A human multi-lineage hepatic organoid model for liver fibrosis

Yuan Guan1, Annika Enejder2, Meiyue Wang1, Zhuoqing Fang1, Lu Cui3, Shih-Yu Chen4, Jingxiao Wang1, Yalun Tan1, Manhong Wu1, Xinyu Chen1, Patrik K. Johansson2, Issra Osman1, Koshi Kunimoto3, Pierre Russo5, Sarah C. Heilshorn2 & Gary Peltz1✉

To investigate the pathogenesis of a congenital form of hepatic fibrosis, human hepatic organoids were engineered to express the most common causative mutation for Autosomal Recessive Polycystic Kidney Disease (ARPKD). Here we show that these hepatic organoids develop the key features of ARPKD liver pathology (abnormal bile ducts and fibrosis) in only 21 days. The ARPKD mutation increases collagen abundance and thick collagen fiber production in hepatic organoids, which mirrors ARPKD liver tissue pathology. Transcriptomic and other analyses indicate that the ARPKD mutation generates cholangiocytes with increased TGFβ pathway activation, which are actively involved stimulating myofibroblasts to form collagen fibers. There is also an expansion of collagen-producing myofibroblasts with markedly increased PDGFRB protein expression and an activated STAT3 signaling pathway. Moreover, the transcriptome of ARPKD organoid myofibroblasts resemble those present in commonly occurring forms of liver fibrosis. PDGFRB pathway involvement was confirmed by the anti-fibrotic effect observed when ARPKD organoids were treated with PDGFRB inhibitors. Besides providing insight into the pathogenesis of congenital (and possibly acquired) forms of liver fibrosis, ARPKD organoids could also be used to test the anti-fibrotic efficacy of potential anti-fibrotic therapies.
Liver fibrosis is a pathologic condition that results from extracellular matrix (ECM) accumulation in response to chronic liver injury. This leads to loss of liver parenchymal cells, reduced liver function, and has severe complications. Although it is most commonly an acquired condition caused by viral infection or chronic alcohol exposure, a few genetic diseases can cause liver fibrosis. While the rate of progression and histological features can vary in response to the different causes, excess production of an altered ECM underlies all forms of liver fibrosis. This fibrotic state results from an interaction between parenchymal and nonparenchymal liver cells, and possibly involves infiltrating immune cells. The key nonparenchymal cell is the hepatic stellate cell (HSC), which is activated by a fibrogenic stimulus to transdifferentiate into a myofibroblast with increased expression of α-smooth muscle actin (SMA), desmin (DES), and type I collagen (COL1A1). Under normal conditions, the liver ECM consists of laminins, collagens (types I, III, and IV), and various proteoglycans, which provide important signals to maintain liver cell homeostasis. However, myofibroblasts increase their production of fibril-forming collagen types I and III, collagen fibers become the most abundant component in the fibrotic liver ECM. Thus, activated myofibroblasts and the collagens they produce are essential mediators of liver fibrogenesis. No available treatments can prevent or reverse its progression if the underlying cause cannot be treated.

ARPKD is a monogenic disorder that causes kidney and liver pathology. The kidney disease progresses to renal failure and perinatal death in 30%, but for those that survive the perinatal period, liver disease becomes progressively more severe and becomes the major cause of morbidity and mortality. ARPKD liver disease is characterized by dilated intrahepatic bile ducts and a biliary fibrosis that is referred to as congenital hepatic fibrosis (CHF). ARPKD is caused by dysfunction of primary cilia due to mutations within poly cystic kidney and hepatic disease-1 (PKHD1), which encodes a 4074 amino acid multi-domain transmembrane protein (fibrocytin/polyductin, FPC) that is expressed in the primary cilia of renal tubular epithelial cells and cholangiocytes. Of the >800 PKHD1 mutations identified, the most common causative mutation is Thr36Met, which accounts for 20% of all mutated alleles and frequently appears in unrelated families of different ethnic origins. FPC is part of a protein complex that is a mechanotransducer of environmental signals.

We present an in vitro model system for human liver fibrosis where induced pluripotent stem cells (iPSCs) differentiate into human hepatic organoids (HOs). We demonstrate that this organoid system, when combined with genome editing technologies, reproduces ARPKD liver pathology, which includes biliary abnormalities and extensive fibrosis, that develops in HOs in only 21 days. The ARPKD organoids have an expanded population of activated collagen-producing myofibroblasts, which have transcriptomic similarities with myofibroblasts in liver tissue obtained from patients with commonly occurring forms of liver fibrosis.

**Results**

**Characterization of a multi-lineage HO.** We have previously shown that iPSCs differentiate into hepatoblasts, which then differentiate into HOs in response to the sequential application of specific growth factor combinations that are added to the culture media. The HOs have sheets of hepatocytes, as well as cholangiocytes that are organized into epithelia around the lumina of bile duct-like structures. We have also previously shown that these HOs can mediate many of the biosynthetic and drug metabolism functions characteristic of human liver.

Trichrome staining and anti-collagen (COL1A1) immunostaining staining of liver sections indicated that the mesenchymal tissue within HOs resembled that in liver (Fig. 1f). Overall, the images indicate that HOs have a more complex pattern of antigen expression and form more complex structures than we previously appreciated. The organoids have ductal structures that are surrounded by mesenchymal cells and a collagen-enriched ECM, and CDS1 staining indicated that some type of vascular structures were indeed present in HOs (Fig. 1g). HOs have primary cilia in their bile ducts (Fig. 1h), which is an essential structure for analyzing ciliopathies and single-cell RNA sequencing (scRNA-Seq) has become a powerful method for characterizing cell types within tissues. Therefore, scRNA-Seq analysis was performed on iPSC, hepatoblast, and HO cultures. As described in Supplementary note 1, the cells at these differentiation stages could be separated into distinct clusters (Supplementary Fig. 1a-c). As expected, the organoids had hepatocytes, cholangiocytes, and the previously characterized bi-potential progenitor cells (Supplementary Fig. 1h) that give rise to hepatocytes and cholangiocytes (Supplementary Fig. 1d). In addition, the scRNA-Seq data (Fig. 1i, Supplementary Fig. 1e, f) and high-dimensional time of flight mass cytometry (CyToF) (Supplementary Fig. 1i, j) data indicated that the organoids also contain other cell types, which expressed markers found on endothelial cells and hepatic stellate cells (HSC). The presence of HSC in the hepatic organoids was confirmed by immunostaining (Supplementary Fig. 1g). Thus, HOs contain a more complex array of cells of different lineages than was noted in our prior studies. Second Harmonic Generation (SHG) microscopy was used to demonstrate that HOs can synthesize pro-collagen and have the enzymatic machinery required for cross-linking collagen to form thick fibers (Supplementary note 2, Fig. 1j).

**An organoid model for ARPKD liver pathology.** PKHD1 mutations that cause amino acid substitutions are generally associated with a non-lethal presentation, while neonatal death tends to be associated with frame shift or splice variant alleles. Consistent with these clinical observations, we could not produce an iPSC line with an engineered homozygous Ashkenazi founder frame shift mutation (c.3761_3762delCCinsG) in PKHD1. However, we successfully engineered homozygous PKHD1M36 mutations into three different iPSC lines (C1–C3) that were produced from different control individuals. Inter-individual variation is responsible for a large percentage of the phenotypic differences observed in different iPSC lines. However, phenotypic differences that commonly occur in lines with the ARPKD mutation (but not in isogenic control lines) can be unequivocally ascribed to the mutation. The morphology of HOs prepared from all three PKHD1M36 iPSC lines (which will be referred to as ARPKD lines) was altered in a characteristic manner (Fig. 2b, c). ARPKD organoids have an increased number of irregular bile ducts: bile duct structures occupied 30–40% of the area in ARPKD organoids versus 10–15% in control HOs. ARPKD organoids also had a markedly increased amount of ECM, which occupied 25–30% of the area in ARPKD HOs versus 0.3–0.5% of control HOs (Fig. 2d, e). Immunostaining confirmed that an increased amount of collagen 1A (COL1A1) was diffusely deposited in ARPKD organoids (Fig. 2f). Also, in contrast to the simple columnar morphology of the ductal epithelium in control organoids, ARPKD organoids had a disorganized ductal epithelium (Fig. 2g). The basis for the abnormal ductular morphology was investigated by immunofluorescence staining. In control organoids, zonula occludens protein 1 (ZO-1) and EZRIN were expressed in a characteristic manner on the apical side of the cholangiocytes surrounding the ductal lumens (Fig. 2h). This pattern indicates that ductal epithelial
cells formed tight junctions that were properly oriented with respect to the ductal plane\textsuperscript{36}, which explains why control organoids had a normal tubular architecture. In contrast, ZO-1 expression was decreased in ARPKD organoids, and was present in a non-oriented manner within the ductal structures. Also, the characteristic expression pattern of a cell polarity-determining protein (Vang-Like 1, VANGL1) in CK19\textsuperscript{+} cholangiocytes was similarly altered in ARPKD organoids (Fig. 2f). Consistent with the immunostaining results, the expression levels for multiple mRNAs associated with the primary ciliary or with planar cell polarity cell were decreased in ARPKD HO\textsubscript{s} relative to isogenic control HO\textsubscript{s} (Fig. 2j). Thus, the orientation and polarity of bile duct cholangiocytes were disrupted in ARPKD organoids. Also, primary ciliary is ∼2-fold more abundant in ARPKD (vs control) organoids (Fig. 2k, l and Supplementary Fig. 2f); but their length was not affected by the ARPKD mutation, which was measured at different stages of differentiation (Supplementary Fig. 2g-h). Thus, the ARPKD mutation does not interfere with the formation of primary cilia, but it is possible that it could impact their function.

**Quantitative assessment of fibrosis.** ARPKD Liver tissue had enlarged bile ducts, and a markedly increased amount of ECM (including collagen fibers) was deposited throughout the liver tissue (Fig. 3a). Hence, ARPKD organoid morphology mirrors the major pathologic features observed in ARPKD liver disease, which includes the increased ECM deposition that is characteristic of CHF. As described in Supplementary note 2, second harmonic generation (SHG) microscopy has been used to analyze liver fibrosis\textsuperscript{37,38}. Therefore, 3D SHG images were obtained to quantitatively evaluate collagen fiber abundance and structure within control and ARPKD hepatic organoids. Control HO\textsubscript{s} organoids had a confluent network of thick collagen fibers throughout the entire organoid (Fig. 3b). A quantitative analysis confirmed the marked increase in cross-linked collagen fibers in ARPKD organoids (average volume fraction 17.0% ± 6.8%, n = 43 organoids) versus control organoids (average volume fraction 3.2% ± 2.9%, n = 40 organoids); and the increase was noted irrespective of whether ARPKD organoids prepared from the 3 donors were evaluated individually (p < 0.05) or as a mixture (p = 0.0001) relative to isogenic controls (Fig. 3b). Moreover, the abundance of thicker collagen fibers (diameter ≥ 6.0 μm) in ARPKD organoids was significantly greater than in isogenic control organoids (individual p < 0.05, mixture p < 0.001).

Liver tissue samples obtained from normal and ARPKD patients were also analyzed by SHG microscopy. The collagen...
Fiber volume in ARPKD liver tissue (13.3% ± 10.6%) was much greater than control liver tissue (1.7% ± 1.2%; p = 0.0001, Fig. 3c).

Also, collagen fibers in the ARPKD liver tissue formed thicker, aligned bundles that were much larger than the ~1 micron-sized, isolated fibers present in normal liver tissue. A significantly higher volume fraction of thick collagen bundles (diameter ≥ 6.0 μm) was present in ARPKD tissue (p < 0.0001). Of note, we did not observe fluid-filled cysts in ARPKD organoids, nor in the ARPKD liver tissue examined here. Polycystic liver disease is a characteristic of Autosomal Dominant Polycystic Kidney Disease (ADPKD), while ARPKD liver disease is characterized by congenital fibrosis and ductal abnormalities. The SHG analyses...
indicated that the pathologic changes observed in ARPKD organoids mirrored those in ARPKD liver tissue.

Pathogenesis of ARPKD liver disease. A multiplex analysis of scRNA-Seq data generated from ARPKD and isogenic control HOs prepared from three unrelated individuals (C1–C3) was performed (Supplementary Fig. 2b). In total, the transcriptomes of 7461 cells in control organoids (average 61,000 reads and 2884 genes per cell), and of 11,960 cells in corresponding ARPKD organoids (average 36,000 reads and 2079 genes per cell) were analyzed (Supplementary Fig. 2c). We identified 15 clusters in control and ARPKD organoids (Fig. 4a, b, Supplementary Fig. 2d). While each of the different cell types was tightly clustered, there were very significant differences in the cellular composition of control and ARPKD organoids. The transcriptome of cell clusters in the HOs was compared with cells in control and cirrhotic human livers, and the initial analysis indicated that various types of mesenchymal cells (clusters 0, 1, 3, 4, 6, 7), mesothelial cells (cluster 2), hepatocyte precursor cells (cluster 5), early endothelia (clusters 8, 9), cholangiocytes (cluster 10), and endothelia (cluster 14) were present in the organoids.
Cluster identities were confirmed by expression of canonical mRNA markers (Supplementary Fig. 2e). To identify cells that could contribute to ARPKD liver pathology, cluster cell abundance within control and ARPKD organoids was calculated (Fig. 4a). Multiple clusters (5, 8, and 12–14) had a similar percentage of cells in control and ARPKD organoids, which indicates that differences observed in other clusters were not caused by a preparation artifact. There was a dramatic shift in the mesenchymal populations present in ARPKD organoids. The percentage of cells within clusters 0, 1, and 7 were 9-, 7-, and 7.8-fold increased, respectively, in ARPKD organoids relative to control organoids, while the percentage of clusters 3 and 4 were 10- and 5-fold increased, respectively, in control organoids. Cluster 6 was present in control organoids

**Fig. 4 scRNA-Seq analysis of HOs.** a A combined t-SNE plot of the ARPKD and control organoid scRNA-Seq data shows the cell types within the 15 different clusters identified by k-nearest neighbor analysis, which include: various types of mesenchymal cells (Mes; clusters 0, 1, 3, 4, 6, 7), mesothelia (Meso; cluster 2), hepatocyte precursors (Hep pre; cluster 5), early endothelia (eEC, clusters 8, 9), cholangiocytes (Cho; cluster 10), and endothelia (EC; cluster 14). The heatmap compares the percentages of cells within each cluster present in control and ARPKD organoids. The colors represent the log-transformed percentage of cells within each cluster. The cell types are listed in Table S1 and were annotated as described in Table S2. b t-SNE plots separately show the 15 clusters identified in control and ARPKD organoids. c The developmental trajectories of ARPKD and isogenic control cells are similar at the iPSC (day 0) and hepatoblast (HB, day 9) stages, but differ significantly at the organoid (day 21) stage. The left graph shows the cells in ARPKD and isogenic control cultures at the iPSC and HB and hepatic organoid stages. The ARPKD and isogenic control cells are separately graphed (middle, right). Each cell cluster in the organoids is indicated by a color, while the iPSC and HB cells are gray. The ARPKD and isogenic control cells are very similar at the iPSC and HB stages, but significantly diverge at the organoid stage. As examples, clusters 0 and 1 are far more abundant in ARPKD than in control organoids; while clusters 3 and 4 are far more abundant in control organoids; and cluster 9 is present in ARPKD, but not in control organoids. d Violin plots show the increased level of expression of mRNAs for liver progenitor cell (SOX9, 2.1-fold; KRT17, 5.3-fold) and cell proliferation (CDKN1A, 3.2-fold; CDKN2B, 1.9-fold) markers in ARPKD relative to control cholangiocytes (cluster 10). e, f Control and ARPKD HOs were immunostained with antibodies to SOX9 and HNF4A or with anti-CK8 and anti-Ki67 antibodies. Scale bars are 50 μm. The dashed lines separate areas with hepatocytes (Hep+) or cholangiocytes (Cho+) from other regions. The boxplots show the fraction of Ki67+ cholangiocytes within 3 pairs of ARPKD and isogenic control HOs (n = 20 per group, C1 vs A1 p < 2.22 × 10−16, C2 vs A2 p = 2.4 × 10−16, C3 vs A3 p < 2.22 × 10−16, ***p < 0.001). Box plots in f (center line, median; box limits, upper and lower quartiles; whiskers, 1.5 × interquartile range). g CK19 and JAG1 expression increased in ARPKD organoids. bhSNE maps of CyToF data generated using ARPKD and control HOs. The dotted circle shows the JAG1+ cell population.

(Fig. 4a, Supplementary Table 1). Cluster identities were confirmed by expression of canonical mRNA markers (Supplementary Fig. 2e). To identify cells that could contribute to ARPKD liver pathology, cluster cell abundance within control and ARPKD organoids was calculated (Fig. 4a). Multiple clusters (5, 8, and 12–14) had a similar percentage of cells in control and ARPKD organoids, which indicates that differences observed in other clusters were not caused by a preparation artifact. There was a dramatic shift in the mesenchymal populations present in ARPKD organoids. The percentage of cells within clusters 0, 1, and 7 were 9-, 7-, and 7.8-fold increased, respectively, in ARPKD organoids relative to control organoids, while the percentage of clusters 3 and 4 were 10- and 5-fold increased, respectively, in control organoids. Cluster 6 was present in control organoids
(14.5%) but was virtually absent (<0.1%) in ARPKD organoids. A series of scRNA-seq (Fig. 4c, Supplementary Fig. 3a-e), and RT-PCR (Supplementary Fig. 3f, g) analyses of developing organoid cultures indicated that the ARPKD mutation’s effect on the mesenchymal populations probably occurred when hepatoblasts differentiate into the cells present the mature liver organoid (see Supplementary note 3).

**ARPKD mutation effects on cholangiocytes.** To investigate the pathogenesis of ARPKD ductal abnormalities, we compared the transcriptomes of cholangiocytes (cluster 10) in ARPKD and control organoids. Pathway analysis of the most differentially expressed genes (Supplementary Table 2, Supplementary Fig. 4A) indicated that the most highly enriched pathway was extracellular matrix organization (Supplementary Fig. 4b, c). ARPKD cholangiocytes had an increased level of expression of mRNAs that regulate the cell cycle and mitosis, collagens and proteins involved in collagen fiber assembly, and of genes expressed in liver progenitor cells (Fig. 4d, Supplementary Fig. 4d). While the cholangiocytes in ARPKD organoids are HNF4A and SOX9 double-positive cells, cholangiocytes in control organoids are either HNF4A or SOX9 positive, but these mRNAs are not co-expressed in the same cells (Fig. 4e). Ki67 staining shows that ARPKD HOs have many more proliferating cholangiocytes (Fig. 4f). CytTOF analysis also showed that JAG1+ cells, which are CK19+ cholangiocytes, were 1.5-fold more abundant in ARPKD than in control HO (Fig. 4g). The transcriptome of ARPKD cholangiocytes was enriched for mRNAs within TGFβ-associated signaling pathways (TGFβ signaling in EMT, signaling by TGFβ Receptor Complex). Also, ARPKD cholangiocytes had a 1.3-fold increased level of TGFβ1 mRNA expression (p = 3.1 × 10^-8) and a 4.8-fold increase in the level of a TGFβ inducible mRNA (TGFβI) (Supplementary Fig. 4e). ARPKD cholangiocytes also express reduced levels of mRNAs encoding planar cell polarity (DVL2, FZD6) and ductal epithelium/tight junction (CLDNI, CLDH1, TJP1, EPCAM) proteins than control organoids (Supplementary Fig. 4f). Thus, the transcriptome analysis indicates that ARPKD cholangiocytes are less mature, more proliferative, and have a reduced level of expression of mRNAs regulating cell polarity and epithelial cell function, but have increased TGFβ pathway activation, and are more actively involved in collagen fiber generation than control cholangiocytes.

**Myofibroblast expansion in ARPKD.** Of the 15 cell clusters, cluster 0 was of particular interest because: (i) it had the largest increase in ARPKD (27%) versus control (2.9%) organoids, and (ii) pathway enrichment analysis indicated that they expressed mRNAs associated with protein digestion, the JAK-STAT pathway, and ECM-receptor interactions (Fig. 5a–c). Most importantly, (iii) of the six mesenchymal cell clusters, the cluster 0 transcriptome was most similar to that of myofibroblast-like cells found in cirrhotic human liver tissue39 (Supplementary Fig. 5a), and (iv) it had the highest level of PDGFRB mRNA, which is a myofibroblast marker (Supplementary Fig. 5b). Immunostaining and CytTOF results indicated that the cells in ARPKD organoids had a marked increase in PDGFRB, SMA, PDGFRα, and VIM protein expression, and PDGFRB+ cells in ARPKD organoids were 4.5-fold increased versus control organoids (Fig. 5c–f, Supplementary Fig. 5c). While a small number of SMA+ cells were occasionally present in control organoids, ARPKD organoids had an increased number of SMA+ cells, which were present in clusters located near ductal structures (Fig. 5f). Similarly, ARPKD liver tissue also had a markedly increased level of PDGFRB and SMA protein expression (Fig. 5g). According to multiple criteria, the cluster 0 transcriptome resembles that of myofibroblasts; and their number was markedly increased in ARPKD organoids.

The cluster 0 cells in ARPKD organoids (27% of total) were distinctly different from cluster 3 cells that were the most abundant in control organoids (22% of total). Relative to cluster 3, cluster 0 cells have increased levels of mRNAs for multiple JAK-STAT3 signaling pathway components and for its effector molecules such as PIM-1,40,41 and PDGFRB, and for multiple RNAs that regulate stem cell plurioty and fibroblast activation (Supplementary Fig. 6a, b). For example, cluster 0 cells had an increased level of leukemia inhibitory factor receptor (LIFR) mRNA, which is of interest because LIF induces an invasive and activated state in fibroblasts in a STAT3-dependent manner42. LIF mRNA was expressed at a low level within multiple different cell clusters, which was similar to its pattern in human liver tissue; and clusters 0 and 3 expressed equivalent LIF mRNA levels (Supplementary Fig. 6c). Similarly, PDGFA/B mRNAs were expressed by multiple cell types within the organoids (Supplementary Fig. 6c). In contrast, cluster 3 cells had increased levels of mRNAs that are associated with cell cycle arrest or cellular senescence (Supplementary Fig. 6d). Thus, the ARPKD mutation promotes the production of myofibroblast-like cells that have the characteristics of activated and proliferative cells, and their transcriptome is quite distinct from those in control organoids. Also,STAT3 mRNA expression increased during HO differentiation (Supplementary Fig. 6e); and the active phosphorylated form of STAT3 was present in ARPKD organoid myofibroblasts and in ARPKD liver tissue. Interestingly, Phospho-STAT3 was also found in cholangiocytes in ARPKD organoids and in areas with abnormal ducts in ARPKD liver tissue (Fig. 5h–j).

**Similarities with commonly occurring forms of human liver fibrosis.** To investigate whether the ARPKD organoid fibrosis mechanistically resembled that in the commonly occurring forms of human liver fibrosis, 254 genes whose expression was increased in the cluster 0 cells were used to form a myofibroblast-specific expression signature (Supplementary Table S3). As described in Supplementary note 4, Gene Set Enrichment Analysis (GSEA) assesses whether this signature was present in other types of fibrotic liver tissue by calculating a normalized expression score (NES). The myofibroblast expression signature was very strongly associated with hepatitis C virus infection-induced cirrhotic liver (NSE 2.56, false discovery rate (FDR) 0), but not normal liver (NES = 2.55, FDR 0) (Fig. 5k). The myofibroblast signature was also associated with non-alcoholic steatohepatitis (NASH) (NSE 1.65, FDR 0), which is now the most common cause of chronic liver disease, but it was not associated with normal liver tissue (NES = 1.64, FDR 0), obesity (NES 0.98, FDR 0.55), or hepatocellular carcinoma (NES 0.24, FDR 0.4). Also, the myofibroblast gene signature found in human cirrhotic liver tissue was strongly correlated with ARPKD organoids but not with control organoids. The results obtained from two different types of GSEA analyses indicate that ARPKD organoid myofibroblasts resemble those causing the commonly occurring forms of human liver fibrosis.

To determine whether the PDGFRB-STAT3 pathway is activated in liver cancer-induced fibrosis, tumor and the adjacent normal liver tissues resected from subjects with liver fibrosis, which was caused by hepatocellular carcinoma (HCC) or cholangiocarcinoma (CCC), were immunostained with PDGFRB and pSTAT3 antibodies. PDGFRB expression was markedly elevated in the tumor tissues (vs adjacent normal liver tissue). Moreover, while pSTAT3 expression was undetectable in adjacent normal liver tissue, low levels of pSTAT3 expression within/near...
Fig. 5 ARPKD HOs have an expanded population of myofibroblasts that resemble those in fibrotic human liver tissue. 

**a** KEGG pathway enrichment analysis identifies 3 gene networks whose expression is most increased in cluster 0 cells: protein digestion and adsorption, JAK-STAT signaling, and extracellular matrix (ECM)-receptor interactions. 

**b** A volcano plot showing the differentially expressed genes (fold change >1.5, adjust p-value <0.05) when the transcriptomes of cluster 0 and cluster 3 are compared. The 20 genes whose mRNAs exhibit the highest level of differential expression are highlighted, which include STAT3, PDGFRB, and SOCS3. The gray dashed line indicates the 1.5-fold cutoff for a differentially expressed gene. A red (or blue) color indicates that a gene is up (or down) regulated in cluster 0 (vs. cluster 3); a black color indicates that the expression level was not significantly different. The differentially expressed genes among the groups were determined using the Wald test in DESeq2.

**c** Violin plots showing three genes (PDGFRB, STAT3, SOCS3) whose mRNAs were markedly increased in cluster 0 cells relative to cluster 3.

**d** Day 21 control and ARPKD organoids were immunostained with an anti-PDGFRB antibody. PDGFRB is highly expressed in ARPKD, but not in control, HOs. Scale bar, 50 μm.

**e** A scatter plot of CyToF data performed with anti-PDGRFβ and anti-CD56 antibodies shows that ARPKD organoids have a markedly increased amount of PDGFRB+ cells (9.0%) relative to control organoids (2.1%).

**f** tSNE maps (upper) generated from CyToF data indicate that ARPKD HOs have an increased amount of SMA+ cells. (Lower) Representative images of HOs immunostained with anti-smooth muscle actin (SMA) and anti-Ck8 antibodies. The yellow dotted lines indicate ductal structures. While only a few SMA+ cells were present at a limited number of sites within control organoids (indicated by an arrow); large clusters of SMA+ cells were present within ARPKD organoids (indicated by arrowhead) and were located near ductal structures. Scale bar, 10 μm.

**g** Normal and ARPKD liver tissue was immunostained with antibodies to PDGFRB, Ck8, and SMA. PDGFRB expression was much higher in ARPKD than in control liver tissue. Also, SMA+ cells were only located within perivascular regions of normal liver tissue; but they infiltrate the parenchyma of ARPKD liver tissue (as shown in the inset). The white dashed line separates epithelial and mesenchymal areas. Scale bar, 50 μm.

**h** ARPKD and Control HOs were immunostained with anti-CK8 and anti-phospho-STAT3 (Tyr705) antibodies. Phospho-STAT3 was extensively expressed in bile duct and mesenchymal cells in ARPKD HOs. The right panels are enlargements of the boxed yellow area in the ARPKD organoid; the arrows indicate mesenchymal cells with Phospho-STAT3 within the nucleus. Scale bar, 5 μm.

**i** Normal and ARPKD liver tissues were immunostained with antibodies to PDGFRB, CK8, and SMA. The bar graph shows the GSEA results (normalized enrichment score (NES)) obtained from each analysis; and the false discovery rate (FDR) for each comparison is shown at the top of each bar. The myofibroblast gene expression signature was very strongly associated with cirrhotic and NASH liver tissue, but this signature was not induced by obesity alone or by the presence of a HCC. GSEA was also performed by comparing the cluster 0 cell transcriptome with overlapping genes present in the expression signatures defined from the scRNA-Seq analysis of myofibroblasts and B cells present in cirrhotic human livers. Bottom: GSEA results for the correlation of the ARPKD myofibroblast gene expression signature with that in liver tissues obtained from: 10 normal and 10 cirrhotic subjects (Cirrhosis); 98 normal and 126 NASH subjects (NASH).
PDGFRB+ cells were seen in the CCC tumor tissue (Fig. 6). The immunostaining results indicate that the PDGFRB-STAT3 pathway is also activated in liver tumor-induced fibrosis.

**Inhibitor effect.** To determine if the PDGFRB/STAT3 pathway was essential to the pathogenesis of ARPKD fibrosis, we examined the effect of three inhibitors of the PDGFR tyrosine kinase (Crenolanib, Sunitinib, and Imatinib) on the extent of ARPKD organoid fibrosis using three different methods: quantitative collagen immunostaining in whole organoids, and measurement of hydroxyproline and quantitative collagen mRNA assessment. Immunostaining of whole-mount organoids indicated that treatment with 10 μM concentrations of PDGFR inhibitors (Imatinib, Crenolanib) significantly decreased collagen formation in ARPKD organoids (p < 0.001) (Fig. 7a). The fibrosis scores in the drug-treated ARPKD organoids were close to that of control hepatic organoids (p-value = 0.068) (Fig. 7b). The inhibitors did not have a very significant effect on bile duct formation in ARPKD organoids (Fig. 7c, d). Although imatinib had a marginal effect, the bile duct area was not reduced to that of control organoids; and its effect on bile ducts was small relative to the very substantial effect that these drugs had on the extent of fibrosis developing in ARPKD organoids.

**COL1A1** mRNA levels in ARPKD organoids were decreased ~17-fold after treatment with the three PDGFR inhibitors (p < 0.001) (Fig. 7e). We also measured 4-hydroxyproline (4-OH Pro) levels, which increase in fibrotic tissue. All three PDGFR inhibitors at 10 μM (p < 0.005) and 0.05 μM (p < 0.005) concentrations significantly
increased 4-OH Pro levels in ARPKD organoids (Fig. 7f, g). Thus, FLT3 inhibition cannot explain the anti-fibrotic efficacy of these drugs. The anti-fibrotic effect of these inhibitors confirms that the PDGFR pathway contributes to the development of fibrosis in ARPKD organoids. Even though these PDGFR inhibitors also inhibit the FLT3 tyrosine kinase; FLT3 mRNA is not expressed above background in any of the 15 cell clusters identified in the hepatic organoids (Supplementary Fig. 7). Of importance, the viability of control organoids was not altered after 1 week of treatment with 10 μM concentrations of the PDGFR inhibitors (Supplementary Fig. 8).

**Discussion**

ARPKD liver pathology, which includes bile duct abnormalities and fibrosis, develops in hepatic organoids with an engineered mutation in *PKHD1* (which encodes a mutated FPC) that is the most common cause of ARPKD. Discordant phenotypes appearing within families has led to suggestions that genetic modifiers could affect ARPKD disease expression. However, ARPKD pathology developed in organoids independently of donor genetic background, and comparison with isogenic controls confirmed that the pathology is induced by the ARPKD mutation. Also, since bile fluid does not flow in the organoid cultures, mutation-induced abnormalities in the mechanosensory function of primary cilium do not appear to contribute to the pathogenesis of ARPKD liver disease. A detailed analysis of ARPKD HOs indicates that four mutation-induced alterations are essential to the pathogenesis of ARPKD liver disease: (i) ARPKD cholangiocytes are less mature, their transcriptional program is associated and of PDGFB-STAT-3 pathways in ARPKD and of PDGFB-STAT-3 pathways in ARPKD; (ii) there are signifi-

**Fig. 7 PDGFR inhibitors reduce fibrosis in ARPKD organoids.** Control or ARPKD organoid cultures were treated with vehicle or PDGFR inhibitors, and the extent of fibrosis developing in the cultures was measured by three different methods. a Whole-mount immunostaining of control HOs and of drug (10 μM) or vehicle-treated ARPKD organoids. Representative images of organoids stained with COL1A and CK8 antibodies, and counterstained with Hoechst 33342. Scale bar, 50 μm. b The fibrosis score was determined by measuring the ratio of the volume of COL1A and CK8 in 100 stacked images obtained from 4 whole-mount organoids per group. The PDGFR inhibitors (Crenolanib, Imatinib) decreased the fibrosis score in ARPKD organoids to levels that were similar to control organoids (n = 4 per group, ARPKD vs Imatinib p = 0.00089, ARPKD vs Crenolanib p = 0.0042, ARPKD vs DAPT p = 0.0014). c Representative images of COL1A and CK8 immunostained ARPKD organoids before and after treatment with 10 μM of the indicated inhibitor. For comparison, an image of an immunostained control organoid is also shown. Scale bar is 50 μm. d The fraction of the total area occupied by bile ducts (CK8+ area) within isogenic control (normal) and ARPKD organoids after treatment with 10 μM of the indicated drug (n = 5 per group, DMSO vs Crenolanib p = 0.37, DMSO vs Sunitinib p = 0.58, DMSO vs Imatinib p = 0.015, n.s. no significant difference). e COL1A1 mRNA levels in ARPKD organoids are significantly higher than in control organoids, and were decreased 17-fold after treatment with the PDGFR inhibitors. RT-PCR measurements made on 3 organoids per treatment group; and the values were normalized relative to simultaneously measured GAPDH mRNA levels (DMSO vs Crenolanib p = 0.00048, DMSO vs Sunitinib p = 0.0021, DMSO vs Imatinib p = 0.00071, ARPKD vs DAPT p = 0.00068). Box plots in b, d, and e (center line, median; box limits, upper and lower quartiles; whiskers, 1.5 × interquartile range). Statistical differences between the groups were determined by unpaired two-tailed t-test. *p < 0.05; **p < 0.01; and ***p < 0.001. f, g 4-OH Pro levels in ARPKD organoids are decreased by treatment with 10, 2, or 0.5 μM concentrations of PDGFRB inhibitors (Control vs DMSO p = 0.00097, DMSO vs Crenolanib p = 0.0046, DMSO vs Sunitinib p = 0.00082, DMSO vs Imatinib p = 0.0013; DMSO vs 2 μM Crenolanib p = 0.00047, DMSO vs 0.5 μM Crenolanib p = 0.0019, DMSO vs 2 μM Sunitinib p = 0.37, DMSO vs 0.5 μM Sunitinib p = 0.0014, DMSO vs 2 μM Imatinib p = 0.0064, DMSO vs 0.5 μM Imatinib p = 0.004). Box plots in b, d, and e (center line, median; box limits, upper and lower quartiles; whiskers, 1.5 × interquartile range). Statistical differences between the groups were determined by unpaired two-tailed t-test. *p < 0.05; **p < 0.01; and ***p < 0.001.
organoids is consistent with prior observations in ARPKD rodent models and in cultured cells. The ARPKD organoid data generated here along with prior observations enabled us to generate a tentative model for the pathogenesis of ARPKD; a self-sustaining circuit is generated, which maintains the fibrotic state (Supplementary note 5, Supplementary Fig. 9).

This is the first liver organoid system that models the key feature of human liver fibrosis, which is the formation of an excess of assembled collagen fibers, and we demonstrate that its features mirror that of fibrotic human liver tissue. Although ARPKD fibrosis is genetically induced, myofibroblasts are the key mediators of all types of liver fibrosis, including the commonly occurring forms that are caused by chronic alcohol exposure or viral diseases5,31–33. Irrespective of whether a liver fibrosis is congenital or acquired, aside from treatments aimed at the underlying cause, no treatments can prevent or reverse its progression. A major limitation on therapeutic development has been the model systems. Prior to the emergence of organoid methodology, in vitro cellular models lack the three-dimensional architecture of liver, they do not have the different cell types involved in fibrogenesis34, and most require injury-induced activation34,35 that induces variability. Commonly used rodent models also require an injury-inducing agent (carbon tetrachloride, bile duct ligation, or dietary modulation), and are time consuming and expensive. Since fibrosis is an intrinsic feature of the ARPKD organoid, no exogenous disease-inducing agent is required, which eliminates the variability caused by the injury-inducing agent. Moreover, any conclusions drawn from animal models are limited by concerns about interspecies differences and about their fidelity with the processes mediating human liver fibrosis. In contrast, this HO model is a human-based in vitro system that generates a 3D multi-lineage liver-like tissue with bile ducts, which has all of the enzymes required to form cross-linked collagen fibers. Furthermore, the myofibroblast gene expression signature in ARPKD organoids, along with the activation of the PDGFRB-STAT3 pathway, resembles that found in the commonly occurring acquired forms of human fibrotic liver diseases. These attractive features indicate that use of this HO could increase our ability to produce new treatments for fibrotic liver disease.

Methods

iPSC and HO generation. Human iPSC lines (C1, C2, and C3) were prepared as previously described26. The methods for differentiating iPSCs into hepatic organoids (HOs) via culture in a series of media containing different growth factors were set to one only if connected to a collection of pixels with values above the high threshold. SHG and CARS microscopy. Under confocal fluorescence microscopy, the volume fraction of thick collagen (individual donor C1, C2, or C3, and mixed, respectively). The stacks covered a field-of-view of 149.9 × 149.9 μm, spanned by 1504 × 1504 pixels or 2048 × 2048 pixels. It resulted in a pixel size of ~100 nm or 73 nm. Each volume image (z-stack) was composed by 15-71 two-dimensional images. The following papers provide a review of CARS microscopy52: a detailed description of CARS microscopy of intracellular lipid stores53, and describe its use for measurement of lipid stores in liver tissue54,55. The CARS and SHG signals were each detected by an analog photomultiplier tube (Hamamatsu). The spatial resolution was estimated to ~0.3 μm in the x-y plane, and ~1 μm in depth. Control and ARPKD HO samples were analyzed using the Fiji Imaris software (v. 9.2.0) to determine the diameters of the collagen fibers, the fibers in the SHG images were binarized using the Imaris Hysteresis thresholding plugin. To test whether the volume fractions obtained for ARPKD and control organoids were significantly different, the unpaired Student’s t-test was applied.

scRNA-seq. scRNA-seq was performed using the 10X Genomics Chromium system.
(Pleasanton, CA). Cellular suspensions (1 x 10^6/ml) were then loaded onto the 10× Chromium instrument (10× Genomics), and were then sequenced according to the manufacturer’s instructions. Three ARPKD and 3 control organoid samples were separately sequenced on two lanes of an Illumina HiSeq2500 Rapid Mode instrument, yielding 889 M total reads.

scRNA-seq for trajectory construction. iPSC, HB, and HO cells were collected on days 0, 9, and 30, respectively, from isogenic control and ARPKD organoid cultures (from three different donors), and were prepared from according to previously described methods. One million cells at each differentiation stage and group (ARPKD or isogenic control) were labeled using a total of 6 cell capture antibodies according to the manufacturer’s protocol (BioLegend TotalSeq™). Then, the pooled cells were captured using two 10X Single Cell 3’ Reagent Kit v3 runs, and sequenced on two lanes of an Illumina HiSeq2500 instrument, which yielded 660 M total reads that covered 10,000 cells in total.

scRNA-seq data analysis. Each sequencing lane generated data from 3 control or 3 ARPKD organoids. The datasets for each organoid were deconvoluted using the ‘demuxlet’ method. To do this, we analyzed previously obtained scRNA-Seq datasets, and identified 3.4 M SNP alleles that could distinguish C1, C2, and C3 donors. These were used to generate a reference vCard file (VCF) that was analyzed by the demuxlet program. The donor for each cell within a lane was identified using the demuxlet program, which was run with the following parameters: the minor allele frequency was set to >1% to identify CNVs; we set ‘–alpha 0 –alpha 0.5’, which assumes that the expected proportion of a 50% genetic mixture from two individuals; and we set ‘–group-list’ to a list of 10X barcodes. By this method, the data in each lane was deconvoluted to determine the donor identity for each transcript.

After the identity of each cell was determined, an expression matrix was generated from 10X Cellranger, and the data was imported into ‘Seurat’ for subsequent analysis. First, cells with unique gene counts <200 or where the percentage of mitochondrial mRNAs was >20% were removed. Then, a total of 4046 variable genes were identified using the default settings. Unwanted sources of variation were removed using the regression analysis was performed to mitigate the effect of signals caused by mRNAs with unique molecular identifiers and mitochondrial expression. Finally, a nonlinear dimension reduction (t-SNE) program in ‘Seurat’ was used to identify 15 unique clusters for all of the 12,580 cells analyzed. Fifteen principal components were used to construct the shared nearest neighbor (SNN) graph, and the parameter regulating the resolution of the ‘FindClusters’ program was set to 0.6. PathfindR (github: egeulgen/pathfindR) was used to perform the pathway enrichment analysis through identification of active subnetworks. Differentially expressed genes (cluster biomarkers) identified from the Seurat analysis were used as the input genes. For cluster 0, 254 marker mRNAs were input for the analysis. In brief, a data frame consisting of the Gene Symbol, log-fold-change, and adjusted-p values were used to perform the active subnetwork search. Pathway enrichment analyses were then performed using the genes identified within each of the identified active subnetworks using either the KEGG or Reactome database.

The gene signatures for the different cell types present in control and fibrotic human liver tissues were obtained from the Gene Expression Omnibus (GEO) database. Tissue samples were isolated from liver lobes that were resected due to liver cancer, and these samples were obtained from patients with and without cirrhosis. The disease status was determined based on the presence of liver cirrhosis (as defined by the presence of fibrosis and cirrhosis). The gene signatures for the different cell types present in control and fibrotic human liver tissues were analyzed using the scRNA-Seq data generated from hepatic organoids.

Transcriptome correlation. We assessed the correlation between the transcriptomes of mesenchymal cell clusters (0, 1, 3, 4, 6, and 7) in hepatic organoids of the four types of mesenchymal cells identified in control and cirrhotic human liver tissue in ref. 49. Mesenchyme 1, vascular smooth muscle cells; Mesenchyme 2, hepatic stellate cells; Mesenchyme 3, myofibroblasts (which were also referred to as scar associated mesenchymal cells (SAMe); and Mesenchyme 4, Mesothelia. To do this, the mean expression level for each gene in the cells within each cluster was determined, and these variable features were ranked based upon the number of clusters that they appeared in. The top 3000 features that were present in the organoid and liver tissue datasets were used for the correlation analysis.

Patient samples and human ethics approval. All tissues in this study were obtained after informed consent was provided, and the human sample studies adhered to relevant ethical guidelines. The de-identified ARPKD liver tissues used in this study were obtained from Children’s Hospital of Philadelphia (CHOP). Liver tissues were collected from two subjects with ARPKD liver disease that were treated at CHOP: ARPKD1: A05-56, 3-month fetus; and ARPKD2: CA9-04, age 34 weeks. De-identified normal liver tissues and cancer tissues used in this study were isolated from liver lobes that were resected due to liver cancer, and these samples were obtained after informed consent was provided, and the studies were performed according to procedures approved by the Stanford University Medical Center IRB (IRB approval #42968). Liver tissues were flash-frozen on dry ice in Tissue-Tek® O.C.T™ (Sakura Finetek U.S.A., Torrance, CA). Tissue blocks were sectioned using a Leica CM3050 S Cryostat into 10-μm sections, and serially obtained sections were stained.

Histology and immunostaining. Fresh organoids were harvested, after settling by gravity, and were embedded in low melting point agarose (IBI Scientific, Dubuque, IA). Organoid blocks were then processed by sectioning the paraffin-embedded tissue into 10-μm sections. The human liver tissues in 4-micron sections were fixed in 4% paraformaldehyde for 10 min, followed by permeabilization with 1% Triton X-100 (Sigma-Aldrich, St. Louis, MO), and blocking was performed with a solution containing 10% chicken serum (Jackson Immunoresearch, Bar Harbor, MA) for 30 min. Then, sections were incubated with the primary antibodies listed in Table S4. When needed, secondary antibody-staining was performed using an Alexa Fluor labeled chicken anti IgG (H + L), which was cross-adsorbed with a secondary antibody in 10% chicken serum (Invitrogen, Pleasanton, CA).

Cilia imaging. Human ARL13B was cloned using the In-Fusion HD Cloning system (Clontech, Takara Bio USA, CA). The full-length ARL13B cDNA (428AA, NM_001174150) was cloned into pRRLSIN.C_pp.PPKGFP.WPRE. The primers used for RT-PCR amplification of ARL13B were Fragment 1.FOR tcccaacgggattcctgccagctgtgagctggagt, Fragment 1.REV cctaccaagttgacacatgatgacactgt, Fragment 2.FOR gatcaatgtgacgacgacgacgacg, Fragment 2.REV cttg aatcctagcagttgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacg
Single-cell mass cytometry (CyTOF). Single-cell preparations from hepatic organoids were generated using the methods described for scRNA-seq analysis. Then, the cell preparations were fixed with 2% paraformaldehyde at room temperature for 20 min, and then washed twice with PBS containing 0.5% BSA. The paraformaldehyde-fixed cells were incubated with metal-conjugated antibodies that were reacted with cell surface antigens (Table S1) for 1 h, washed twice with PBS containing 0.5% BSA. The cells were then permeabilized with methanol for 15 min at 4 °C, washed twice with PBS containing 0.5% BSA, and then incubated with metal-conjugated antibodies against intracellular antigens for 1 h. The cells were then washed with PBS containing 0.5% BSA, and incubated at room temperature for 20 min with an iodine-containing DNA intercalator (Fluidigm) in PBS containing 2% paraformaldehyde. After intercalation/fixation, the cell samples were washed once with PBS containing 0.5% BSA, and twice with water before measurements were made using a CyTOF mass cytometer (Fluidigm). Normalization and de-barcoding were performed using previously described Matlab programs.7,56. After measurement and normalization, the individual data files were analyzed by first gating out doublets, debris, and dead cells based on cell length, DNA content, and cospot staining. t-SNE maps were generated with software tools Cyt, which is a graphical software package obtained from Dana Pe’er.

RT-PCR analyses. Total RNA was extracted using TRIzol® RNA Isolation Reagents (Ambion, Grand Island, NY), and 2 μg RNA was reverse-transcribed using the ‘Script’ cDNA Synthesis Kit (H-RNAD, Hershey, PA) according to the manufacturer’s guidelines. The following TaqMan primer sets (Life Technologies Grand Island, NY) were used for the analyses: COL1A1 Hs00164004_m1, PDGFBR Hs01019589_m1, VIMETIN Hs00985111_m1, ACTA2 (SMA), Hs00426835_g1, CDH1 Hs01023895_m1, KRT19 Hs00716767_s1, GAPDH Hs00727626_x1.

Image segmentation and quantification. Immunostained and trichrome-stained image processing and analysis were performed using the Fiji (2.1.0) implementation of Imagej. Quantification of the area of positive staining was calculated using the ‘Trainable Weka Segmentation’ plug-in.17. For training purposes, each measurement in 5–10 areas of each analyzed element in every image (total 5 images) was used as positive and the classifiers by manually labeling all of the positive spots. Then, for experimental image analysis, the saved classifier was used to generate the probability map for each image. Target channels were then isolated, thresholded, and binarized. For each treatment group, the ‘Area Fraction’ measured in 4 to 6 images were assessed. An unpaired Student’s t-test was used to test whether the measurements were significantly different.

For the quantification of collagen volume in 3-dimensional whole-mount organoids that were stained with anti-COL1A antibodies, split channels including >100 stack of images from 100 to 200 μm segments were used to calculate the volume. This was accomplished by multiplying the sum of the area in each stack by the depth (volume = total area x depth). The collagen volume measured in 4 areas of each type was used to compare the collagen volume difference resulting from treatment with each drug. The macro script that we used for this analysis is available upon request.

Single organoid proline and hydroxyproline quantitation. Individual organoids were formed in Nunclon “Spera” 96-well microplates by seeding 5000 hepatoblasts per well. The plates were then centrifuged at 300 × g for 5 min. The drugs tested were PDGER2B tyrosine kinase inhibitors (Cremadom, Sunthimib, and Ima- tinib), and a NOTCH inhibitor (DAPT). All anti-fibrotic reagents purchased from Selleckchem (TX, USA) or Tocris (MN, USA). The indicated concentration of each drug (10, 2 or 0.5 μM) was added to the microwell on day 12, and the drug was present for the last 10 days of organoid culture. The individual organoids were isolated and transferred to 0.2 ml PCR tube containing 30 μl 6 N HCl, and hydrolysis was performed at 120 °C for 100 min. The clear supernatant was transferred for further dilution and LC-MS analysis. Hydrolyzed organoid samples were brought to a 100 μl volume by addition ddH2O. A 10 μl aliquot from each sample was mixed with an equal volume of an internal standard solution and then dried under SpeedVac. Hydroxyproline and Proline standards (50 μM) and D7-L proline (Cambridge Isotope Laboratories, MA) were used as internal standards. The dried mixture was resuspended in 50 μl water and derivatized with dansyl chloride (Sigma) using our previously described modification28–30 of the method developed by Guo and Lee30. LC–MS analysis was then performed using an Agilent QTOF 6545 (Agilent, Santa Clara) coupled with a 1290 infinity I UHPLC (Agilent, Santa Clara). The samples were run on a Phenomenex C18 kinetic column. A mobile phase was 0.1% formic acid in water and B was 100% acetonitrile. The acquired data were analyzed using MassHunter Quantitative analysis software and the proline and 4-hydroxyproline concentrations were calculated based on a 10-point calibration curve.

Statistics and reproducibility. All images shown were obtained from differentiation experiments using three pairs of iPSC cell lines (control and mutated) that were derived from different three subjects, and at least three independent differentiation were performed for each experiment. The statistical analyses were performed using the two-tailed unpaired Student’s t-test using RStudio software (R version 4.0.2) using the following parameters: ggplot2 (ver: 3.3.2), ggubr (ver: 4.0.0), ggsignif (ver: 0.6.0). All p values lower than 0.05 were considered statistically significant.

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

Data availability. All raw single-cell RNA-seq data and the processed data were deposited in the Gene Expression Omnibus (GEO) and are available under accession GSE154883. An additional single-cell RNA-seq dataset, which was used for assessing organoid differentiation, that was previously generated is available at GSE139832. Three publicly available liver gene expression datasets (cirrhosis, GSE6764; NAFLD, GSE83452; Obesity, GSE126848), were obtained from the Gene Expression Omnibus using the ‘GEOquery’ as described39. Human fetal liver tissue scRNA-Seq dataset used in assessing the hepatoblast cluster was obtained from the Genome Sequencing Center (GSC, CRA002443). R scripts and macros used for performing the analysis are available from the corresponding authors upon request. Source data are provided with this paper.

Received: 19 March 2021; Accepted: 28 September 2021; Published online: 22 October 2021

References
1. Yoon, Y. J., Friedman, S. L. & Lee, Y. A. Antifibrotic therapies: where are we now? Semin Liver Dis. 36, 87–98 (2016).
2. Hernandez-Gea, V. & Friedman, S. L. Pathogenesis of liver fibrosis. Annu. Rev. Pathol. 6, 425–456 (2011).
3. Panizzi, M. & Luong, T. V. Pathogenesis of biliary fibrosis. Biochim. Biophys. Acta Mol. Basis Dis. 1864, 1279–1283 (2018).
4. Panizzi, M. Pathophysiology of liver fibrosis. Dig. Dis. 33, 492–497 (2015).
5. Zhang, C. Y., Yuan, W. G., He, P., Lei, J. H. & Wang, C. X. Liver fibrosis and hepatic stellate cells: etiology, pathological hallmarks and therapeutic targets. World J. Gastroenterol. 22, 10512–10522 (2016).
6. Tsuchida, T. & Friedman, S. L. Mechanisms of hepatic stellate cell activation. Nat. Rev. Gastroenterol. Hepatol. 14, 397–411 (2017).
7. Yin, C., Evasion, K. J., Ashina, K. & Stainer, D. Y. Hepatic stellate cells in liver development, regeneration, and cancer. J. Clin. Invest. 123, 1902–1910 (2013).
8. Zhang, D. Y. et al. A hepatic stellate cell gene expression signature associated with outcomes in hepatitis C cirrhosis and hepatocellular carcinoma after curative resection. Gut 65, 1754–1764 (2016).
9. Mederacke, I. et al. Fate tracing reveals hepatic stellate cells as dominant contributors to liver fibrosis independent of its aetiology. Nat. Commun. 4, 2823 (2013).
10. Aycock, R. S. &eyer, J. M. Collagens of normal and cirrhotic human liver. J. Hepatocell. Tissue Res. 23, 1–31 (1987).
11. Olaso, E. et al. DDR2 receptor promotes MMP-2-mediated proliferation and invasion by hepatic stellate cells. J. Clin. Invest. 108, 1369–1378 (2001).
12. Hartung, E. A. & Guay-Woodford, L. M. Autosomal recessive polycystic kidney disease: a hepatoportal fibrolytic disorder with pleiotropic effects. Pediatr. Nephrol. 11, 1910–1915 (1997).
13. Bucher, A. et al. Clinical manifestations of autosomal recessive polycystic kidney disease (ARPKD): kidney-related and non-kidney-related phenotypes. Pediatr. Nephrol. 29, 1915–1925 (2014).
14. Wilson, P. D. Polycystic kidney disease. N. Engl. J. Med. 350, 151–164 (2004).
15. Hildebrandt, F., Benzing, T. & Katsnian, S. Ciliopathies. N. Engl. J. Med. 364, 1463–1464 (2011).
16. Ward, C. J. et al. The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. Nat. Genet. 30, 259–269 (2002).
17. Onuchic, L. F. et al. PKHD1, the polycystic kidney and hepatic disease 1 gene, encodes a novel large protein containing multiple immunoglobulin-like domains.
plexin-transcription-factor domains and parallel beta-helix 1 repeats. Am. J. Hum. Genet. 70, 1305–1317 (2002).

18. Bergmann et al. Clinical consequences of PKHD1 mutations in 164 patients with autosomal-recessive polycystic kidney disease (ARPKD). Kidney Int. 67, 829–848 (2005).

19. Bergmann, C. Genetics of autosomal recessive polycystic kidney disease and its differential diagnoses. Front Pediatr. 5, 221 (2017).

20. Harris, P. C. & Rossetti, S. Molecular genetics of autosomal recessive polycystic kidney disease. Mol. Genet. Metab. 81, 75–85 (2004).

21. Kim, I. et al. Fibrocystin/polyductin modulates renal tubular formation by regulating polycystin-2 expression and function. J. Am. Soc. Nephrol. 19, 455–468 (2008).

22. Wang, S. et al. Fibrocystin/polyductin, found in the same protein complex with polycystin-2, regulates calcium responses in kidney epithelia. Mol. Cell Biol. 27, 3241–3252 (2007).

23. Rohatgi, R. et al. Mechanoregulation of intracellular Ca2+ in human autosomal recessive polycystic kidney disease cyst-lining renal epithelial cells. Am. J. Physiol. Ren. Physiol. 294, F890–F899 (2008).

24. Banales, J. M. et al. The cAMP effectors Epac and protein kinase a (PKA) are involved in the hepatic cystogenesis of an animal model of autosomal recessive polycystic kidney disease (ARPKD). Hepatology 49, 160–174 (2009).

25. Spiri, C. et al. Protein kinase A-dependent 8-fos(675) -beta-catenin, a novel signaling defect in a mouse model of congenital hepatic fibrosis. Hepatology 51, 1713–1723 (2010).

26. Guan, Y. et al. Human hepatic organoids for the analysis of human genetic diseases. J. Insect Sci. 2, pii: J9459 (2017).

27. Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science 356, eaah4573 (2017).

28. Zeisel, A. et al. Brain structure. Cell types in the mouse cortex and thalamus revealed by a combination of broad and narrow molecular markers. Neuron 78, 559–584 (2015).

29. Montoro, D. T. et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature 560, 319–324 (2018).

30. Plasschaert, L. W. et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature 560, 377–381 (2018).

31. Halpern, K. B. et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature 542, 352–356 (2017).

32. Guan, Y. et al. The phosphatidylyethanolamine biosynthesis pathway provides a new target for cancer chemotherapy. J. Hepatol. 72, 746–760 (2019).

33. Furu, L. et al. Milder presentation of recessive polycystic kidney disease requires presence of amino acid substitution mutations. J. Am. Soc. Nephrol. 14, 2004–2014 (2003).

34. Quint, A. et al. An Ashkenazi founder mutation in the PKHD1 gene. Eur. J. Med. Genet. 59, 86–90 (2016).

35. Kilpinen, H. et al. Common genetic variation drives molecular heterogeneity in human iPSCs. Nature 546, 370–375 (2017).

36. Huang, R. Y., Guilford, P. & Thiery, J. P. Early events in cell adhesion and morphogenesis. Adv. Drug Deliv. Rev. 121, 147–157 (2017).

37. Sun, W. et al. Nonlinear optical microscopy: use of second harmonic generation and two-photon microscopy for automated quantitative liver analysis and two-photon microscopy for automated quantitative liver analysis. J. Biomed. Opt. 14, D487 (2016).

38. Rasband, W. S. ImageJ (U.S. NIH, 2019).

39. Hildebrand, T. & Ruegsegger, P. A new method for the model-independent assessment of thickness in three-dimensional images. J. Micr. 185, 67–75 (1996).

40. Kang, H. M. et al. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. Nat. Biotechnol. 36, 89–94 (2018).

41. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satij, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36, 411–420 (2018).

42. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30 (2000).

43. Fabregat, A. et al. The reactome pathway knowledgebase. Nucleic Acids Res. 44, D481–D487 (2016).

44. Stuart, T. et al. Comprehensive integration of single-cell data. Cell 177, 1888–1902 (2019). e1821.

45. Wolf, F. A. et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. Genome Biol. 20, 59 (2019).

46. Davis, S. & Meltzer, P. S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 23, 1846–1847 (2007).

47. Wurmbach, E. et al. Genome-wide molecular proﬁles of HCV-induced dysplasia and hepatocellular carcinoma. Hepatology 45, 938–947 (2007).

48. Lefebvre, P. et al. Interspecies NASH disease activity whole-genome proﬁling identifies a ﬁbrogenic role of PPARalpha-regulated dermatopontin. JCI Insight 2, 92364 (2017).

49. Surolia, M. P. et al. Hepatic transcriptome signatures in patients with varying degrees of nonalcoholic fatty liver disease compared with healthy normal-weight individuals. Am. J. Physiol. Gastrointest. Liver Physiol. 316, G462–G472 (2019).

50. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression proﬁles. Proc. Natl Acad. Sci. USA 102, 15545–15550 (2005).

51. Zunder, E. R. et al. Palladium-based mass tag cell barcoding with a double-filtering scheme and single-cell deconvolution algorithm. Nat. Protoc. 10, 316–333 (2015).

52. Finck, R. et al. Normalization of mass cytometry data with bead standards. Bioinformatics 33, 2157–2165 (2017).

53. Arciga-Carreras, I. et al. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classiﬁcation. Bioinformatics 33, 2424–2426 (2017).
of Opiate-induced Changes in Brain Adenosine Levels and Narcotic Drug Responses. *Neuroscience* 228, 235-242 (2013).
29. Park, W. et al. Metabolomic markers differentiate mucinous and non-mucinous pancreatic cysts. *Gastrointest. Endosc.* 78, 295-302 (2013).
30. Wu, M. et al. Increased dipeptide abundance in non-small cell lung cancer. *Rapid Commun. Mass Spectrom.* 27, 2091-2098 (2013).
31. Guo, K. & Li, L. Differential 13C-/13C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal. Chem.* 81, 3919-3932 (2009).

Acknowledgements

We thank Dr. Hyun Min Kang for his generous advice on the use of the demuxlet program for deconvoluting multiplexed scRNA-Seq data, and Dr. Bob Lewis for helpful advice and thoughtful review of this manuscript. Y.G. and G.P. were partially supported by awards (1R01DK102182-01A1 NIDDK, 5U01DA04439902 NIDA) made to G.P. A.E. and S.C.H. were partially supported by a Stanford Bio-X Interdisciplinary Initiatives Seed Grant and by awards (R01EB027171, R01HL142718) made to S.C.H., G.P., and S.C.H. were partially supported by awards (National Science Foundation CBET 203302) made to G.P. and S.C.H.

Author contributions

Y.G., A.E., S.C.H. and G.P. wrote the paper; Y.G., A.E., P.K.J., I.O., L.C., and K.K. performed experiments; Y.G., J.W., M.W., F.Z., A.E., S.-Y.C. and S.C.H. analyzed data; and P.R. obtained patient samples.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-26410-9.

Correspondence and requests for materials should be addressed to Gary Peltz.

Peer review information Nature Communications thanks Hans Clevers and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021