Human liver organoids model progressive inflammatory and fibrotic injury in non-alcoholic fatty liver disease

Anja Hess[1,2], Stefan D. Gentile[1,3], Amel Ben Saad[1,3], Raza-Ur Rahman[1,3], Tim Habboub[1] and Alan C. Mullen[1,3,4,5,6*]

1. Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA
2. Max Planck Institute for Molecular Genetics, Department of Genome Regulation, Berlin, Germany
3. Harvard Medical School, Boston, MA, USA
4. Harvard Stem Cell Institute, Cambridge, MA, USA
5. Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA
6. Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Boston, MA, USA

* Corresponding author. GRJ7, 55 Fruit Street, Massachusetts General Hospital, Boston, MA 02114, USA. (617) 726-6342. Email: acmullen@mgh.harvard.edu

Abstract

Chronic liver injury promotes inflammation, which can progress to fibrosis and cirrhosis, a major cause of mortality. Resolving the transcriptional changes orchestrating this transformation remains challenging. Here, we derive single-cell transcriptomic maps of progressive liver injury in human pluripotent stem cell (hPSC) - derived liver organoids (HLOs), modeling steatohepatitis with palmitic acid and fibrosis through TGF-β1 treatment. We observe that palmitic acid drives inflammation and non-alcoholic fatty liver disease (NAFLD) expression signatures, while TGF-β1 expands hepatic stellate-like populations, remodels cell cycle patterning, and induces extracellular matrix pathways. Analysis of receptor-ligand expression defines the induction of genes regulating Notch and fatty acid signaling with palmitic acid treatment, while the TGF-β1 response is shaped by the co-expression of *COL1A1* and integrins to promote crosstalk between hepatocytes, cholangiocytes, and stellate cells. Finally, inflamed and fibrotic HLOs sequentially induce genes predicting disease progression in NAFLD. Our findings highlight HLOs as dynamic human *in vitro* systems to study evolving liver injury, providing a single-cell transcriptomic reference that will facilitate benchmarking future organoid-based liver injury models.

Introduction

Chronic liver injury promotes sustained inflammation, leading to liver fibrosis, which can progress to cirrhosis, a major cause of morbidity and mortality worldwide[1,2]. Non-alcoholic fatty liver disease (NAFLD) is among the most common causes of chronic liver disease and the most rapidly increasing indication for liver transplantation in the US[3–5]. The majority of cases of NAFLD that lead to end-stage liver disease progress from simple steatosis to non-alcoholic steatohepatitis (NASH, inflammation due to steatosis) and then fibrosis[1,6]. While the progression from steatosis to fibrosis is well documented, there are currently few treatment options to disrupt this process other than weight loss. More broadly, there are no approved treatments available that target common inflammatory or fibrotic pathways arising from chronic liver injury, which could prevent the progression of NASH or other chronic liver diseases. Thus, the development of human cell-based models of liver inflammation and fibrosis to investigate molecular pathways and the activity of these pathways in individual cell types is critical for creating and testing new approaches to prevent liver failure.

Conceptually, the prerequisite for chronic liver injury is continued exposure to hepatotoxic agents, such as the saturated free fatty acid palmitic acid (PA)[7,8], which accumulates in NAFLD[9–11]. Oleic acid (OA) is another NAFLD-associated saturated free fatty acid[12] with protective[13–15] and toxic[16] effects having been described. As a consequence of toxic agent accumulation, hepatocytes and other resident liver cell types may undergo apoptosis or activate pro-inflammatory signaling pathways leading to the secretion of cytokines such as tumor necrosis factor alpha (TNFα)[17–20]. This subsequently triggers the activation of resident liver macrophages (Kupffer cells)[21], the recruitment of systemic immune cells[22], and their common secretion of cytokines, including transforming growth factor beta (TGF-β)[23]. Finally, TGF-β activates hepatic stellate cells (HSCs) and promotes their differentiation towards myofibroblasts[24,25], driving the acquisition of the fibrotic phenotype. Hallmarks of fibrosis include the up-regulation of transcripts *COL1A1, COL3A1, ACTA2, TIMP1* and receptor-ligand pairs coding *TNFS12/TNFRSF12A* and *PDGFB/PDGFR*[26,27], as well as the accumulation of myofibroblast-secreted collagen types I and III in the extracellular matrix (ECM) and an expansion of the ECM[28,29].

Generation of human pluripotent stem cell (hPSC) - derived multi-lineage liver organoid systems (HLOs)[30–36] has allowed aspects of liver inflammation[33,37,38] and fibrosis[39] to be modeled in individual systems *in vitro*. However, a multi-lineage liver organoid model effectively recapitulating evolving liver injury would further require the demonstration of specific transcriptional responses corresponding to different stages of liver injury. This would include the potential to create different degrees of liver injury displaying signatures indicative of the underlying injury mechanism and ideally recapitulate sequential phenotypes covering normal, inflammatory, and finally fibrotic stages. Single-cell RNA sequencing (scRNA-seq) has emerged as a key tool for benchmarking and deciphering such transcriptional dynamics, and efforts are underway to build single-cell atlases of multi-cellular organoids as promising 3D models to comprehensively evaluate their potential to recapitulate human diseases[40].

Here, we present an HLO model to profile multiple stages of liver injury, including PA-induced steatohepatitis and TGF-β1-induced fibrosis, providing a single-cell reference of the HLO injury landscape. We apply morphological, histological, and transcriptional analysis to evaluate the generation of liver disease phenotypes and optimal 3D culture conditions for this purpose. We compare the transcriptional landscape of differential injury conditions and demonstrate their ability to align along the sequential trajectory of liver fibrosis development. Finally, we apply clinically-validated liver fibrosis staging gene signatures that correctly classify the severity of liver injury in HLOs across disease progression of NAFLD.

Results

Human liver organoids recapitulate the transcriptional landscape of major cell types in the adult human liver

We first differentiated hPSCs into HLOs as previously described[33,34] (**Fig. 1a**). We confirmed loss of the canonical pluripotency genes *SOX2*, *NANOG,* and *POU5F1*[41] by day 16 of differentiation and induction of the human liver genes *ASGR1* and *HNF4A* (hepatocyte markers[42,43]), *KRT19/CK19* and *SOX9* (cholangiocyte markers[44,45]) as well as *VIM* and *DES* (HSC markers[46,47]) in day 21 HLOs (**Fig.**

1b). To evaluate HLO identity at the protein and histologic level, we performed hematoxylin and eosin (H&E) staining and immunohistochemistry (IHC) and found that HLOs at day 21 display an organized, sphero-luminal structure, and multiple cells express the nuclear hepatocyte marker CEBPα[48] (**Fig. 1c**, Supplementary Fig. 1). These results are in agreement with previous observations on the emergence of distinct liver cell types from at least two different lineages in HLOs after day 20[33].

To further validate the cellular composition of HLOs, we performed 10X scRNA-seq on day 21 and analyzed a total of 17,898 cells after quality control (QC) (**Fig. 1d**, Supplementary Fig. 2, Methods). Annotation of organoid cell populations can create difficulties since differentiating systems may contain transient cell states and retain premature features[49]. To address these challenges, we initially evaluated three different annotation strategies[50]. We first annotated cell clusters based on marker genes from the literature for fetal and adult liver cell types (Supplementary Table 1, Supplementary Fig. 2d-e), rendering five main clusters of hepatocyte-like, HSC-like, two cholangiocyte-like populations, and a small fraction of embryonic stem cell (ESC)-like cells expressing *NANOG* and *POU5F1*, but not *SOX2* (**Fig. 1d**, top). In a second approach, we annotated cell clusters based on genes in cell-type marker databases[51,52], composed from a wide range of scRNA-seq data, including one human liver study[53] (Supplementary Table 2, Supplementary Fig. 2d-e). This resulted in annotations closely related to the previous approach (**Fig. 1d**, center). Here, the cholangiocyte-like I cluster was annotated as hepatoblast (HB)-like, the common progenitor giving rise to cholangiocytes and hepatocytes[54,55], since HB-markers *KRT8* and *KRT18* and cholangiocyte markers *KRT19* and *CLDN4* co-occurred across the top differentially expressed genes defining this cluster (Supplementary Table 3, Supplementary Fig. 2d). Increasing the clustering coarseness resolved a small cholangiocyte population with the HB annotation being largely preserved, suggesting that the majority of the previously cholangiocyte-like I labeled cells retain precursor features. In contrast, the literature based annotation preserved the cholangiocyte annotation throughout cluster coarseness levels, suggesting this method being potentially less sensitive discriminating cholangiocytes from their precursor states (Supplementary Fig. 2e). A fifth cluster (cholangiocyte-like II / smooth muscle cell-like) showed ambiguous annotations and thus could not be confidentially annotated by the two strategies. To complement this analysis with a third method, we applied SingleCellNet (SCN)[56], a random forest-

based classifier, to compare HLOs to scRNA-seq data from fetal and adult liver tissue[57]. We chose this approach since SCN has previously been reported to yield high inter-dataset and inter-experimental accuracy for benchmarking data of human pancreatic cell populations[58], which developmentally are closely related to the liver[59]. The SCN annotations for hepatocyte- and HSC-like cells are consistent with the previous methods, and the cholangiocyte-I/ HB-like cluster is mainly assigned to the normal adult liver tissue-specific *KRT8+*/*KRT19+* population, labeled as bi-potent cells according to the terminology in the reference study[57] (**Fig. 1d**, bottom, Supplementary Fig. 3a-b). Overall, the attributions of cellular identities for the main clusters of hepatocyte- and HSC-like cells are overlapping across annotation strategies, and two out of three strategies indicated a cholangiocyte-precursor population with bipotent features constituting the third major cell type in day 21 HLOs. Based on this comparative analysis we decided to utilize the database annotation method for all subsequent exploratory analyses since it showed sensitivity to progenitor states at the cholangiocyte-HB interface and allows for the annotation of potentially emerging cell types beyond the repertoire of a single reference study.

To further ensure cellular identities we confirmed the expression of canonical marker genes for each cell type and found they overlap with the consensus annotations (**Fig. 1e**). *EPCAM, CLDN4, KRT7* and *KRT19* expression has been reported for adult cholangiocytes as well as hybrid (i.e. bi-potent) hepatic progenitors[44], and these genes are characteristic for the cholangiocyte I/HB-like cluster. These results support the consensus annotations from the previous strategies.

We next performed pathway enrichment analysis, revealing upregulated signatures for liver-characterizing metabolic processes and lipid metabolism in hepatocyte-like cells. HSC-like cells were enriched for pathways related to extracellular matrix and focal adhesion. Cholangiocyte-I/ HB-like cells showed enrichment for prostaglandin synthesis genes, in line with prostaglandins being involved in favoring liver fate decisions in early progenitors[60], as well as genes associated with epithelial-mesenchymal-transition (EMT) (**Fig. 1f**, Supplementary Fig. 3c). Together, these results indicate the presence of hepatocyte-, HSC- and emerging cholangiocyte-like populations in day 21 HLOs across multiple cell type annotation strategies.

Inducible liver injury phenotypes in HLOs upon TGF-β1 and fatty acid treatment

We next evaluated conditions to induce a fibrotic phenotype in the HLO model. TGF-β1 drives liver fibrogenesis *in vivo*[61,62], and we treated HLOs with TGF-β1 (10 ng/ml and 25 ng/ml) in four different 3D culture systems. The final differentiation steps for HLOs are performed in Matrigel, and HLOs were either left in Matrigel or removed from Matrigel and cultured on i) ultra-low attachment plates (commercially available cell culture plates coated with a hydrogel layer to avoid attachment) as previously described[34], ii) 1% agarose coated plates, and iii) an orbital shaker (**Fig. 2a**).

We observed morphological changes in HLOs cultured in Matrigel and on the orbital shaker after five days of treatment including HLO tissue consolidation, along with surface roughening in TGF-β1-treated HLOs (**Fig. 2b**, Supplementary Fig. 4a). For comparison, we also evaluated HLOs treated with OA, which did not demonstrate the same compaction in Matrigel as observed with TGF-β1 but was associated with a darker appearance on light microscopy when cultured in the orbital shaker, as previously described[33,34]. These results suggest HLOs are reacting in an injury-specific manner to the applied treatments.

We further quantified the contractile effect of TGF-β1 by culturing HLOs in Matrigel drops and measuring the Matrigel drop area after TGF-β1 application. This analysis demonstrates a significant reduction in droplet size with TGF-β1 treatment (**Fig. 2c**). To further investigate alterations in HLOs we stained for H&E and observed the intraluminal accumulation of hyaline, monomorph structures upon TGF-β1 treatment (**Fig. 2d**), providing further evidence for an injury response to TGF-β1 treatment.

We next performed Sirius red staining (**Fig. 2e**), a standard method for quantification of type I and III collagen deposition in liver fibrosis[63]. We optimized an existing pipeline (Methods) for the computational quantification of Sirius red staining in the human liver for sections with low tissue

amounts and options to select individual HLO areas of interest for the calculation of Sirius red percentage per tissue and per area separately (Methods, github.com/anjahess/sr_organoids). We utilized our pipeline to analyze HLOs cultured via the four previously described methods that were treated with TGF-β1 (10 ng/ml) for five days. We found a significant increase of collagen deposition only in Matrigel and orbital shaker-cultured HLOs (**Fig. 2f**), suggesting the latter culture methods lead to accumulation of extracellular type I and III collagen as observed in liver fibrosis.

We next evaluated the gene expression response across the four culture methods in HLOs when treated with TGF-β1. Canonical transcriptional changes in fibrosis include the induction of the alpha-1 subunit of type 1 collagen (*COL1A1*)[26] and *TNFA*[64], while TGF-β1 has been reported to reduce alpha-fetoprotein (*AFP*) levels in hepatic progenitor cells[65], foregut endoderm[66], and hepatocellular carcinoma (HCC) cell lines[67]. We analyzed mRNA levels for the respective genes. We found that only HLOs cultured on an orbital shaker with TGF-β1 showed a significant increase in *COL1A1* and *TNFA* expression (**Fig. 2g**), while *AFP* reduction was significant in HLOs cultured on Matrigel and 1% agarose, potentially reflecting different levels of TGF-β1-responsive progenitor-like cells across culture conditions. These results show that culturing in the orbital shaker provides the conditions under which HLOs demonstrate the most robust fibrotic and inflammatory response to TGF-β1.

We therefore focused on the orbital shaker method to screen for an inflammatory response to fatty acids. In H&E, both PA and OA treatments resulted in the increase of stain-free vacuoles as a correlate of lipid droplets (**Fig. 2h**, arrow heads). Treatment with PA was associated with an increase in *TNFA* mRNA levels, but *COL1A1* mRNA and protein levels did not change significantly (**Fig. 2i,** Supplementary Fig. 4c). Treatment of HLOs with OA was not associated with an increase in *TNFA* or *COL1A1* expression under the same conditions (Supplementary Fig. 4b) and did not lead to increased collagen deposition at the protein level (Supplementary Fig. 4c). These results show that under the observed treatment conditions, PA but not OA induced an inflammatory response and neither induced clear fibrotic injury.

Evaluating fibrotic and inflammatory responses at the single cell level

To dissect cell-type-specific transcriptional injury-response patterns, we treated HLOs cultured on an orbital shaker with either TGF-β1 or PA for five days (**Fig. 3a**) before performing scRNA-seq (*n* = 8 HLO samples, with replicates for each control and treatment condition). We analyzed 44,275 cells after QC (Methods), annotating cell types as described previously and identified hepatocyte-, cholangiocyte-, and HSC-like populations along with cells that more closely align with premature, cycling cells (MKI67+ progenitors), myoepithelial cells and HBs. Database and literature-based strategies overlapped in their annotation of cholangiocyte-like cells in the orbital shaker system, suggesting a less precursor-like state for this population compared to native HLOs at day 21 (**Fig. 1**, Supplementary Fig. 5d). Cell clusters were also identified near the HSC-like population with expression patterns associated with premature HSCs (PHSCs) and smooth muscle cells (SMC), the latter of which may represent a myofibroblast-like population (**Fig. 3b**, Supplementary Fig. 5-6). These results indicate a diversification of mesenchymal cell types in the HLO injury system suggestive of treatment-induced changes on the single-cell level.

To understand which general injury gene expression signatures are activated with each treatment, we evaluated injury-response scores based on gene ontology terms[68]. TGF-β1 treatment resulted in increased scores for fibroblast activation and smooth-muscle and contractile fiber activity along with alterations in cell cycle phase distributions, while these scores remained unchanged in the PA-treated HLOs. In contrast, PA-treated HLOs showed induction of immune response scores when compared to their controls and were enriched for chemokine-activity associated genes. Both treatments resulted in the reduction of liver-regeneration scores (**Fig. 3c,** Supplementary Table 4). These findings demonstrate the allocation of cells with induced inflammatory scores under PA treatment to the epithelial clusters (hepatocyte- and cholangiocyte-like cells) and the up-regulation of pro-fibrotic scores with TGF-β1 in mesenchymal clusters (HSC- and PHSC-like cells).

We then evaluated how cell type distributions change with each condition. TGF-β1-treated HLOs displayed an increase in HSC-related populations. HSC-like cells increased from less than 10% in

controls (5.7% and 4.9%) to greater than 15% (22.7% and 16.5%) with TGF-β1 treatment, while PHSC-like cells increased from less than 10% (4.4 % and 7.5%) in the controls to around 70% (66.4% and 76.4%) with TGF-β1 treatment. In contrast, hepatocyte-like populations decreased by 10-fold from 19.6% to 1.9%, and a near-total loss of cycling MKI67^{+} progenitor was observed (**Fig. 3d**, Supplementary Table 5), in line with previously reported cell cycle arrest in TGF-β1-treated hepatocytes and other cell types[69,70]. While the analysis shows an expansion of HSC-like cells in response to TGF-β1, as observed in the development of liver fibrosis[71], no significant alterations in cell population proportions were observed with PA treatment.

We also compared the distribution of cell types in the orbital shaker controls and the original HLO differentiation conditions (**Fig. 1**) and observed that culturing for five days in the orbital shaker was associated with a relative increase in hepatocyte- and cholangiocyte-like cells and a relative reduction in HSC-like cells (Supplementary Fig. 6c). These results suggest that the orbital shaker culturing method improves the ratio of epithelial to mesenchymal lineages resembling more closely the physiological proportions observed in the human liver[26,53,57].

PHSC-, SMC-, and myoepithelial-like cells cluster near HSCs (**Fig. 3b**), and these clusters share expression of many of the genes characteristic of HSCs (**Fig. 3e**). The cluster annotated as PHSCs is characterized by an HSC-like signature but does not express collagen genes at the level observed for HSC-like cells, instead marker genes of early progenitor cells *CD59, EPCAM, ITGB1, PERP* and *KRT8* are among the differentially expressed genes (Supplementary Table 6). These cells also express *CLDN4, KRT7* and *KRT19* (**Fig. 3e**), genes that in the liver specify cholangiocytes. In the context of a general upregulation of immaturity markers in this cluster and the physiological presence of the respective transcripts in placental cells[72–74], this is in line with the acquisition of a premature phenotype of HSC-like cells in response to TGF-β1.

Differential gene expression analysis per cell cluster further displayed injury-specific signatures depending on the respective cell type. The fibrotic module of differentially expressed genes included *COL1A1, TAGLN2* and *TGFBI* and was induced across TGF-β1 treated cell types, with HSC- and

PHSC-like cells showing the strongest relative induction (**Fig. 3f**). The inflammatory module of differential genes included Interleukins *IL1* and *IL32* (the top up-regulated gene in NAFLD[75]) and other canonical inflammatory genes including *NFKBIA* and *CCL20* (**Fig. 3f**, Supplementary Tables 7-13), showing the strongest relative induction in hepatocyte- and cholangiocyte-like cells treated with PA, followed by HSC- and SMC-like populations (Supplementary Fig. 7). To contextualize differential gene expression more broadly, we performed cell-type resolved pathway analysis and found TGF-β1 hepatocyte-like cells enriching for TGF-β- and virus-associated pathways (being related to differentially expressed genes *ITGB1, APP, CDKN2B, FN1, JUNB, SKIL* (TGF-β) and *ITGB1, GSN, NPC2, CLTB, HLA-C, CLTA, RHOB* (virus), respectively) while the PHSC-like cells and cholangiocyte-like cells enriched for ECM- and cytoskeleton-associated pathways (**Fig. 3g**, Supplementary Fig. 8). Strikingly, PA-treated hepatocyte-, cholangiocyte-, and HSC-like cells displayed the nonalcoholic fatty liver disease (NAFLD) pathway among their top four ranked enriched pathways, associated with genes including *NDUFA13, NDUFA11, UQCRQ, UQCR11* and cyclooxygenases *COX7A2* and *COX7C*. These analyses show that TGF-β1 and PA each induce different patterns of gene expression across different cell types that mirror fibrotic injury and fatty liver-associated inflammation, respectively.

HLOs shift crosstalk towards HSC- and cholangiocyte-like cells inducing specific interactomes with TGF-β1 and PA treatment

To understand the cell-cell interactions in TGF-β1 and PA HLO injury models we utilized CellPhoneDB[76] and identified ligand-receptor pairs exclusive for each injury type. In TGF-β1-treated HLOs, receptor-ligand expression was observed involving *TNFSF12* and *PDGFB*, which recapitulates findings from analyses on scar-associated macrophages in fibrotic human liver[26] (**Fig. 4a**, Supplementary Table 14). With PA treatment, receptor-ligand expression was observed that involves *WNT7B, LGALS9* and *CD55*, reflecting the induction of previously described molecular mediators of cellular responses to injury and inflammation in the liver[77–79]. We further identified *ICAM1* and *JAM2* containing ligand-receptor pairs shared between both liver injury entities. This general evaluation of the interactome reveals groups of receptor-ligand pairs changing in response to inflammatory and

fibrotic injury in HLOs overlapping with previously reported molecular mediators *in vivo*.

We assessed interactions between each cell type and found HSC-like cells to constitute the most interactive cluster based on receptor/ligand co-expression (**Fig. 4b**). In controls and following treatment with TGF-β1 and PA, HSC- and cholangiocyte-like cells showed the strongest interaction. The number of interactions also increased with TGF-β1 but not PA treatment. We also evaluated condition-specific patterns of receptor-ligand pair expression (**Fig. 4c,** Supplementary Fig. 9-10). This analysis allowed us to explore "hot-spots" of injury-specific interactions observed between HSC-like and cholangiocyte-, hepatocyte-, and other HSC-like cells. These interactions revealed new collagen-integrin pairs with TGF-β1 treatment. This includes multiple collagens that can interact with integrin α1β1 on cholangiocyte-like cells and a shift from collagen-integrin α1β1 receptor-ligand interactions between HSC- and hepatocyte-related cells to collagen-integrin α2β1 interactions (**Fig. 4d**). These findings suggest that TGF-β1 signaling induces expression of genes that can promote cross talk between HSCs and cholangiocytes and between HSCs and hepatocytes through integrin signaling.

Expression of members of the Ephrin receptor A (EPHA) family and its ligands, ephrines (EFN) characterizes a subpopulation of cholangiocytes with migration potential and has been linked to filopodia formation in cholangiocytes[80-81]. We found that expression of receptor/ligand pairs for ephrin signaling are also regulated by TGF-β1. *EPHA4* expression is paired with multiple ephrin ligands in cholangiocyte-like cells following TGF-β1 treatment (**Fig. 4e**), potentially relating to the oberserved surface alterations in these HLOs (Supplementary Fig. 4a). In addition, TGF-β1 treatment leads to expression of *EPHB2* in cholangiocyte-like cells and expression of the gene encoding the ligand EFNA5 in HSC-like cells, indicating the presence of a receptor pathway previously reported to promote hepatic fibrogenesis *in vivo*[82,83]. These results suggest that TGF-β1 induces expression of genes in the ephrin pathway that can promote autocrine signaling in cholangiocytes and communication between HSCs and cholangiocytes.

With PA treatment, hepatocyte-like cells express the Notch-ligand *DLL1* in conjunction with *NOTCH1-3* in HSC- and cholangiocyte-like cells (**Fig. 4f**). We found PA treatment induced expression of free-

fatty acid receptor *FFAR2* on hepatocyte-like cells in relation to the ligand *FAM3C* in hepatocyte-like cells themselves, HSC-, and cholangiocyte-like cells. Together, these findings suggest that PA induces crosstalk between multiple cell types involving pathways linked to liver disease[84–89] and reveal potential molecular targets in fatty acid-induced liver injury.

Trajectory inference reconstructs the emergence of major HLO lineages and injury-specific terminal states

The development of liver fibrosis from steatohepatitis can be understood as a sequential process over time with fibrosis following steatohepatitis[6]. Therefore, we investigated the projection of our steatohepatitis model (PA) and our fibrosis model (TGF-β1) on a force-directed layout for generating continuous trajectories between discrete time points[90]. We first visualized our previously annotated cell types across the trajectories (**Fig. 5a**, I.) before analyzing differential trajectories in control-, PA- and TGF-β1-treated HLOs (**Fig. 5a**, II.-IV.). PA-treated HLOs displayed a distinct trajectory towards hepatocyte-like cells (**Fig. 5a**, III., dashed circle). TGF-β1 treated MKI67+ progenitor-like cells related to the HSC- and SMC-like clusters mainly via the PHSC population (**Fig. 5a**, IV., dashed circle), while connections to HSC-like cells through the hepatocyte- and cholangiocyte-like subfractions shared among healthy and PA-treated HLOs were not observed, adding evidence that TGF-β1 not only changes cell type proportions favoring HSC-like progenitors, but also the developmental axis through which they emerge.

To further investigate the directionality of these trends, we overlaid Palantir[91] selecting an *MKI67*-high progenitor-like cell as the early cell. As expected from the immature gene expression signature of the PHSCs, cells corresponding to this cluster aligned earlier along the pseudotime representation than their HSC counterparts. SMC- and hepatocyte-like cells corresponded to regions at the end of the pseudotime path (**Fig. 5a**, V., arrow heads). Unsupervised analysis inferred six terminal states (**Fig. 5a**, VI.) among which one hepatocyte-like cluster was dominated by PA-treated cells and the PHSC cluster represented mostly TGF-β1-treated cells. We next identified cells with high differentiation potential (**Fig. 5a**, VII.), showing enhanced activity in four defined regions. Cells with the highest

differentiation potential were detected at the interface of premature cells identified as HB-like cells close to the hepatocyte-like terminal states and mesenchymal populations (**Fig. 5a**, VII., arrow head 1), potentially pointing at a hepatomesenchymal transition state previously described in normal development[92]. A second cluster of intermediate-high differentiation potential aligns with the early bidirectional fate decision towards either the hepatocyte-like or the mesenchymal, PHSC-like state (**Fig. 5a**, VII., arrow head 2). A third region of intermediate-high differential potential is present at the PHSC-like – HSC-like interface (**Fig. 5a**, VII., arrow head 3), and a fourth at the transition from HSCs towards SMC-like cells, whose transcriptional signature is closely related to myofibroblasts (**Fig. 5a**, VII., arrow head 4), consistent with a canonical feature of fibrosis pathology[24,71].

To examine the differentiating regions as potential representations of healthy-to-injury transitions, we next applied MAGIC[93] to generate imputed data and visualize the expression of relevant genes along the differentiation trajectory with a focus on the acquisition of “pro-fibrotic” and “pro-inflammatory” gene expression signatures (**Fig. 5b**). Inflammatory transcripts *IL32, NFKBIA*, *CCL20* and *CXCL1* showed a bimodal expression signature increasing towards inflamed hepatocyte-like cells with PA treatment and PHSC-like cells with TGF-β1 treatment. The PA-treated hepatocyte terminal state showed enrichment for fetal hepatocyte transcripts *AFP, APOM, APOA1* and *APOB*. We observed a targeted diffusion distribution of pro-fibrotic genes towards HSC-like cells, corresponding to the TGF-β1-treatment enriched trajectory via the PHSC-like cluster (e.g., *COL1A1*, *TGFBI*, *IGFBP3*). Specifically, myofibroblast-characterizing transcripts *ACTA2* and *TIMP1* increased towards the SMC-like cluster, again suggesting these cells could represent an HSC-derived myofibroblasts-like population. These transcripts are also characteristic for activated HSCs, a hallmark of fibrosis *in vivo*[71], again suggesting the trajectory with enhanced differentiation potential connecting HSC- and myofibroblast-like populations (**Fig. 5a,** VII., arrow head 4) represents the activation of these HSC-like cells.

To better understand gene expression signatures accounting for deviating terminal states observed in injured HLOs (**Fig 5a**, V.-VI.), we next inferred gene expression trends and opposed hepatocyte-like and HSC-like cells with their injured counterparts (**Fig. 5c**, left: hepatocyte gene trends, right: HSC

gene trends). As expected, *MKI67* decreased towards all terminal states (**Fig. 5c**, I.). We observed a clear shift towards inflammatory (**Fig. 5c**, II.) and fibrotic genes (**Fig. 5c**, III.) along pseudotime in PA-specific hepatocyte-like cells and TGF-β1-specific PHSC-like cells (red arrows) when comparing them to their general population counterparts, which included the control cells (black arrows). A relative increase in inflammatory and fibrotic gene expression is evident in terminal states specific for PA and TGF-β1 treatment. The general population of hepatocyte-like cells used as a control does contain a mix of control and PA-treated cells, and the general population of HSC-like cell controls contain a mix of control and TGF-β1 treated cells (**Fig. 5a**), which could also reduce the differences observed between control and injury conditions. Projecting lineage-specific gene expression trends over pseudotime (**Fig. 5d**) allowed for the identification of inflammatory genes (**Fig. 5d**, left column) plateauing or slightly decreasing in TGF-β1-specific PHSC-like cells (brown line), while fibrosis-related genes (**Fig. 5d**, right column) peaked towards the end of pseudotime, suggesting a transition through inflammatory states towards the fibrotic phenotype in these cells. PA-treated hepatocyte-like cells (red line) showed a constant increase of inflammatory gene expression with only small increases in fibrotic signatures towards the end of pseudotime, suggesting an inflammatory terminal state with an emerging fibrotic phenotype for this population. Pseudotime is indicated on the x-axis, with a value of 1 representing the greatest pseudotime (SMCs). Other cell populations terminate before 1 based on pseudotime calculations for each terminal cell state. Together, these results support the presence of intermediate inflammatory stages on the path towards fibrosis.

Candidate gene signatures predicting disease progression in NAFLD progressively increase with treatment of PA and TGF-β1

We next investigated how changes in gene expression following treatment of HLOs with PA and TGF-β1 relate to the development of fibrosis in patients with NAFLD. We applied the 26 and 98 gene signatures established to predict fibrosis in NAFLD[94] to bulk gene expression in HLOs. This analysis revealed a significant increase in the score for both gene signatures with PA and TGF-β1 treatment, with the mean score increasing to the greatest extent with TGF-β1 treatment (**Fig. 6a**). We then

plotted expression signatures for individual genes that demonstrated the greatest dynamic range in expression across clusters (**Fig. 6b**). Many of these genes show a trend of induction from control, to PA, to TGF-β1, including *CXCL6*, *GSN*, *IL32 and TAGLN2*, while other genes, such as *S100A4* show the highest expression with PA treatment. This indicates the gradual induction of NAFLD disease progression scores in the HLO system.

We then investigated the 26 and 98 gene signatures at the single cell level in HLOs to understand the cell types responsible for each signature. The 26 gene signature was increased in hepatocyte-, smooth muscle-, myoepithelial-, cholangiocyte-, HSC-, and MKI67^{+} progenitor-like cells, suggesting these clusters all contribute to the fibrosis signature induced by PA or TGF-β1 treatment of HLOs. The 98-gene signature is significantly up-regulated in all TGF-β1 treated cell types, while PA treatment was associated with an increase in the 98 gene signature only in MKI67^{+} progenitor- and myoepithelial-like cell types (**Fig. 6c**). This again is in line with a gradual acquisition of fibrosis severity, where PA treatment does not induce the full spectrum of fibrosis-related genes observed with TGF-β1.

To further dissect the contribution of individual cell types to the score enrichment, we next plotted the relative expression of individual genes per cell cluster across treatment conditions (Supplementary Fig. 11). Hepatocyte-like cells in HLOs treated with TGF-β1 showed higher expression of *BTG2, EHD4* and *GSN*, while PA induced *IL32* and *S100A4* in this population. HSC-like cells showed induction of *COL4A1*, *COL4A2, COL5A1, PKM, VIM,* and *TPM4* with TGF-β1 treatment, while PA treatment resulted in reduced expression of these signature genes. Cholangiocyte-like cells showed an increased expression of signature genes *CXCL6* and *IL32* in both injury scenarios while *MEAF6, GSN, S100A11, PGP, TMP4, VIM,* and *TAGLN2* showed higher expression only in the TGF-β1 condition. A gradual increase in expression levels from control to PA to TGF-β1 in PHSC-like cells was observed for *CDC42SE1, CXCL6, GSN, PGP, PKM, S100A11, TAGLN2, TMEM51, TNFAIP8,* and *VOPP1*. These results identify the individual cell types that contribute to the gene signature linked to disease progression in NAFLD.

Discussion

Here, we report that multicellular human liver organoids (HLOs) can model progression of steatohepatitis and liver fibrosis when maintained under specific 3D culture conditions. While studies have started to decipher the contribution of individual cell types to the response to liver fibrosis humans[26,95], our understanding of how cell types interact during evolving liver injury, including their temporal order of activation, is limited. Systematic phenotypical, histological, and scRNA-seq analysis in the HLO system in this study indicates a high degree of specificity and temporal orchestration in the responses of individual cell populations to different injury types.

PA-treated HLOs as models of liver inflammation display altered phenotypes and induced *TNFA* expression. The general cell type distribution does not change with PA treatment, but gene scoring reveals enhanced signals for transcripts related to immune response and chemokine activity, while fibrosis-related scores are not induced. This is mainly driven by the hepatocyte- and cholangiocyte-like populations, followed by HSC-like cells, up-regulating not only the highly specific NAFLD cytokine *IL32*[75] but also canonical inflammatory pathways related to NFκB. All subpopulations except for the myoeptihelial-like cells additionally are enriched for NAFLD related transcripts including cyclooxygenase-coding genes *COX6* and *COX7* and NADH-ubiquinone oxidoreductase subunits *NDUFA3, NDUFA11* and *NDUFA13*. Unique cell-cell cross talk defining interactions including *FFAR2/FAM3C* and *DLL1/NOTCH1-3* are induced. While most inflamed HLO cells align on developmental trajectories with their respective controls, PA-treated hepatocyte-like cells deviate, giving rise to a distinct terminal state. Grading PA-treated HLOs with clinically validated NAFLD-fibrosis scores shows a moderate, but significant increase driven by gene expression across cellular subpopulations, including hepatocyte- and SMC-like (most likely myofibroblasts-reflecting) populations. Taken together, transcriptomic data generated from PA-treated HLOs capture an inflammatory transition at single-cell resolution that is driven primarily by hepatocyte-like cells.

TGF-β stimulation in the same HLO system triggers more pronounced changes in HLO morphology with increased collagen deposition and an expansion of the HSC-like population consistent with

observations in chronic human liver disease, and scRNA-seq profiling revealed a premature transcriptional signature of the most significantly expanding HSC-subpopulation (PHSC), defined by the upregulation of *CD59, EPCAM, ITGB1* and *KRT8*. *TNFSF12* and *PDGFB*-driven receptor ligand interactions are induced as previously shown for human fibrotic livers[26], and cholangiocyte-like and HSC-like cells are the most interactive partners in the fibrotic HLO. PHSC-like cells emerge as a distinct terminal state on the liver injury trajectory. The progression of fibrosis is classified as more severe across all cell populations when applying the clinically validated gene signature score on TGF-β1 treated HLOs.

While many genes shaping the trajectories towards the injured terminal states overlap between PA- and TGF-β1 mediated injury, sequentially induced expression levels and the bifurcated anatomy of the pseudotime trajectory representation stress the different cell types (hepatocyte- versus PHSC-like cells) and transdifferentiation events that preserve the injury-response specificity in the HLO system and correspond well with *in vivo* observations finding maintained inflammatory hepatocyte activation preceding HSC-to-myofibroblast differentiation events in liver fibrosis development[1,2]. Taken together, these analyses suggest that PA and TGF-β1 treatment in HLOs recapitulate the sequentially evolving cell states of inflammatory and fibrotic liver injury established for NAFLD pathology. In addition, HLOs not only develop fibrosis-like pathologies, but are also capable of mirroring gradual transcriptome-wide injury responses corresponding to fibrosis severity classes in NAFLD patients. However, the current HLO system lacks immune cells, and it is possible that the introduction of cells such as macrophages would also promote the transition from inflammation to fibrosis.

To conclude, we present a comprehensive scRNA-seq evaluation of a stem-cell derived multi-cellular human liver organoid system capable of mimicking differential injury-responses in a clinical disease progression-like manner. We build on previous developments in human liver organoid models yielding multi-cellular systems demonstrating the feasibility of deriving injury scenarios of genetically-programmed, ADPKD-like fibrosis[39] and oleate-mediated steatohepatitis[33,37]. We explore multiple culture conditions and identify their contribution to generating physiological cell-type distributions towards the inducibility of multiple, stressor-specific liver injury responses. We utilize scRNA-seq to

validate our multi-injury system and contribute to current efforts of establishing frameworks for the systematic benchmarking of HLOs as state-of-the art *in vitro* model systems overcoming previous limitations in 2D culture and non-human animal models[40].

Our multi-injury liver organoid system demonstrates the temporal emergence of transcriptionally altered cell populations in the course of liver injury and defines specific molecular pathways involved in the early cell-cell interactions preceding broad pro-fibrotic alterations of the HLO cellular landscape. This flexible, perturbable system will facilitate the study of liver injury development, enable stressor-specific testing of therapeutic candidates, and provide a baseline reference atlas for the benchmarking of novel liver injury organoid models in the future.

Methods

hPSC and HLO culture. H1 hPSC cells (WA01) were obtained from WiCell. The ESCRO approval was received from Massachusetts General Hospital. hPSCs were maintained in mTeSR medium (StemCell Technologies) on Matrigel (Corning, 354230) as previously described[96] and split with Accutase (Thermo Fisher Scientific) upon 80-90% confluence. For passage and thawing, Rho-associated kinase (ROCK)-Inhibitor (StemCell Technologies, Y-27632) was added to mTeSR at 10 µM. For the differentiation of human liver organoids (HLOs), cells were split at a ratio of 1:10 and cultured until ~85% confluent before HLOs were differentiated as previously described[33,34]. On day 20, HLOs were either kept in Matrigel domes or isolated. To perform isolation, HLOs were incubated in DPBS(-/-) on ice for 15 minutes followed by repeated manual dissociation with a P1000 at 4°C followed by centrifugation at 190 x *g* for three minutes, removal of supernatant, and resuspension in 10 µL final medium per receiving well of a 24-well-plate. Isolated HLOs in media were either plated to i) 1% agarose coated plates ii) plates on an orbital shaker at 80 revolutios per minute, or iii) hydrogel plastics ultra low attachment surface 24-well plates (Corning, 3473).

Liver injury induction. Injury media solutions based on Hepatocyte Culture Medium (HCM, Lonza, CC-3198) were prepared to obtain solutions of TGF-β1 (10 ng/mL and 25 ng/mL, R&D Systems 240-B-002), palmitic acid (PA, 500 µM, Sigma Aldrich P0500), and oleic acid (OA, 400 and 800 µM, Sigma Aldrich O1383). HLO media was changed to injury media on day 20-21. Isolated HLOs were resuspended in the final injury media at the final step of the isolation process (see above) and kept in solution for five days. Controls for TGF-β1 were cultured in HCM. PA was dissolved into HCM with 10% BSA and 1% ethanol before dilution to final concentration in HCM. OA was dissolved into DPBS(-/-) with 12.5 mM NaOH and 1.67% BSA at 8 mM before dilution to final concentration in HCM. PA and OA controls were prepared accordingly, omitting the initial step of dissolving the active agent in the carrier solutions.

Contraction assay. Day 20-21 HLOs in 50 µL Matrigel domes cultured in HCM were exposed to TGF-β1 at 10 ng/mL and 25 ng/mL. Medium was changed every three days. On day 0 and 5 of incubation, images of plates with scale bars were taken. Images were loaded into ImageJ (version 1.53a), with the scale set based on the known distance. Subsequently, the circular HLO/matrix drop areas were manually selected and measured. All distances in mm^2 were normalized to the mean of the control areas in mm^2, and two-tailed Kruskal-Wallis-statistics were performed on all groups with a Conover post hoc test.

qPCR. Total RNA was isolated from HLOs to perform qPCR. Briefly, 1 mL medium was removed from 6 well plates. Matrigel domes containing HLOs were carefully detached from the wells with a cell scraper and transferred to a 1.5 mL tube including medium. If isolated HLOs served as starting material, the detachment step was not needed. Matrigel-embedded HLOs were left at 4°C for 15 minutes to dissolve the Matrigel. Then, HLOs were centrifuged for three min at ~190 x *g* at 4°C (Matrigel-embedded HLOs) or at RT (non-matrigel HLOs or 2D cell layers). Supernatant was discarded, and the pellet was resuspended in 300 µL Trizol (Invitrogen, 15596026), rigorously vortexed or mechanically processed until fully dissociated, and finally incubated for 10 minutes at room temperature and then stored at -80°C. RNA was prepared via phenol-chloroform extraction. For cDNA synthesis, 200 or 500 ng RNA were reverse-transcribed using the iScript gDNA Clear cDNA Synthesis Kit (Biorad, 1725034), and no-RT-controls and mastermix controls were prepared. qPCR reactions were prepared from cDNA at 1:5 dilution with SYBR Green iTaq Universal SYBR Green Supermix (Biorad, 1725120) and qPCR primers at 10 µM in a total volume of 10 µL in a ≥ 40-cycle and melt curve reaction cycler (Biorad, Base #CT009383, Optical Head #786BR2648). All biological samples were measured in technical triplicates.

qPCR data analysis. All gene Ct means were calculated from three technical replicates. Only samples with housekeeping Ct values below 35 were considered for analysis. Mean Ct values for housekeeping genes β-Actin (*ACTB*) or glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) were subtracted from target gene means to generate the ΔCt value. The control ΔCt value average was calculated from three biological replicates and subtracted from all experimental ΔCts to render ΔΔCt values. The log2 fold-change was calculated as $2^{(-\Delta\Delta Ct)}$.

Primer design and validation. Primers were designed using PrimerBlast. Briefly, transcripts of interest were selected by NCBI Reference Sequence ID and the 'pick primers' hyperlink function was used to import the cDNA sequence into the PrimerBlast mask. The following settings varied from the defaults: Exon junction span → primer must span an exon-exon junction, PCR-product size → 50-200, allowing primers to bind to variant transcripts of the same gene. Blast was performed against both, RefSeq mRNA (Homo sapiens) and RefSeq representative genomes (Homo sapiens) in order to obtain primers specific to transcript sequence and mRNA

rather than gDNA. Primers with i) no additional match in RefSeq mRNA or at least > 4 mismatches and ii) no additional genomic hits or hits with at least > 800 bp product size, were selected for further validation. Next, primers were tested on human liver tissue RNA at a final concentration of 10 µM as described above, and 2% agarose gels of the primer product were prepared with a 100 bp ladder. Primers were selected for further experiments if they showed a single band in the expected size range. The band was cut, cDNA was extracted from agarose and sent for Sanger sequencing, and the product identity was assured by re-blasting. A primer was selected for experiments when the sequence matched the selected sequence of the initially selected transcripts (NCBI BLAST). Primer sequences are provided in **Supplementary Table 15**.

Histology and immunohistochemistry. Medium was removed from day 16, day 21 or day 25-26 HLOs treated for five days with pro-fibrotic substances, washed twice with warm DPBS(-/-) and fixed overnight at 4°C with 4% PFA in DPBS(-/-). PFA was removed and the HLOs were washed 2x with DPBS(-/-) before being transferred to 70% ethanol for storage and paraffin embedding. H&E and Sirius red staining was performed on cut and deparaffinized HLOs in line with standard protocols. Anti-human-CEBPα (Sigma Aldrich, HPA052734) and the secondary antibody (HPR anti-mouse) were used at a dilution of 1:200.

Organoid Sirius red quantification pipeline. The pipeline for HLO-specific Sirius red staining analysis is written in python (≥ version 3), ImageJ (version 1.53a) macro and bash and will be made publicly available upon release at https://github.com/anjahess/sirius_red. First, full slides were cleared from black artifacts. Then, for whole slide quantification, all image areas containing tissue were selected and saved for quantification. The Sirius red quantification module was adopted from the NIH-Image J macro “Quantifying Stained Liver Tissue“ (available at https://imagej.nih.gov/ij/docs/examples/stained-sections/index.html, requested 2020-11-14). Briefly, RGB stacks were generated from images, and the tissue containing area was defined and measured by the *setAutoThreshold("Default stack")* function. The Sirius red staining threshold reached from minimum to maximum as retrieved from the *setAutoThreshold* and *getThreshold(min, max)* functions measured in the green channel of the RGB stack, and divided by an experiment-specifc brightness factor (0.95-1.3). For single HLO analysis, 200 x 200 µm selection squares were placed around HLOs in the interactive, user-supervised mode and resulting images were processed as described for whole slide scans. Finally, csv-formatted result tables were called from python (version 3.7) and the Sirius red stained area was calculated either per total area or per tissue area (measures as the auto-thresholded area in the blue channel). If normalization was performed, values were calculated as percentage of mean of all control samples of the respective experiment.

Single-cell RNA sequencing. For day 21 and 34 HLO single-cell RNA sequencing, HLOs underwent a passage step at day 16 and remained in HCM medium with 10% Matrigel on ultra-low attachment plates as previously described[34]. At day 21 and 34, one well of a six well plate of free-floating HLOs was collected, washed 1X with warm DPBS(-/-) and briefly dissociated by incubation with 300 µL 0.05-0.25% Trypsin-EDTA for 10 minutes at 37°C (Corning, Ref. 25-052-CI RT). Day 34 and day 25-26 orbital shaker-cultured HLOs underwent an additional DBPS (-/-) wash and 10 minutes of trypsinization to allow complete HLO digestion. For each replicate, full single cell dissociation was confirmed manually by light microscopy. After a final DPBS(-/-) wash, HLOs were resuspended in DPBS(-/-), counted after Trypan blue staining with the TC20™ Automated Cell Counter according to a previously determined optimal gating of 7-20 µm and transferred on ice. Library preparation was performed on biological replicates with the highest viability counts. Libraries were prepared according to the 10X ChromiumSingle Cell 3'Reagent Kits v3 instructions. The library was sequenced by paired-end sequencing on an Illumina® NextSeq 2000 P3 flow cell or the NovaSeq 6000 system according to the manufacturer's recommendations.

Single-cell RNA sequencing analysis. Raw bcl files were converted to fastq by the command *bcl2fastq --use-bases-mask Y26,I8,Y98*. Fastq quality was assessed with multiqc[97]. Fastq files were aligned to the GRCh38 genome with *cellranger count,* cellranger (=> version 3.0.2). Doublet scores were calculated with Scrublet[98] assuming 0.06 percent doublets at a calling threshold of 0.2-0.22 on the cellranger filtered count matrices (manually adjusted based on bimodal histogram distribution). Downstream analysis was performed with scanpy[99] (version 1.7.2, for functions abbreviated with "*sc*" from here on) in python (version 3.7). The following cellular quality control criteria were applied to each replicate after manual inspection of violin and scatter plots of the unfiltered inputs: (i) 100 features per cell, (ii) a maximum mitochondrial gene fraction of 50% (iii) genes-by-count ratio maximum of 8500. Transcripts were accepted if present in at least three cells. Subsequently, technical replicates and, when appropriate, replicates of the conditions to be compared were merged using the scanpy *concatenate* function with default settings. Concatenated datasets were total-count normalized (target sum 10,000), logarithmized (X=log(X+1), natural logarithm), and a total of 4000 highly variable genes were selected for downstream analysis. The data was scaled to a maximum value of 10.

Cell cycle scoring. Cell cycle scores (S-, G2M-Score) and cell cycle phase (S, G2M, G1) were assigned based on a previously published cell cycle defining gene list[100] with the *sc.score_genes_cell_cycle* function.

Dimensionality reduction, embedding, clustering, and marker gene identification. Normalized, log-transformed and scaled data were objected to calculation of principal component analysis (PCA) coordinates,

loadings, and variance via *sc.tl.pca*. The neighborhood graph was calculated using the first 50 principal components and embedded utilizing the Uniform Manifold Approximation and Projection (UMAP)[101] algorithm via *sc.api.tl.umap*. Clusters were identified with the Leiden algorithm[102] at resolution 0.1 or 0.5. Coarsenesses lower than the default parameters were chosen to reproduce the underlying biology of approximately five cell types experimentally validated for HLOs[33]. Finally, cluster-characterizing genes were defined using scanpy's implementation of the Wilcoxon rank-sum test followed by Benjamini–Hochberg correction (*sc.tl.rank_genes_group*).

Batch correction. After merging and normalization, samples were manually controlled for batch effects by overlap inspection of PCA and UMAP plots. Batch effects were corrected with the scanpy implementation of Harmony[103] applying the *sce.pp.harmony_integrate* function followed by recalculation of the neighborhood graph and UMAP based on the Harmony representation. Even replicate distributions were assured by generation and inspection of barplots and UMAP plots indicating replicate proportions by cell cluster (Supplementary Figures 2,5-6).

Cell cluster annotation. *Automated cluster annotation based on literature markers.* Literature research was performed to assemble a set of specific marker genes for cell types potentially emerging in HLOs, including mature liver cells and progenitors (**Supplementary Table 1**). Wilcoxon Rank sum test derived marker gene lists (cf. above) were selected for the top 70 marker genes per cluster and forwarded to the *gseapy.enrichr*[104] function together with the literature reference list and the number of genes in the dataset was provided as background parameter. Results from *enrichr* with a *p*-value below 0.05 were assigned as the new cluster identity based on the number of significantly enriched marker genes. If no hits were found that met the *p*-value criteria, the label with the lowest adjusted *p*-value was chosen as the identity. No gene hits resulted in the label "unannotated". *Automated annotation based on curated databases.* The PanglaoDB[51] marker gene set (version 27 March 2020) was downloaded and reduced to human transcripts and subgroups relevant in the context of human liver. Wilcoxon Rank sum test derived marker gene lists were selected and annotated as described above. One main cluster remained unannotated and showed enriched expression of *CENPF, MKI67, UBE2C* and *TOP2A*. Comparing against an additional reference, CellMarker[52] (wget http://biocc.hrbmu.edu.cn/CellMarker/download/Human_cell_markers.txt), resulted in successful annotation as MKI67+ progenitor cells and the respective marker gene set complemented the reference for all subsequent annotations (**Supplementary Table 2**).

SingleCellNet annotation. *Generation of a reference data object.* For comparing HLO cells to human liver cells, human 10X scRNA-seq data were chosen as a reference[57]. Respective *cellranger count* outputs were downloaded via: wget ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE156nnn/GSE156625/suppl/GSE156625%5FHCCFgenes%2Etsv%2Egz; wget ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE156nnn/GSE156625/suppl/GSE156625%5FHCCFmatrix%2Emtx%2Egz; wget ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE156nnn/GSE156625/suppl/GSE156625%5FHCCFbarcodes%2Etsv%2Egz and loaded into scanpy via *sc.read_10x_mtx* to obtain an *AnnData* object with raw count data. To add cell types defined by the authors, the public *HCCF1F2.h5ad* file of the respective dataset was downloaded from: https://data.mendeley.com/public-files/datasets/6wmzcskt6k/files/8ecdfe82-955a-40bd-b6b7-50edd67e9d31/file_downloaded and relevant metadata columns (*"louvain", "NormalvsTumor"*) were extracted from the *adata.obs* dataframe, integrated into the previously generated *AnnData* object, and were added to the raw count data composing the new *AnnData* object. *SingleCellNet classification of HLO-cell types based on human reference data.* Using pysinglecellnet (version 0.1), the created *AnnData* object was defined as the reference dataset. A classifier was trained with the top 15 genes and gene pairs, respectively. Genes were not limited to highly variable genes. Subsequently, each HLO sample was classified. To this end, each cellranger filtered feature matrix raw output (*matrix.mtx, features.tsv, barcodes.tsv*) was loaded into scanpy as an individual query sample and underwent classification. The annotated cell type was stored as a categorical named *"SCN_class"* in the *adata.obs* slot from where it could be retrieved for downstream analyses.

Inflammatory gene and fibrosis scoring. To render gene lists for biological processes related to fibrotic and inflammatory injury scenarios, publicly available GO term[68,105] gene lists were browsed via Amigo2[106] filtering for Homo sapiens genes only and the *"bioentity_label"* column was exported. All GO terms are listed in **Supplementary Table 4**. Cell cycle scores were calculated as described above. Duplicate values were removed from lists prior to scoring. CYP450 superfamily scores were calculated based on a custom gene list and included CYP450 members and the transcript coding for the Cytochrome P450 Oxidoreductase (*POR*). Finally, the scanpy function *sc.tl.score_genes* using default options was used for scoring cells.

DGE and pathway enrichment analysis. *Differential gene expression (DGE).* DGE was performed on raw count data for each defined cluster separately based on batch-corrected datasets including all technical replicates of the respective time points or conditions that were compared (Control vs. TGF-β1 10 ng/mL vs. PA 500 µM). For each cluster, differentially expressed genes were rendered applying the Wilcoxon rank-sum test as described above, after exclusion of mitochondrial and ribosomal genes. *Gene Set Analysis.* For each cluster a

hypergeometric enrichment test was performed on the top 150 differentially expressed genes per condition. Subsequently, the python implementation of GSEA/EnrichR[104] in gseapy (version 0.10.7) was applied to render enriched terms. Gene sets used were *"GO_Biological_Process_2018"* and *"WikiPathways_2019_Human"*, the organism was set to *"Human"*.

CellPhoneDB cell-cell interactome. CellPhoneDB[76] (version 2.1.7) with rpy2 (version 3.0.5) was run on *h5ad* files containing log-transformed count data from OS-cultured HLOs treated with TGF-β1 (10ng/ml) or PA (500 µM) and their respective controls to infer cell-cell interactions. Metadata from previously harmony-integrated and cell-type annotated *AnnData* objects of the same conditions were exported to *.txt* files and provided as an input to the *statistical_analysis* function along with *"gene_name"* set as an identifier for genes in the counts data. To generate the count network data and inputs for dotplots, the *plot_heatmap_plot function* was applied, and all plots were created in python (version 3.7) using seaborn[107]. Mean expression indicated by dot size in all figures refers to CellPhoneDB output *"means"*, and is the mean of the respective two individual partner's average expression. To generate Venn diagrams displaying overlapping and distinct significant (p-value < 0.05) receptor-ligand interaction groups, the CellPhoneDB output file *"pvalues"* served as an input, containing p-values for the enrichment of the interacting ligand-receptor pair in each of the interacting pairs of cell types. The list was filtered for pairs with at least one p-value below 0.05 (meaning at least one significant interaction) for each condition. Next, the list of receptor-ligand pairs was grouped by the first interacting partner, which resulted in a total of 142 possible groups for each condition. The overlapping and distinct fractions were calculated (**Supplementary Table 14**), and Venn diagrams generated using *pyvenn* (Adam Labadorf, https://github.com/adamlabadorf/pyvenn).

Harmony timeseries forced layout. Concatenated, quality criteria filtered (cf. above) data from OS-HLOs controls (n = 4), OS-HLOs + TGF-β1 (n = 2), and OS-HLOs + PA (n = 2) were total-count normalized (10E4), log-transformed, and 5000 highly variable genes were selected. Subsequently, an augmented affinity matrix and force-directed layout was calculated with harmony_timeseries via the *sce.tl.harmony_timeseries* function with 1000 components.

Palantir trajectory inference for time course reconstruction of liver injury stages. Harmony augmented affinity matrices served as distances for Palantir[91] trajectory analysis as inputted to *sce.tl.palantir*. The first cell from the MKI67+ progenitor-like cluster in the sorted *adata.var* indices was chosen as an early cell. No terminal or start cell was predefined. Palantir analysis was performed with 30 nearest-neighbors, 25 maximum iterations, scaling of components enabled and 1200 waypoints. Additional plots were generated with palantir (version 1.0.0).

Retrieval of human NAFLD signatures in injured HLOs. Previously published 26 and 98 gene signatures predicting fibrosis in NAFLD[94] were transferred to lists, and each cell from the integrated dataset of all orbital-shaker cultured HLOs (*n* = 4 controls, *n* = 2 PA, *n* = 2 TGF-β1, total 44,275 cells, quality control and processing as described above) obtained a score for the respective gene lists based on raw count data as described for other inflammatory and fibrosis-associated scores. Dot plots were generated with scanpy based on scaled count data for the entire dataset or for each cell type individually splitting into the three treatment groups. For statistical analysis, lists of numerical score values of all cells from each group (control, PA, TGF-β1), either in the full dataset or per cell cluster calculated with the leiden algorithm at resolution 0.5 and annotated with the database approach were forwarded to statstical testing described below.

Statistics and reproducibility. All statistical tests were performed in python (version 3.6) utilizing scipy.stats[108] (version 1.8.0) and scikit-posthocs[109] (version 0.6.7). Statistics comparing two groups were performed using the two-tailed Mann-Whitney U test followed by a Bonferroni correction. For three or more groups, the two-tailed Kruskal-Wallis non-parametric test was applied followed by a post hoc Conover's test with Bonferroni correction. A difference of mean between groups was considered significant at a *p*-value below 0.05. *N* numbers for individual HLO samples in qPCR and immunohistochemistry experiments indicate individual experiments (individual differentiation cycles starting from day 0 hPSCs). For scRNA-seq, *n* numbers or samples refer to individual HLO single cell suspensions that arose from the same differentiation experiment if performed in replicate, e.g. for day 21 HLOs 3 replicates (*n* = 3) were sequenced, which in this case means from one HLO experiment, single cell suspensions from three different wells from the same treatment and time point were harvested and underwent individual library preps.

Data availability. Raw data for single-cell RNA sequencing are available at Gene Expression Omnibus (GEO) under the accession GSE207889 upon release.

Code availability. The following scripts are available upon request and will be made publicly available upon release: Sirius red quantification pipeline: https://github.com/anjahess/sirius_red ; Single-cell RNA sequencing scripts: https://github.com/anjahess/scrnaseq.

Supplementary information

Supplementary information includes 11 figures and 15 tables:

Supplementary Figure 1, Representative images of fixed HLOs, related to Figure 1c.

Supplementary Figure 2, Replicate distribution and marker gene expression from scRNA-seq, related to Figure 1d.

Supplementary Figure 3, SingleCellNet classification of HLO cells and analysis of WikiPathways, related to Figure 1d.

Supplementary Figure 4, Induction of fibrotic phenotypes in OS-cultured HLOs, related to Figure 2b, f, i.

Supplementary Figure 5, Marker gene expression and replicate distribution in 44,275 single cells from control and injured HLOs, related to Figure 3b.

Supplementary Figure 6, Contributions per replicate for TGF-β1 and PA treatment with respecitive controls, related to Figure 3b.

Supplementary Figure 7, Differential gene expression (DGE) for each major cell type after treatment with TGF-β1 and PA compared to controls and gene modules for fibrosis and inflammation, related to Figures 3e-f.

Supplementary Figure 8, Pathway enrichment across cell clusters, related to Figures 3g.

Supplementary Figure 9, Overview of differential cell-cell interactions in OS-HLOs upon TGF-β1 treatment, related to Figure 4c.

Supplementary Figure 10, Overview of differential cell-cell interactions in OS-HLOs upon PA treatment, related to Figure 4c.

Supplementary Figure 11, Expression of genes composing signatures predicting fibrosis stages in NAFLD, related to Figure 6c.

Supplementary Table 1, Literature-based marker genes, related to Figure 1.

Supplementary Table 2, Cell types for database annotation, related to Figures 1 and 2.

Supplementary Table 3, Top 100 cluster-defining genes in day 21 HLOs, related to Figure 1.

Supplementary Table 4, GO terms for scoring inflammation and fibrosis, related to Figure 3.

Supplementary Table 5, Cell type distributions for control, TGF-β1 and PA-treated HLOs, related to Figure 3.

Supplementary Table 6, Database annotation results for PHSC-like cells, related to Figure 3.

Supplementary Table 7, DGE for hepatocyte-like cells, related to Figure 3.

Supplementary Table 8, DGE for cholangiocyte-like cells, related to Figure 3.

Supplementary Table 9, DGE for HSC-like cells, related to Figure 3.

Supplementary Table 10, DGE for PHSC-like cells, related to Figure 3.

Supplementary Table 11, DGE for MKI67+ progenitor-like cells, related to Figure 3.

Supplementary Table 12, DGE for myoepithelial-like cells, related to Figure 3.

Supplementary Table 13, DGE for smooth muscle cell-like cells, related to Figure 3.

Supplementary Table 14, CellPhoneDB receptor-ligand pair groups, related to Figure 4.

Supplementary Table 15, Sequences of human qPCR primers.

Acknowledgments

The authors thank Ramnik Xavier for helpful discussions, Cristin McCabe and Jacques Deguine for assistance with data management, and Adam Slamin, Dan Dubinsky, and the Broad Genomics Platform for help with generation of single cell sequencing data. The authors also thank the Massachusetts General Hospital (MGH) NextGen Sequencing Core for additional assistance with generation of single cell sequencing data and Annika Gabriel for designing color palettes used in the manuscript. A.H. was supported by the Studienstiftung des deutschen Volkes. A.C.M. was supported through internal funding at MGH.

Author contributions

A.H. and A.C.M. conceived and designed the study with additional assistance from S.D.G. Bench experiments were performed by A.H. and S.D.G. with A.B.S. and T.H. providing additional assistance. Computational analyses were performed by A.H. with assistance from R.R. The manuscript was written by A.H. and A.C.M with input from all other authors.

Competing interests

A.C.M. receives research funding from Boehringer Ingelheim, Bristol-Myers Squibb, and Glaxo Smith Klein for other projects and is also a consultant for Third Rock Ventures.

Fig. 1

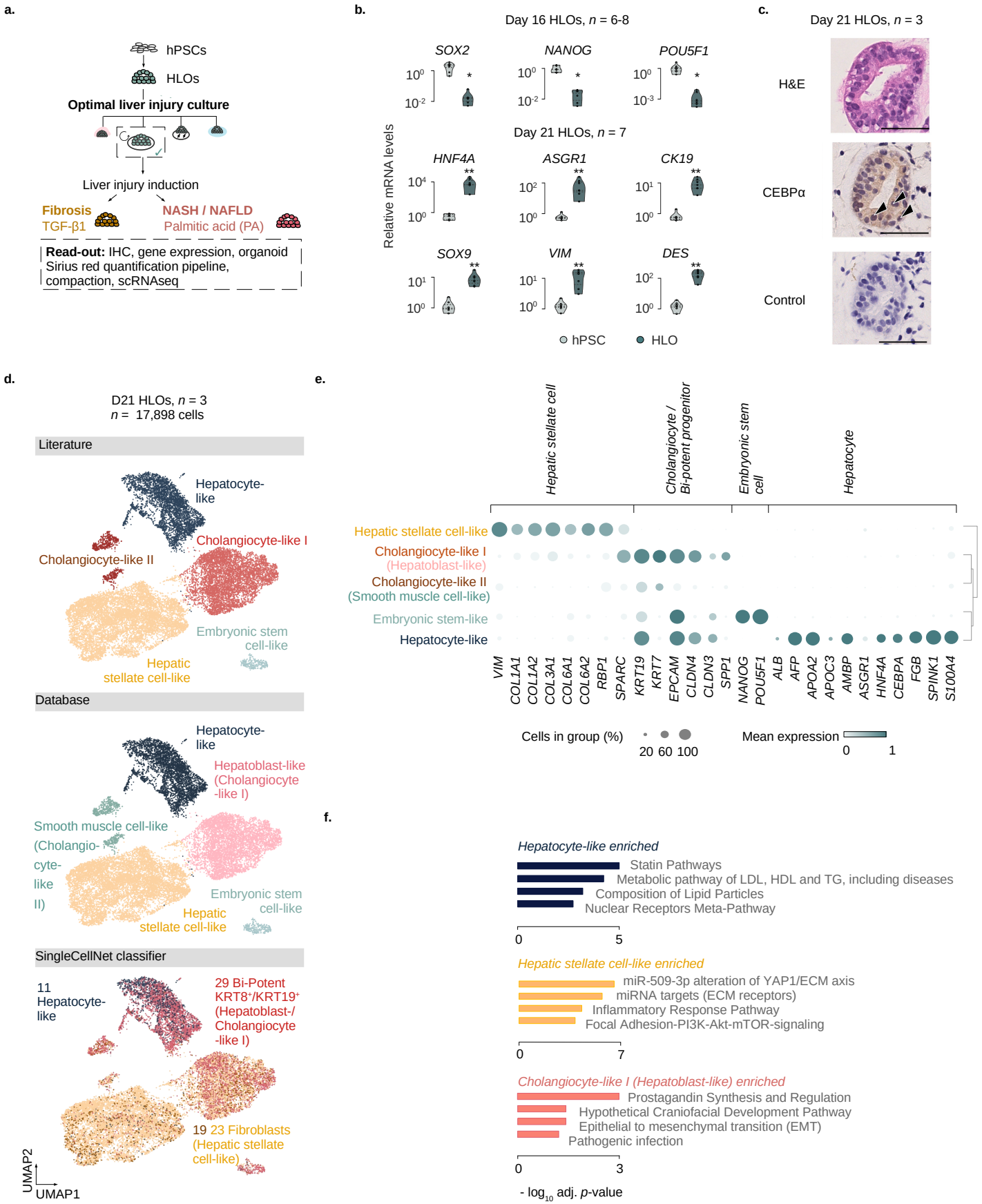

a.
hPSCs
HLOs
Optimal liver injury culture
Liver injury induction
Fibrosis
TGF-β1
NASH / NAFLD
Palmitic acid (PA)
Read-out: IHC, gene expression, organoid Sirius red quantification pipeline, compaction, scRNAseq
b.
Day 16 HLOs, n = 6-8
Day 21 HLOs, n = 7
Relative mRNA levels
SOX2
NANOG
POU5F1
HNF4A
ASGR1
CK19
SOX9
VIM
DES
hPSC
HLO
c.
Day 21 HLOs, n = 3
H&E
CEBPα
Control
d.
D21 HLOs, n = 3
n = 17,898 cells
Literature
Hepatocyte-like
Cholangiocyte-like I
Cholangiocyte-like II
Embryonic stem cell-like
Hepatic stellate cell-like
Database
Hepatocyte-like
Hepatoblast-like (Cholangiocyte -like I)
Smooth muscle cell-like (Cholangio-cyte-like II)
Embryonic stem cell-like
Hepatic stellate cell-like
SingleCellNet classifier
11 Hepatocyte-like
29 Bi-Potent KRT8+/KRT19+ (Hepatoblast-/Cholangiocyte -like I)
19 23 Fibroblasts (Hepatic stellate cell-like)
UMAP2
UMAP1
e.
Hepatic stellate cell
Cholangiocyte / Bi-potent progenitor
Embryonic stem cell
Hepatocyte
Hepatic stellate cell-like
Cholangiocyte-like I (Hepatoblast-like)
Cholangiocyte-like II (Smooth muscle cell-like)
Embryonic stem-like
Hepatocyte-like
VIM COL1A1 COL1A2 COL3A1 COL6A1 COL6A2 RBP1 SPARC KRT19 KRT7 EPCAM CLDN4 CLDN3 SPP1 NANOG POU5F1 ALB AFP APOA2 APOC3 AMBP ASGR1 HNF4A CEBPA FGB SPINK1 S100A4
Cells in group (%)
20 60 100
Mean expression
0 1
f.
Hepatocyte-like enriched
Statin Pathways
Metabolic pathway of LDL, HDL and TG, including diseases
Composition of Lipid Particles
Nuclear Receptors Meta-Pathway
0 5
Hepatic stellate cell-like enriched
miR-509-3p alteration of YAP1/ECM axis
miRNA targets (ECM receptors)
Inflammatory Response Pathway
Focal Adhesion-PI3K-Akt-mTOR-signaling
0 7
Cholangiocyte-like I (Hepatoblast-like) enriched
Prostagandin Synthesis and Regulation
Hypothetical Craniofacial Development Pathway
Epithelial to mesenchymal transition (EMT)
Pathogenic infection
0 3
- log10 adj. p-value

Fig. 1 HLOs recapitulate major components of the adult human liver landscape. a. *Overview of the experimental design for this study.* Human liver organoids (HLOs) are differentiated from human pluripotent stem cells (hPSCs), cultured in four different three-dimensional conditions, and tested for their respective potential to model liver injury (fibrosis induced by TGF-β1 and steatohepatitis by PA). HLOs acquiring an injured phenotype along with untreated HLOs are subjected to scRNA-seq and other readouts. **b.** *Expression of pluripotency genes is reduced with HLO differentiation, while cell-type-specific adult liver markers are induced.* Relative gene expression was measured by quantitative reverse transcriptase (qRT) PCR and normalized to housekeeping gene *ACTB* and displayed relative to hPSC controls (2^{-ddCt}). Comparison between hPSCs and HLOs from the same experiment were performed at day 16 or 21 as indicated. *N* = 6-8 individual experiments as indicated by circles. Mann-Whitney-U-statistics (two-tailed) hPSCs vs. HLOs: $p\text{-values}^{\text{day 16 HLOs}}$ = 0.0024, Bonferroni-adjusted level of significance: 0.0125 (*); $p\text{-values}^{\text{day 21 HLOs}}$ = 0.002165, Bonferroni-adjusted level of significance: 0.0071 (**). **c.** *HLOs at day 21 show an organized, luminal structure and express adult liver proteins.* Representative immunohistochemistry images of HLOs stained in hematoxylin & eosin (H&E), CEBPα and negative control with secondary antibody on day 21. Arrows indicate sites of nuclear protein staining. Scale bars indicate 50 µm, *n* = 3 individual experiments. **d.** *Single-cell RNA sequencing (scRNA-seq) analysis of day 21 HLOs identifies the major cell types of the human liver across three different annotation strategies.* UMAP plots show 17,898 cells from 10X scRNA-seq in day 21 HLOs (*n* = 3 technical replicates). Cell type annotations generated by three different methods are displayed: Literature-based annotation (top), database-annotation (center), and SingleCellNet[56] annotation comparing HLOs to human liver single-cell data[57] (bottom). **e.** *Dotplot showing the expression of canonical cell-type-specific marker genes expected in the adult human liver across clusters.* Cell types defined by single-cell data are displayed on the y-axis, and canonical genes (bottom) are sorted by cell type (top). The fraction of cells expressing a gene is indicated by the size of the circle, and the mean expression of a gene is indicated by color. Hierarchical clustering is represented by the dendrogram on the right. **f.** *Enrichment of human WikiPathways 2019[110] for three major cell types in HLOs.* Terms have been shortened for readability, full lists are provided in Supplementary Fig. 3c.

Fig. 2

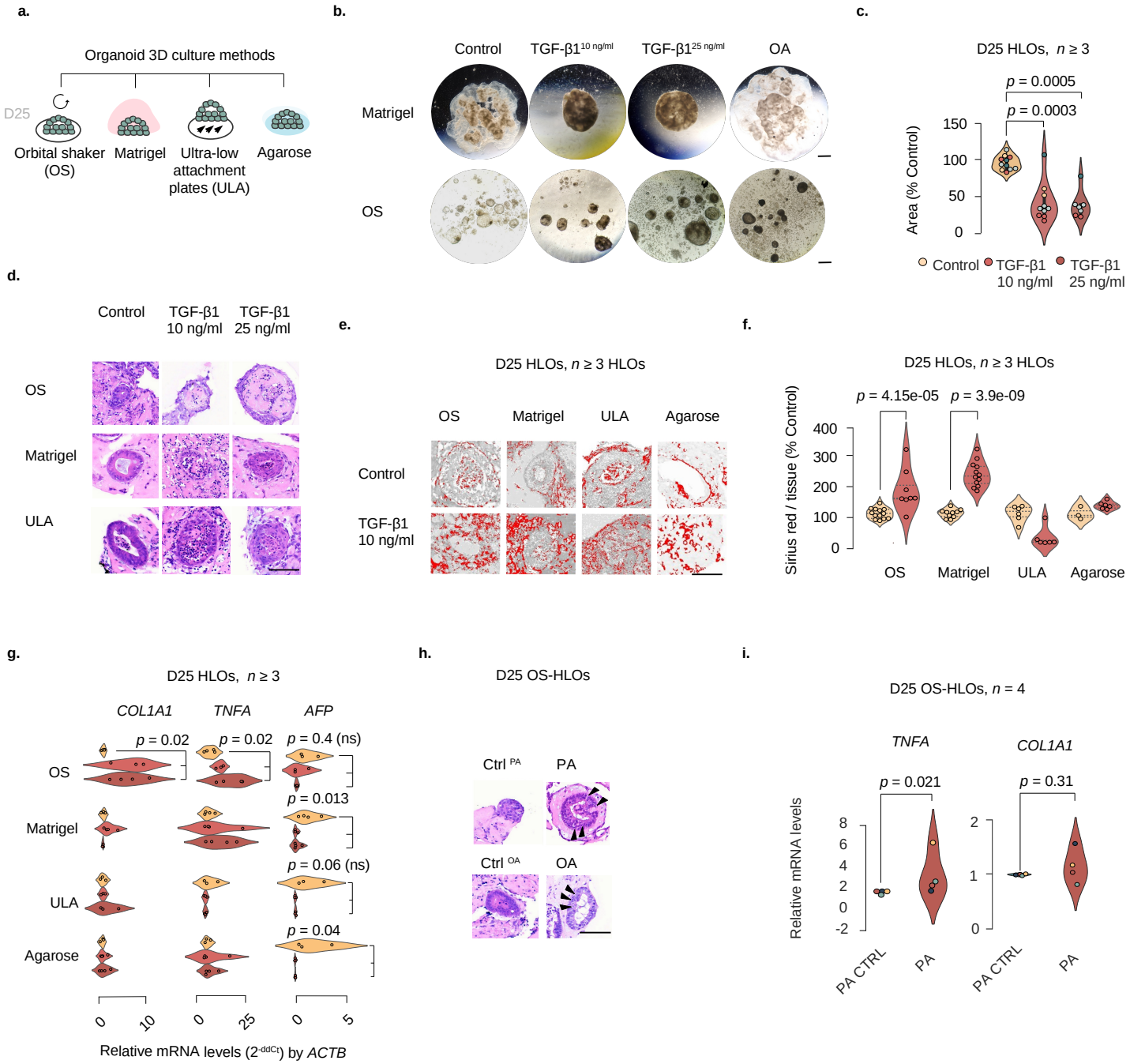

a.
Organoid 3D culture methods
D25
Orbital shaker (OS)
Matrigel
Ultra-low attachment plates (ULA)
Agarose
b.
Control
TGF-β1 10 ng/ml
TGF-β1 25 ng/ml
OA
Matrigel
OS
c.
D25 HLOs, n ≥ 3
p = 0.0005
p = 0.0003
Area (% Control)
150
100
50
0
Control
TGF-β1 10 ng/ml
TGF-β1 25 ng/ml
d.
Control
TGF-β1 10 ng/ml
TGF-β1 25 ng/ml
OS
Matrigel
ULA
e.
D25 HLOs, n ≥ 3 HLOs
OS
Matrigel
ULA
Agarose
Control
TGF-β1 10 ng/ml
f.
D25 HLOs, n ≥ 3 HLOs
p = 4.15e-05
p = 3.9e-09
Sirius red / tissue (% Control)
400
300
200
100
0
OS
Matrigel
ULA
Agarose
g.
D25 HLOs, n ≥ 3
COL1A1
TNFA
AFP
p = 0.02
p = 0.02
p = 0.4 (ns)
p = 0.013
p = 0.06 (ns)
p = 0.04
OS
Matrigel
ULA
Agarose
0
10
0
25
0
5
Relative mRNA levels (2-ddCt) by ACTB
h.
D25 OS-HLOs
Ctrl PA
PA
Ctrl OA
OA
i.
D25 OS-HLOs, n = 4
TNFA
COL1A1
p = 0.021
p = 0.31
Relative mRNA levels
8
6
4
2
0
-2
2
1
0
PA CTRL
PA
PA CTRL
PA

Fig. 2 Inducible liver injury phenotypes in HLOs upon TGF-β1 and fatty acid treatment. a. Overview of the four different 3D culture methods tested for HLO injury induction. **b.** HLOs undergo injury-specific morphologic changes when treated with TGF-β1 and oleic acid (OA) including contraction of the Matrigel dome with TGF-β1 treatment and darkening of color with each treatment when cultured in an orbital shaker (OS) after removal from Matrigel. Representative brightfield images of whole Matrigel domes (each 50 µl) with embedded HLOs (upper panel, scale bar 3 mm) and isolated HLOs cultured on an orbital shaker (lower panel, scale bar 0.1 mm) after five days of treatment with TGF-β1 (10, 25 ng/ml), OA (400 µM), and control conditions. **c.** Compaction assay for Matrigel cultured HLOs indicates a significant surface area reduction for HLOs treated with TGF-β1. Kruskal-Wallis test (two-tailed) followed by a post hoc Conover's test with Bonferroni correction. Each color represents an individual experiment, and each circle represents a separate biological replicate. *N* = 3-4 individual experiments. **d.** H&E staining of HLOs cultured in an OS (upper row), Matrigel (middle row), and ultra-low attachment plate (ULA, lower row) after five days of treatment with TGF-β1 (10 or 25 ng/ml) display distinct structural changes. Representative H&E staining images showing HLO wall thickening in Matrigel- and ULA-TGF-β1 HLOs while hyaline-like intraorganoidal mass accumulation is present in OS-TGF-β1 HLOs. Scale bar 100 µm. *N* = 1-3 individual experiments. **e.** Representative thresholded images of Sirius red staining of HLOs cultured in four different 3D systems and stimulated with TGF-β1 (10 ng/ml) for five days showing collagen fiber deposition in Matrigel-cultured HLOs and OS-HLOs. Scale bars 100 µm. *N* ≥ 4 HLOs from one individual experiment. **f.** Sirius red staining quantification indicates significant collagen fiber deposition occurs only in Matrigel- and OS-cultured HLOs. Shown is the percentage of Sirius red positive tissue by total tissue normalized to the mean of the untreated control HLOs of the respective culture condition. Quantification is based on thresholded 200 x 200 µm squares around single HLOs. Kruskal-Wallis (two-tailed) test followed by a post hoc Conover's test with Bonferroni correction. *N* ≥ 4 HLOs from one individual experiment. **g.** qPCR for *COL1A1, TNFA* and *AFP* transcripts indicates signficant induction of *TNFA* and *COL1A1* only in HLOs incubated in an orbital shaker. Shown are relative mRNA levels normalized to *ACTB* for each individual culture condition (2^{-ddCt}). Kruskal-Wallis (two-tailed) test followed by a post hoc Conover's test with Bonferroni correction comparing three treatment groups per culture method. *P*-values < 0.05 are indicated in the figure. Non-significant (ns) *p*-values: <u>Matrigel:</u> *COL1A1* - control vs. TGF-$\beta1^{10\ \&\ 25ng/ml}$: $p = 0.12$, *TNFA* - Control vs. TGF-$\beta1^{10\ \&\ 25ng/ml}$: $p = 0.94$. <u>1% agarose:</u> *COL1A1* - control vs. TGF-$\beta1^{10\ \&\ 25ng/ml}$: $p = 1.0$, *TNFA* - Control vs. TGF-$\beta1^{10\ \&\ 25ng/ml}$: $p = 0.7$. <u>ULA:</u> *COL1A1* - control vs. TGF-$\beta1^{10\ \&\ 25ng/ml}$: $p = 0.97$, *TNFA* - Control vs. TGF-$\beta1^{10\ \&\ 25ng/ml}$: $p = 0.74$. *N* ≥ 3 individual experiments. **h.** H&E staining of HLOs cultured on OS after five days of treatment with PA (500 µM) or OA (400 µM). Both culture methods display vacuole enrichment.

Scale bar, 100 µm. **i.** *TNFA* expression is induced in OS-HLOs treated with PA (500 µM) for five days. Shown are qPCR results for relative mRNA levels of *TNFA* and *COL1A1* normalized to *ACTB*. Mann-Whitney-U test (two-tailed) followed by Bonferroni correction (adjusted level of significance = 0.025). *N* = 4 individual experiments (represented by individual dot colors).

Fig. 3

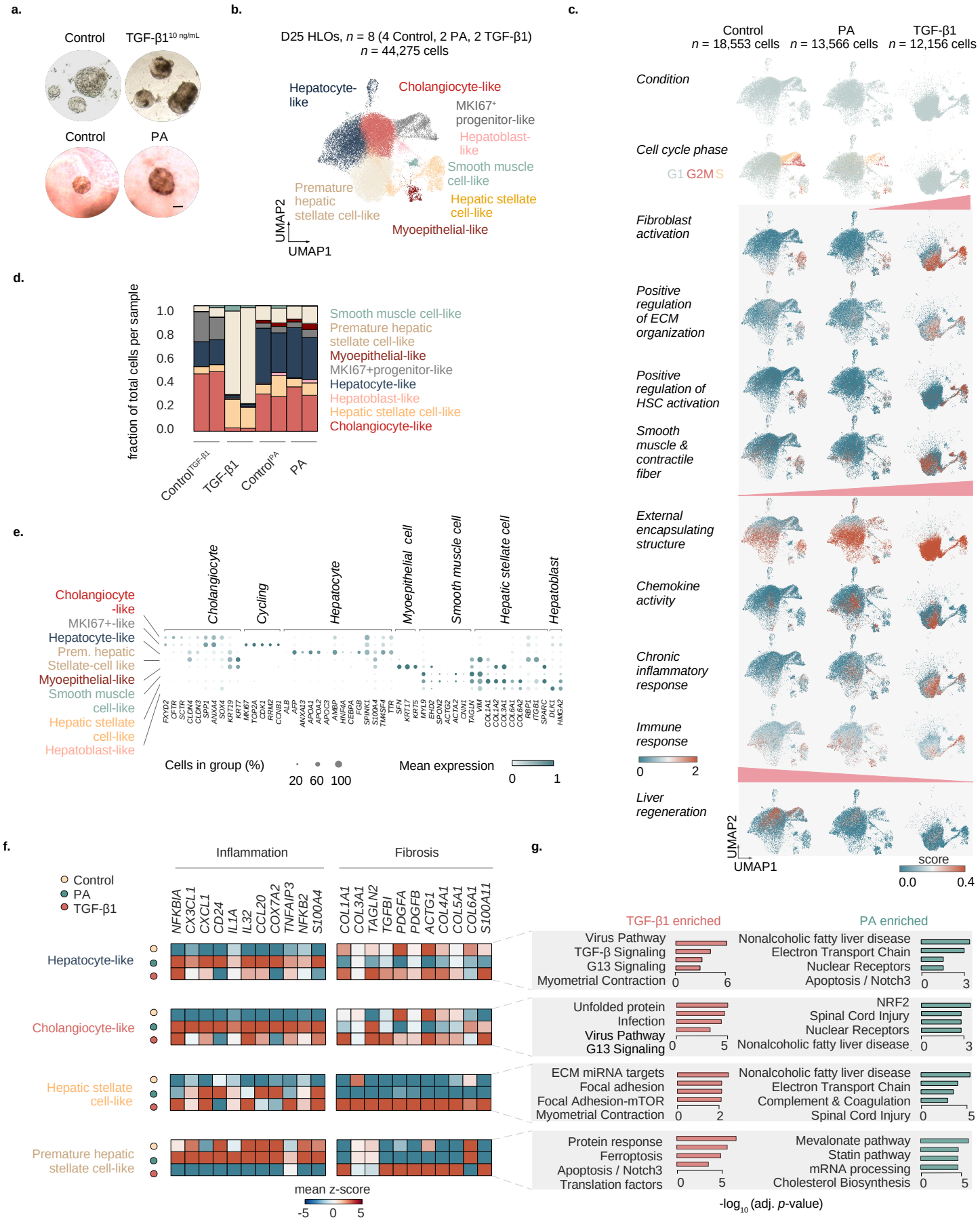

a.
Control
TGF-β1 10 ng/mL
Control
PA
b.
D25 HLOs, n = 8 (4 Control, 2 PA, 2 TGF-β1)
n = 44,275 cells
Hepatocyte-like
Cholangiocyte-like
MKI67+ progenitor-like
Hepatoblast-like
Smooth muscle cell-like
Premature hepatic stellate cell-like
Hepatic stellate cell-like
Myoepithelial-like
UMAP2
UMAP1
c.
Control n = 18,553 cells
PA n = 13,566 cells
TGF-β1 n = 12,156 cells
Condition
Cell cycle phase
G1 G2M S
Fibroblast activation
Positive regulation of ECM organization
Positive regulation of HSC activation
Smooth muscle & contractile fiber
External encapsulating structure
Chemokine activity
Chronic inflammatory response
Immune response
Liver regeneration
0 2
UMAP2
UMAP1
score
0.0 0.4
d.
fraction of total cells per sample
1.0
0.8
0.6
0.4
0.2
0.0
Control TGF-β1
TGF-β1
Control PA
PA
Smooth muscle cell-like
Premature hepatic stellate cell-like
Myoepithelial-like
MKI67+progenitor-like
Hepatocyte-like
Hepatoblast-like
Hepatic stellate cell-like
Cholangiocyte-like
e.
Cholangiocyte
Cycling
Hepatocyte
Myoepithelial cell
Smooth muscle cell
Hepatic stellate cell
Hepatoblast
Cholangiocyte -like
MKI67+-like
Hepatocyte-like
Prem. hepatic Stellate-cell like
Myoepithelial-like
Smooth muscle cell-like
Hepatic stellate cell-like
Hepatoblast-like
FXYD2 CFTR SCTR CLDN4 CLDN3 SPP1 ANXA4 SOX4 KRT19 KRT7 MKI67 TOP2A CDK1 RRM2 CCNB1 ALB AFP ANXA13 APOA1 APOA2 APOC3 AMBP HNF4A CEBPA FGB SPINK1 S100A4 TMSB4 TTR SFN KRT17 KRT5 MYL9 FHL2 SPON2 ACTG2 ACTA2 CNN1 TAGLN VIM COL1A1 COL1A2 COL3A1 COL6A1 COL6A2 RBP1 ITGB1 SPARC DLK1 HMGA2
Cells in group (%)
20 60 100
Mean expression
0 1
f.
Control
PA
TGF-β1
Inflammation
Fibrosis
NFKBIA CX3CL1 CXCL1 CD24 IL1A IL32 CCL20 COX7A2 TNFAIP3 NFKB2 S100A4
COL1A1 COL3A1 TAGLN2 TGFBI PDGFA PDGFB ACTG1 COL4A1 COL5A1 COL6A1 S100A11
Hepatocyte-like
Cholangiocyte-like
Hepatic stellate cell-like
Premature hepatic stellate cell-like
mean z-score
-5 0 5
g.
TGF-β1 enriched
Virus Pathway
TGF-β Signaling
G13 Signaling
Myometrial Contraction
0 6
Unfolded protein
Infection
Virus Pathway
G13 Signaling
0 5
ECM miRNA targets
Focal adhesion
Focal Adhesion-mTOR
Myometrial Contraction
0 2
Protein response
Ferroptosis
Apoptosis / Notch3
Translation factors
0 5
PA enriched
Nonalcoholic fatty liver disease
Electron Transport Chain
Nuclear Receptors
Apoptosis / Notch3
0 3
NRF2
Spinal Cord Injury
Nuclear Receptors
Nonalcoholic fatty liver disease
0 3
Nonalcoholic fatty liver disease
Electron Transport Chain
Complement & Coagulation
Spinal Cord Injury
0 5
Mevalonate pathway
Statin pathway
mRNA processing
Cholesterol Biosynthesis
0 5
-log10 (adj. p-value)

Fig. 3 Evaluating fibrotic and inflammatory responses at the single cell level. a. Brightfield images of day 25 HLOs after five days of treatment with TGF-β1 (10 ng/ml, upper panel) or PA (500 µM, lower panel) and respective controls. Scale bar indicates 50 µm. **b.** UMAP representations of 44,275 clustered cells from 8 HLO samples after five days of treatment with TGF-β1 (10 ng/ml) or PA (500 µM) and their respective controls. *N* = 2 replicates per condition, total *n* = 8, OS-cultured HLOs at day 25. **c.** TGF-β1 treatment induces fibrotic signatures while PA induces inflammatory signals. UMAP plots compare GO term[105] gene set score distributions for selected categories related to fibrosis and inflammation (labeled, left) for HLOs treated with TGF-β1 (10 ng/ml) and PA (500 µM) and their respective controls (labeled, top). Each condition contains two replicates, *n*-numbers as indicated. Increased expression is indicated by a shift from blue to red. Cell cluster annotations are provided in b. The scale bar in the lower right indicates the expression score for all plots except for immune response, which has a separate score. **d.** Cell cluster proportions as the fraction of total cells per sample are shown for each of the 8 individual replicates, color encodes the cell type as annotated in b., and cell types are listed in the order they are displayed. **e.** Dotplot showing canonical marker gene expressions for each cluster from d. Scaled mean expression is displayed, dot sizes correspond to the percentage of cells in each cluster expressing the marker gene. Canonical genes (bottom) are sorted by cell type (top). **f.** TGF-β1 treatment induces genes associated with fibrosis while PA induces genes associated with inflammation. Cell clusters were separated for expression analysis (left), and colors indicate the culture conditions for each cell. Differential expression (right) is displayed for selected genes associated with fibrosis and inflammation for TGF-β1 treatment, PA treatment, and controls. Differential expression is displayed as mean z-score calculated separately for each cell cluster. **g.** Top enriched Wiki pathways 2019[110] for differential expressed genes in each treatment for each major HLO population as labeled in f. Terms have been shortened for readability, full lists are provided in Supplementary Fig. 8.

Fig. 4

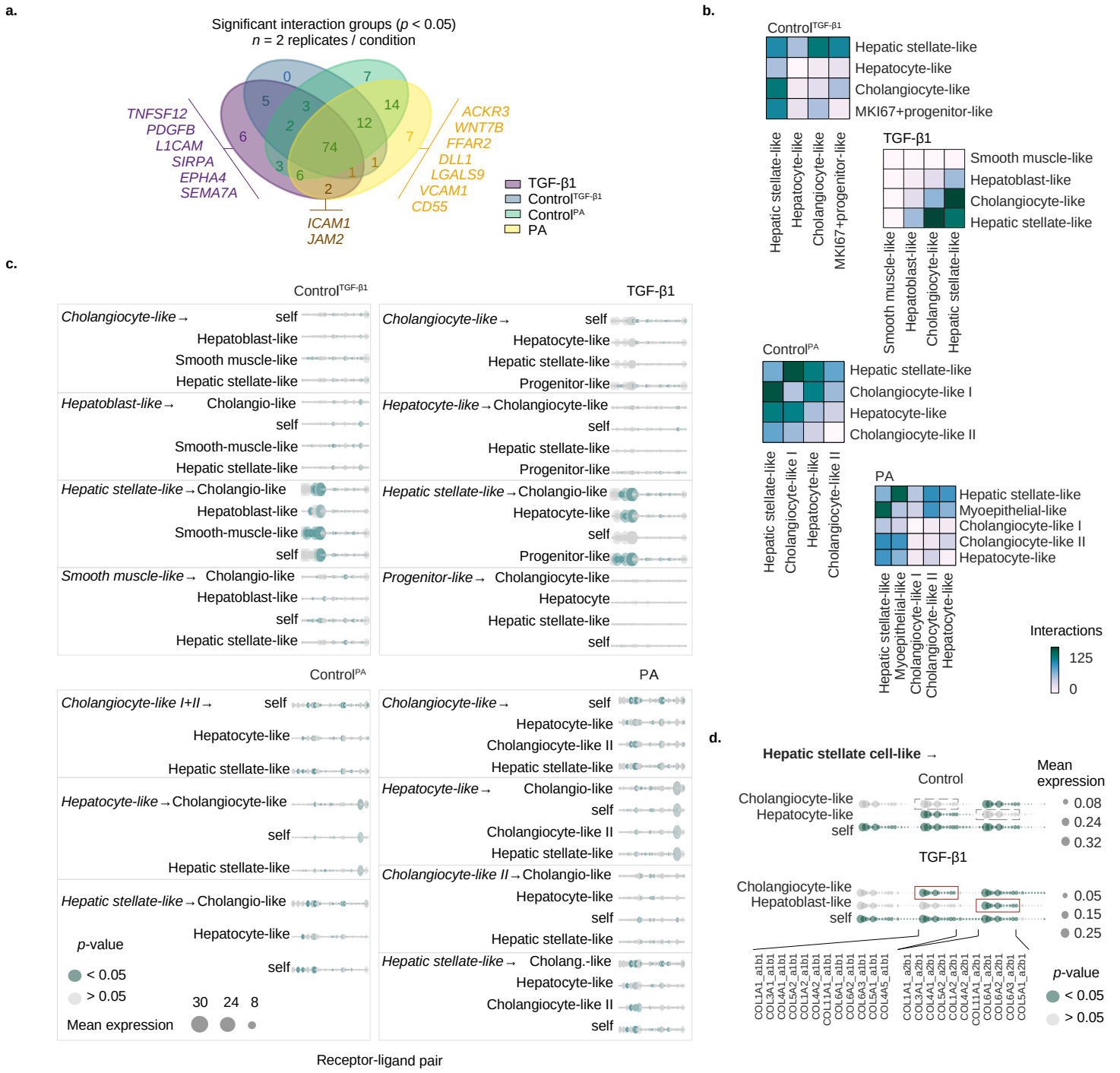

a.
Significant interaction groups (p < 0.05)
n = 2 replicates / condition
TNFSF12
PDGFB
L1CAM
SIRPA
EPHA4
SEMA7A
ICAM1
JAM2
ACKR3
WNT7B
FFAR2
DLL1
LGALS9
VCAM1
CD55
TGF-β1
Control^TGF-β1
Control^PA
PA
b.
Control^TGF-β1
Hepatic stellate-like
Hepatocyte-like
Cholangiocyte-like
MKI67+progenitor-like
TGF-β1
Smooth muscle-like
Hepatoblast-like
Cholangiocyte-like
Hepatic stellate-like
Control^PA
Hepatic stellate-like
Cholangiocyte-like I
Hepatocyte-like
Cholangiocyte-like II
PA
Hepatic stellate-like
Myoepithelial-like
Cholangiocyte-like I
Cholangiocyte-like II
Hepatocyte-like
Interactions
125
0
c.
Control^TGF-β1
TGF-β1
Control^PA
PA
p-value
< 0.05
> 0.05
Mean expression
30 24 8
Receptor-ligand pair
d.
Hepatic stellate cell-like →
Control
TGF-β1
Mean expression
0.08
0.24
0.32
0.05
0.15
0.25
p-value
< 0.05
> 0.05
e.
Cholangiocyte-like →
Control
TGF-β1
Mean expression
2
6
10
f.
Hepatocyte-like →
Control
PA
p-value
< 0.05
> 0.05
Mean expr.
0
0.15
0.30
0.45
DLL1/ NOTCH1
DLL1/ NOTCH2
DLL1/ NOTCH3
FFAR2/ FAM3C

Fig. 4 HLOs show increased crosstalk between HSC- and cholangiocyte-like cells inducing specific interactomes with TGF-β1 and PA treatment. **a.** Venn diagram displays the intersecting and unique significant interacting groups from CellPhoneDB[76] analysis for day 25 OS-cultured HLOs treated with TGF-β1, PA, and their respective controls, n = 44,275 cells. **b.** Heatmaps showing the total number of interactions between cell clusters in day 25 OS-cultured HLOs treated with TGF-β1, PA, and their respective controls. Supplementary Fig. 6a provides cluster distributions across replicates. **c.** Dotplots showing all significant ($p < 0.05$) receptor-ligand pairs per cell-cell cluster pair from CellPhoneDB analysis for OS-cultured HLOs. Receptor-ligand pairs are shown on the x-axes, and cell clusters are shown on the y-axes. Mean expression is indicated by the size of the circle, and significant *p*-values are indicated by blue color. **d.** Differential receptor-ligand pair interactions with HSC-like cells in HLOs after exposure to TGF-β1. Dotplots showing differential *COL1A1* receptor-ligand interactions per cell-cell cluster for OS-HLOs treated with TGF-β1 and corresponding controls. Mean expression (mean of all partner‘s individual average expression, Methods) is indicated by the size of the circle, and significant *p*-values are indicated by blue color. “Self“ indicates co-expression of receptor and ligand on the same cell types, which in this case is co-expression on HSC-like cells. **e.** Example of differential receptor-ligand pair interactions with cholangiocyte-like cells after exposure to TGF-β1. Dotplots showing differential *EPHA4* receptor-ligand interactions per cell-cell cluster for OS-HLOs treated with TGF-β1 and corresponding controls. Mean expression is indicated by the size of the circle, and significant *p*-values are indicated by blue color. **f.** Differential receptor-ligand pair interactions in HLOs after exposure to PA. Dotplots showing significant hepatocyte-like-driven *DLL1-NOTCH1-3* (left) and hepatocyte-like-driven *FFAR2-FAM3C* (right) receptor-ligand pair interactions in OS-HLOs treated with PA compared to their respective controls. Mean expression is indicated by the size of the circle, and significant *p*-values are indicated by blue color.

Fig. 5

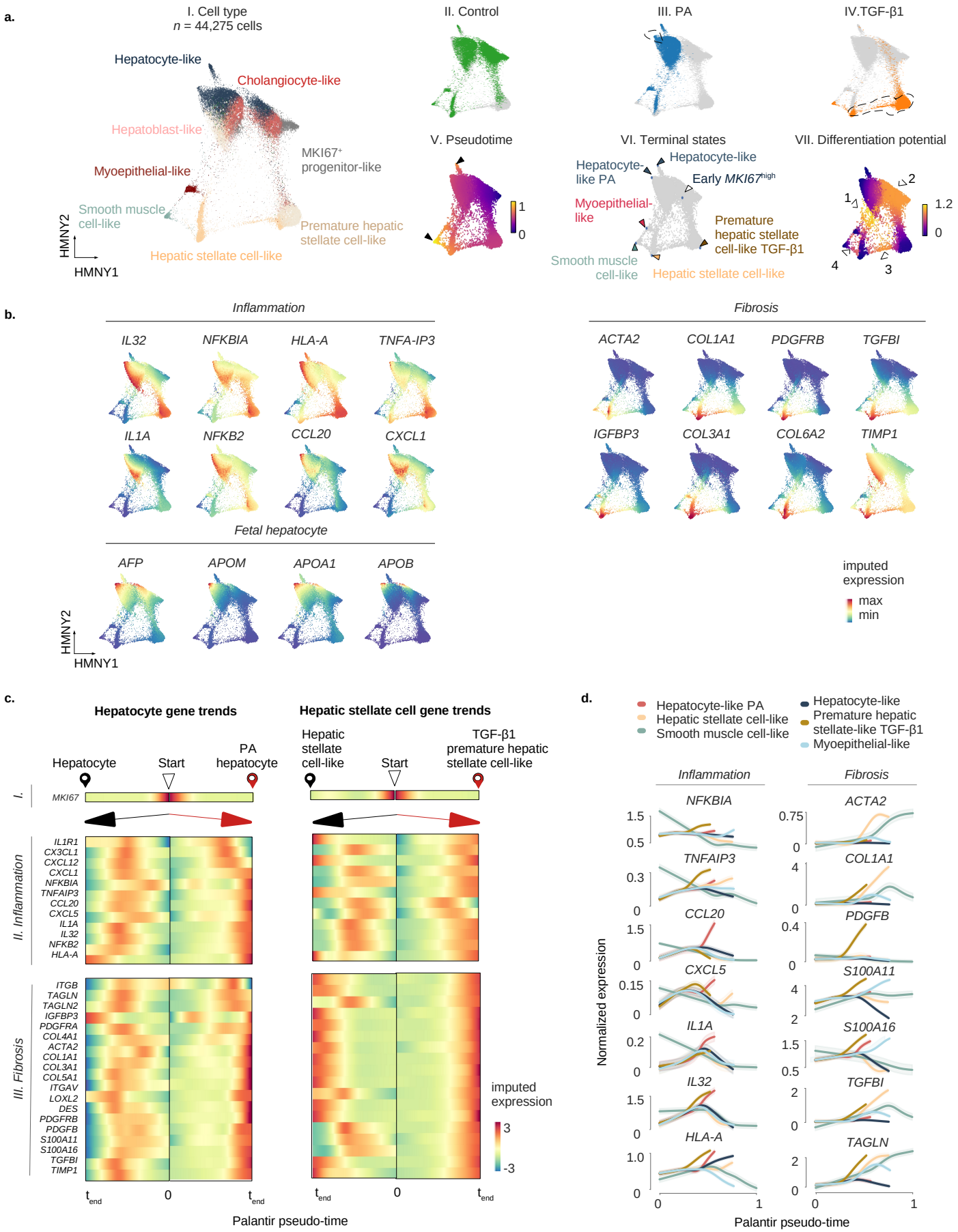

a.
I. Cell type
n = 44,275 cells
Hepatocyte-like
Cholangiocyte-like
Hepatoblast-like
MKI67+ progenitor-like
Myoepithelial-like
Smooth muscle cell-like
Premature hepatic stellate cell-like
Hepatic stellate cell-like
HMNY2
HMNY1
II. Control
III. PA
IV.TGF-β1
V. Pseudotime
1
0
VI. Terminal states
Hepatocyte-like
Hepatocyte-like PA
Early MKI67high
Myoepithelial-like
Premature hepatic stellate cell-like TGF-β1
Smooth muscle cell-like
Hepatic stellate cell-like
VII. Differentiation potential
2
1
1.2
0
4
3
b.
Inflammation
IL32
NFKBIA
HLA-A
TNFA-IP3
IL1A
NFKB2
CCL20
CXCL1
Fetal hepatocyte
AFP
APOM
APOA1
APOB
Fibrosis
ACTA2
COL1A1
PDGFRB
TGFBI
IGFBP3
COL3A1
COL6A2
TIMP1
imputed expression
max
min
c.
Hepatocyte gene trends
Hepatic stellate cell gene trends
Hepatocyte
Start
PA hepatocyte
Hepatic stellate cell-like
Start
TGF-β1 premature hepatic stellate cell-like
I.
MKI67
II. Inflammation
IL1R1
CX3CL1
CXCL12
CXCL1
NFKBIA
TNFAIP3
CCL20
CXCL5
IL1A
IL32
NFKB2
HLA-A
III. Fibrosis
ITGB
TAGLN
TAGLN2
IGFBP3
PDGFRA
COL4A1
ACTA2
COL1A1
COL3A1
COL5A1
ITGAV
LOXL2
DES
PDGFRB
PDGFB
S100A11
S100A16
TGFBI
TIMP1
imputed expression
3
-3
tend
0
tend
Palantir pseudo-time
d.
Hepatocyte-like PA
Hepatic stellate cell-like
Smooth muscle cell-like
Hepatocyte-like
Premature hepatic stellate-like TGF-β1
Myoepithelial-like
Inflammation
Fibrosis
NFKBIA
TNFAIP3
CCL20
CXCL5
IL1A
IL32
HLA-A
ACTA2
COL1A1
PDGFB
S100A11
S100A16
TGFBI
TAGLN
Normalized expression
Palantir pseudo-time

Fig. 5 Trajectory inference reconstructs the emergence of major HLO lineages and injury-specific terminal states. a. Palantir results based on harmony force directed layout display the lineage relationships between control, inflammatory (PA-treated), and fibrotic (TGF-β1-treated) HLO cell populations, revealing increased pseudotime towards hepatocyte-, smooth muscle cell-, and HSC-like populations, while PHSC-like cells and PA-treated hepatocyte-like cells show lower pseudotime than their healthy counterparts and differentiation events specific to the injured phenotype. (**I.**) Harmony augmented affinity matrix-based force-directed layout colored by cell types corresponding to Fig. 3b. (**II.**) Control-HLO cells mapped to harmony force directed layout. (**III.**) PA-HLO cells mapped to harmony force directed layout, dashed circle (upper left) indicates clusters enriched relative to the control condition. (**IV.**) TGF-β1-treated HLO cells mapped to harmony force directed layout, dashed circle (lower right) indicates clusters enriched relative to the control condition. (**V.-VII.**) Palantir results include pseudotime (**V.**, arrow heads indicate pseudotime maxima), terminal states (**VI.**, white arrow head: MKI67-high early cell, colored arrow heads: terminal states), and regions of high differentiation potential (**VII.**, arrow heads 1-4). *n* = 8 samples and 44,275 cells. **b.** MAGIC-imputed gene expression of canonical inflammation- and fibrosis-related genes overlaps with injury-specific terminal states. Fetal hepatocyte-related transcripts enrich along the trajectory towards the PA-treated hepatocyte-like cells. Projection of the imputed gene expression of transcripts related to inflammation (**upper left**), fibrosis (**upper right**), and developmental stages of fetal hepatocytes (**lower left**). **c.** Gene expression trends reveal the increase of inflammatory and pro-fibrotic transcripts along the pseudotime ordering towards injury-specific (PA and TGF-β1) terminal states. Heatmaps show imputed gene expression over Palantir pseudotime for canonical inflammation (II.) and fibrosis (III.) related genes for hepatocyte-like (left column) and HSC-like (right column) terminal states. Opposed are terminal states specific to injury conditions (PA-treated hepatocyte-like and TGF-β1-treated PHSC-like cells, indicated by red arrows). Imputed expression levels are indicated by color, and x-axes display pseudo-time. MKI67 expression, an indication of proliferation that marks early cells, is also quantified (I.) **d.** Inflammatory genes (**left panel**) plateau or decrease in TGF-β1-specific PHSC-like cells (**brown line**), while fibrosis-related (**right panel**) genes peak towards the end of pseudotime, suggesting a transition trough inflammatory states towards the fibrotic phenotype in these cells. PA-treated hepatocytes (**red line**) show a constant increase of inflammatory gene expression with only small increases in fibrotic signatures towards the end of pseudotime, suggesting an inflammatory terminal state with a beginning fibrotic phenotype for these cells. Gene expression trends of representative genes from 5c. across lineages and pseudotime. Pseudo-time is indicated on the x-axes and normalized expression is shown on the y-axes. Cell clusters are

indicated by color (top) and can be identified in 5a (I). Pseudo-time is determined in relation to terminal states (5a. V.), and not all cell clusters reach a value of 1.

Fig. 6

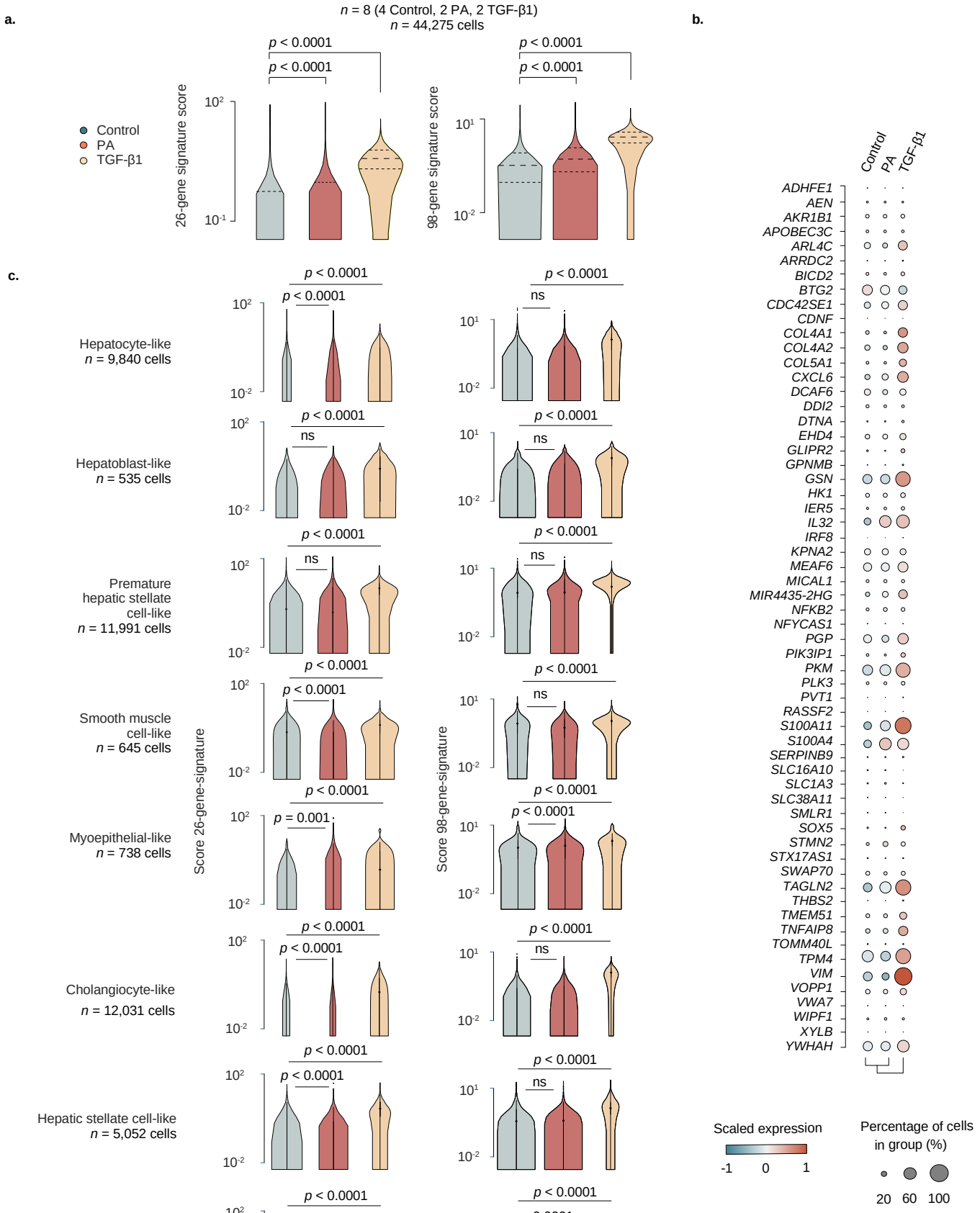

a.
n = 8 (4 Control, 2 PA, 2 TGF-β1)
n = 44,275 cells
Control
PA
TGF-β1
26-gene signature score
98-gene signature score
p < 0.0001
p < 0.0001
b.
Control
PA
TGF-β1
ADHFE1
AEN
AKR1B1
APOBEC3C
ARL4C
ARRDC2
BICD2
BTG2
CDC42SE1
CDNF
COL4A1
COL4A2
COL5A1
CXCL6
DCAF6
DDI2
DTNA
EHD4
GLIPR2
GPNMB
GSN
HK1
IER5
IL32
IRF8
KPNA2
MEAF6
MICAL1
MIR4435-2HG
NFKB2
NFYCAS1
PGP
PIK3IP1
PKM
PLK3
PVT1
RASSF2
S100A11
S100A4
SERPINB9
SLC16A10
SLC1A3
SLC38A11
SMLR1
SOX5
STMN2
STX17AS1
SWAP70
TAGLN2
THBS2
TMEM51
TNFAIP8
TOMM40L
TPM4
VIM
VOPP1
VWA7
WIPF1
XYLB
YWHAH
Scaled expression
-1
0
1
Percentage of cells in group (%)
20
60
100
c.
Hepatocyte-like
n = 9,840 cells
Hepatoblast-like
n = 535 cells
Premature hepatic stellate cell-like
n = 11,991 cells
Smooth muscle cell-like
n = 645 cells
Myoepithelial-like
n = 738 cells
Cholangiocyte-like
n = 12,031 cells
Hepatic stellate cell-like
n = 5,052 cells
MKI67+ progenitor-like
n = 3,451 cells
Score 26-gene-signature
Score 98-gene-signature
p < 0.0001
ns
p = 0.001

Fig. 6 Gene signatures predicting fibrosis stages in NAFLD gradually increase from control HLOs to those treated with PA to those treated with TGF-β1. a. Application of 26- and 98-gene signatures[94] to predict fibrosis stages across NAFLD to control and injured OS-HLOs. *P*-values derived from Kruskal-Wallis test (two-tailed) with post hoc Conover's test followed by Bonferroni correction. Log scale; ns, non-significant. *N* = 8 samples, *n* = 44,275 cells. **b.** Dotplot showing scaled mean expression of cross-condition intersecting genes from the 98-signature score across treatment conditions. Genes displayed are expressed in at least 3 cells in each HLO sample and are present in each individual replicate. Dot sizes correspond to the percentage of cells expressing the respective gene in each condition, and the level of expression is indicated by color. Full differential expression data can be found in **Supplementary Tables 7-13**. **c.** Deconvolution of the scores from a. to individual cell types as annotated in Fig. 3b. *P*-values derived from Kruskal-Wallis test (two-tailed) followed by a post hoc Conover's test with Bonferroni correction. *N* numbers are indicated below cell cluster names. Log scale, ns, non-significant.

References

1. Kisseleva, T. & Brenner, D. Molecular and cellular mechanisms of liver fibrosis and its regression. *Nature Reviews Gastroenterology & Hepatology* **18**, 151–166 (2021).
2. Koyama, Y. & Brenner, D. A. Liver inflammation and fibrosis. *J. Clin. Invest.* **127**, 55–64 (2017).
3. Shetty, A. *et al.* Nonalcoholic Fatty Liver Disease after Liver Transplant. *J. Clin. Transl. Hepatol.* **9**, 428–435 (2021).
4. Wang, S., Toy, M., Hang Pham, T. T. & So, S. Causes and trends in liver disease and hepatocellular carcinoma among men and women who received liver transplants in the U.S., 2010-2019. *PLoS ONE* **15**, 1–14 (2020).
5. Wong, R. J. *et al.* Nonalcoholic Steatohepatitis Is the Second Leading Etiology of Liver Disease Among Adults Awaiting Liver Transplantation in the United States. *Gastroenterology* **148**, 547–555 (2015).
6. Huby, T. & Gautier, E. L. Immune cell-mediated features of non-alcoholic steatohepatitis. *Nat Rev Immunol* **22**, 429–443 (2022).
7. Malhi, H., Bronk, S. F., Werneburg, N. W. & Gores, G. J. Free Fatty Acids Induce JNK-dependent Hepatocyte Lipoapoptosis. *J. Biol. Chem.* **281**, 12093–12101 (2006).
8. Natarajan, S. K. *et al.* Saturated free fatty acids induce cholangiocyte lipoapoptosis. *Hepatology* **60**, 1942–1956 (2014).
9. Miura, K. *et al.* Toll-like receptor 2 and palmitic acid cooperatively contribute to the development of nonalcoholic steatohepatitis through inflammasome activation in mice. *Hepatology* **57**, 577–589 (2013).
10. Ogawa, Y. *et al.* Palmitate-induced lipotoxicity is crucial for the pathogenesis of nonalcoholic fatty liver disease in cooperation with gut-derived endotoxin. *Sci. Rep.* **8**, 1–14 (2018).
11. Hirsova, P., Ibrabim, S. H., Gores, G. J. & Malhi, H. Lipotoxic lethal and sublethal

stress signaling in hepatocytes: relevance to NASH pathogenesis. *J Lipid Res* **57**, 1758–1770 (2016).

12. Araya, J. *et al.* Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to hepatic steatosis in patients with non-alcoholic fatty liver disease. *Clin. Sci. Lond. Engl. 1979* **106**, 635–643 (2004).
13. Ducheix, S. *et al.* Dietary oleic acid regulates hepatic lipogenesis through a liver X receptor-dependent signaling. *PLOS ONE* **12**, e0181393 (2017).
14. Juárez-Hernández, E., Chávez-Tapia, N. C., Uribe, M. & Barbero-Becerra, V. J. Role of bioactive fatty acids in nonalcoholic fatty liver disease. *Nutr. J.* **15**, 72 (2016).
15. Pardo, V., González-Rodríguez, Á., Muntané, J., Kozma, S. C. & Valverde, Á. M. Role of hepatocyte S6K1 in palmitic acid-induced endoplasmic reticulum stress, lipotoxicity, insulin resistance and in oleic acid-induced protection. *Food Chem. Toxicol.* **80**, 298–309 (2015).
16. Rafiei, H., Omidian, K. & Bandy, B. Dietary Polyphenols Protect Against Oleic Acid-Induced Steatosis in an in Vitro Model of NAFLD by Modulating Lipid Metabolism and Improving Mitochondrial Function. *Nutrients* **11**, 541 (2019).
17. Son, G. *et al.* Selective inactivation of NF-κB in the liver using NF-κB decoy suppresses CCl4-induced liver injury and fibrosis. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **293**, 631–639 (2007).
18. Wandrer, F. *et al.* TNF-Receptor-1 inhibition reduces liver steatosis, hepatocellular injury and fibrosis in NAFLD mice. *Cell Death Dis.* **11**, 1–9 (2020).
19. Zhao, S. *et al.* The concentration of tumor necrosis factor-α determines its protective or damaging effect on liver injury by regulating Yap activity. *Cell Death Dis.* **11**, 1–13 (2020).
20. Yang, Y. M. & Seki, E. TNFα in Liver Fibrosis. *Curr. Pathobiol. Rep.* **3**, 253–261 (2015).
21. Sunami, Y. *et al.* Hepatic activation of IKK/NFκB signaling induces liver fibrosis via

macrophage-mediated chronic inflammation. *Hepatology* **56**, 1117–1128 (2012).

22. Zigmond, E. *et al.* Infiltrating Monocyte-Derived Macrophages and Resident Kupffer Cells Display Different Ontogeny and Functions in Acute Liver Injury. *J. Immunol.* **193**, 344–353 (2014).
23. Lodyga, M. *et al.* Cadherin-11–mediated adhesion of macrophages to myofibroblasts establishes a profibrotic niche of active TGF-β. *Sci. Signal.* **12**, eaao3469 (2019).
24. Hellerbrand, C., Stefanovic, B., Giordano, F., Burchardt, E. R. & Brenner, D. A. The role of TGFb1 in initiating hepatic stellate cell activation in vivo. *J. Hepatol.* **30**, 77–87 (1999).
25. Seki, E. *et al.* TLR4 enhances TGF-β signaling and hepatic fibrosis. *Nat. Med.* **13**, 1324–1332 (2007).
26. Ramachandran, P. *et al.* Resolving the fibrotic niche of human liver cirrhosis at single- cell level. *Nature* **575**, 512–518 (2019).
27. Ying, H. Z. *et al.* PDGF signaling pathway in hepatic fibrosis pathogenesis and therapeutics (Review). *Mol. Med. Rep.* **16**, 7879–7889 (2017).
28. Ortiz, C. *et al.* Extracellular Matrix Remodeling in Chronic Liver Disease. *Curr. Tissue Microenviron. Rep.* **2**, 41–52 (2021).
29. Drew, L. Tipping the balance. *Nature* **564**, S74–S75 (2018).
30. Guan, Y. *et al.* Human hepatic organoids for the analysis of human genetic diseases. *JCI Insight* **2**, 94954 (2017).
31. Wu, F. *et al.* Generation of hepatobiliary organoids from human induced pluripotent stem cells. *J. Hepatol.* **70**, 1145–1158 (2019).
32. Mun, S. J. *et al.* Generation of expandable human pluripotent stem cell-derived hepatocyte-like liver organoids. *J Hepatol* **71**, 970–985 (2019).
33. Ouchi, R. *et al.* Modeling Steatohepatitis in Humans with Pluripotent Stem Cell-Derived Organoids. *Cell Metab.* **30**, 374-384.e6 (2019).
34. Thompson, W. L. & Takebe, T. Generation of multi-cellular human liver organoids

from pluripotent stem cells. *Methods Cell Biol* **159**, 47–68 (2020).

35. Kulkeaw, K., Tubsuwan, A., Tongkrajang, N. & Whangviboonkij, N. Generation of human liver organoids from pluripotent stem cell-derived hepatic endoderms. *PeerJ* **8**, e9968 (2020).
36. Sekine, K. *et al.* Generation of human induced pluripotent stem cell-derived liver buds with chemically defined and animal origin-free media. *Sci. Rep.* **10**, 17937 (2020).
37. Ramli, M. N. B. *et al.* Human Pluripotent Stem Cell-Derived Organoids as Models of Liver Disease. *Gastroenterology* **159**, 1471-1486.e12 (2020).
38. Sinton, M. C. *et al.* A human pluripotent stem cell model for the analysis of metabolic dysfunction in hepatic steatosis. *iScience* **24**, 101931 (2021).
39. Guan, Y. *et al.* A human multi-lineage hepatic organoid model for liver fibrosis. *Nat Commun* **12**, 6138 (2021).
40. Bock, C. *et al.* The Organoid Cell Atlas. *Nat. Biotechnol.* **39**, 13–17 (2021).
41. Li, M. & Belmonte, J. C. I. Ground rules of the pluripotency gene regulatory network. *Nat. Rev. Genet.* **18**, 180–191 (2017).
42. Aizarani, N. *et al.* A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* **572**, 199–204 (2019).
43. Limaye, P. B. *et al.* Expression of specific hepatocyte and cholangiocyte transcription factors in human liver disease and embryonic development. *Lab. Invest.* **88**, 865–872 (2008).
44. Segal, J. M. *et al.* Single cell analysis of human foetal liver captures the transcriptional profile of hepatobiliary hybrid progenitors. *Nat Commun* **10**, 3350 (2019).
45. Antoniou, A. *et al.* Intrahepatic Bile Ducts Develop According to a New Mode of Tubulogenesis Regulated by the Transcription Factor SOX9. *Gastroenterology* **136**, 2325–2333 (2009).

46. Uyama, N. *et al.* Fascin, a novel marker of human hepatic stellate cells, may regulate their proliferation, migration, and collagen gene expression through the FAK-PI3K-Akt pathway. *Lab. Invest.* **92**, 57–71 (2012).
47. Yokoi, Y. *et al.* Immunocytochemical Detection of Desmin in Fat-Storing Cells (Ito Cells). *Hepatology* **4**, 709–714 (1984).
48. Tao, L.-L. *et al.* The role of C/EBP-α expression in human liver and liver fibrosis and its relationship with autophagy. *Int. J. Clin. Exp. Pathol.* **8**, 13102–13107 (2015).
49. Bhaduri, A., Andrews, M. G., Kriegstein, A. R. & Nowakowski, T. J. Are Organoids Ready for Prime Time? *Cell Stem Cell* **27**, 361–365 (2020).
50. Clarke, Z. A. *et al.* Tutorial: guidelines for annotating single-cell transcriptomic maps using automated and manual methods. *Nat. Protoc.* **16**, 2749–2764 (2021).
51. Franzén, O., Gan, L. M. & Björkegren, J. L. M. PanglaoDB: A web server for exploration of mouse and human single-cell RNA sequencing data. *Database* **2019**, 1–9 (2019).
52. Zhang, X. *et al.* CellMarker: A manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.* **47**, D721–D728 (2019).
53. MacParland, S. A. *et al.* Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nature Communications* **9**, 4383 (2018).
54. Ober, E. A. & Lemaigre, F. P. Development of the liver: Insights into organ and tissue morphogenesis. *J. Hepatol.* **68**, 1049–1062 (2018).
55. Shin, D. & Monga, S. P. S. Cellular and molecular basis of liver development. *Compr Physiol* **3**, 799–815 (2013).
56. Tan, Y. & Cahan, P. SingleCellNet : A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Brief Report SingleCellNet : A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms and Across Species. *Cell Syst.* **9**, 207-213.e2 (2019).
57. Sharma, A. *et al.* Onco-fetal Reprogramming of Endothelial Cells Drives

Immunosuppressive Macrophages in Hepatocellular Carcinoma. *Cell* **183**, 377-394.e21 (2020).

58. Abdelaal, T. *et al.* A comparison of automatic cell identification methods for single-cell RNA sequencing data. *Genome Biol.* 1–19 (2019).
59. Zaret, K. S. Genetic programming of liver and pancreas progenitors: lessons for stem-cell differentiation. *Nat. Rev. Genet.* **9**, 329–340 (2008).
60. Nissim, S. *et al.* Prostaglandin E2 Regulates Liver versus Pancreas Cell-Fate Decisions and Endodermal Outgrowth. *Dev. Cell* **28**, 423–437 (2014).
61. Kanzler, S. *et al.* TGF-beta1 in liver fibrosis: an inducible transgenic mouse model to study liver fibrogenesis. *Am J Physiol* **276**, G1059-1068 (1999).
62. Kim, K. K., Sheppard, D. & Chapman, H. A. TGF-β1 Signaling and Tissue Fibrosis. *Cold Spring Harb Perspect Biol* **10**, a022293 (2018).
63. Huang, Y. *et al.* Image analysis of liver collagen using sirius red is more accurate and correlates better with serum fibrosis markers than trichrome. *Liver Int.* **33**, 1249–1256 (2013).
64. Niederreiter, L. & Tilg, H. Cytokines and fatty liver diseases. *Liver Res.* **2**, 14–20 (2018).
65. Yang, A.-T. *et al.* TGF-β1 Induces the Dual Regulation of Hepatic Progenitor Cells with Both Anti- and Proliver Fibrosis. *Stem Cells Int.* **2016**, 1492694 (2016).
66. Funa, N. S. *et al.* TGF-β modulates cell fate in human ES cell-derived foregut endoderm by inhibiting multiple endogenous signaling pathways. *bioRxiv* 2021.07.15.452587 Preprint at https://doi.org/10.1101/2021.07.15.452587 (2021).
67. Damdinsuren, B. *et al.* TGF-beta1-induced cell growth arrest and partial differentiation is related to the suppression of Id1 in human hepatoma cells. *Oncol. Rep.* **15**, 401–408 (2006).
68. The Gene Ontology Consortium. Gene Ontology : tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

69. Laiho, M., DeCaprio, J. A., Ludlow, J. W., Livingston, D. M. & Massague, J. Growth inhibition by TGF-β linked to suppression of retinoblastoma protein phosphorylation. *Cell* **62**, 175–185 (1990).

70. Sheahan, S., Bellamy, C. O., Dunbar, D. R., Harrison, D. J. & Prost, S. Deficiency of G1 regulators P53, P21Cip1 and/or pRb decreases hepatocyte sensitivity to TGFbeta cell cycle arrest. *BMC Cancer* **7**, 215 (2007).

71. Tsuchida, T. & Friedman, S. L. Mechanisms of hepatic stellate cell activation. *Nat. Rev. Gastroenterol. Hepatol.* **14**, 397–411 (2017).

72. Ahn, C., Yang, H., Lee, D., An, B. & Jeung, E.-B. Placental claudin expression and its regulation by endogenous sex steroid hormones. *Steroids* **100**, 44–51 (2015).

73. Maldonado-Estrada, J., Menu, E., Roques, P., Barré-Sinoussi, F. & Chaouat, G. Evaluation of Cytokeratin 7 as an accurate intracellular marker with which to assess the purity of human placental villous trophoblast cells by flow cytometry. *J. Immunol. Methods* **286**, 21–34 (2004).

74. Tamai, Y. *et al.* Cytokeratins 8 and 19 in the Mouse Placental Development. *J. Cell Biol.* **151**, 563–572 (2000).

75. Baselli, G. A. *et al.* Liver transcriptomics highlights interleukin-32 as novel NAFLD-related cytokine and candidate biomarker. *Gut* **69**, 1855–1866 (2020).

76. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).

77. Okabe, H. *et al.* Wnt signaling regulates hepatobiliary repair following cholestatic liver injury in mice. *Hepatology* **64**, 1652–1666 (2016).

78. Golden-Mason, L. & Rosen, H. R. Galectin-9: Diverse roles in hepatic immune homeostasis and inflammation. *Hepatology* **66**, 271–279 (2017).

79. Wang, H., Mehal, W., Nagy, L. E. & Rotman, Y. Immunological mechanisms and therapeutic targets of fatty liver diseases. *Cell. Mol. Immunol.* **18**, 73–91 (2021).

80. Ueno, Y., Fukushima, K., Nakagome, Y., Kido, O. & Shimosegawa, T. Bioinformatic approach for cholangiocyte pathophysiology. *Hepatol. Res.* **37**, 444–448 (2007).
81. Fukushima, K., Ueno, Y., Inoue, J., Kanno, N. & Shimosegawa, T. Filopodia formation via a specific Eph family member and PI3K in immortalized cholangiocytes. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **291**, 812–819 (2006).
82. Mimche, P. N. *et al.* EphB2 receptor tyrosine kinase promotes hepatic fibrogenesis in mice via activation of hepatic stellate cells. *Sci. Rep.* **8**, 2532 (2018).
83. Wu, B., Rockel, J. S., Lagares, D. & Kapoor, M. Ephrins and Eph Receptor Signaling in Tissue Repair and Fibrosis. *Curr Rheumatol Rep* **21**, 23 (2019).
84. Zhu, C., Tabas, I., Schwabe, R. F. & Pajvani, U. B. Maladaptive regeneration — the reawakening of developmental pathways in NASH and fibrosis. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 131–142 (2021).
85. Zhu, C. *et al.* Hepatocyte Notch activation induces liver fibrosis in nonalcoholic steatohepatitis. *Sci. Transl. Med.* **10**, eaat0344 (2018).
86. Zhang, X. *et al.* FAM3 gene family: A promising therapeutical target for NAFLD and type 2 diabetes. *Metab. - Clin. Exp.* **81**, 71–82 (2018).
87. Secor, J. D., Fligor, S. C., Tsikis, S. T., Yu, L. J. & Puder, M. Free Fatty Acid Receptors as Mediators and Therapeutic Targets in Liver Disease. *Front Physiol* **12**, 656441 (2021).
88. Gentilini, A. *et al.* CXCR7 contributes to the aggressive phenotype of cholangiocarcinoma cells. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1865**, 2246–2256 (2019).
89. Fukusato, T., Gerber, M. A., Thung, S. N., Ferrone, S. & Schaffner, F. Expression of HLA Class I antigens on hepatocytes in liver disease. *Am. J. Pathol.* **123**, 264–270 (1986).
90. Nowotschin, S. *et al.* The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature* **569**, 361–367 (2019).

91. Setty, M. *et al.* Characterization of cell fate probabilities in single-cell data with Palantir. *Nat Biotechnol* **37**, 451–460 (2019).

92. Lotto, J. *et al.* Single-Cell Transcriptomics Reveals Early Emergence of Liver Parenchymal and Non-parenchymal Cell Lineages. *Cell* **183**, 702-716 (2020).

93. van Dijk, D. *et al.* Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716-729.e27 (2018).

94. Pantano, L. *et al.* Molecular characterization and cell type composition deconvolution of fibrosis in NAFLD. *Sci. Rep.* **11**, 1–14 (2021).

95. Brazovskaja, A. *et al.* Cell atlas of the regenerating human liver after portal vein embolization. *bioRxiv* 2021.06.03.444016 Preprint at https://doi.org/10.1101/2021.06.03.444016 (2021).

96. Daneshvar, K. *et al.* lncRNA DIGIT and BRD3 protein form phase-separated condensates to regulate endoderm differentiation. *Nat. Cell Biol.* **22**, 1211–1222 (2020).

97. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

98. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **8**, 281-291.e9 (2019).

99. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

100. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

101. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018).

102. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 1–12 (2019).

103. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289–1296 (2019).

104. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).

105. Carbon, S. *et al.* The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).

106. Carbon, S. *et al.* AmiGO: Online access to ontology and annotation data. *Bioinformatics* **25**, 288–289 (2009).

107. Waskom, M. L. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 1–4 (2021).

108. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

109. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

110. Martens, M. *et al.* WikiPathways: connecting communities. *Nucleic Acids Res.* **49**, 613–621 (2021).

Supplementary information

Human liver organoids model progressive inflammatory and fibrotic injury in non-alcoholic fatty liver disease

Anja Hess[1,2], Stefan D. Gentile[1,3], Amel Ben Saad[1,3], Raza-Ur Rahman[1,3], Tim Habboub[1] and Alan C. Mullen[1,3,4,5,6*]

1 Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA

2 Max Planck Institute for Molecular Genetics, Department of Genome Regulation, Berlin, Germany

3 Harvard Medical School, Boston, MA, USA

4 Harvard Stem Cell Institute, Cambridge, MA, USA

5 Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA

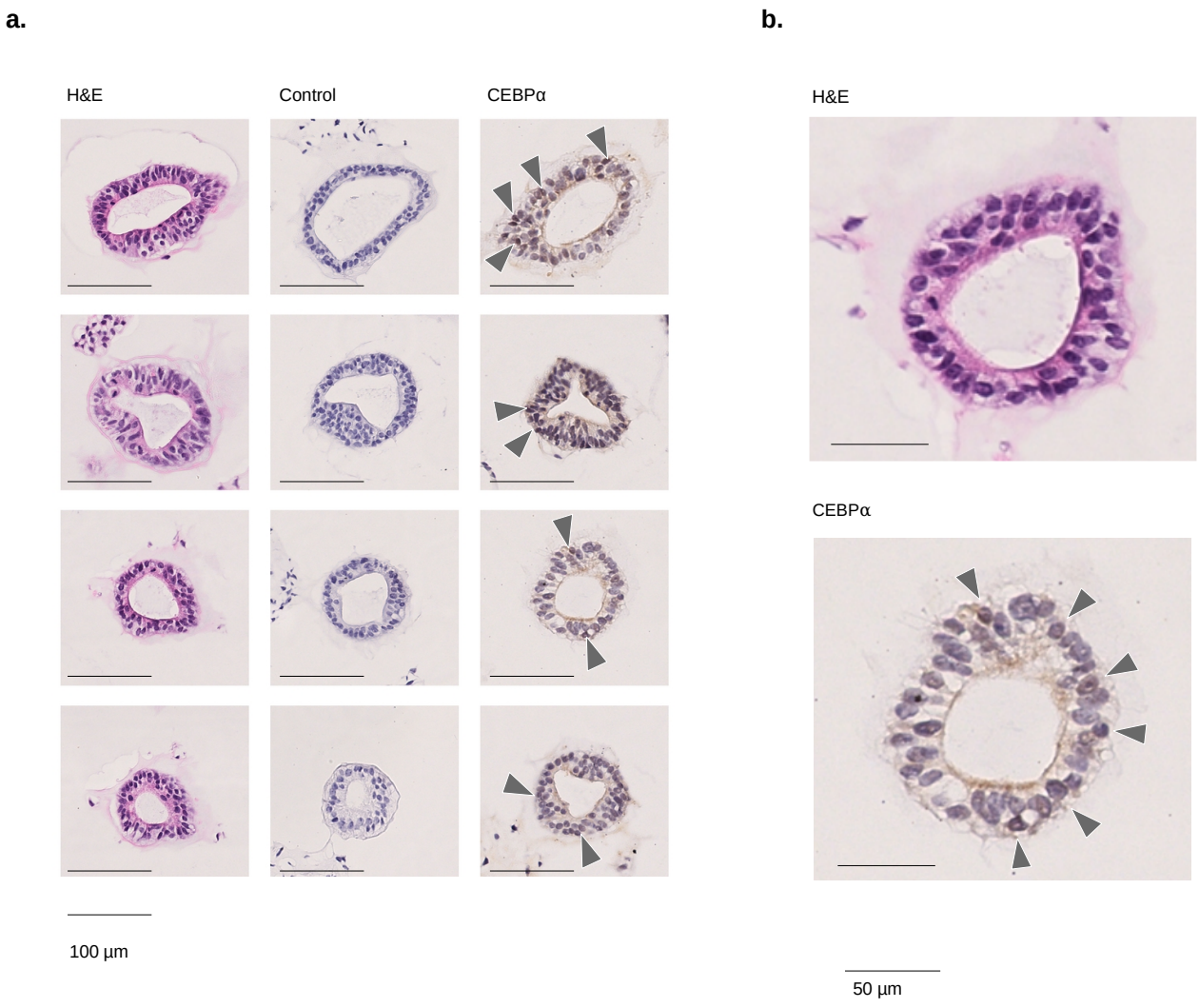
6 Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Boston, MA, USA

* Corresponding author. GRJ7, 55 Fruit Street, Massachusetts General Hospital, Boston, MA 02114, USA. (617) 726-6342. Email: acmullen@mgh.harvard.edu

Abbreviations:

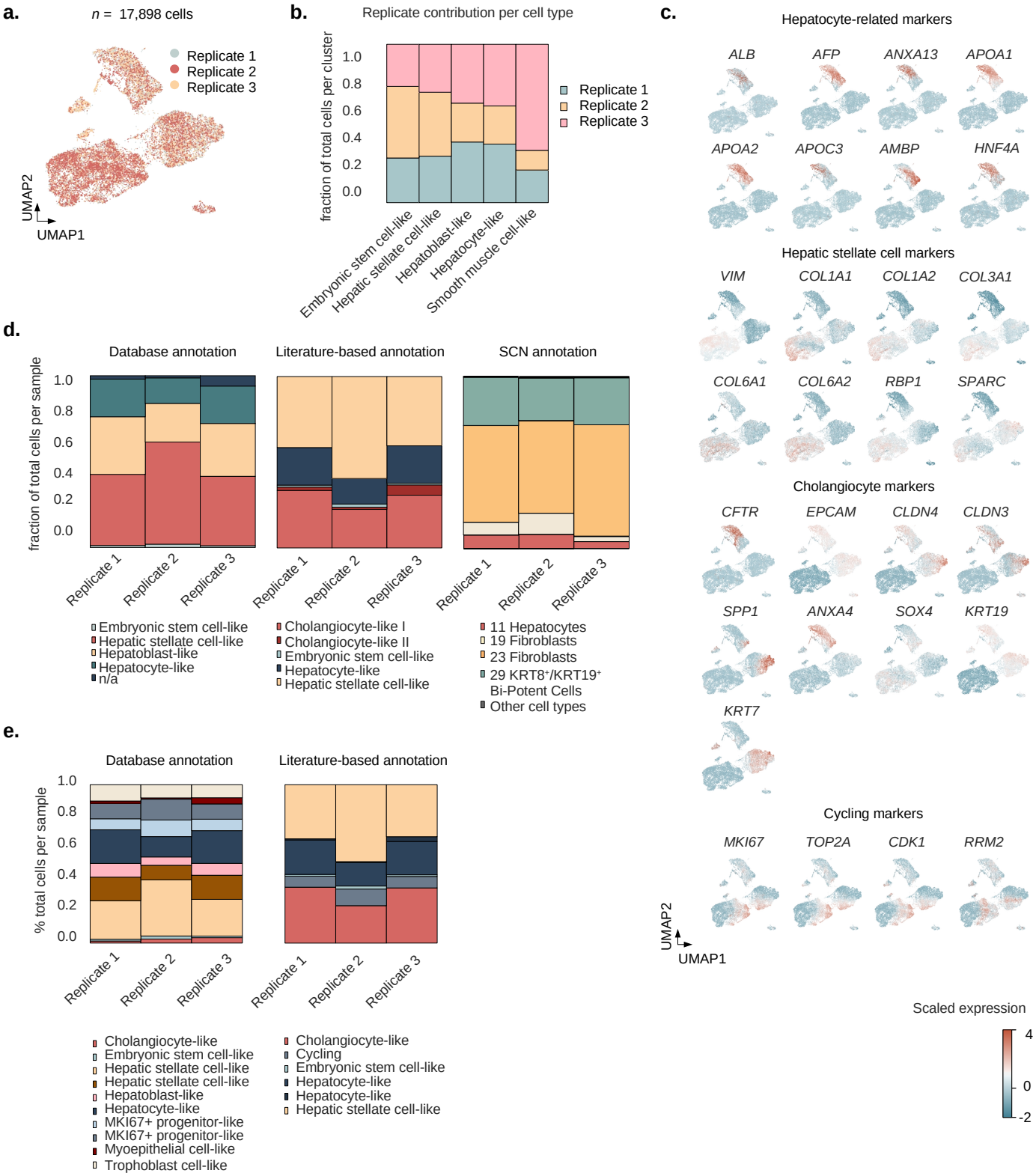
HLOs	Human liver organoids
OS	Orbital shaker – cultured HLOs
ULA	Ultra-low attachment plate – cultured HLOs
D21	Day 21 of differentiation
D25	Day 25 of differentiation
D34	Day 34 of differentiation
PA	Palmitic Acid at 500 µM
TGF-β1	TGF-β1 at 10 ng/µl

S1



Supplementary Figure 1. Representative images of fixed HLOs. Corresponding to Fig. 1c.

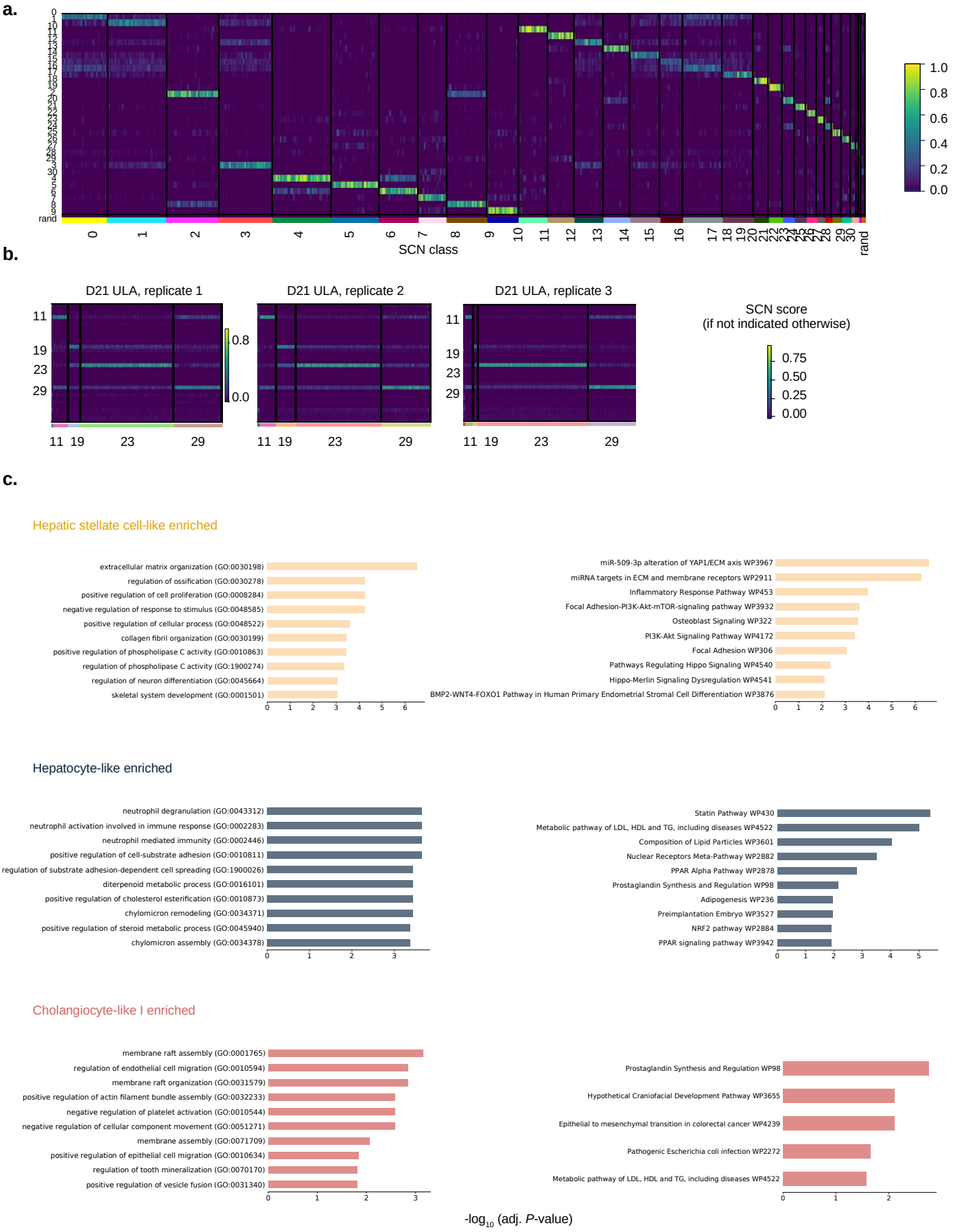
a. Panels showing individual HLOs from left to right: H&E, secondary antibody only (Control), anti-CEBPα (nuclear staining indicated by arrowheads). Scale bars 100 µm. **b.** Two HLOs from a., increased magnification for representative indication of nuclear CEBPα staining (arrowheads). Scale bars 50 µm.


a.
n = 17,898 cells
Replicate 1
Replicate 2
Replicate 3
UMAP2
UMAP1
b.
Replicate contribution per cell type
fraction of total cells per cluster
1.0
0.8
0.6
0.4
0.2
0.0
Replicate 1
Replicate 2
Replicate 3
Embryonic stem cell-like
Hepatic stellate cell-like
Hepatoblast-like
Hepatocyte-like
Smooth muscle cell-like
c.
Hepatocyte-related markers
ALB
AFP
ANXA13
APOA1
APOA2
APOC3
AMBP
HNF4A
Hepatic stellate cell markers
VIM
COL1A1
COL1A2
COL3A1
COL6A1
COL6A2
RBP1
SPARC
Cholangiocyte markers
CFTR
EPCAM
CLDN4
CLDN3
SPP1
ANXA4
SOX4
KRT19
KRT7
Cycling markers
MKI67
TOP2A
CDK1
RRM2
UMAP2
UMAP1
Scaled expression
4
0
-2
d.
Database annotation
Literature-based annotation
SCN annotation
fraction of total cells per sample
1.0
0.8
0.6
0.4
0.2
0.0
Replicate 1
Replicate 2
Replicate 3
Embryonic stem cell-like
Hepatic stellate cell-like
Hepatoblast-like
Hepatocyte-like
n/a
Cholangiocyte-like I
Cholangiocyte-like II
Embryonic stem cell-like
Hepatocyte-like
Hepatic stellate cell-like
11 Hepatocytes
19 Fibroblasts
23 Fibroblasts
29 KRT8+/KRT19+ Bi-Potent Cells
Other cell types
e.
Database annotation
Literature-based annotation
% total cells per sample
1.0
0.8
0.6
0.4
0.2
0.0
Replicate 1
Replicate 2
Replicate 3
Cholangiocyte-like
Embryonic stem cell-like
Hepatic stellate cell-like
Hepatic stellate cell-like
Hepatoblast-like
Hepatocyte-like
MKI67+ progenitor-like
MKI67+ progenitor-like
Myoepithelial cell-like
Trophoblast cell-like
Cholangiocyte-like
Cycling
Embryonic stem cell-like
Hepatocyte-like
Hepatocyte-like
Hepatic stellate cell-like

Supplementary Figure 2. Replicate distribution and marker gene expression from scRNA-seq. Corresponding to Fig. 1d.

a. UMAP plot mapping 17,898 single cells from day 21 ULA-HLOs from Fig. 1 colored by replicate. **b.** Barplot indicating the replicate distribution as the proportion of total cells per cluster for each cluster calculated with the Leiden algorithm at resolution 0.1 and database annotation. **c.** UMAP plots are colored for the scaled expression of canonical cell type marker genes in HLOs after quality control (QC). **d.** Barplots showing cell type distributions across replicates as the proportion of each cell type per total cells per sample. Annotation approach (Methods) is indicated (from left to right: database-annotation utilizing PanglaoDB[1] and CellMarker[2] and statistical testing via GSEA-py enrichr[3], literature-based marker sets and statistical testing via GSEA-py enrichr[3], SingelCellNet[5] using human liver scRNAseq data[4] as a reference). Leiden resolution = 0.1. **e.** Barplots showing the cell type distibutions for database and literature annotations as described for e, this time at increased Leiden resolution of 0.5.

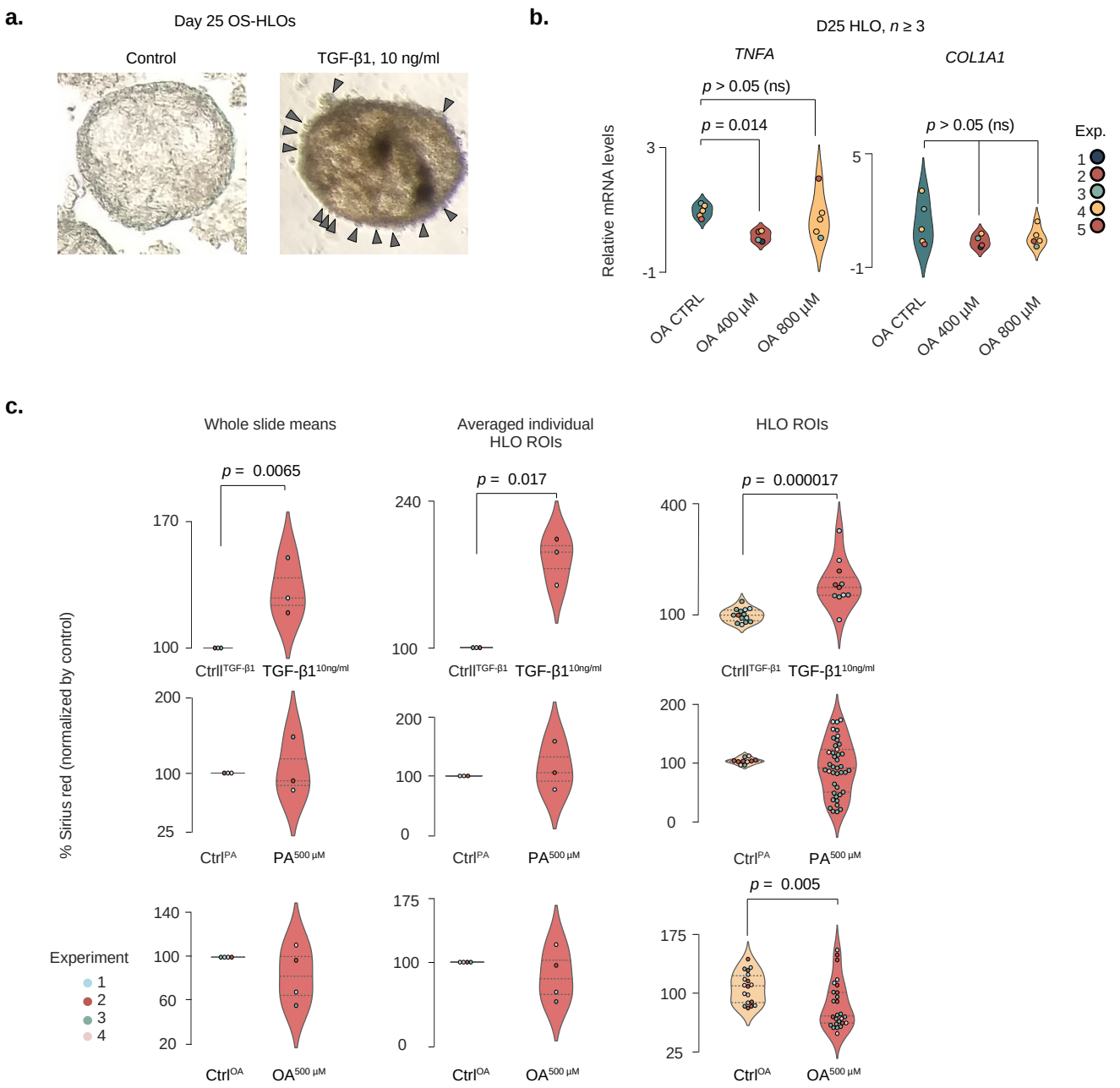
S3


a.
SCN class
rand
1.0
0.8
0.6
0.4
0.2
0.0
b.
D21 ULA, replicate 1
D21 ULA, replicate 2
D21 ULA, replicate 3
SCN score
(if not indicated otherwise)
0.75
0.50
0.25
0.00
c.
Hepatic stellate cell-like enriched
extracellular matrix organization (GO:0030198)
regulation of ossification (GO:0030278)
positive regulation of cell proliferation (GO:0008284)
negative regulation of response to stimulus (GO:0048585)
positive regulation of cellular process (GO:0048522)
collagen fibril organization (GO:0030199)
positive regulation of phospholipase C activity (GO:0010863)
regulation of phospholipase C activity (GO:1900274)
regulation of neuron differentiation (GO:0045664)
skeletal system development (GO:0001501)
miR-509-3p alteration of YAP1/ECM axis WP3967
miRNA targets in ECM and membrane receptors WP2911
Inflammatory Response Pathway WP453
Focal Adhesion-PI3K-Akt-mTOR-signaling pathway WP3932
Osteoblast Signaling WP322
PI3K-Akt Signaling Pathway WP4172
Focal Adhesion WP306
Pathways Regulating Hippo Signaling WP4540
Hippo-Merlin Signaling Dysregulation WP4541
BMP2-WNT4-FOXO1 Pathway in Human Primary Endometrial Stromal Cell Differentiation WP3876
Hepatocyte-like enriched
neutrophil degranulation (GO:0043312)
neutrophil activation involved in immune response (GO:0002283)
neutrophil mediated immunity (GO:0002446)
positive regulation of cell-substrate adhesion (GO:0010811)
positive regulation of substrate adhesion-dependent cell spreading (GO:1900026)
diterpenoid metabolic process (GO:0016101)
positive regulation of cholesterol esterification (GO:0010873)
chylomicron remodeling (GO:0034371)
positive regulation of steroid metabolic process (GO:0045940)
chylomicron assembly (GO:0034378)
Statin Pathway WP430
Metabolic pathway of LDL, HDL and TG, including diseases WP4522
Composition of Lipid Particles WP3601
Nuclear Receptors Meta-Pathway WP2882
PPAR Alpha Pathway WP2878
Prostaglandin Synthesis and Regulation WP98
Adipogenesis WP236
Preimplantation Embryo WP3527
NRF2 pathway WP2884
PPAR signaling pathway WP3942
Cholangiocyte-like I enriched
membrane raft assembly (GO:0001765)
regulation of endothelial cell migration (GO:0010594)
membrane raft organization (GO:0031579)
positive regulation of actin filament bundle assembly (GO:0032233)
negative regulation of platelet activation (GO:0010544)
negative regulation of cellular component movement (GO:0051271)
membrane assembly (GO:0071709)
positive regulation of epithelial cell migration (GO:0010634)
regulation of tooth mineralization (GO:0070170)
positive regulation of vesicle fusion (GO:0031340)
Prostaglandin Synthesis and Regulation WP98
Hypothetical Craniofacial Development Pathway WP3655
Epithelial to mesenchymal transition in colorectal cancer WP4239
Pathogenic Escherichia coli infection WP2272
Metabolic pathway of LDL, HDL and TG, including diseases WP4522
-log10 (adj. P-value)

Supplementary Figure 3. SingleCellNet classification of HLO cells and analysis of WikiPathways[6]. Corresponding to Fig. 1d and f.

a. SingleCellNet (SCN)[5] classifier output heatmap trained on 31 clusters from human liver scRNA-seq data[4]. Heatmap showing the score calculated by SCN for each cell of the reference data (x-axis) for each of the 31 possible clusters and the 'random' category (y-axis). **b.** Individual SCN[5] results for each day 21 (D21) replicate. Heat maps showing classification score for each individual cell with the final SCN-annotation in the sample (x-axis) for each of the 31 possible clusters and the 'random' category (y-axis). Cluster labels of cell types accounting for < 5% of all cells were removed for readability. **c.** Enrichment of human GO-terms[9] (left) and WikiPathways 2019[6] (right) for the three major cell types in HLOs as labeled in main figure 1f. A maximum of ten significantly (p-value < 0.05) enriched terms was plotted. Leiden resolution = 0.1.

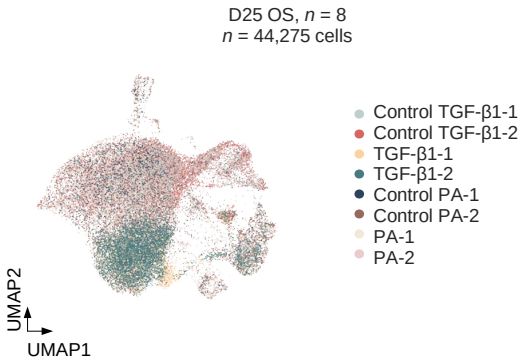
S4


a.
Day 25 OS-HLOs
Control
TGF-β1, 10 ng/ml
b.
D25 HLO, n ≥ 3
TNFA
COL1A1
p > 0.05 (ns)
p = 0.014
p > 0.05 (ns)
Relative mRNA levels
3
-1
5
-1
Exp.
1
2
3
4
5
OA CTRL
OA 400 µM
OA 800 µM
OA CTRL
OA 400 µM
OA 800 µM
c.
Whole slide means
Averaged individual HLO ROIs
HLO ROIs
p = 0.0065
p = 0.017
p = 0.000017
% Sirius red (normalized by control)
170
100
240
100
400
100
CtrlTGF-β1
TGF-β110ng/ml
200
100
25
200
100
0
200
100
0
CtrlPA
PA500 µM
Experiment
1
2
3
4
140
100
60
20
175
100
0
p = 0.005
175
100
25
CtrlOA
OA500 µM

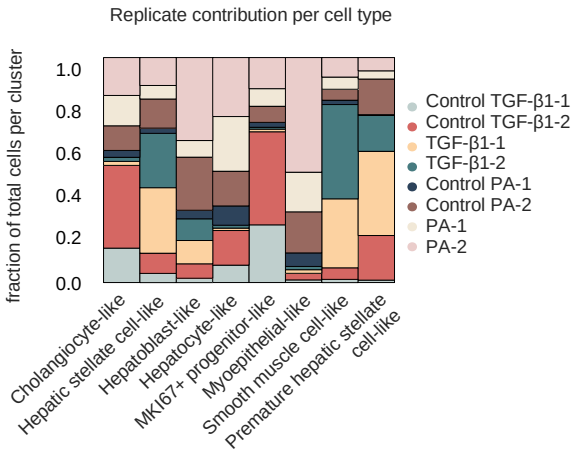
Supplementary Figure 4. Induction of fibrotic phenotypes in OS-cultured HLOs. Corresponding to Fig. 2b, f, i.

a. Enlarged representative brightfield images from Fig. 2, showing day 25 HLOs cultured on an orbital shaker (control, left) and treated with TGF-β1 (10 ng/ml, right). Arrowheads indicate areas of surface roughening. **b.** qPCR results for *COL1A1* and *TNFA.* Shown are qPCR results for relative mRNA levels normalized to *GAPDH* (2^{-ddCt}). Kruskal Wallis test (two-tailed) followed by Conover's post hoc test with Bonferroni correction. Ns, not significant. $N \geq 3$ individual experiments. **c.** Sirius red quantification in HLOs treated with TGF-β1, PA, and OA. Shown is the percentage of Sirius red positive tissue by total tissue (normalized by the mean of the untreated control HLOs in each experiment). Quantification based on the whole slide Sirius red percentage mean (left), mean of averaged 200 x 200 µm regions of interest (ROIs) around single HLOs (center), or the Sirius red percentage per tissue for each HLO ROI (right). Kruskal-Wallis (two-tailed) test followed by a post hoc Conover's test with Bonferroni correction. $N \geq 4$ HLOs from one individual experiment (organoid ROIs), $n \geq 3$ individual experiments (all plots).

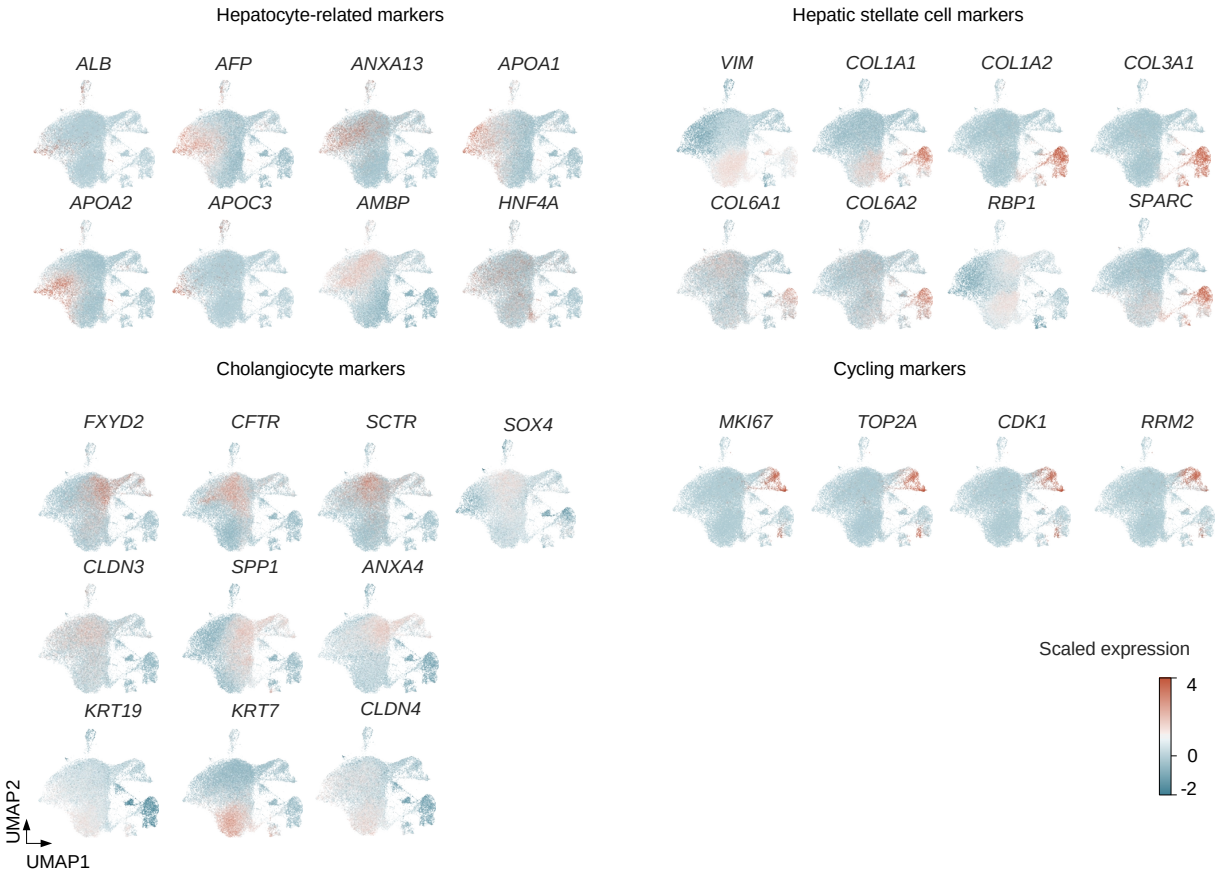
a.


D25 OS, n = 8
n = 44,275 cells
Control TGF-β1-1
Control TGF-β1-2
TGF-β1-1
TGF-β1-2
Control PA-1
Control PA-2
PA-1
PA-2
UMAP2
UMAP1

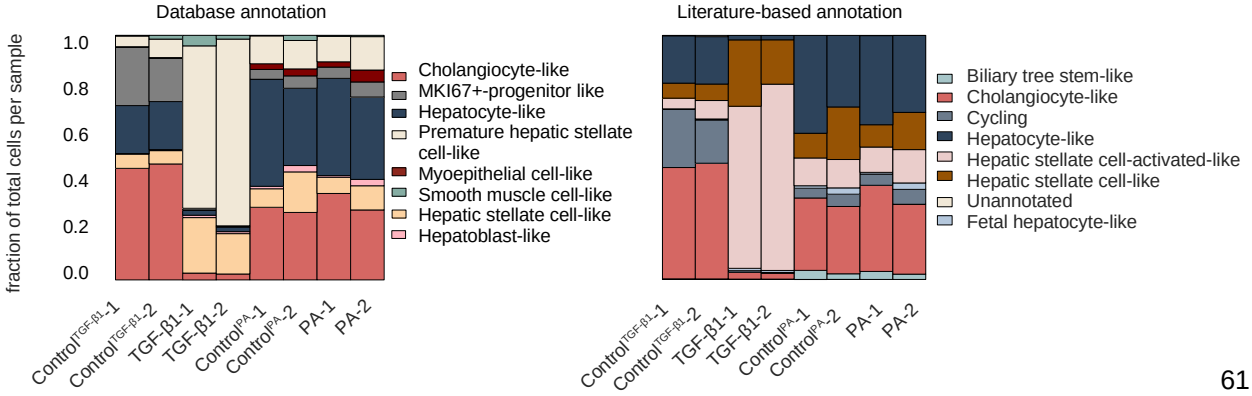
b.


Replicate contribution per cell type
fraction of total cells per cluster
1.0
0.8
0.6
0.4
0.2
0.0
Cholangiocyte-like
Hepatic stellate cell-like
Hepatoblast-like
Hepatocyte-like
MKI67+ progenitor-like
Myoepithelial-like
Smooth muscle cell-like
Premature hepatic stellate cell-like
Control TGF-β1-1
Control TGF-β1-2
TGF-β1-1
TGF-β1-2
Control PA-1
Control PA-2
PA-1
PA-2

c.


Hepatocyte-related markers
ALB
AFP
ANXA13
APOA1
APOA2
APOC3
AMBP
HNF4A
Hepatic stellate cell markers
VIM
COL1A1
COL1A2
COL3A1
COL6A1
COL6A2
RBP1
SPARC
Cholangiocyte markers
FXYD2
CFTR
SCTR
SOX4
CLDN3
SPP1
ANXA4
KRT19
KRT7
CLDN4
Cycling markers
MKI67
TOP2A
CDK1
RRM2
Scaled expression
4
0
-2
UMAP2
UMAP1

d.

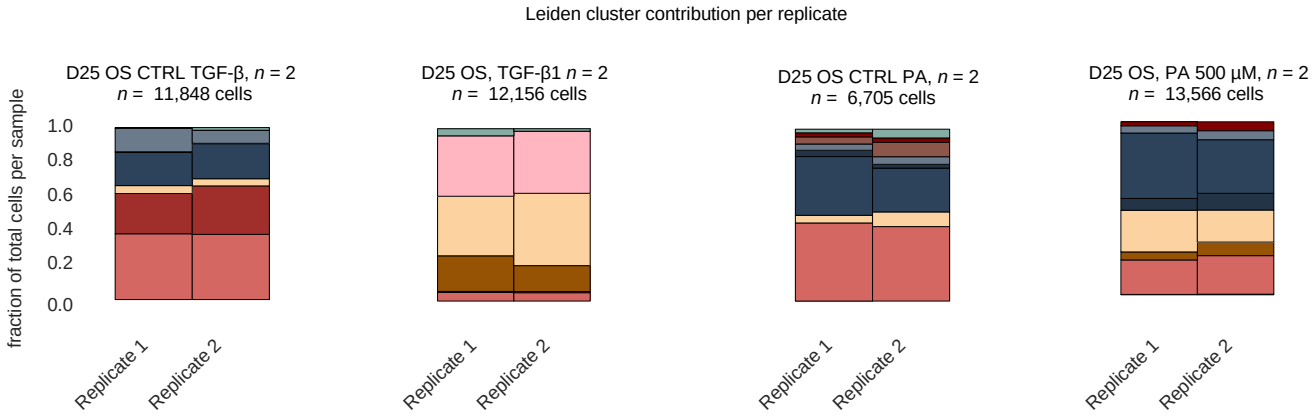

Database annotation
fraction of total cells per sample
1.0
0.8
0.6
0.4
0.2
0.0
Cholangiocyte-like
MKI67+-progenitor like
Hepatocyte-like
Premature hepatic stellate cell-like
Myoepithelial cell-like
Smooth muscle cell-like
Hepatic stellate cell-like
Hepatoblast-like
Control TGF-β1-1
Control TGF-β1-2
TGF-β1-1
TGF-β1-2
Control PA-1
Control PA-2
PA-1
PA-2
Literature-based annotation
Biliary tree stem-like
Cholangiocyte-like
Cycling
Hepatocyte-like
Hepatic stellate cell-activated-like
Hepatic stellate cell-like
Unannotated
Fetal hepatocyte-like

Supplementary Figure 5. Marker gene expression and replicate distribution in 44,275 single cells from control and injured HLOs. Corresponding to Fig. 3b.

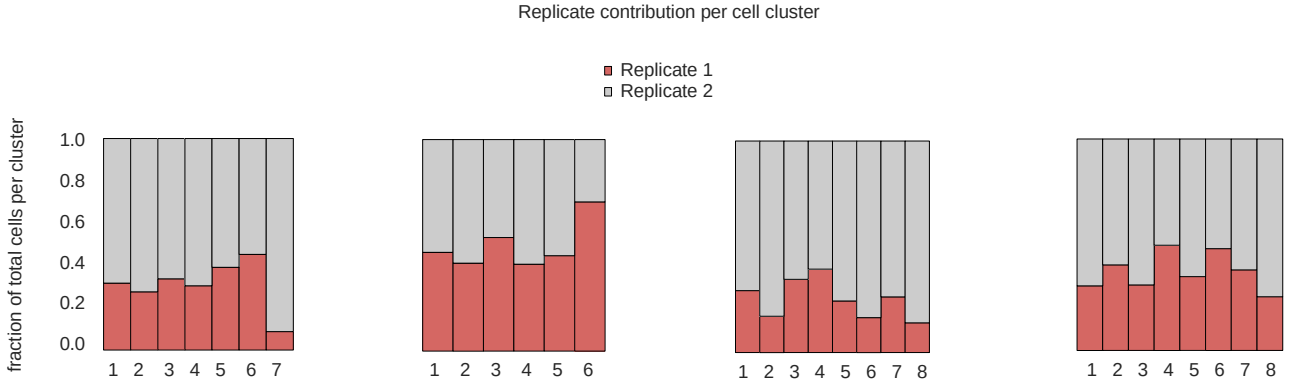
a. UMAP plot mapping 44,275 single cells from OS-HLOs treated with TGF-β1 (10 ng/ml, *n* = 2) or PA (500 µM, *n* = 2) and their respective controls (*n* = 4) colored by replicate. **b.** Barplot indicating the replicate distribution as the proportion of total cells per cluster for each cluster calculated with the Leiden algorithm at resolution 0.5, cluster labels as annotated by database approach (Methods). **c.** UMAP plots are colored for the scaled expression of canonical cell type marker genes expressed in HLOs after quality control (QC). **d.** Barplots showing cell type distributions across replicates as the proportion of each cell type per total cells per sample. Annotations generated by database annotation (left) and literature-based annotation (right), Leiden cluster resolution = 0.5.

S6

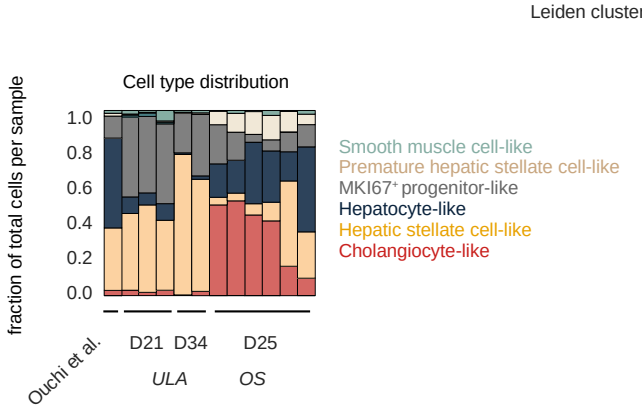
a.


Leiden cluster contribution per replicate
D25 OS CTRL TGF-β, n = 2
n = 11,848 cells
D25 OS, TGF-β1 n = 2
n = 12,156 cells
D25 OS CTRL PA, n = 2
n = 6,705 cells
D25 OS, PA 500 µM, n = 2
n = 13,566 cells
fraction of total cells per sample
1.0
0.8
0.6
0.4
0.2
0.0
Replicate 1
Replicate 2

b.


Replicate contribution per cell cluster
Replicate 1
Replicate 2
fraction of total cells per cluster
1.0
0.8
0.6
0.4
0.2
0.0
1 2 3 4 5 6 7
1 2 3 4 5 6
1 2 3 4 5 6 7 8
1 2 3 4 5 6 7 8

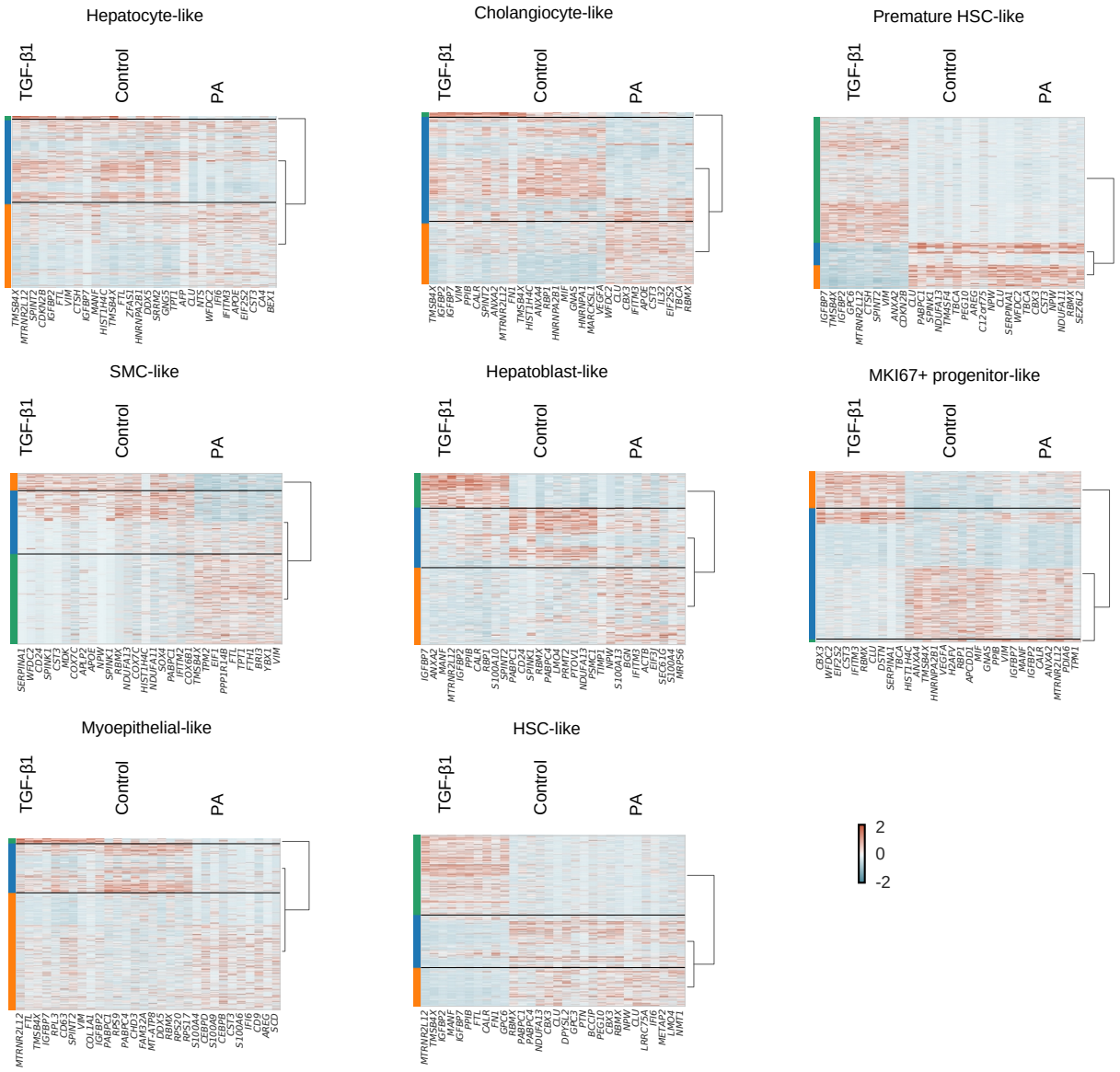
c.


Leiden cluster (resolution = 0.5)
Cell type distribution
fraction of total cells per sample
1.0
0.8
0.6
0.4
0.2
0.0
Ouchi et al.
D21 D34
D25
ULA
OS
Smooth muscle cell-like
Premature hepatic stellate cell-like
MKI67⁺ progenitor-like
Hepatocyte-like
Hepatic stellate cell-like
Cholangiocyte-like

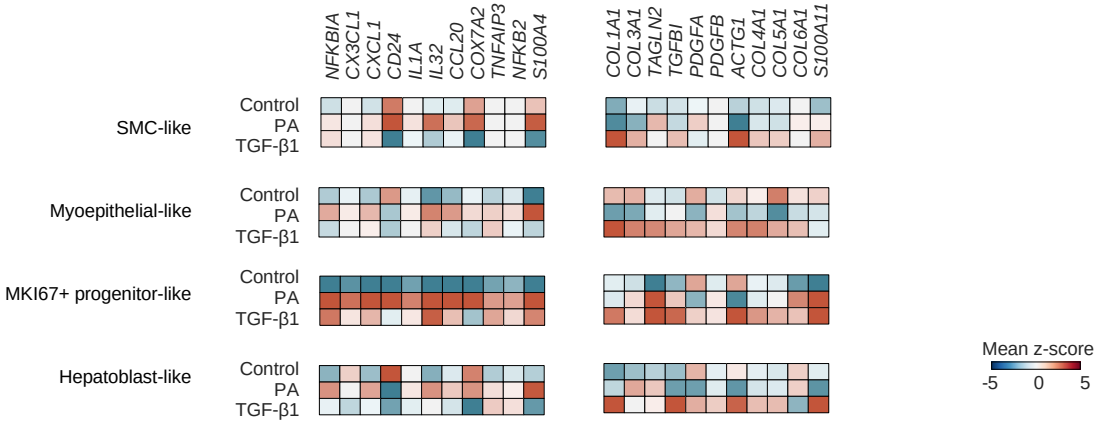
Supplementary Figure 6. Contributions per replicate for TGF-β1 and PA treatment with respecitive controls. Corresponding to Fig. 3b.

a. Barplots indicating the contribution of each cell cluster (each in an individual color) per replicate, displayed as the fraction of total cells per sample. Distributions are shown for all OS-HLO conditions performed in duplicate: TGF-β1 (10 ng/ml, *n* = 2) or PA (500 µM, *n* = 2) and their respective controls (*n* = 2 for each condition). **b.** Barplots display the replicate contribution as the fraction of total cells per cluster for each cell cluster for all samples. **c.** Barplot shows the cell cluster proportions for control samples and indicates a relative reduction of HSC-like and an induction of hepatocyte- and cholangiocyte-like cells in OS-HLOs. Cell type contribution for each individual replicate is indicated as the fraction of total cells per sample for each cluster calculated with the Leiden algorithm at resolution 0.5 across harmony-integrated 94,753 single cells from 16 individual samples representing different HLO conditions and time points: day 21 ultra-low attachment (ULA-) cultured HLOs: *n* = 3 replicates and 17,854 cells, day 34 ULA-HLOs: *n* = 2 replicates and 16,670 cells, day 25 orbital shaker HLOs: *n* = 6 replicates and 30,569 cells, day 25 orbital shaker HLOs with treatments: *n* = 4 replicates and 25,609 cells, not shown, and one HLO reference: *n* = 1 replicate and 4,051 cells (Ouchi et al.)[7].

a.


Hepatocyte-like
Cholangiocyte-like
Premature HSC-like
SMC-like
Hepatoblast-like
MKI67+ progenitor-like
Myoepithelial-like
HSC-like
TGF-β1
Control
PA
2
0
-2

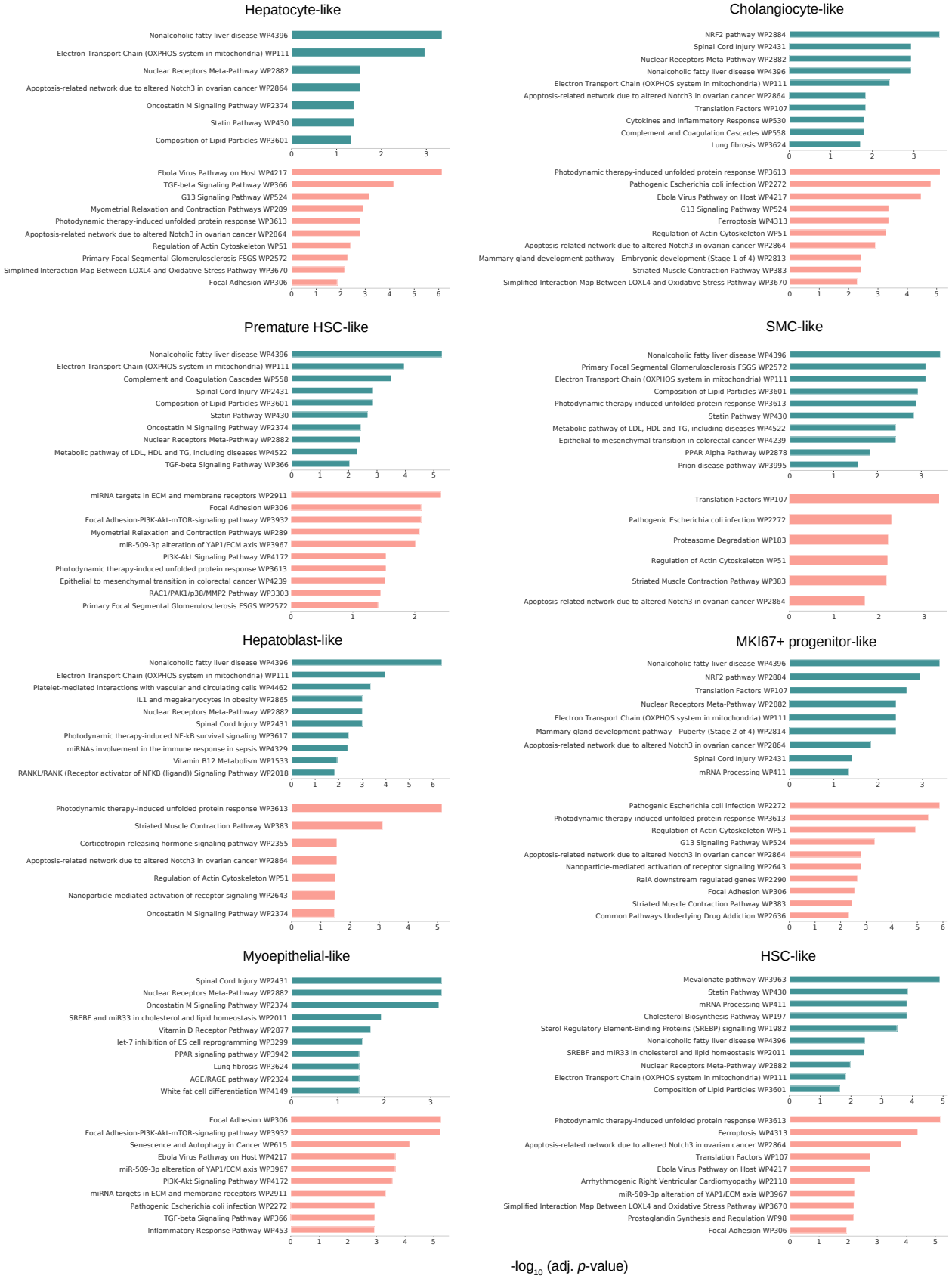
b.


NFKBIA
CX3CL1
CXCL1
CD24
IL1A
IL32
CCL20
COX7A2
TNFAIP3
NFKB2
S100A4
COL1A1
COL3A1
TAGLN2
TGFBI
PDGFA
PDGFB
ACTG1
COL4A1
COL5A1
COL6A1
S100A11
SMC-like
Myoepithelial-like
MKI67+ progenitor-like
Hepatoblast-like
Control
PA
TGF-β1
Mean z-score
-5
0
5

Supplementary Figure 7. Differential gene expression (DGE) for each major cell type after treatment with TGF-β1 and PA compared to controls and gene modules for fibrosis and inflammation. Corresponding to Fig. 3e-f.

a. Heatmaps showing top differentially expressed genes based on Wilcoxon rank sum test calculated on raw count data for each cluster separately. Raw count data for each cell type was extracted, DGE was performed and expression was normalized and scaled again per cluster for visualization only. **b.** Differential expression is displayed for selected genes associated with fibrosis and inflammation for PA treatment, TGF-β1 treatment, and controls for all cell types indicated on the left. Differential expression is displayed as the mean z-score calculated separately for each cell cluster. Data are displayed for cell types not shown in Fig 3. Full lists of differentially expressed genes for each cell type are supplied in **Supplementary Tables 7-13**.

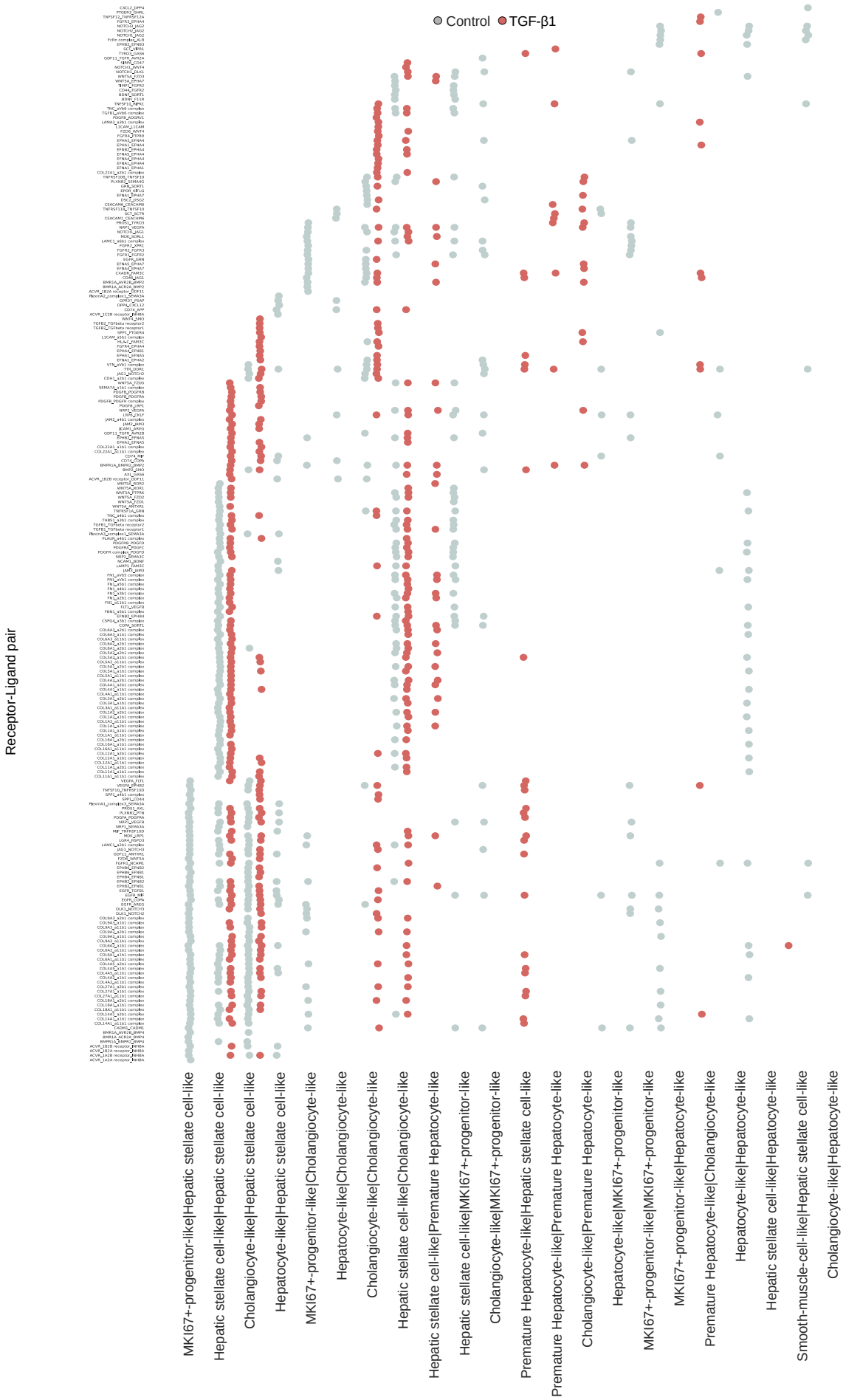
S8


Hepatocyte-like
Nonalcoholic fatty liver disease WP4396
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Nuclear Receptors Meta-Pathway WP2882
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Oncostatin M Signaling Pathway WP2374
Statin Pathway WP430
Composition of Lipid Particles WP3601
Ebola Virus Pathway on Host WP4217
TGF-beta Signaling Pathway WP366
G13 Signaling Pathway WP524
Myometrial Relaxation and Contraction Pathways WP289
Photodynamic therapy-induced unfolded protein response WP3613
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Regulation of Actin Cytoskeleton WP51
Primary Focal Segmental Glomerulosclerosis FSGS WP2572
Simplified Interaction Map Between LOXL4 and Oxidative Stress Pathway WP3670
Focal Adhesion WP306
Cholangiocyte-like
NRF2 pathway WP2884
Spinal Cord Injury WP2431
Nuclear Receptors Meta-Pathway WP2882
Nonalcoholic fatty liver disease WP4396
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Translation Factors WP107
Cytokines and Inflammatory Response WP530
Complement and Coagulation Cascades WP558
Lung fibrosis WP3624
Photodynamic therapy-induced unfolded protein response WP3613
Pathogenic Escherichia coli infection WP2272
Ebola Virus Pathway on Host WP4217
G13 Signaling Pathway WP524
Ferroptosis WP4313
Regulation of Actin Cytoskeleton WP51
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Mammary gland development pathway - Embryonic development (Stage 1 of 4) WP2813
Striated Muscle Contraction Pathway WP383
Simplified Interaction Map Between LOXL4 and Oxidative Stress Pathway WP3670
Premature HSC-like
Nonalcoholic fatty liver disease WP4396
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Complement and Coagulation Cascades WP558
Spinal Cord Injury WP2431
Composition of Lipid Particles WP3601
Statin Pathway WP430
Oncostatin M Signaling Pathway WP2374
Nuclear Receptors Meta-Pathway WP2882
Metabolic pathway of LDL, HDL and TG, including diseases WP4522
TGF-beta Signaling Pathway WP366
miRNA targets in ECM and membrane receptors WP2911
Focal Adhesion WP306
Focal Adhesion-PI3K-Akt-mTOR-signaling pathway WP3932
Myometrial Relaxation and Contraction Pathways WP289
miR-509-3p alteration of YAP1/ECM axis WP3967
PI3K-Akt Signaling Pathway WP4172
Photodynamic therapy-induced unfolded protein response WP3613
Epithelial to mesenchymal transition in colorectal cancer WP4239
RAC1/PAK1/p38/MMP2 Pathway WP3303
Primary Focal Segmental Glomerulosclerosis FSGS WP2572
SMC-like
Nonalcoholic fatty liver disease WP4396
Primary Focal Segmental Glomerulosclerosis FSGS WP2572
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Composition of Lipid Particles WP3601
Photodynamic therapy-induced unfolded protein response WP3613
Statin Pathway WP430
Metabolic pathway of LDL, HDL and TG, including diseases WP4522
Epithelial to mesenchymal transition in colorectal cancer WP4239
PPAR Alpha Pathway WP2878
Prion disease pathway WP3995
Translation Factors WP107
Pathogenic Escherichia coli infection WP2272
Proteasome Degradation WP183
Regulation of Actin Cytoskeleton WP51
Striated Muscle Contraction Pathway WP383
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Hepatoblast-like
Nonalcoholic fatty liver disease WP4396
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Platelet-mediated interactions with vascular and circulating cells WP4462
IL1 and megakaryocytes in obesity WP2865
Nuclear Receptors Meta-Pathway WP2882
Spinal Cord Injury WP2431
Photodynamic therapy-induced NF-kB survival signaling WP3617
miRNAs involvement in the immune response in sepsis WP4329
Vitamin B12 Metabolism WP1533
RANKL/RANK (Receptor activator of NFKB (ligand)) Signaling Pathway WP2018
Photodynamic therapy-induced unfolded protein response WP3613
Striated Muscle Contraction Pathway WP383
Corticotropin-releasing hormone signaling pathway WP2355
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Regulation of Actin Cytoskeleton WP51
Nanoparticle-mediated activation of receptor signaling WP2643
Oncostatin M Signaling Pathway WP2374
MKI67+ progenitor-like
Nonalcoholic fatty liver disease WP4396
NRF2 pathway WP2884
Translation Factors WP107
Nuclear Receptors Meta-Pathway WP2882
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Mammary gland development pathway - Puberty (Stage 2 of 4) WP2814
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Spinal Cord Injury WP2431
mRNA Processing WP411
Pathogenic Escherichia coli infection WP2272
Photodynamic therapy-induced unfolded protein response WP3613
Regulation of Actin Cytoskeleton WP51
G13 Signaling Pathway WP524
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Nanoparticle-mediated activation of receptor signaling WP2643
RalA downstream regulated genes WP2290
Focal Adhesion WP306
Striated Muscle Contraction Pathway WP383
Common Pathways Underlying Drug Addiction WP2636
Myoepithelial-like
Spinal Cord Injury WP2431
Nuclear Receptors Meta-Pathway WP2882
Oncostatin M Signaling Pathway WP2374
SREBF and miR33 in cholesterol and lipid homeostasis WP2011
Vitamin D Receptor Pathway WP2877
let-7 inhibition of ES cell reprogramming WP3299
PPAR signaling pathway WP3942
Lung fibrosis WP3624
AGE/RAGE pathway WP2324
White fat cell differentiation WP4149
Focal Adhesion WP306
Focal Adhesion-PI3K-Akt-mTOR-signaling pathway WP3932
Senescence and Autophagy in Cancer WP615
Ebola Virus Pathway on Host WP4217
miR-509-3p alteration of YAP1/ECM axis WP3967
PI3K-Akt Signaling Pathway WP4172
miRNA targets in ECM and membrane receptors WP2911
Pathogenic Escherichia coli infection WP2272
TGF-beta Signaling Pathway WP366
Inflammatory Response Pathway WP453
HSC-like
Mevalonate pathway WP3963
Statin Pathway WP430
mRNA Processing WP411
Cholesterol Biosynthesis Pathway WP197
Sterol Regulatory Element-Binding Proteins (SREBP) signalling WP1982
Nonalcoholic fatty liver disease WP4396
SREBF and miR33 in cholesterol and lipid homeostasis WP2011
Nuclear Receptors Meta-Pathway WP2882
Electron Transport Chain (OXPHOS system in mitochondria) WP111
Composition of Lipid Particles WP3601
Photodynamic therapy-induced unfolded protein response WP3613
Ferroptosis WP4313
Apoptosis-related network due to altered Notch3 in ovarian cancer WP2864
Translation Factors WP107
Ebola Virus Pathway on Host WP4217
Arrhythmogenic Right Ventricular Cardiomyopathy WP2118
miR-509-3p alteration of YAP1/ECM axis WP3967
Simplified Interaction Map Between LOXL4 and Oxidative Stress Pathway WP3670
Prostaglandin Synthesis and Regulation WP98
Focal Adhesion WP306
-log10 (adj. p-value)
PA enriched
TGF-β1 enriched

Supplementary Figure 8. WikiPathway 2019[7] enrichment analysis across cell clusters. Corresponding to Fig. 3e and f.

Barplots show the negative decadic logarithm of adjusted *p*-values for Wiki-Pathway 2019[6] terms enriched among differential genes for each condition for each cell cluster. Cut-off for plots is an adjusted *p*-value below 0.05.

Significant interactions for TGF-β1 ($p < 0.05$)
n = 2 replicates / condition



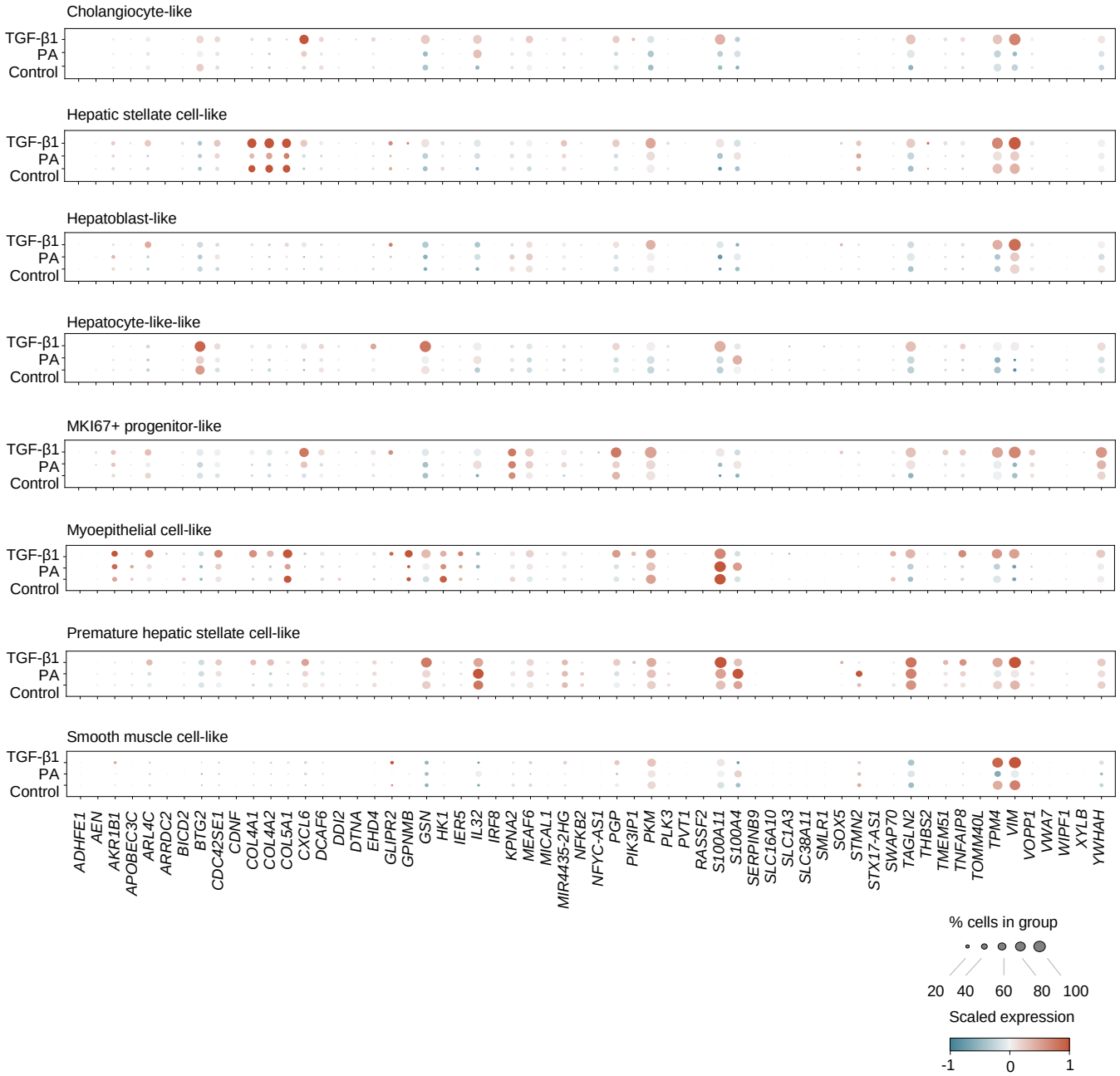
Supplementary Figure 9. Overview of differential cell-cell interactions in OS-HLOs upon TGF-β1 treatment. Corresponding to Fig. 4c. Dotplots showing all significant ($p < 0.05$) receptor-ligand pairs per cell-cell cluster pair from CellPhoneDB analysis for OS-cultured injured HLOs (red) and their respective controls (grey). Y-axes denote the receptor-ligand pair, each column (x-axis) represents a cell-cell pair. TGF-β1 ($n = 2$) and TGF-β1-control ($n = 2$) interactome. Only cell-cell pairs with significant interactions are displayed.

Significant interactions for PA (p < 0.05)
n = 2 replicates / condition
Control
PA 500 µM
Receptor-Ligand pair
Cholangiocyte-like|Hepatic stellate cell-like
Cholangiocyte-like|Cholangiocyte-like
Hepatocyte-like|Cholangiocyte-like
Hepatic stellate cell-like|Cholangiocyte-like
Hepatic stellate cell-like|Hepatic stellate cell-like
Hepatocyte-like|Hepatic stellate cell-like
Hepatic stellate cell-like|Hepatocyte-like-like
Cholangiocyte-like|Hepatocyte-like-like
Premature Hepatocyte-like-like|Hepatic stellate cell-like
Premature Hepatocyte-like-like|Cholangiocyte-like
Cholangiocyte-like|Premature Hepatocyte-like-like
Premature Hepatocyte-like-like|Premature Hepatocyte-like-like
Premature Hepatocyte-like-like|Hepatocyte-like-like
Hepatocyte-like-like|Hepatocyte-like-like
Hepatocyte-like-like|Premature Hepatocyte-like-like
Hepatic stellate cell-like|Premature Hepatocyte-like-like

Supplementary Figure 10. Overview of differential cell-cell interactions in OS-HLOs upon PA treatment. Corresponding to Fig. 4c.

Dotplots showing all significant ($p < 0.05$) receptor-ligand pairs per cell-cell cluster pair from CellPhoneDB analysis for OS-cultured injured HLOs (red) and their respective controls (grey). Y-axes denote the receptor-ligand pair, each column (x-axis) represents a cell-cell pair. PA ($n = 2$) and PA-control ($n = 2$) interactome. Only cell-cell pairs with significant interactions are displayed.

S11



Supplementary Figure 11. Expression of cell type- and injury- specific candidate genes composing signatures predicting fibrosis stages in NAFLD gradually change from healthy to PA-treated to TGF-β1-treated HLOs. Corresponding to Fig. 6c.

Dot plots of scaled expression values are displayed for genes from 98-gene signature[8] to predict fibrosis stages in a cohort of 143 patients across different stages of NAFLD. Each panel shows the scaled expression (calculated previously on all clusters jointly) for the individual clusters among the database-annotated cell types corresponding to Fig. 3b. The gene list was reduced to genes present in each individual replicate (n = 4 controls, n = 2 PA, n = 2 TGF-β1, total 44,275 cells) and expressed in more than 3 cells. Dot sizes correspond to the fraction of cells in a group expressing the respective gene, and color indicates expression level.

References:

1. Franzén, O., Gan, L. M. & Björkegren, J. L. M. PanglaoDB: A web server for exploration of mouse and human single-cell RNA sequencing data. *Database* 1–9 (2019).
2. Zhang, X. et al. CellMarker: A manually curated resource of cell markers in human and mouse. *Nucleic Acids Res.* 47, D721–D728 (2019).
3. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44, W90–W97 (2016).
4. Sharma, A. et al. Onco-fetal Reprogramming of Endothelial Cells Drives Immunosuppressive Macrophages in Hepatocellular Carcinoma. *Cell* 183, 377-394 (2020).
5. Tan, Y. & Cahan, P. SingleCellNet : A Computational Tool to Classify Single Cell RNA-Seq Data Across Platforms. *Cell Syst.* 9, 207-213 (2019).
6. Martens, M. et al. WikiPathways: connecting communities. *Nucleic Acids Res.* 49, D613–D621 (2021).
7. Ouchi, R. et al. Modeling Steatohepatitis in Humans with Pluripotent Stem Cell-Derived Organoids. *Cell Metab.* 30, 374-384.e6 (2019).
8. Pantano, L. et al. Molecular characterization and cell type composition deconvolution of fibrosis in NAFLD. *Sci. Rep.* 11, 1–14 (2021).
9. The Gene Ontology Consortium. Gene Ontology : tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).