentially treated with hydrogen peroxidase for 5 minutes, citrate-based heat induced (95 °C) antigen retrieval for 40 minutes, protein block for 10 minutes, primary antibody for 30 minutes, one-step polymer-HRP conjugated (mouse on rat HRP) secondary antibody for 30 minutes, DAB chromogen for 5 minutes, CAT Hematoxylin for 5 seconds, and bluing solution for 10 seconds.
**RNAscope**
Detection of Adamts2, Rsopo, and Col1a1 was performed using the RNAscope® Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics (ACD), Cat, 323100) in accordance with the manufacturer’s instructions. Briefly, 5µm liver sections were dewaxed, incubated with endogenous enzyme block, boiled in pretreatment buffer and treated with protease, followed by target probe hybridization using the RNAscope® Mm-Adamts2-No-XHs (Cat, 465521; Lot, 19086B; ACD), Mm-Rsopo (Cat, 402011; Lot, 18338A; ACD) or Mm-Col1a1-C2 (Cat, 31937-C2; Lot, 19086C; ACD) probes. Target RNA was detected with Cy3 (Adams2 or Rsopo) or Fluorescein (Col1a1) tyramide (Perkin-Elmer) at 1:750 dilution. For combined RNAscope and immunofluorescence staining sections were processed as for multiplex immunofluorescence staining (as above) after the RNAscope protocol. Slides were imaged using a slide scanner (AxioScan.Z1, Zeiss) at 40X (for RNAscope) magnification or a Zeiss LSM780 inverted confocal microscope. Images were processed using Zen Blue (Zeiss) and Fiji image software.

**EdU Click-IT Immunofluorescence staining**
EdU incorporation into DNA was detected using the Click-IT EdU Alexa Fluor Imaging kit (Invitrogen/Molecular Probes, C10640). Formalin-fixed paraffin-embedded 5µm sections were dewaxed, rehydrated, then incubated in 4% neutral-buffered formalin for 20 minutes. Sections were washed in 0.1% Triton X-100 in PBS (PBSTX) for 10 minutes followed by heat-mediated antigen retrieval in pH6 sodium citrate (microwave; 15 minutes), washed for 10 minutes in PBSTX and blocked for 1 hour using protein block (GeneTex, GTX30963). The Click-IT solution was then made according to manufacturer’s instructions. Slides were incubated in the EdU cocktail for 30 minutes and rinsed three times in PBS. The azide used was coupled to an Alexa Fluor 647 fluorophore. Upon completion of the EdU Click-IT reaction, slides were processed as above for multiplex staining. Sections were imaged using a slide scanner (AxioScan.Z1, Zeiss) at 20X magnification. Images were processed using Zen Blue (Zeiss) and Fiji image software.

**Picrosirius Red Staining**
Picrosirius red (PSR) staining was performed using 0.1% Direct Red 80 (Sigma) in 1.3% picric acid solution (Sigma, 239801). Formalin-fixed paraffin-embedded 5µm sections were dewaxed, rehydrated, incubated in 0.4% phosphomolybdic acid for 5 minutes, and washed with PBS. Sections were then stained with picrosirius red for two hours, before washing twice with agitation for 30 s in acidified water. Slides were placed in 0.1% Fast Green (ThermoFisher Scientific, F/P025/46) for 30 seconds followed by two 30 seconds washes with agitation in acidified water. Following dehydration (100% ethanol), slides were cleared in xylene and mounted using DPX. Sections were imaged using a slide scanner (AxioScan.Z1, Zeiss) at 20X magnification.

**Image Quantification**
Cell counts for zonation of HSC in uninjured liver, Lhx2+ cell expansion, zonation of HSC following acute liver injury, Pdgfrb-BAC-eGFP reporting efficiency and specificity, and NGFR and Adamts2 marker specificity were counted manually from multiple high-powered images per sample. For zonation in uninjured liver the peri-central and peri-portal regions were defined as areas of positive Cyp2e1 or E-cadherin staining, respectively. All areas were processed using Zen Blue software to calculate cell count/mm². In uninjured liver and following acute liver injury (72 hours post single CCl4) NGFR was used as a marker for PaHSC and Adamts2 as a marker for CaHSC. For quantification of proliferation following acute injury PaHSC were identified as NGFR+/PDGFRα+ cells and CaHSC as NGFR+/PDGFRβ+ cells.

To quantify PSR staining digital morphometric pixel analysis was performed using the Trainable Weka Segmentation (TWS) plugin (Arganda-Carreras et al., 2017) in Fiji software (Schindelin et al., 2012). Briefly, the TWS plugin was trained to produce a classifier segmenting images into areas of positive staining, tissue background and white space. The same trained classifier was applied to all images to produce a percentage area of positive staining for each tissue section.

Quantitative image analysis of PSR staining in rat tissue was performed using Visiopharm v2017.2. The Tissue Find APP was used to find the tissue regions in the images, before positive expression was categorized as low, medium, or high based on the level of staining intensity. Percentage area was calculated by summing the values for low, medium, and high expression, dividing by the value for staining on the entire tissue area, and multiplying by 100.

**Hydroxyproline Assay**
Liver samples were dehydrated overnight at 62°C, followed by homogenization in water (50µL water/1mg dry tissue weight) using a bead-based TissueLyser. Total protein was measured using a BCA protein assay (Pierce (Thermo Fisher) BCA Protein Assay Kit; 23227). Homogenates were hydrolyzed overnight in 6 N HCl at 110°C. Samples and orthohydroxyproline standards were added in duplicate to microplate wells and dried. Chloramine T (Sigma; 857319) was added to all wells and the plate was incubated at room temperature for 30 minutes with shaking. Ehrlich’s Reagent (Fisher; D71-25) was added to all wells and the plate sealed and incubated at 60°C for 40 minutes. Optical density was measured at 560nm on Molecular Devices’ SpectraMAX PLUS Microplate Reader. OH-P content was calculated for all samples and normalized to total protein and compared to an 8-point standard curve.

**Immunocytochemistry**
TWNT4 cells (Nakagawa et al., 2016) (courtesy of Bryan Fuchs, PhD) were plated on 96-well optical plates (Greiner Bio-One; 655946) in complete media (DMEM, (GIBCO; 15-018-CM), 10% FBS, (Hyclone; SH30088.03), Penicillin-streptomycin-glutamate
(GIBCO; 10378), serum-starved overnight, was pretreated with DMSO or 1µM LPAR1 antagonist (BMS-986020) (Palmer et al., 2018) (Medchem Express; HY-100619) for 30 minutes, then treated with 0.1% BSA (control) or 10µM 18:2 LPA (Avanti Polar Lipids; 857138) for 20 minutes. Cells were then fixed in 4% paraformaldehyde for 20 minutes at room temperature, washed with PBS, permeabilized with 0.3% TritonX-100/PBS, blocked with 2% BSA/PBS, and stained with Alexa647-phalloidin, mouse anti-pMLC2 (mAb3675, Cell Signaling Technology), and Hoechst 33342 (ThermoFisher; H3570). Images were acquired with ImageXpress Pico automated imaging system (Molecular Devices).

Contraction Assay
Collagen gel contraction assay was performed using the CytoSelect 48-Well Cell Contraction Assay Kit (CBA-5021). In brief, TWNT4 cells (200K cells/well) were mixed with collagen solution and allowed to polymerize at 37°C in a CO2 incubator for 1 hour. After collagen gel polymerization, 0.5ml media (Serum free- DMEM, (GIBCO; 15-018-CM) containing either either 3 µM LPAR1 antagonist (BMS-986020) or DMSO were added atop the collagen lattice for 30 minutes, followed by addition of 50 µM LPA (Avanti Polar Lipids) and incubation at 37°C and 5% CO2. Media was changed daily by carefully removing 250µl of media and replacing with 250µl (with /without contraction mediators). Collagen gel contraction was measured after 96 hours using light microscopy or on a Celigo imaging cyrometer platform (Nexcelom Biosciences) using the bright-field channel. Contracted gel area was quantified using ImageJ analysis.

qPCR
Primary HSC or TWNT4 cells were pre-treated for 30 minutes with 3µM LPAR1 antagonist (BMS-986020) or DMSO, followed by addition of 10µM 18:2 LPA (Avanti Polar Lipids) for 2 hours. RNA was isolated using the RNeasy Plus Mini Kit (QIAGEN) and cDNA was made using the High Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific; 4368814). RT-qPCR for CTGF and HPRT1 was performed using TaqMan Gene Expression Assays (ThermoFisher Scientific, Hs00170014_m1 and Hs02800695_m1).

Primary cell isolation
Digestion protocol 1
Mouse liver was perfused through the inferior vena cava with phosphate buffered saline (PBS). The liver was excised, minced with a scalpel, digested in 5mg/ml pronase (Sigma, P5147), 2.84mg/ml collagenase B (Roche, 11088815001; 0.188U/mg) and 0.019mg/ml DNase 1 (Roche, 10104159001) at 37°C for 20 minutes with agitation (200–250 rpm), and then strained through a 120µm nylon mesh. The cell suspension was centrifuged at 400g for 7 minutes, supernatant removed, cell pellet resuspended in PEB buffer (PBS, 2% FBS, and 2mM EDTA), and DNase I added (0.02mg/ml). Following red blood cell lysis with RBC lysis buffer (BioLegend; Cat:420301), the cell suspension was again centrifuged at 400g for 7 minutes, supernatant removed, cell pellet resuspended in PEB buffer and DNase I added (0.02mg/ml).

Digestion protocol 2
Hepatic stellate cells (HSC) were isolated from mice as described previously (Mederacke et al., 2015). Mice were anaesthetized via inhalation of isoflurane (1%–3%). Following cannulation of the inferior vena cava, the portal vein was cut to allow retrograde stepwise perfusion of EGTA (0.19mg/ml; 2 minutes), pronase (0.4mg/ml; 5 minutes; Sigma, P5147) and collagenase D (0.185U/ml; 7 minutes; Roche, 11088882001). Liver was then excised and minced with a scalpel, digested in 5mg/ml pronase (Sigma, P5147), 2.84mg/ml collagenase B (Roche, 11088815001; 0.188U/mg) and 0.019mg/ml DNase 1 (Roche, 10104159001) at 37°C before incubation with antibodies CD45-PE/Cy7 (1:100; BioLegend; Clone: 30-F11; Cat. 420404), DAPI was replaced with 7-AAD viability stain (BioLegend; Cat. 420404), the cell suspension was again centrifuged at 400g for 7 minutes, supernatant removed, cell pellet resuspended in PEB buffer and DNase I added (0.02mg/ml). Following red blood cell lysis with RBC lysis buffer (BioLegend; Cat:420301), the cell suspension was again centrifuged at 400g for 7 minutes, supernatant removed, cell pellet resuspended in PEB buffer and DNase I added (0.02mg/ml). For cells isolated from digestion protocol 2 DAPI was replaced with 7-AAD viability stain (BioLegend; Cat. 420404; Lot. B251165). Cell sorting was performed on a FACS Aria II (Becton Dickinson, Basel, Switzerland).

Cell sorting
Cells were blocked in 1% purified anti-mouse CD16/32 (BioLegend; Clone. 93 Cat. 101324; Lot. B254979) and 10% normal mouse serum (Sigma, M5905) for 10 minutes at 4°C before incubation with antibodies CD45-PE/Cy7 (1:100; BioLegend; Clone: 30; F11; Cat. 103114; Lot. B243728) and CD102-AF647 (1:100; BioLegend; Clone: 3C4 (MIC2/4); Cat. 105612; Lot. B227625) for 20 minutes at 4°C. For cells isolated from digestion protocol 1 live/dead (DAP 1:1000) staining was performed immediately prior to running the samples. For cells isolated from digestion protocol 2 DAPI was replaced with 7-AAD viability stain (BioLegend; Cat. 420404; Lot. B251165). Cell sorting was performed on a FACS Aria II (Becton Dickinson, Basel, Switzerland).

Human liver tissue RNA-seq
RNA sequencing was performed by Q2 Solutions (Morrisville, North Carolina). For human liver samples, total RNA was isolated from three 11µm FFPE curls per sample in one tube (33µm total). All samples had > 100ng of input RNA. Sequencing libraries were created using the TruSeq RNA Access Target enrichment and library preparation methodology which provides high data quality data even from degraded or FFPE-derived RNA samples. Libraries were sequenced on an Illumina HiSeq2500 with 50bp paired-end sequencing and a total read depth of 40M reads per sample. R packages edgeR and limma were used to normalize sequence count data and
conduct differential gene-expression analysis. False discovery rate (FDR) was calculated using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995).

**Single-cell workflows**

**10X Chromium**

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 + 2% FBS, washed twice and counted using a Bio-Rad TC20. Approximately 10,769 cells were added to each lane of a 10X chip and partitioned into Gel Beads in Emulsion in the Chromium™ instrument, where cell lysis and barcoded reverse transcription of RNA occurred, followed by amplification, fragmentation and 5’ adaptor and sample index attachment. Libraries were sequenced on an Illumina HiSeq 4000.

**Smart-seq2**

Single cells were processed by SciLifeLab – Eukaryotic Single cell Genomic Facility (Karolinska Institute). Before shipping single cells were sorted into wells of a 384-well plate containing pre-prepared lysis buffer. Libraries were sequenced on an Illumina HiSeq 4000.

**Pre-processing scRNA-seq data**

**Mouse 10X Chromium:**

We aligned to the mm10 reference genome (Ensembl 84) and estimated cell-containing partitions and associated UMIs using the Cell Ranger v2.1.0 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, prior to scaling and centering the resulting value by subtracting the mean expression of each gene and then calculating \( \ln(10^4 E_n + 1) \). Highly variable genes were identified using Seurat’s `FindVariableGenes` function with default parameters. Non-mesenchymal mouse scRNA-seq data following chronic CCl4-induced liver fibrosis were analyzed from our previously-obtained datasets (Ramachandran et al., 2019).

**Human 10X Chromium:**

We aligned to the GRCh38 reference genome (Ensembl 84) and processed our single-cell transcriptomic data as above. Mesenchymal cells were isolated based on PDGFRB expression.

**Mouse Smart-seq2**

The single-cell transcriptomic data was initially processed at the Eukaryotic Single-Cell Genomics Facility at the Science for Life Laboratory in Stockholm, Sweden: obtained reads were mapped in STAR to the mm10 build of the mouse genome (concatenated with transcripts for eGFP and the ERCC spike-in set), and then processed via rpkmforgenes, MULTo, and RefSeq to yield a count for each endogenous gene, spike-in, and eGFP transcript per cell. We performed quality control in R packages `scran` v1.6.9 (Lun et al., 2016), removing cells with library size or features less than, or with ERCC percentage greater than, 3 median absolute deviations from the dataset median. We then computed normalized expression using sum factors (with separate calculation of spike-in factors), before transferring these values to Seurat to identify highly variable genes as above.

**Dimensionality reduction, clustering, and DE analysis**

We performed unsupervised clustering and differential gene expression analyses in the Seurat R package v2.3.0 (Satija et al., 2015). In particular we used SNN graph-based clustering, where the SNN graph was constructed using from 2 to 10 principal components as determined by dataset variability shown in principal components analysis (PCA); the resolution parameter to determine the resulting number of clusters was also tuned accordingly. In total, we present scRNA-seq data from nine mouse liver samples in 10X (three uninjured, three acute CCl4, three chronic CCl4) and three mouse liver samples in SmartSeq2 (all uninjured).

All heatmaps, t-SNE visualizations, and violin plots were produced using Seurat functions in conjunction with the `ggplot2`, `pheatmap`, and `grid` R packages. t-SNE visualizations were constructed using the same number of principal components as the associated clustering, with perplexity ranging from 100 to 300 according to the number of cells in the dataset. We conducted differential gene expression analysis in Seurat using an AUC classifier to assess significance, retaining only those genes with a log-fold change of at least 0.25 and expression in at least 25% of cells in the cluster under comparison.

**Defining cell expression signatures**

Signature scores were defined per cell as the geometric mean of the expression of the associated signature genes, scaled to a range of 0 to 1 across the dataset. For a signature of fibrillar collagen production we aggregated expression of the following genes: Col1a1, Col1a2, Col3a1. For a signature of proliferation we aggregated expression of the following genes: Mki67, Cdca8, Cdc20, Ccn2a, Ccnb1. To perform unbiased thresholding on these scores we used k-means clustering (using the `threshold` function from the `rmand` R package) and binarised the results.
Identifying and applying an HSC zonation signature

We used unsupervised Independent Component Analysis (ICA) in Seurat to generate components of variability in uninjured HSCs. Analyzing each component in turn, we identified and isolated the one with highest correlation to observed Ngfr heterogeneity. We then thresholded on the gene weight loadings along this component of interest to extract an 81-gene signature, including 52 genes associated with Ngfr and 29 genes associated with Spon2.

Using this 81-gene signature as the input to supervised Seurat clustering, we clustered the homeostatic HSCs into two subpopulations: Ngfrhi and Ngfrlo/neg. We classified acute and chronic CCl4 HSCs into the same two subpopulations in the same manner. To assess whether the zonation profiles of these signature genes remained consistent across acute and chronic injury, we ordered the cells at each time point by the strength at which they expressed the 52-gene Ngfr-associated signature and manually inspected the profile of each gene across this ordering.

Analyzing functional phenotypes of mesenchymal cells

For further analysis of the function related gene expression profile we adopted the self-organizing maps (SOM) approach as implemented in the SCRAT R package v1.0.027. For each lineage of interest we constructed a SOM in SCRAT using default input parameters and according to its clusters. We defined the signatures expressed in a cell by applying a threshold criterion (ethresh = 0.95 × e7) 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 SOMs 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 (http://pantherdb.org).

Inferring injury dynamics and transcriptional regulation

To generate cellular trajectories (pseudotemporal dynamics) we used the monocle R package v2.6.1 (Trapnell et al., 2014). We ordered cells (Ngfrhi versus Ngfrlo/neg) in an unsupervised manner, scaled the resulting pseudotime values from 0 to 1, and mapped these onto the t-SNE visualizations generated by Seurat. We removed mitochondrial and ribosomal genes from the geneset 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 1e-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 v0.6.035, 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 kNN 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 visualizations as generated previously. 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 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 (Zhang et al., 2015). To further analyze transcription factor regulons, we adopted the SCENIC v0.1.7 workflow in R (Aibar et al., 2017), 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.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using GraphPad Prism (GraphPad Software, USA). Comparison of changes in histological cell counts, topographical localization of counted cells, morphometric pixel analysis, and gene expression were performed using a Mann-Whitney test (unpaired; two-tailed). Comparison of RNA-seq data from human NASH patients was performed using a Kruskal-Wallis test with Dunn’s multiple comparisons test. Analysis of correlation was performed using a Spearman correlation coefficient. All statistical tests used, exact value of n, and P values obtained are displayed in the figure legends. P values < 0.05 were considered statistically significant.
DATA AND CODE AVAILABILITY

All mouse mesenchymal data is deposited in the Gene Expression Omnibus. The accession number for the data is GEO: GSE137720. All human mesenchymal data, as well as mouse leucocyte data, is available from the Gene Expression Omnibus (GEO: GSE136103).

R markdown scripts enabling the main steps of the analysis are available from the Lead Contact upon reasonable request.

Additional Resources

Our uninjured and 6 week CCl₄ expression data is freely available for user-friendly interactive browsing online: http://livermesenchyme.hendersonlab.mvm.ed.ac.uk