UC San Diego

UC San Diego Previously Published Works

Title

Temporal analyses of postnatal liver development and maturation by single-cell transcriptomics

Permalink

https://escholarship.org/uc/item/96b0d4d8

Journal

Developmental Cell, 57(3)

ISSN

1534-5807

Authors

Liang, Yan
Kaneko, Kota
Xin, Bing
et al.

Publication Date

2022-02-01

DOI

10.1016/j.devcel.2022.01.004

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Dev Cell. Author manuscript; available in PMC 2023 February 07.

Published in final edited form as:

Dev Cell. 2022 February 07; 57(3): 398–414.e5. doi:10.1016/j.devcel.2022.01.004.

Temporal Analyses of Postnatal Liver Development and Maturation by Single Cell Transcriptomics

Yan Liang[1], **Kota Kaneko**[1], **Bing Xin**[1], **Jin Lee**[1], **Xin Sun**[2], **Kun Zhang**[3], **Gen-Sheng Feng**[1,4,*]

[1]Department of Pathology, Division of Biological Sciences, and Moores Cancer Center, University of California at San Diego, La Jolla, CA 92093;

[2]Department of Pediatrics, University of California at San Diego, La Jolla, CA 92093, USA

[3]Department of Bioengineering, University of California at San Diego, La Jolla, CA 92093, USA

[4]Lead Contact

SUMMARY

The postnatal development and maturation of liver, the major metabolic organ, are inadequately understood. We have analyzed 52,834 single cell transcriptomes and identify 31 cell types or states in mouse livers at postnatal day 1, 3, 7, 21 and 56. We observe unexpectedly high levels of hepatocyte heterogeneity in the developing liver and progressive construction of the zonated metabolic functions from pericentral to periportal hepatocytes, which is orchestrated with development of sinusoid endothelial, stellate and Kupffer cells. Trajectory and gene regulatory analyses capture 36 transcription factors, including a circadian regulator *Bhlhe40*, in programming liver development. Remarkably, we identify a special group of macrophages enriched at day 7 with a hybrid phenotype of macrophages and endothelial cells, which may regulate sinusoidal construction and Treg cell function. This study provides a comprehensive atlas that covers all hepatic cell types instrumental for further dissection of liver development, metabolism and disease.

Graphical Abstract

[*]**Correspondence:** gfeng@health.ucsd.edu.

AUTHOR CONTRIBUTIONS

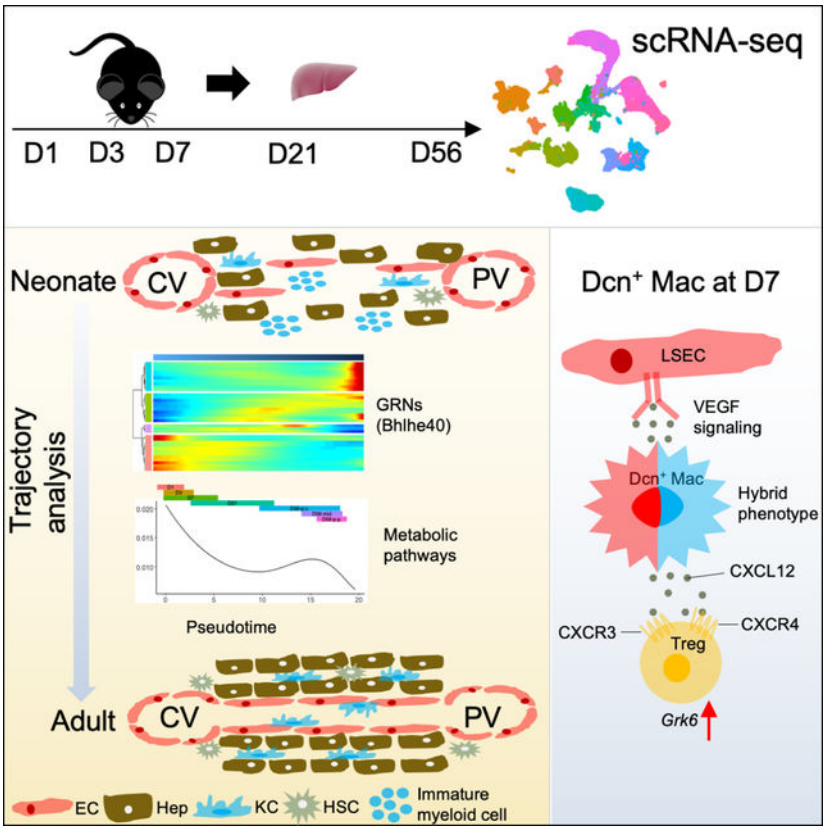
Conceptualization: Y.L., and G.S.F.; Methodology: Y.L., K.Z., and G.S.F.; Software: Y.L.; Investigation: Y.L., K.K., J.L., and B.X.; Data Analysis and Interpretation: Y.L., X.S., K.Z., and G.S.F; Writing – Original Draft: Y.L., and G.S.F.; Writing - Reviewing and Editing: Y.L., K.K., J.L., X.S., K.Z., and G.S.F.; Supervision and Funding Acquisition: G.S.F.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DECLARATION OF INTERESTS

The authors declare no competing interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



In Brief

Liang et al. present an atlas of mouse liver development at single cell resolution from newborn to adult. They demonstrate functional maturation of hepatocytes in concert with endothelial, stellate and Kupffer cells and transient emergence of a macrophage group.

INTRODUCTION

The liver executes vital functions, including metabolism, detoxification, bile secretion, and production of plasma proteins. Hepatocytes are the main parenchymal cells, accounting for 80% of the liver volume and carrying out most of the metabolic and synthetic functions (Miyajima et al., 2014). In adult liver, the basic architectural unit is the zonated liver lobule, in which portal vein (PV) and hepatic artery and bile duct form a portal triad in the corner of hexagonal lobules, with the central vein (CV) in the lobule center. Blood enters the lobules from PV and hepatic artery, and exchanges nutrients and oxygen with hepatocytes while flowing through sinusoids. The hepatocyte heterogeneity along with gradients of oxygen, nutrients and hormones forms metabolic zones between the CV and PV areas (Jungermann and Kietzmann, 1996). Hepatocytes locating in different layers along the lobule axis express different receptors, translocators, and enzymes.

Liver development is initiated around E8.5–E9.0 in mouse embryos by outgrowth of the primary liver bud from the ventral wall of the foregut (Zaret, 2002; Zhao and Duncan, 2005), driven by concerted activities of fibroblast growth factors (FGFs) from the adjacent

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cardiac mesoderm and bone morphogenetic proteins (BMPs) from septum transversum mesenchyme (STM). By E9.5, the parenchymal progenitor cells delaminate from the bud and invade the surrounding STM as cords of hepatoblasts that can differentiate into hepatocytes and cholangiocytes, the biliary epithelium. Subsequent liver organogenesis that generates the complex architecture involves extensive differentiation of parenchymal and non-parenchymal cell (NPC) types, development of the biliary tract, sinusoidal capillaries and vasculature, and organization of extracellular matrix. Remarkably, the fetal liver is the main site of hematopoiesis, harboring hematopoietic stem cells, from E10.5 to E16.5 (Crawford et al., 2010; Sasaki and Matsumura, 1986). During the late stage of embryogenesis, the liver transitions to a primary organ of metabolism, following migration of hematopoietic stem cells to bone marrow (Gordillo et al., 2015; Zaret, 2002).

Single cell RNA-sequencing (scRNA-seq) has been used to characterize the landscape of the mouse gut endoderm and gastrulation (Nowotschin et al., 2019; Pijuan-Sala et al., 2019), and to dissect differentiation of the bipotential hepatoblasts into hepatocyte and cholangiocyte lineages (Yang et al., 2017). A comprehensive scRNA-seq analysis characterized the onset of liver development from specification of endoderm progenitors to parenchymal and non-parenchymal cell lineage diversification and establishment in E7.5-E10.5 mouse embryos (Lotto et al., 2020). Another study interrogated single cell transcriptomes in endodermal and hepatic cells from E7.5 to E15.5 in Foxa2^{eGFP} mouse line (Mu et al., 2020).

After birth, hepatocytes organize into hepatic lobules, establishing the zonation structure. However, the driving factors of zonal construction and how hepatocytes adopt specific metabolic functions remain to be determined. One microarray analysis demonstrated four developmental stages from E11.5 to normal adult liver with distinct transcriptome profiles (Li et al., 2009). The dividing points of four stages are E14.5, E17.5 and Day 3. Another study using bulk RNA sequencing focused more on postnatal development from E17.5 to day 60 (Gunewardena et al., 2015). However, the bulk data analyses provided only the average values of different cell types, thus confounding the critical information on changes of cell type composition, cell-cell interaction, and functional heterogeneity. Combined use of scRNA-seq with single-molecule fluorescence *in situ* hybridization allowed reconstruction of liver zonation and inference of lobule coordinates in fasted adult mouse, with over half of the liver genes shown to be significantly zonated, albeit the analysis was focused on hepatocytes only (Halpern et al., 2017).

In the present study, we isolated all major hepatic cell types for comprehensive analysis of progressive liver development in mice after birth at single cell resolution. We revealed the process of hepatocytes and sinusoidal endothelial cell zonation establishment and identified candidate regulators in hepatocyte development and their putative roles in tumorigenesis. We performed computational predictions of the circadian rhythm development in the liver as well as response of several critical metabolic pathways in hepatocytes after birth, including glycolysis, fatty acid β-oxidation, and lipid and cholesterol biosynthesis. Remarkably, we identified a special group of macrophages, enriched on postnatal day 7, and inferred their role in regulating sinusoidal vascularization and regulatory T (Treg) cell activity via cell-cell

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

communication analysis. Herein we present a detailed single-cell atlas on postnatal liver development.

RESULTS

scRNA-seq identifies distinct hepatocyte subpopulations in the early postnatal liver

To investigate liver development at single cell resolution, we established a protocol for efficient isolation of all hepatic cell types, including hepatocytes and non-parenchymal cells (NPC), from mouse liver at high quality (see Methods). We collected cells at postnatal day 1, 3, 7, 21, and 56 (marked as D1, D3, D7, D21, and D56), which cover the whole period of liver development from newborn to adult. All isolated single cells were used for scRNA-seq on a 10x Genomics platform (Table S1, Figure S1A). Of a total of 65,891 cells subjected to RNA sequencing, 52,834 cells passed quality control based on the number of genes expressed, the count of raw reads and the mitochondrial gene percentage. Cells from the five time points were normalized and scaled for clustering and dimensional reduction using UMAP (Figures 1A, 1C). We identified a total of 31 cell types or states from all ages, based on expression of manually selected well-known markers (Figure 1B, Table S2). Within the 31 clusters, we identified three groups of hepatocytes (Figure 1A). Hep-neonatal consisted of hepatocytes collected at the early stages, including D1, D3, and D7, while Hep-D21 and Hep-D56 included hepatocytes at D21 and D56, respectively.

We interrogated 9137 assigned hepatocytes from D1 to D56, expressing the biomarkers *Alb* and *Hnf4a* (Figure 1B). These cells were clustered primarily by time point (Figure 1D), with hepatocytes from D1, D3 and D7 locating close, relative to the D21 and D56 cells. Several hepatoblast or hepatocellular carcinoma (HCC) markers, such as *Afp*, *Ahsg* and *H19* (Spear et al., 2006; Su et al., 2017), were highly expressed by hepatocytes at the early time points (Figure 1E), suggesting that they are not mature hepatocytes. Indeed, hepatocytes at D1-D7 were actively proliferating, with high expression of *Mki67*, encoding the proliferation marker Ki67 (Figure 1F), as well as other markers of proliferation (Whitfield et al., 2006), including *Mybl2*, *Top2a*, and *Ccnd1* (Figure S1B). Hepatocytes from D21 were at a transition status, with low expression of *Mki67*, *Afp*, *Ahsg* and *H19* (Figure 1E, 1F), and also low levels of *Hnf4a*, *Cyp1a2*, and *Cyp3a11* (Figure 1G), mature hepatocyte markers in adult liver (D56).

At each time point, we identified 2–3 subgroups (Figures 2A–2E), with a few marker genes shared at multiple time points. D1-Hep1, D3-Hep1, and D7-Hep1 highly expressed *Rpl10*, *Fabp5* and other ribosomal proteins, suggesting active protein translation. D1-Hep2, D3-Hep2, and D7-Hep2 highly expressed metabolism-related genes, *Cyp4a14* and *Cps1* (Figures 2A–2C, S1C). At later stages, D21-Hep2 and D56-Hep2 highly expressed *Neat1*, *Mlxipl* and *Malat1*; D21-Hep3 and D56-Hep3 expressed *Snca* and *Cd24a* (Figures 2D–2E, S2A). Interestingly, D1-Hep3 was featured by unique expression of *Scd2* (Figures 2F, S1C), encoding an isoform of stearoyl-CoA desaturase (SCD) a key enzyme for biosynthesis of monounsaturated fatty acids (Miyazaki et al., 2005; Miyazaki et al., 2003). *Scd1* expression was abundant at D21 and D56, while *Scd2* was mainly expressed at D1 (Figures 2G–2H, S1D–S1E). We validated the distinct *Scd1* and *Scd2* expression patterns in isolated hepatocytes by qRT-PCR (Figures 2I–2J, S1L). Similar age-related expression profiles of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Scd1 and *Scd2* were observed in HSC (Figures S1F–S1G) and Kupffer cells (Figures S1H–S1I), as validated by qRT-PCR (Figures S1J–S1K, S1L). *Scd2* expression peaked in Kupffer cells at D7 (Figure S1I), similarly detected by qRT-PCR (Figure S1K).

We identified two rare hepatocyte subgroups at D21 and D56 (Figure 2D–E). Cells in Hep3 highly expressed *Cd24a* (Figure S2A), a hepatoblast marker (Ochsner et al., 2007) and hepatocyte progenitors in adult liver (Qiu et al., 2011). *Cd24a* expression was scattered in all subpopulations at neonatal stages but was constrained in Hep3 at D21 (Figure S2B). Hep2 cells were featured by high expression of *Neat1*, *Malat1* and *Mlxipl* (Figure S2A). *Neat1* and *Malat1* represent the most abundant lncRNAs in nuclei, and *Mlxipl* is one of the few predominantly spliced polyadenylated protein-coding mRNAs showing nuclear retention (Bahar Halpern et al., 2015). *Malat1* has been frequently detected in scRNA-seq datasets, with its expression associated with mitochondrial RNA (mtRNA) proportions and cell death (Alvarez et al., 2020). We observed high correlation between *Malat1* expression and mtRNA proportions in Hep1 and Hep3 cells but not in Hep2 (Figure 2K), supporting that Hep2 is not a group of poor-quality cells. Next, we performed small molecule fluorescence in situ hybridization (smFISH) for *Neat1* and *Malat1* at D7 and D21 (Figure 2L). Pearson's correlation analysis showed highly correlated expression of *Malat1* and *Neat1*, identifying a group of hepatocytes at D21 featured by their co-expression.

Progressive establishment of the zonation pattern along the lobule axis

We explored distribution of hepatocytes in a lobule, by examining expression of four zonation markers, *Cyp2f2*, *Cyp2e1*, *Cdh1*, and *Glul* (Figure S2C) (Braeuning et al., 2006). As reported earlier (Doi et al., 2007; Halpern *et al.*, 2017), *Cyp2e1* was most abundant in the CV area and decreased gradually toward the PV region, with *Cyp2f2* having an opposite pattern. *Glul*, encoding glutamine synthetase (GS), was only expressed in pericentral hepatocytes, while *Cdh1*, encoding E-Cadherin, was expressed in periportal hepatocytes. The spatial location of hepatocytes at D56 was clearly visible based on expression of these four markers, which were inexplicit at earlier time points (Figure S2C). A previous study divided hepatic lobules into nine layers based on scRNAseq and smFISH data (Halpern *et al.*, 2017). We trained that dataset with assigned layers and inferred a spatial location of single cells from hepatocyte clusters (see Methods, Table S3). Layer 1 represents hepatocytes in the CV area, with layer 9 near PV. The predicted layer distribution at D56 (Figure 2M) was consistent with the prior probabilities of proposed adult liver zonation model (Halpern *et al.*, 2017), with the density peaking in layer 8. However, the predicted zonation profiles at earlier time points did not show similar patterns (Figure 2M), suggesting progressive zonal construction in postnatal liver. The heatmaps of 18 most significant zonation hallmark genes in hepatocytes demonstrate gradually distinguishable CV to PV transition from D1 to D56 (Figure S2E). *Glul* expression was high in pericentral hepatocytes since D1, albeit with gradual restriction to the CV area, while percentages of PV gene-positive hepatocytes increased steadily toward D56. Consistently, immunostaining of E-Cadherin and CYP2E1 showed reciprocally increasing expression in PV and CV hepatocytes, respectively (Figure 2N). A previous study reported a similar pattern of E-Cadherin and GS in hepatocytes during postnatal development (Gola et al., 2021). We mapped the predicted layers back to tSNE visualization at D21 and D56 (Figures 2O,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

S2D). Although immunostaining detected opposite expression patterns of E-Cadherin and CYP2E1 in the CV and PV areas at D21, relative to D3 and D7 (Figure 2N), transcriptomic profiling showed that most hepatocytes from different layers were still mixed (Figure S2D), suggesting that the metabolic zones are not fully developed yet in D21 hepatocytes. Consistent with the expression of zone markers (Figure S2C), zonated hepatocytes could only be distinguished at D56 (Figure 2O).

We established developing trajectory and calculated the pseudotime for each single cell (Figures 3A–3B), using Monocle (Qiu et al., 2017; Trapnell et al., 2014). The D56 hepatocytes were divided into three groups, D56-p.c., D56-mid and D56-p.p., based on the predicted layer, and randomly down-sampled, to be comparable to other time points. Since liver samples were collected from D1 to D56, the real time points could be used to evaluate accuracy of the pseudotime calculation (Figure 3C). At D56, the pericentral hepatocytes (D56-p.c.) showed up first in the trajectory followed by the mid-lobular (D56-mid), and then the periportal (D56-p.p.). Consistently, zonation prediction also showed steady increase of periportal hepatocytes toward D56 (Figure 2M), suggesting that the periportal hepatocytes are functionally matured later.

Identification of transcription factors in programming liver development

We ordered single hepatocytes from all time points by pseudotime and performed differential expression analysis, to identify genes whose expression changed significantly as a function of pseudotime. These selected genes were then grouped into six clusters by pseudotemporal expression patterns (Figure 3D, Table S4, Figure S10). For each cluster, we performed pathway enrichment analysis with simplified gene set collection from Molecular Signatures Databases (MSigDB) (Subramanian et al., 2005), with the significantly enriched pathways in each cluster listed in Table S4. The peripheral circadian clock-related pathway was identified in cluster 1, where genes were gradually increased from D1 to D56.

To search for transcription factors (TFs) as developmental regulators, we performed single-cell regulatory network inference and clustering (SCENIC) (Aibar et al., 2017), which allows construction of gene regulatory network (GRN) using the same subset of single cells for trajectory analysis. A TF activity matrix was generated with scores that measure expression levels of the TF's downstream targets (regulon) in a single cell. This approach captured 36 TFs that exhibited significant changes as a function of pseudotime in both expression and activity (Figures S3A–S3B), including *Bhlhe40* (Dec1), a core regulator of circadian rhythm (Honma et al., 2002; Sato et al., 2004). Its expression and activity increased dramatically at the late stage (Figure 3E), consistent with the pathway enrichment data (Figure 3D).

Among the 36 TFs, *Jun*, *Junb*, and *Fos* are known to play important roles in tumorigenesis, and we explored the other TFs in HCC. During post-natal development, *Klf9* showed steadily increasing expression and activity after birth, while an opposite pattern was observed for *Hmgb2*, with the highest level detected at the onset (Figure S3C). RNA-seq data analysis of 371 HCC patients deposited in TCGA demonstrated that patients with higher *Klf9* expression and activity had longer survival (Figure S3D, S3E), with the activity scores steadily decreased with tumor progression (Figure S3F). In contrast, we observed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

shorter survival for patients with high *Hmgb2* expression and activity, which also exhibited increasing trend with the disease progression (Figures S3D–S3E). Thus, patients with higher expression and activity of a TF active in the early postnatal stage showed poorer prognosis.

Development of various metabolic functions in hepatocytes

To explore development of metabolic functions in hepatocytes, we calculated the pathway enrichment scores of gene sets collected from Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa and Goto, 2000) and MSigDB (Subramanian *et al.*, 2005). We then ordered cells along pseudotime and filtered for gene sets with scores significantly changed along the trajectory. Further, we visualized gene sets of interest through fitting a natural spline (see Methods) and investigated detailed changes of genes in each gene set. Hepatocytes use glycolysis as the main energy source in the fetus due to low oxygen and under-developed mitochondria. The metabolic energy source shifts in the newborn, due to rapid increase of mitochondrial oxidative phosphorylation with air breathing (Böhme et al., 1983; Lindgren et al., 2019). We observed steadily decreasing expression of hypoxia-related pathway from D1 to D56 (Figure 3F). The hypoxia pathway was more prominently suppressed in the PV than CV area at D56, with an opposite pattern observed for the tricarboxylic acid (TCA) cycle (Figure 3G). Expression of most glycolysis-specific genes slowly increased when the glucose level restored and decreased subsequently, while gluconeogenesis-specific genes, including *Pcx* and *Fbp1*, were highly expressed right after birth (Figure S4A). This observation was consistent with sudden hypoglycemia at birth due to disruption of glucose supply from mother (Böhme *et al.*, 1983), which could be remedied by enhanced gluconeogenesis. Similarly, glycogenolysis-specific genes also peaked at early stages (Figure S4B), to replenish blood glucose, while glycogenesis showed up later and peaked in periportal hepatocytes at D56.

We observed a peak of fatty acid β-oxidation immediately after birth, followed by rapid decrease over time and slight increase in periportal hepatocytes at D56 (Figure 3H), producing large amounts of acetyl-CoA. A ketone body-related pathway showed a similar pattern as β-oxidation, with a peak right after birth (Figure 3I). The genes involved in mevalonate pathway for cholesterol biosynthesis were modestly expressed in newborns but peaked at D21 (Figure 3J), suggesting that the pathway is not well established in neonatal livers. Short fatty acid elongation in mitochondria might be important for hepatocytes in newborn liver, while ER dominated this process later during development (Figure 3K). Bile secretion peaked in the PV region at D56 (Figure S4C), and enzymes in urea production from ammonia, encoded by *Cps1*, *Otc*, *Ass1*, *Asl* and *Arg1*, increased after birth and peaked in the PV region at D56 (Figures S4D). However, metabolism of xenobiotics showed an opposing pattern (Figure S4E), with the pathway-related genes highly expressed in pericentral hepatocytes and decreased gradually in the mid-lobe and PV regions.

Development of liver endothelial and mesenchymal cells

Liver endothelial cells consist of macrovascular endothelial cells (MaVECs), lymphatic endothelial cells, and liver sinusoid endothelial cells (LSECs). Liver endothelial cells across all five time points were clustered into 11 subpopulations (Figure 4A), with non-LSECs sharing similar transcriptomes over the time. LSECs at D56 constituted two large and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

continuous groups, periportal and pericentral LSECs (EC1, EC2) (Figures 4B, S5C). In contrast, we observed higher heterogeneity in pre-mature LSECs (EC3-EC6), without clear zonation separation. A small group of EC9 cells that expressed proliferation and cell cycle-related genes were detected at all time points, albeit with a decreasing trend (Figure 4B). All EC9 cells were in G2/M or S phase as predicted by cell cycle marker genes (Figure S5D), suggesting that a pool of LSECs with proliferation potential is maintained in developing liver. MaVECs highly expressed a marker gene *Vwf* but were negative for LSEC markers, *Stab2* and *Lyve1* (Géraud et al., 2010) (Figure S5A); these cells were further divided into two subtypes based on their physical location. Central vein vascular endothelial cells (EC11, CVECs) expressed CV markers *Rspo3* and *Wnt2*, and portal vein vascular endothelial cells (EC7, PVECs) expressed PV markers *Dll4* and *Efnb2* (Figure S5B) (Halpern et al., 2018). These markers effectively distinguished zonation pattern for both LSECs and MaVECs. The lymphatic ECs (EC10) expressed *Thy1* (Jurisic et al., 2010) and also *Stab2* and *Lyve1* (Figure S5A).

Previous single-cell analysis of embryonic liver from E7.5 to E10.5 (Lotto *et al.*, 2020) showed a developmental trajectory from hemangioblasts to embryonic endothelial cells. By combining the embryonic with our neonatal endothelial cell data (Figures 4C–4D, S5E), we constructed a trajectory from E7.5 to D7 (Figure 4E). The hemangioblasts with highest differentiation potential were defined as the root cells. Starting from the root, hemangioblasts differentiated into hematopoietic or endothelial cells at branch point 16. These early endothelial cells further differentiated into LSECs or MaVECs at branch point 8. Around this branching point, proliferating LSECs emerged after birth (EC9), which exhibited similar transcriptomic profiles as early embryonic endothelial cells. Within the LSEC branch, neonatal cells developed from EC3, EC4 to EC6 and EC5. In the branch of embryonic MaVECs, EC11 was the neonatal CV vascular ECs. EC7 emerged later, consisting of neonatal PV vascular ECs, likely differentiated from EC3, the earliest subtype of neonatal LSECs. These analyses suggest that endothelial cells in PV and CV areas take separate developmental paths, and that the PV area is remodeled after birth, with concerted differentiation of hepatocytes and endothelial cells in the region.

Taking similar approaches used for hepatocytes, we constructed a developmental trajectory of LSECs and calculated pseudotime for each single cell (Figures 4F, S5F–S5G). This approach identified differentially expressed genes as a function of pseudotime, clustered by their pseudotemporal expression patterns (Figure S5H), and with the significantly enriched pathways listed in Table S5. Among genes that decreased during postnatal development, we observed enrichment of VEGFA-VEGFR signaling pathway. We also detected several glycolysis- and hypoxia-related pathways, showing that LSECs experienced similar metabolic changes as hepatocytes after birth. For upregulated genes, we observed a significant enrichment of interferon pathways. Several TFs, including *Bhlhe40*, *Tcf4*, *Hmgb1* and *Meis2*, exhibited significant changes in expression and activity along the LSEC developmental trajectory (Figure 4G). These TFs are possible regulators of sinusoidal buildup, of which *Meis2* was mainly expressed in endothelial cells (Figure S5I). The expression of *Hdac1*, encoding a histone deacetylase, was also changed remarkably in the process (Figure 4G). Interestingly, the circadian regulator *Bhlhe40* showed opposite expression patterns between LSECs and hepatocytes (Figure 3E).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Mesenchymal cells have been reported as the source of myofibroblasts after liver injury, especially HSCs (Iwaisako et al., 2014). We identified three main clusters of mesenchymal cells (Figure 1A), including HSCs, fibroblasts, and mesothelial cells (Meso). By segregating these cells, we identified more mesenchymal cell subgroups (Figure 4H), which were initially included within other mesenchymal cells. Two subgroups, Meso1 and Meso2 (Figure 4H), shared mesothelial markers *Gpm6a* (Li et al., 2013), *Upk3b* and *Wt1* (Figure S5J), albeit with distinct transcriptome profiles. *Alcam*, a marker of embryonic mesothelial cells in mouse liver (Li *et al.*, 2013), was expressed in Meso1, but not Meso2. Of note, Meso1 was mainly constituted by cells at the neonatal stage (Figures 4I–4J). A previous lineage tracing experiment showed that these mesothelial cells migrated inward from liver surface and differentiated into HSCs, fibroblasts, and vascular smooth muscle cells (VSMCs) during liver development (Asahina et al., 2011). Meso2 cells expressed *Igfbp4*, *Cd34* and *Clec3b*, and emerged later at D21 and D56, in replacement of Meso1. Additionally, we detected two fibroblast clusters, capsular fibroblast (CF), located beneath the liver surface, and portal fibroblast (PF), near the portal triad (Balog et al., 2020). A group of cells with proliferating potential was also identified (Figure S5K), which expressed HSC markers, *Reln* and *Lrat*, or portal fibroblast markers, *Cd34* and *Clec3b* (Dobie et al., 2019) (Figure S5J).

Dynamic changes of hematopoietic and immune cell populations

Our dataset from D1-D56 showed many clusters of developing hematopoietic cells, including hematopoietic progenitor cells (HPCs), granulocyte-monocyte progenitors (GMPs), B cells, neutrophils, and erythroid lineages (Figure 1A). Both HPCs and GMPs expressed *Cd34*, a hematopoietic precursor marker (Figure 1B). The development trajectories of these cell types were inferred by force-directed graph (FDG) analysis (Figure 5A), which corroborated the predicted relationships shown in UMAP visualization (Figure 1A). HPCs were connected to three main paths, developing towards GMPs, erythroid lineage, and B cells. GMPs then gave rise to monocytes and neutrophils. Based on markers used previously (Bjerregaard et al., 2003; Borregaard and Cowland, 1997), neutrophil lineages exhibited three major stages: a) immature neutrophils (iNP; or neutrophilic promyelocytes), expressing high levels of *Mpo* and *Elane* (encoding elastase); b) intermediate mature neutrophils (imNP; or neutrophilic myelocytes and metamyelocytes), expressing *Ltf* (encoding lactoferrin); and c) mature neutrophils (mNP) (Figure 1B). Of note, this neutrophil lineage was not detected in a recent scRNA-seq study of human fetal liver (Popescu et al., 2019). Similarly, B cell lineage was divided into several stages. First, the immunoglobulin heavy chain gene arrangement occurred in pro-B cells, with a surrogate light chain expressed. This process requires expression of *Rag1/2* for gene rearrangement and *VpreB* for surrogate light chain (Figure 1B). A light chain rearrangement occurred in small pre-B cells expressing *Rag1/2*. Gene rearrangement, and the expression of *Rag1/2*, stopped in a middle stage, called large pre-B cells. Based on reported markers (Elliott and Sinclair, 2012), we divided the erythroid lineage into three stages, which shared similar features of hematopoiesis in adult bone barrow. Hepatic hematopoiesis persisted after birth but receded rapidly, barely detectable after D7 (Figures 5B, S6A–S6D). Development of T cell and NK cell lineages was not observed in postnatal liver. It was reported previously (Rugh, 1990) that T cell development in liver occurred much earlier than B cells, with T

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cell progenitors migrating to thymus around E13. The percentages of T and NK cells as well as dendritic cells (DCs) gradually increased over time (Figures S6E–S6F), suggesting their hepatic recruitment from thymus or bone marrow after birth.

Because of their high levels of heterogeneity, the observed changes in percentages of T or NK cells from D1 to D56 (Figure S6E) might not depict the patterns for individual subtypes. Thus, we segregated these cells and analyzed them separately. Within the eleven T and NK cell subtypes identified (Figures 5C, S6G), Treg cells (Figure 5D) displayed a unique change over time, with a peak at D7 (Figures 5E, S6H). Consistently, a recent report (Li et al., 2020) showed hepatic accumulation of Treg cells between D7 and D14 after birth. We compared differentially expressed genes in Treg cells between D7, D21 and D56. Among the top-ranked genes were *Brd1* and *Gtf2h3* (Figure 5F), likely involved in cell proliferation (Drapkin et al., 1996; Mishima et al., 2011; Mishima et al., 2014). Treg cells also highly expressed *Grk6*, encoding a G-protein-coupled receptor kinase that phosphorylates CXCR4 upon CXCL12 binding (Busillo et al., 2010).

A unique subtype of macrophages emerges transiently around postnatal day 7

A Kupffer cell cluster was located close to dendritic cells (DCs) and monocytes on UMAP (Figure 1A). This cluster expressed Kupffer cell markers *Adgre1* (encoding F4/80) and *Csf1r* (Ginhoux and Guilliams, 2016) (Figures 1B, 6A–6B), with a peak at D7 (Figure 6D). When analyzing Kupffer cells separately (Figures 6C, S7A–S7B), we identified six subtypes, KC1-KC6, with distinct gene expression profiles (Figure S7C) and TF activities (Figure S7D). The KC6 group emerged at D1, with high expression of *Ccl9* and high TF activity of *Nfil3*. KC5 mainly emerged at D56, featured by *Ly6a* and *Cxcl13* expression and high TF activity of *Zbtb7a* and *Irf2*. Other subtypes (KC1-KC4) consisted of Kupffer cells from D1 to D21, with most gene expression patterns conserved over time. KC1 showed high expression of ribosomal proteins and *Myc* activity, and KC4 included actively proliferating cells.

We identified a subcluster of macrophages (*Dcn*$^{+}$ Mac), featured by high expression of decorin (Table S2), a member of small leucine-rich proteoglycan family (Iozzo, 1999). Besides the traditional Kupffer cell markers, *Adgre1* and *Csf1r* (Figures 1B, 6A, 6B, 6E), the *Dcn*$^{+}$ macrophages also expressed general endothelial cell markers *Pecam1*, *Eng*, *Kdr*, and LSEC markers *Lyve1* and *Clec4g*, albeit negative for a MaVEC marker *Vwf* (Figures 1B, 6E–6F). Almost all the *Dcn*$^{+}$ macrophages emerged at D7 (Figure 6D), which prompted a thorough interrogation of this unique subtype. We validated the existence of *Dcn*$^{+}$ Mac and its abundance at D7 by co-immunostaining for the LSEC marker LYVE-1, the endothelial cell-specific TF, ERG, and the macrophage marker F4/80 (Figure 6G). FACS analysis also detected this cell type and validated its peak level at D7 (Figure 6H). RNA velocity analysis showed a developmental trajectory of *Dcn*$^{+}$ Mac from Kupffer cells rather than endothelial cells (Figure 6I). We plotted top 10 genes significantly up-regulated in *Dcn*$^{+}$ Macrophages, via phage portraits (Figure S8A), which requires further validation by cell lineage tracing. The group of Kupffer cells at the transition point (Figure 6J; transitioning KC, labelled by asterisk) was identified at D1, D3 and D7 (Figure 6K), showing differentiation potential at the neonatal stage. Then, we mapped six Kupffer cell subpopulations in this UMAP

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Figure S8B). The transitioning KCs were derived primarily from KC2 and displayed high TF activity of *Pou2f2* and *Tcf7l2* (Figure S7D), which might drive *Dcn*$^+$ Mac transition from KC2. Several recent reports (Chakarov et al., 2019; Lim et al., 2018) addressed macrophages expressing one endothelial cell marker, LYVE1. In contrast, the *Dcn*$^+$ Mac detected in this study exhibited a combination of macrophage and endothelial cell gene expression profiles. To identify genes selectively expressed in transitioning KC, we performed differential gene expression analysis between the transitioning KC and other Kupffer cells located far from *Dcn*$^+$ Mac (Figure 6J). Most genes highly expressed in lung LYVE1-expressing macrophages, including *Fcna*, *Marco*, *Cd163*, *Cd209f*, *Timd4* and *Mrc1* (Chakarov *et al.*, 2019; Lim *et al.*, 2018), were significantly expressed in transitioning KC (Figure S8C).

By ligand-receptor analysis using CellPhoneDB (Efremova et al., 2020), we identified cell types actively interacting with *Dcn*$^+$ Mac, including HSCs, fibroblasts, mesothelial cells, cholangiocytes, VSMCs and all subtypes of endothelial cells, with much less interactions observed between these cell types and Kupffer cells (Figure 7A, Figure S9A). We then examined signaling crosstalk related to *Dcn*$^+$ Mac, as compared to Kupffer cells (Figure 7B). To explore possible influences of *Dcn*$^+$ Mac on other cell types, we focused on pairs with ligands expressed by *Dcn*$^+$ Mac and receptors expressed on other cell types. This analysis detected several VEGF related pairs between *Dcn*$^+$ Mac and LSEC, none of which was identified between Kupffer cell and LSEC. *Dcn*$^+$ Mac also secreted CXCL12 that binds CXCR3 and CXCR4 on Treg cells. Together with high expression of *Grk6* in Treg cells at D7 (Figure 5F), these data may be indicative of the significance of Treg accumulation at D7 (Figure 5E) and potential crosstalk between Treg and *Dcn*$^+$ Mac cells.

Intercellular communications between hepatocytes and NPCs

Using the same method described above, we also interrogated crosstalk between hepatocytes and NPCs, and the dynamic changes of intercellular signaling at the early postnatal stages. We performed ligand-receptor analysis for all five time points individually. To investigate the cell-cell interactions that regulate hepatocyte zonation, we filtered for significant pairs with the receptors expressed on hepatocytes and previously shown to be zonated. This analysis identified RSPO1/3-LGR5 interactions between LSECs and hepatocytes (Figure 7C). RSPO3-LGR5 was significantly expressed across all time points, while RSPO1-LGR5 was only identified at neonatal stage. The binding of RSPO to LGR5, present on pericentral hepatocytes, promotes Wnt signaling, which modulates zonated metabolic activities in pericentral area (Planas-Paz et al., 2016; Rocha et al., 2015; Torre et al., 2010). This signaling event was also identified between Kupffer cells and hepatocytes across all time points (Figure 7D). We also identified ASGR2-F8 interaction between hepatocytes and LSECs. ASGR2 is a marker of periportal hepatocytes (Halpern *et al.*, 2018). We then counted the interactions between hepatocytes and all other NPCs (Figure S9B). For almost all cell types closely interacting with hepatocytes, the most active crosstalk occurred at D7, as compared to other time points. Some interaction pairs, including PDGFRB-PDGFD between hepatocytes and LSECs (Figure 7C) and NOV-NOTCH1 between hepatocyte and Kupffer cells (Figure 7D), were only identified at D7. Therefore, this scRNA-seq analysis revealed a critical time point D7 in the postnatal liver, involving numerous molecular and cellular activities and communications.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

DISCUSSION

This study provides the detailed blueprint that describes stepwise changes at single cell resolution to illustrate how a neonatal liver develops into a major metabolic organ. Trajectory analysis revealed progressive and concerted development and functional maturation of hepatocytes and LSECs in metabolic zone construction. Complex interactions between hepatic cell types are apparently involved in coordinating developmental programs in the early stages. Remarkably, we identified a group of macrophages at D7, which exhibits a hybrid phenotype of macrophages and endothelial cells. Intercellular interaction analysis suggests a putative role of *Dcn*$^{+}$ Mac in regulation of sinusoidal vascularization and Treg cells.

Hepatocytes displayed high levels of heterogeneity in the early developmental stages after birth. In neonatal liver, hepatocytes highly expressed hepatoblast or HCC markers, including *Afp*, *Ahsg* and *H19*. We also identified a unique *Scd2*$^{+}$ hepatocyte subtype at D1, whose expression was gradually replaced by *Scd1*, suggesting a specific role of *Scd2* in neonatal liver metabolism. A rare *Cd24a*$^{+}$ hepatocyte group was identified at D21 and D56, representing a candidate for hepatocyte progenitors, in agreement with previous data (Qiu *et al.*, 2011). By combining these hepatocyte subpopulations, pseudotemporal analysis enabled us to explore the developmental trajectory based on the samples collected from multiple time points. The unbiased analysis of upstream regulators and enriched pathways predicted an establishment of liver circadian clock after birth regulated by factors, such as Bhlhe40 (Dec1), and also inferred dynamic changes of several critical metabolic pathways in adaptation to dramatic environmental changes in neonatal life. A previous study measured expression of 25 selected metabolism-related genes in liver lysates using real-time PCR (Nakagaki et al., 2018). We found most of these pathways changed in an increasing trend, although a few metabolic pathways peaked in newborns and decreased later. Due to a sudden hypoglycemia environment occurring at birth due to disruption of glucose supply from mother, neonates had two strategies to overcome this crisis, by enhancing gluconeogenesis and glycogenolysis. Another mechanism is the enhanced β-oxidation to generate acetyl-CoA for synthesis of ketone bodies as energy source. We also detected dynamic changes in fatty acid metabolism in ER and mitochondria, which requires further investigation.

The scRNA-seq data analysis suggests that the metabolic zonation profile is not fully developed before weaning at D21. Although immunostaining showed an E-cadherin^{+} layer of hepatocytes surrounding PV at D21 (Figure 2N), combined transcriptome profiling and zonation prediction demonstrated that the D21 hepatocytes had not shared different metabolic labors yet as in adult liver. Furthermore, the periportal area matured later than the pericentral area, as suggested by zonation prediction (Figure 2M) and pseudotemporal analysis (Figure 3C). Pseudotemporal analysis with integrated embryonic (Lotto *et al.*, 2020) and postnatal datasets revealed similar developmental pattern for endothelial cells (Figures 4C–4E). As reported earlier, development of the vascular system in portal triad is initiated soon after birth (Swartley et al., 2016). Consistent with previous data (Planas-Paz *et al.*, 2016; Rocha *et al.*, 2015), we identified RSPO1/3-LGR5 interactions between LSECs and hepatocytes, and between Kupffer cells and hepatocytes (Figures 7C–7D), suggesting an important role in controlling hepatocyte zonation. Other ligands expressed by LSECs

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

or Kupffer cells, including PDGF, HGF, PGF and BMP, might also regulate hepatocyte proliferation and zone remodeling at the early stage.

In previous experiments, we observed highly similar transcriptomic profiles between liver tumors and 1-month-old young liver tissues (Wang et al., 2019). The shared gene expression patterns between actively developing liver tissue and liver tumor will be instrumental for search of new biomarkers in HCC initiation and progression. Indeed, a rare hepatocyte subpopulation highly expressing *Scd2* was enriched only in the D1 liver. Elevated expression of *Scd2* was also detected in Myc-induced HCC (Chen et al., 2021), accompanied by significantly reduced *Scd1* expression, showing a similar pattern as D1 hepatocytes. Several upstream regulators identified in developmental trajectories showed correlation with HCC patients' survival. A pre-mature circadian clock program was observed in neonatal liver, and circadian deregulations drive spontaneous HCC development (Hanley et al., 2021; Kettner et al., 2016).

Starting around E13 to E14, hematopoietic cells migrate from embryonic liver to thymus, spleen and bone morrow (Crawford *et al.*, 2010). We did observe the complete processes of erythrocyte, neutrophil, and B cell lineage differentiation from progenitor cells in the D1, D3, and D7 datasets (Figure 5A). Consistent with previous bulk RNA-seq data of the liver (Gunewardena *et al.*, 2015; Li *et al.*, 2009), D7 is a critical time point, at which we identified highly enriched *Dcn*$^+$ Mac and Treg cells and observed active interactions between hepatocytes and other cell types. *Dcn*$^+$ Mac is a previously unrecognized cell subtype with a hybrid phenotype of macrophages and endothelial cells. The co-expression of LYVE-1, ERG and F4/80 in one single cell visualized under confocal microscope indicates that *Dcn*$^+$ Mac is not a doublet mistakenly detected by scRNA-seq or FACS. The *Dcn*$^+$ Mac cells were likely derived from Kupffer cells and acquired endothelial cell markers, and engaged in interaction with LSEC, HSC, fibroblast, mesothelial and Treg cells. Our data also suggest that CXCL12-GRK6-CXCR4 signaling between *Dcn*$^+$ Mac and Treg cells may promote Treg cell recruitment and activity.

While this work fills in a gap of knowledge on the postnatal stage of development between embryonic and adult livers at single cell resolution, several key results may provide fresh views on previously unappreciated high levels of hepatocyte heterogeneity in the neonatal and postnatal livers, the progressive construction of metabolic zones in hepatocytes from pericentral to periportal regions, and the concerted development of hepatocytes and neighboring endothelial cells, HSC, and Kupffer cells. Further dissection of the newly identified *Dcn*$^+$ Mac cell functions and their interactions with other hepatic cell types may contribute to better understanding of temporal development of the unique immune-tolerant environment in the liver.

Limitations of the Study

For scRNA-seq analysis, we isolated single cells from the neonatal livers from D1 to D7 without perfusion, as it was unnecessary and not feasible, while a two-step perfusion method was used for isolation of single cells from D21 and D56 livers. The two different isolation methods may affect proportions of hepatic and hematopoietic cells captured in the analysis. The conclusions inferred by bioinformatics methods, such as trajectory analysis, RNA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

velocity, and zonation prediction, require further validation by experimental and functional interrogations.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gen-Sheng Feng (gfeng@health.ucsd.edu).

Materials Availability—This study did not generate new unique reagents.

Data and Code Availability—The accession number for the raw data and processed data reported in this paper is GSE171993.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice—All animals used in this study were in C57BL/6J background. The animal protocol (S09108) was approved by the IACUC at the University of California San Diego.

METHOD DETAILS

Liver cell isolation for scRNA-seq—For neonatal samples (D1, D3 and D7), 4 male mice were included in each time point. Decapitation was used for neonate sacrifice because they are resistance to hypoxia at this age. First, fetal livers were removed from the body, cut into small pieces and rinsed using Ca^{2+}-free HBSS to reduce blood. Next, liver pellets were transferred into HBSS buffer with collagenase H and DNase I, further dissociated with gentleMACS program *m_liver_03* and then incubated at 37°C for 30 minutes. After incubation, the additional gentleMACS program *m_liver_04* was used for more complete dissociation. Liver was passed through a 100 μm cell strainer. Cells were then centrifuged at 50 g for 3 minutes, in order to separate hepatocytes (in pellets) from non-parenchymal cells, NPCs (in suspension). In non-parenchymal cell suspension, ACK buffer was added to lyse red blood cells. Hepatocytes and NPCs were further washed with PBS separately, resuspended in DMEM with 10% FBS, counted with hemocytometer and mixed by original ratios of cell numbers.

For D21 and D56 samples, 2 male mice were included for each time point. Mice were first sacrificed using CO_2. Liver was perfused using a 2-step method with Ca^{2+}-free HBSS buffer and then with collagenase H in HBSS buffer containing Ca^{2+}, gently minced and passed through a 100 μm cell strainer. Cells were centrifuged at 50 g for 3 min, in order to separate hepatocytes (in pallets) from NPCs (in suspension). To remove dead cells and debris, the pelleted hepatocytes were resuspended in 45% Percoll and centrifuged at 50 g for 10 min, without brake. Meanwhile, the ACK buffer and Dead Cell Removal Kit were used to lyse red blood cells and eliminate dead cells in non-parenchymal cells. Hepatocytes and NPCs were further washed with PBS separately, resuspended in DMEM with 10% FBS and counted with hemocytometer.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Liver single cell preparations were made within the same four-hour period of a day for collection of samples at all five time points.

Single cell library construction and sequencing—The isolated single cells were immediately loaded onto 10x Chromium Controller, and then partitioned into nanoliter-scale Gel Beads-In-Emulsion (GEMs). Cells in GEMs were lysed, and RNAs released from cells were immediately captured by barcoded beads in the same GEMs, followed by reverse transcription, amplification, fragmentation, adaptor ligation and index PCR. Libraries were constructed using Chromium Single Cell 3' Reagent Kits (V2 chemistry, 10x Genomics). Sequencing was performed on Illumina HiSeq 4000 at IGM Genomics Center, UCSD, with the following read length: Read 1, 26 bp, including 16 bp cell barcode and 12 bp unique molecular identifier (UMI); Read 2, 98 bp transcript insert; i7 sample index, 8 bp.

scRNA-seq data pre-processing, dimensionality reduction and clustering—Sequenced reads were aligned to mouse reference genome GRCm38 using CellRanger package (v3.0.2). All libraries were then aggregated for batch effect correction and sequencing depth normalization. An expression matrix including all cells from D1 to D56 livers was generated, with each row representing a gene and each column representing a cell, and then loaded into the R package, Seurat. Next, low quality cells and genes were filtered for downstream analysis. In brief, genes expressed in less than 3 cells were removed; cells failed to meet following criteria were removed: 1) the number of genes detected in each cell should be more than 200 but less than 6500; 2) the UMIs of mitochondrial genes should be less than 10% of total UMI. A total of 52834 cells and 24057 genes passed the filter. After filtering, the raw expression matrix was normalized by the total expression, multiplied by scale factor 10,000, and log transformed. Next, we regressed on total numbers of UMIs per cell as well as mitochondrial gene percentages. The z-scored residuals calculated by normalization and scaling were stored for downstream analysis. We then performed dimensionality reduction and unsupervised clustering using Seurat functions. In brief, PCA was performed first, and 75 PCs were loaded for UMAP and tSNE dimensionality reduction. To find clusters, the same PCs were imported into FindClusters, a SNN graph-based clustering algorithm, with resolution = 3.0. Next, the Wilcoxon Rank Sum test was performed to identify markers (FDR < 0.05) in each cluster. Cell types were then assigned based on these markers manually.

Further, to find subpopulations in hepatocyte, endothelial cells, mesenchymal cells, NK cells, T cells and Kupffer cells, cells from corresponding groups were segregated and re-analyzed for higher resolution.

For combined analysis of embryonic (Lotto *et al.*, 2020) and fetal (this work) endothelial cells, an additional dataset integration using Seurat was perform to remove batch effect.

Zonation profile prediction of hepatocytes—We predicted spatial layers (zonation) of hepatocytes using linear regression method. The training dataset was downloaded from previously published work (Halpern *et al.*, 2017), with processed expression matrix and zonation information for each cell. We normalized, scaled and integrated this dataset with our hepatocyte subpopulations to remove batch effect. Next, the Wilcoxon Rank Sum test

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

was performed to identify markers (FDR < 0.05) in each layer. We combined markers for 9 layers to make a zonation signature with 210 genes in total. With zonation signature gene set, the linear model was trained using original dataset (Halpern *et al.*, 2017) and then used to predict layer for single hepatocyte in this study. The predicted layer was then normalized from 1 to 9, representing spatial distribution of hepatocytes from pericentral to periportal area.

Trajectory analysis—We performed trajectory analysis using different methods listed below, depending on different situations and purposes.

RNA velocity analysis.: As described in original publication (La Manno *et al.*, 2018), this algorithm was designed based on a simple model for transcriptional dynamics, in which the RNA velocity was estimated from spliced and un-spliced mRNA, and then used to predict the future state of individual cells. We first generated loom files of spliced and un-spliced reads counts from 10x samples using velocyto *run10x* pipeline. We then visualized RNA velocity on specified tSNE or UMAP embedding. For gene of interest, we also plotted (1) spliced or (2) un-spliced counts on the same embedding, (3) phase portraits against steady-state (un-spliced levels on y-axis above steady-state represented gene of interest being induced, while un-spliced levels below steady-state represented being repressed), and (4) residual levels u, with positive (red) and negative (blue) values indicating induced and repressed, respectively, embedded onto the same dimensionality reduction.

Monocle.: This method was used to perform trajectory analysis and calculate pseudotime for a relatively small group of cells. In brief, we selected differentially expressed genes between different time points as ordering genes and applied DDRTree, which learned the structure of dataset manifold as developmental trajectory and ordered cells onto that manifold with a calculated pseudotime starting from a given root cell. Next, we selected genes significantly changed as a function of pseudotime along trajectory. In brief, for each gene, the expression level was fitted with natural spline (sm.ns) to describe the smooth change of this gene along trajectory. The smooth values were then employed to fit a negative binomial model via vglm() function from VGAM package. This was the full model. Next, to test for significance, a chi-square ratio test was performed to compare this full model against null model. The whole process was wrapped in *differentialGeneTest* function in monocle. We adapted this method here and also for analysis described below. Finally, we made heatmap with significantly changed genes, clustered them by pseudotemporal expression pattern and then performed pathway enrichment analysis with gene sets from Msigdb v7.0 (Subramanian *et al.*, 2005).

Monocle 3.: For a much larger dataset, we performed trajectory analysis for integrated embryonic and fetal endothelial cells using monocle 3 with UMAP visualization for a better performance.

Force-directed graph analysis.: This method was performed to visualize connections between immature hematopoietic cells and their progenitor cells, with pseudotime calculated (not shown). The code was adapted from published work on fetal liver hematopoiesis (Popescu *et al.*, 2019).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Gene regulatory network analysis—SCENIC (Single Cell Regulatory Network Inference and Clustering) was performed to identify upstream regulators important for hepatocyte, LSEC and Kupffer cell development individually. With segregated raw expression matrix for cell type of interest, we further filtered for genes expressed at least with a count of 3 in 1% of all cells and genes found in RcisTarget's mouse databases (mm9-500bp-upstream-7species.mc9nr.feather, mm9-tss-centered-10kb-7species.mc9nr.feather). Next, the filtered expression matrix was normalized as log2(exprMat + 1). SCENIC was then performed to identify and score regulons (TFs) based on the expression of their regulated target genes. With regulon scores, we adopted differential expression function from Monocle to test for TFs with significantly changed activities as a function of pseudotime. We fitted regulon score with gamma distribution instead and also performed chi-square ratio test for significance. A similar heatmap was plotted for TF activities using regulon scores. Also, for a TF of interest, we fitted natural spline to its scaled expression values (Expression) and regulon scores (Activity) and plotted the smooth change of said TF along pseudotime, which was stretched from 0 to 100.

Single cell pathway enrichment—We collected metabolism-related pathways from Kyoto Encyclopedia of Genes and Genomes (KEGG) and Msigdb. To find metabolic pathways significantly changed during postnatal development, we first calculated enrichment score for individual cell i of each gene set j. The score was defined as follows:

$$S_{ij} = n_j / \sum_{k=1}^{n_j} rank_{ik},$$

where n_j represented the length of gene set j; $rank_{ik}$ represented the rank of cell i ordered by expression levels of gene k (in gene set j) over all cells included in this analysis. Next, for each gene set j, we ordered cells along trajectory and tested if S_{ij} changed significantly as a function of pseudotime. The statistical test method was also adapted from monocle as described above. We fitted gene set enrichment scores with gamma distribution and performed chi-square ratio test for significance. Similarly, we plotted smooth scores after fitting the natural spline.

Cell-cell interaction analysis—We performed CellPhoneDB with our dataset to identify important ligand-receptor interactions. All cell types were included for the analysis at each time point. The genes encoding ligands and receptors were kept only if they were expressed by at least 10% of cells. The significance of an interaction was calculated by permutation test with iterations = 1000.

TCGA survival analysis—The TCGA dataset of HCC was downloaded from TCGA-LIHC project, including processed results of RNA-sequencing and clinical information for 371 patients. For each gene of interest, all samples were divided into three groups based on gene expression levels: Low (patients with expression levels lower than 33.3% percentile); High (patients with expression levels higher than 66.7% percentile); Mid (other patients). Survival analysis was performed using survival and survminer R package, including the Low and High groups divided based on gene of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Immunostaining and FISH—To prepare fresh frozen tissue sections, all liver tissues were collected at indicated time points and were immediately embedded Tissue-Tek O.C.T compound (Sakura) and stored at −80°C. Fresh frozen tissue sections were fixed by cold acetone overnight, followed by cold 4% PFA overnight. After primary and secondary antibody incubation, fresh frozen tissue sections were mounted with VECTASHIELD mounting medium with DAPI or Anti-Fade Fluorescence Mounting Medium (for confocal microscope). For single-cell resolution, we checked slides under Leica SP8 Confocal with Lightning Deconvolution at microscopy core, UCSD. Quantification on the fluorescence intensity across the zonation was performed using imageJ. HCR FISH (Molecular Instruments) was performed according to manufacturer's instruction. Mean fluorescence intensity for individual nucleus was analyzed with ImageJ, and Pearson's correlation analysis was performed with GraphPad Prism.

Flow cytometry analysis—Non-parenchymal cells were isolated from liver as described above (see Liver cell isolation for scRNA-seq). Cells were first stained for LIVE/DEAD fixable Aqua, anti-CD16/CD32 for Fc blocking, and then panel antibodies. Cells were fixed using 1% PFA in PBS overnight at 4°C, resuspended in PBS, and analyzed on BD LSRFortessa X-20 Cell Analyzer at UCSD Human Embryonic Stem Cell Core facility at Sanford Consortium for Regenerative Medicine. Data were then analyzed using FlowJo 10.6.2.

qRT-PCR—RNAs were extracted from different cell types at indicated time points using RNeasy kit (Qiagen, cat #73304) following manufacturer's instruction. Real-time PCR was conducted to check mRNA levels of genes at indicated time points. *gapdh* was used as a housekeeping control. The primers used in this study were listed in Key Resource Table.

Statistical analysis—The statistical analysis was performed using R or GraphPad Prism 9 (for FACS analysis only). The details of test and significance are specified in figure legends.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank our lab members for helpful discussion. Single cell RNA-seq was conducted at the IGM Genomics Core at UCSD funded by NIH SIG grant (#S10 OD026929). This work was supported by NIH grants (R01DK128320, R01CA236074 and R01CA239629) to G.S.F. J.L. was supported by a postdoc fellowship from ILCA and a pinnacle award from AASLD. Funding source for microscopy core: NINDS P30NS047107.

REFERENCES

Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods 14, 1083–1086. 10.1038/nmeth.4463. [PubMed: 28991892]

Alvarez M, Rahmani E, Jew B, Garske KM, Miao Z, Benhammou JN, Ye CJ, Pisegna JR, Pietiläinen KH, Halperin E, and Pajukanta P (2020). Enhancing droplet-based singlenucleus RNA-seq resolution using the semi-supervised machine learning classifier DIEM. Sci Rep 10, 11019. 10.1038/s41598-020-67513-5. [PubMed: 32620816]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Asahina K, Zhou B, Pu WT, and Tsukamoto H (2011). Septum transversum-derived mesothelium gives rise to hepatic stellate cells and perivascular mesenchymal cells in developing mouse liver. Hepatology 53, 983–995. 10.1002/hep.24119. [PubMed: 21294146]

Bahar Halpern K, Caspi I, Lemze D, Levy M, Landen S, Elinav E, Ulitsky I, and Itzkovitz S (2015). Nuclear Retention of mRNA in Mammalian Tissues. Cell Rep 13, 2653–2662. 10.1016/j.celrep.2015.11.036. [PubMed: 26711333]

Balog S, Li Y, Ogawa T, Miki T, Saito T, French SW, and Asahina K (2020). Development of Capsular Fibrosis Beneath the Liver Surface in Humans and Mice. Hepatology 71, 291–305. 10.1002/hep.30809. [PubMed: 31206736]

Bergen V, Lange M, Peidli S, Wolf FA, and Theis FJ (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol 38, 1408–1414. 10.1038/s41587-020-0591-3. [PubMed: 32747759]

Bjerregaard MD, Jurlander J, Klausen P, Borregaard N, and Cowland JB (2003). The in vivo profile of transcription factors during neutrophil differentiation in human bone marrow. Blood 101, 4322–4332. 10.1182/blood-2002-03-0835. [PubMed: 12560239]

Borregaard N, and Cowland JB (1997). Granules of the human neutrophilic polymorphonuclear leukocyte. Blood 89, 3503–3521. [PubMed: 9160655]

Braeuning A, Ittrich C, Köhle C, Hailfinger S, Bonin M, Buchmann A, and Schwarz M (2006). Differential gene expression in periportal and perivenous mouse hepatocytes. FEBS J 273, 5051–5061. 10.1111/j.1742-4658.2006.05503.x. [PubMed: 17054714]

Busillo JM, Armando S, Sengupta R, Meucci O, Bouvier M, and Benovic JL (2010). Site-specific phosphorylation of CXCR4 is dynamically regulated by multiple kinases and results in differential modulation of CXCR4 signaling. J Biol Chem 285, 7805–7817. 10.1074/jbc.M109.091173. [PubMed: 20048153]

Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411–420. 10.1038/nbt.4096. [PubMed: 29608179]

Böhme HJ, Sparmann G, and Hofmann E (1983). Biochemistry of liver development in the perinatal period. Experientia 39, 473–483. 10.1007/bf01965164. [PubMed: 6133774]

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature 566, 496–502. 10.1038/s41586-019-0969-x. [PubMed: 30787437]

Chakarov S, Lim HY, Tan L, Lim SY, See P, Lum J, Zhang XM, Foo S, Nakamizo S, Duan K, et al. (2019). Two distinct interstitial macrophage populations coexist across tissues in specific subtissular niches. Science 363. 10.1126/science.aau0964.

Chen WS, Liang Y, Zong M, Liu JJ, Kaneko K, Hanley KL, Zhang K, and Feng GS (2021). Single-cell transcriptomics reveals opposing roles of Shp2 in Myc-driven liver tumor cells and microenvironment. Cell Rep 37, 109974. 10.1016/j.celrep.2021.109974. [PubMed: 34758313]

Crawford LW, Foley JF, and Elmore SA (2010). Histology atlas of the developing mouse hepatobiliary system with emphasis on embryonic days 9.5–18.5. Toxicol Pathol 38, 872–906. 10.1177/0192623310374329. [PubMed: 20805319]

Dobie R, Wilson-Kanamori JR, Henderson BEP, Smith JR, Matchett KP, Portman JR, Wallenborg K, Picelli S, Zagorska A, Pendem SV, et al. (2019). Single-Cell Transcriptomics Uncovers Zonation of Function in the Mesenchyme during Liver Fibrosis. Cell Rep 29, 1832–1847.e1838. 10.1016/j.celrep.2019.10.024. [PubMed: 31722201]

Doi Y, Tamura S, Nammo T, Fukui K, Kiso S, and Nagafuchi A (2007). Development of complementary expression patterns of E- and N-cadherin in the mouse liver. Hepatol Res 37, 230–237. 10.1111/j.1872-034X.2007.00028.x. [PubMed: 17362306]

Drapkin R, Le Roy G, Cho H, Akoulitchev S, and Reinberg D (1996). Human cyclindependent kinase-activating kinase exists in three distinct complexes. Proc Natl Acad Sci U S A 93, 6488–6493. 10.1073/pnas.93.13.6488. [PubMed: 8692842]

Efremova M, Vento-Tormo M, Teichmann SA, and Vento-Tormo R (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc 15, 1484–1506. 10.1038/s41596-020-0292-x. [PubMed: 32103204]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Elliott S, and Sinclair AM (2012). The effect of erythropoietin on normal and neoplastic cells. Biologics 6, 163–189. 10.2147/BTT.S32281. [PubMed: 22848149]

Ginhoux F, and Guilliams M (2016). Tissue-Resident Macrophage Ontogeny and Homeostasis. Immunity 44, 439–449. 10.1016/j.immuni.2016.02.024. [PubMed: 26982352]

Gola A, Dorrington MG, Speranza E, Sala C, Shih RM, Radtke AJ, Wong HS, Baptista AP, Hernandez JM, Castellani G, et al. (2021). Commensal-driven immune zonation of the liver promotes host defence. Nature 589, 131–136. 10.1038/s41586-0202977-2. [PubMed: 33239787]

Gordillo M, Evans T, and Gouon-Evans V (2015). Orchestrating liver development. Development 142, 2094–2108. 10.1242/dev.114215. [PubMed: 26081571]

Gunewardena SS, Yoo B, Peng L, Lu H, Zhong X, Klaassen CD, and Cui JY (2015). Deciphering the Developmental Dynamics of the Mouse Liver Transcriptome. PLoS One 10, e0141220. 10.1371/journal.pone.0141220. [PubMed: 26496202]

Géraud C, Schledzewski K, Demory A, Klein D, Kaus M, Peyre F, Sticht C, Evdokimov K, Lu S, Schmieder A, and Goerdt S (2010). Liver sinusoidal endothelium: a microenvironment-dependent differentiation program in rat including the novel junctional protein liver endothelial differentiation-associated protein-1. Hepatology 52, 313–326. 10.1002/hep.23618. [PubMed: 20578158]

Halpern KB, Shenhav R, Massalha H, Toth B, Egozi A, Massasa EE, Medgalia C, David E, Giladi A, Moor AE, et al. (2018). Paired-cell sequencing enables spatial gene expression mapping of liver endothelial cells. Nat Biotechnol 36, 962–970. 10.1038/nbt.4231. [PubMed: 30222169]

Halpern KB, Shenhav R, Matcovitch-Natan O, Toth B, Lemze D, Golan M, Massasa EE, Baydatch S, Landen S, Moor AE, et al. (2017). Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature 542, 352–356. 10.1038/nature21065. [PubMed: 28166538]

Hanley KL, Liang Y, Wang G, Lin X, Yang M, Karin M, Fu W, and Feng GS (2021). Concurrent Disruption of Ras/MAPK and NF-kappaB Pathways Induces Circadian Deregulation and Hepatocarcinogenesis. Mol Cancer Res. 10.1158/1541-7786.MCR-21-0479.

Honma S, Kawamoto T, Takagi Y, Fujimoto K, Sato F, Noshiro M, Kato Y, and Honma K (2002). Dec1 and Dec2 are regulators of the mammalian molecular clock. Nature 419, 841844. 10.1038/nature01123.

Iozzo RV (1999). The biology of the small leucine-rich proteoglycans. Functional network of interactive proteins. J Biol Chem 274, 18843–18846. 10.1074/jbc.274.27.18843. [PubMed: 10383378]

Iwaisako K, Jiang C, Zhang M, Cong M, Moore-Morris TJ, Park TJ, Liu X, Xu J, Wang P, Paik YH, et al. (2014). Origin of myofibroblasts in the fibrotic liver in mice. Proc Natl Acad Sci U S A 111, E3297–3305. 10.1073/pnas.1400062111. [PubMed: 25074909]

Jungermann K, and Kietzmann T (1996). Zonation of parenchymal and nonparenchymal metabolism in liver. Annu Rev Nutr 16, 179–203. 10.1146/annurev.nu.16.070196.001143. [PubMed: 8839925]

Jurisic G, Iolyeva M, Proulx ST, Halin C, and Detmar M (2010). Thymus cell antigen 1 (Thy1, CD90) is expressed by lymphatic vessels and mediates cell adhesion to lymphatic endothelium. Exp Cell Res 316, 2982–2992. 10.1016/j.yexcr.2010.06.013. [PubMed: 20599951]

Kanehisa M, and Goto S (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28, 27–30. 10.1093/nar/28.1.27. [PubMed: 10592173]

Kettner NM, Voicu H, Finegold MJ, Coarfa C, Sreekumar A, Putluri N, Katchy CA, Lee C, Moore DD, and Fu L (2016). Circadian Homeostasis of Liver Metabolism Suppresses Hepatocarcinogenesis. Cancer Cell 30, 909–924. 10.1016/j.ccell.2016.10.007. [PubMed: 27889186]

La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, et al. (2018). RNA velocity of single cells. Nature 560, 494–498. 10.1038/s41586-018-0414-6. [PubMed: 30089906]

Li M, Zhao W, Wang Y, Jin L, Jin G, Sun X, Wang W, Wang K, Xu X, Hao J, et al. (2020). A wave of Foxp3. Cell Mol Immunol 17, 507–518. 10.1038/s41423-019-0246-9. [PubMed: 31171863]

Li T, Huang J, Jiang Y, Zeng Y, He F, Zhang MQ, Han Z, and Zhang X (2009). Multistage analysis of gene expression and transcription regulation in C57/B6 mouse liver development. Genomics 93, 235–242. 10.1016/j.ygeno.2008.10.006. [PubMed: 19015022]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Li Y, Wang J, and Asahina K (2013). Mesothelial cells give rise to hepatic stellate cells and myofibroblasts via mesothelial-mesenchymal transition in liver injury. Proc Natl Acad Sci U S A 110, 2324–2329. 10.1073/pnas.1214136110. [PubMed: 23345421]

Lim HY, Lim SY, Tan CK, Thiam CH, Goh CC, Carbajo D, Chew SHS, See P, Chakarov S, Wang XN, et al. (2018). Hyaluronan Receptor LYVE-1-Expressing Macrophages Maintain Arterial Tone through Hyaluronan-Mediated Regulation of Smooth Muscle Cell Collagen. Immunity 49, 1191. 10.1016/j.immuni.2018.12.009. [PubMed: 30566884]

Lindgren IM, Drake RR, Chattergoon NN, and Thornburg KL (2019). Down-regulation of MEIS1 promotes the maturation of oxidative phosphorylation in perinatal cardiomyocytes. FASEB J 33, 7417–7426. 10.1096/fj.201801330RR. [PubMed: 30884246]

Lotto J, Drissler S, Cullum R, Wei W, Setty M, Bell EM, Boutet SC, Nowotschin S, Kuo YY, Garg V, et al. (2020). Single-Cell Transcriptomics Reveals Early Emergence of Liver Parenchymal and Non-parenchymal Cell Lineages. Cell 183, 702–716.e714. 10.1016/j.cell.2020.09.012. [PubMed: 33125890]

Mishima Y, Miyagi S, Saraya A, Negishi M, Endoh M, Endo TA, Toyoda T, Shinga J, Katsumoto T, Chiba T, et al. (2011). The Hbo1-Brd1/Brpf2 complex is responsible for global acetylation of H3K14 and required for fetal liver erythropoiesis. Blood 118, 2443–2453. 10.1182/blood-2011-01-331892. [PubMed: 21753189]

Mishima Y, Wang C, Miyagi S, Saraya A, Hosokawa H, Mochizuki-Kashio M, Nakajima-Takagi Y, Koide S, Negishi M, Sashida G, et al. (2014). Histone acetylation mediated by Brd1 is crucial for Cd8 gene activation during early thymocyte development. Nat Commun 5, 5872. 10.1038/ncomms6872. [PubMed: 25519988]

Miyajima A, Tanaka M, and Itoh T (2014). Stem/progenitor cells in liver development, homeostasis, regeneration, and reprogramming. Cell Stem Cell 14, 561–574. 10.1016/j.stem.2014.04.010. [PubMed: 24792114]

Miyazaki M, Dobrzyn A, Elias PM, and Ntambi JM (2005). Stearoyl-CoA desaturase-2 gene expression is required for lipid synthesis during early skin and liver development. Proc Natl Acad Sci U S A 102, 12501–12506. 10.1073/pnas.0503132102. [PubMed: 16118274]

Miyazaki M, Jacobson MJ, Man WC, Cohen P, Asilmaz E, Friedman JM, and Ntambi JM (2003). Identification and characterization of murine SCD4, a novel heart-specific stearoyl-CoA desaturase isoform regulated by leptin and dietary factors. J Biol Chem 278, 33904–33911. 10.1074/jbc.M304724200. [PubMed: 12815040]

Mu T, Xu L, Zhong Y, Liu X, Zhao Z, Huang C, Lan X, Lufei C, Zhou Y, Su Y, et al. (2020). Embryonic liver developmental trajectory revealed by single-cell RNA sequencing in the Foxa2(eGFP) mouse. Commun Biol 3, 642. 10.1038/s42003-020-01364-8. [PubMed: 33144666]

Nakagaki BN, Mafra K, de Carvalho É, Lopes ME, Carvalho-Gontijo R, de CastroOliveira HM, Campolina-Silva GH, de Miranda CDM, Antunes MM, Silva ACC, et al. (2018). Immune and metabolic shifts during neonatal development reprogram liver identity and function. J Hepatol 69, 1294–1307. 10.1016/j.jhep.2018.08.018. [PubMed: 30171870]

Nowotschin S, Setty M, Kuo YY, Liu V, Garg V, Sharma R, Simon CS, Saiz N, Gardner R, Boutet SC, et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. Nature 569, 361–367. 10.1038/s41586-019-1127-1. [PubMed: 30959515]

Ochsner SA, Strick-Marchand H, Qiu Q, Venable S, Dean A, Wilde M, Weiss MC, and Darlington GJ (2007). Transcriptional profiling of bipotential embryonic liver cells to identify liver progenitor cell surface markers. Stem Cells 25, 2476–2487. 10.1634/stemcells.2007-0101. [PubMed: 17641245]

Pijuan-Sala B, Griffiths JA, Guibentif C, Hiscock TW, Jawaid W, Calero-Nieto FJ, Mulas C, Ibarra-Soria X, Tyser RCV, Ho DLL, et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature 566, 490–495. 10.1038/s41586-0190933-9. [PubMed: 30787436]

Planas-Paz L, Orsini V, Boulter L, Calabrese D, Pikiolek M, Nigsch F, Xie Y, Roma G, Donovan A, Marti P, et al. (2016). The RSPO-LGR4/5-ZNRF3/RNF43 module controls liver zonation and size. Nat Cell Biol 18, 467–479. 10.1038/ncb3337. [PubMed: 27088858]

Popescu DM, Botting RA, Stephenson E, Green K, Webb S, Jardine L, Calderbank EF, Polanski K, Goh I, Efremova M, et al. (2019). Decoding human fetal liver haematopoiesis. Nature 574, 365–371. 10.1038/s41586-019-1652-y. [PubMed: 31597962]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Qiu Q, Hernandez JC, Dean AM, Rao PH, and Darlington GJ (2011). CD24-positive cells from normal adult mouse liver are hepatocyte progenitor cells. Stem Cells Dev 20, 21772188. 10.1089/scd.2010.0352.

Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, and Trapnell C (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 14, 979–982. 10.1038/nmeth.4402. [PubMed: 28825705]

Rocha AS, Vidal V, Mertz M, Kendall TJ, Charlet A, Okamoto H, and Schedl A (2015). The Angiocrine Factor Rspondin3 Is a Key Determinant of Liver Zonation. Cell Rep 13, 17571764. 10.1016/j.celrep.2015.10.049.

Rugh R (1990). The mouse : its reproduction and development (Oxford University Press).

Sasaki K, and Matsumura G (1986). Haemopoietic cells of yolk sac and liver in the mouse embryo: a light and electron microscopical study. J Anat 148, 87–97. [PubMed: 3693096]

Sato F, Kawamoto T, Fujimoto K, Noshiro M, Honda KK, Honma S, Honma K, and Kato Y (2004). Functional analysis of the basic helix-loop-helix transcription factor DEC1 in circadian regulation. Interaction with BMAL1. Eur J Biochem 271, 4409–4419. 10.1111/j.1432-1033.2004.04379.x. [PubMed: 15560782]

Spear BT, Jin L, Ramasamy S, and Dobierzewska A (2006). Transcriptional control in the mammalian liver: liver development, perinatal repression, and zonal gene regulation. Cell Mol Life Sci 63, 2922–2938. 10.1007/s00018-006-6258-5. [PubMed: 17041810]

Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888–1902.e1821. 10.1016/j.cell.2019.05.031. [PubMed: 31178118]

Su X, Shi Y, Zou X, Lu ZN, Xie G, Yang JYH, Wu CC, Cui XF, He KY, Luo Q, et al. (2017). Single-cell RNA-Seq analysis reveals dynamic trajectories during mouse liver development. BMC Genomics 18, 946. 10.1186/s12864-017-4342-x. [PubMed: 29202695]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545–15550. 10.1073/pnas.0506580102. [PubMed: 16199517]

Swartley OM, Foley JF, Livingston DP, Cullen JM, and Elmore SA (2016). Histology Atlas of the Developing Mouse Hepatobiliary Hemolymphatic Vascular System with Emphasis on Embryonic Days 11.5–18.5 and Early Postnatal Development. Toxicol Pathol 44, 705–725. 10.1177/0192623316630836. [PubMed: 26961180]

Torre C, Perret C, and Colnot S (2010). Molecular determinants of liver zonation. Prog Mol Biol Transl Sci 97, 127–150. 10.1016/B978-0-12-385233-5.00005-2. [PubMed: 21074732]

Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, and Rinn JL (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32, 381–386. 10.1038/nbt.2859. [PubMed: 24658644]

Wang G, Luo X, Liang Y, Kaneko K, Li H, Fu XD, and Feng GS (2019). A tumorigenic index for quantitative analysis of liver cancer initiation and progression. Proc Natl Acad Sci U S A. 10.1073/pnas.1911193116.

Whitfield ML, George LK, Grant GD, and Perou CM (2006). Common markers of proliferation. Nat Rev Cancer 6, 99–106. 10.1038/nrc1802. [PubMed: 16491069]

Yang L, Wang WH, Qiu WL, Guo Z, Bi E, and Xu CR (2017). A single-cell transcriptomic analysis reveals precise pathways and regulatory mechanisms underlying hepatoblast differentiation. Hepatology 66, 1387–1401. 10.1002/hep.29353. [PubMed: 28681484]

Zaret KS (2002). Regulatory phases of early liver development: paradigms of organogenesis. Nat Rev Genet 3, 499–512. 10.1038/nrg837. [PubMed: 12094228]

Zhao R, and Duncan SA (2005). Embryonic development of the liver. Hepatology 41, 956–967. 10.1002/hep.20691. [PubMed: 15841465]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- Single cell transcriptomics of all hepatic cell types in neonatal and adult livers
- Concerted development of zonated metabolic functions in hepatocytes and NPCs
- Transient emergence of a distinct group of macrophages at postnatal day 7
- Hepatic cell-cell communications in programming postnatal liver development

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

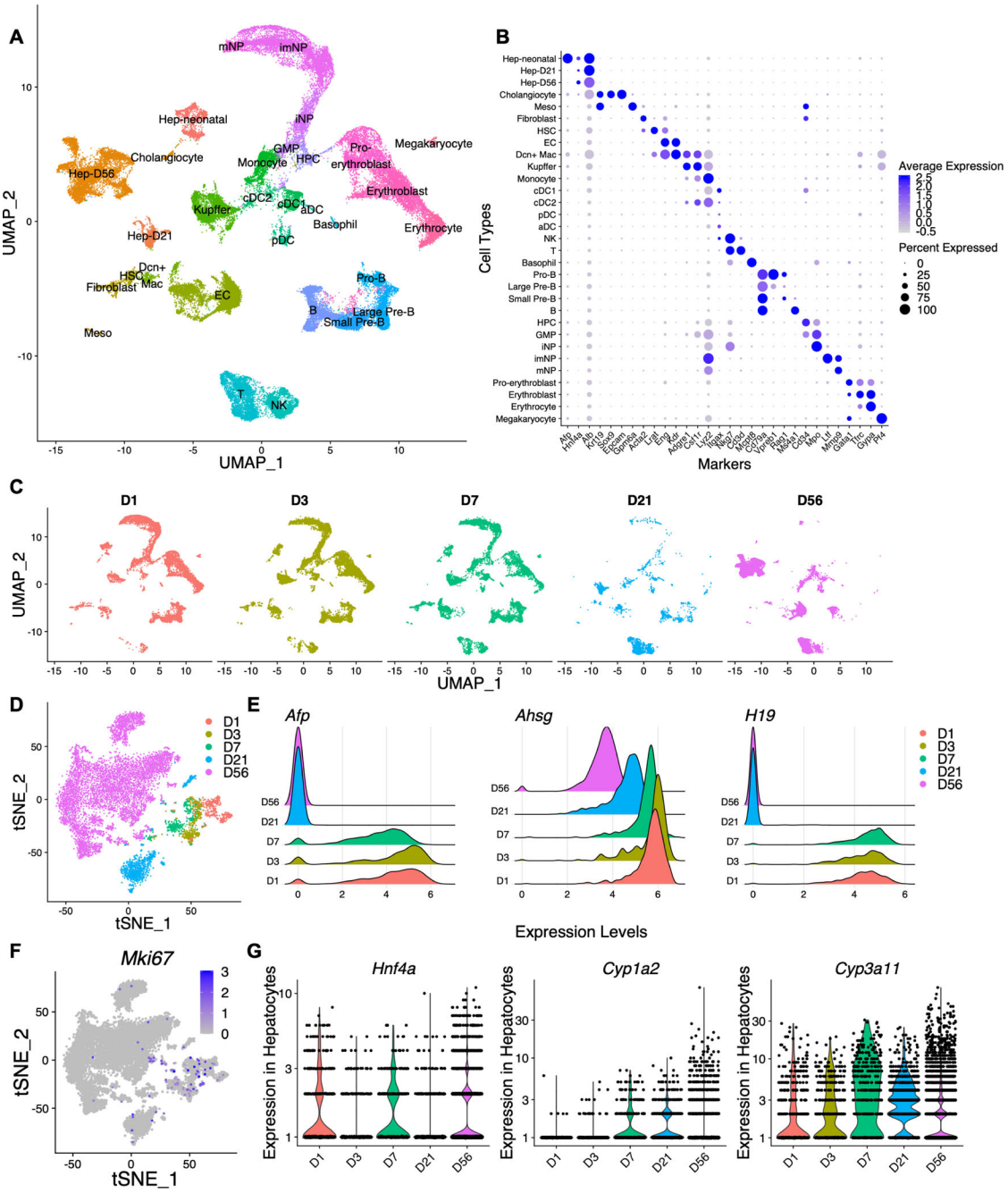


Figure 1. scRNA-seq identifies hepatic cell types in developing and adult liver

(A). UMAP visualization of liver cell types at D1, D3, D7, D21 and D56. Colors indicate cell types, including hepatocyte (Hep-neonatal from D1, D3 and D7; Hep-D21; Hep-D56), endothelial cell (EC), hepatic stellate cell (HSC), cholangiocyte, fibroblast, mesothelial cell (meso), megakaryocyte, erythroid cell (pro-erythroblast, erythroblast, erythrocyte), T cell, natural killer (NK) cell, B cell (Pro-B, large pre-B, small pre-B, B), dendritic cell (classical dendritic cell 1 – cDC1, classical dendritic cell 2 - cDC2, plasmacytoid dendritic cell - pDC, activating dendritic cell - aDC), monocyte, *Dcn*$^{+}$ macrophage (*Dcn*$^{+}$ Mac), Kupffer cell, neutrophils (immature neutrophil – iNP, intermediate mature neutrophil – imNP, mature

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

neutrophil - mNP), basophil, granulocyte-monocyte progenitor (GMP) and hematopoietic progenitor cell (HPC).

(B). Expression of selected markers for cell types. The dot size corresponds to the ratio of cells expressing the gene in the cell type. The color scales correspond to the averaged expression levels. See also Table S2.

(C). Temporal UMAP visualization of hepatic cells from Figure 1A. The UMAP was separated and colored by time points (n = 2~4 for each time point).

(D). tSNE map of hepatocytes from 5 time points (indicated by 5 colors). Cells in Hep-neonatal, Hep-D21 and Hep-D56 from Fig.1A were segregated and re-analyzed.

(E). Ridge plots displaying the expression distributions of indicated markers in hepatocytes from Fig.1D at five time points. X-axis indicates log-normalized expression levels.

(F). tSNE map displaying *Mki67* expression in hepatocytes from Fig.1D.

(G). Violin plots displaying expression levels of indicated markers in hepatocytes from Fig. 1D. Y-axis indicates log-normalized expression levels.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

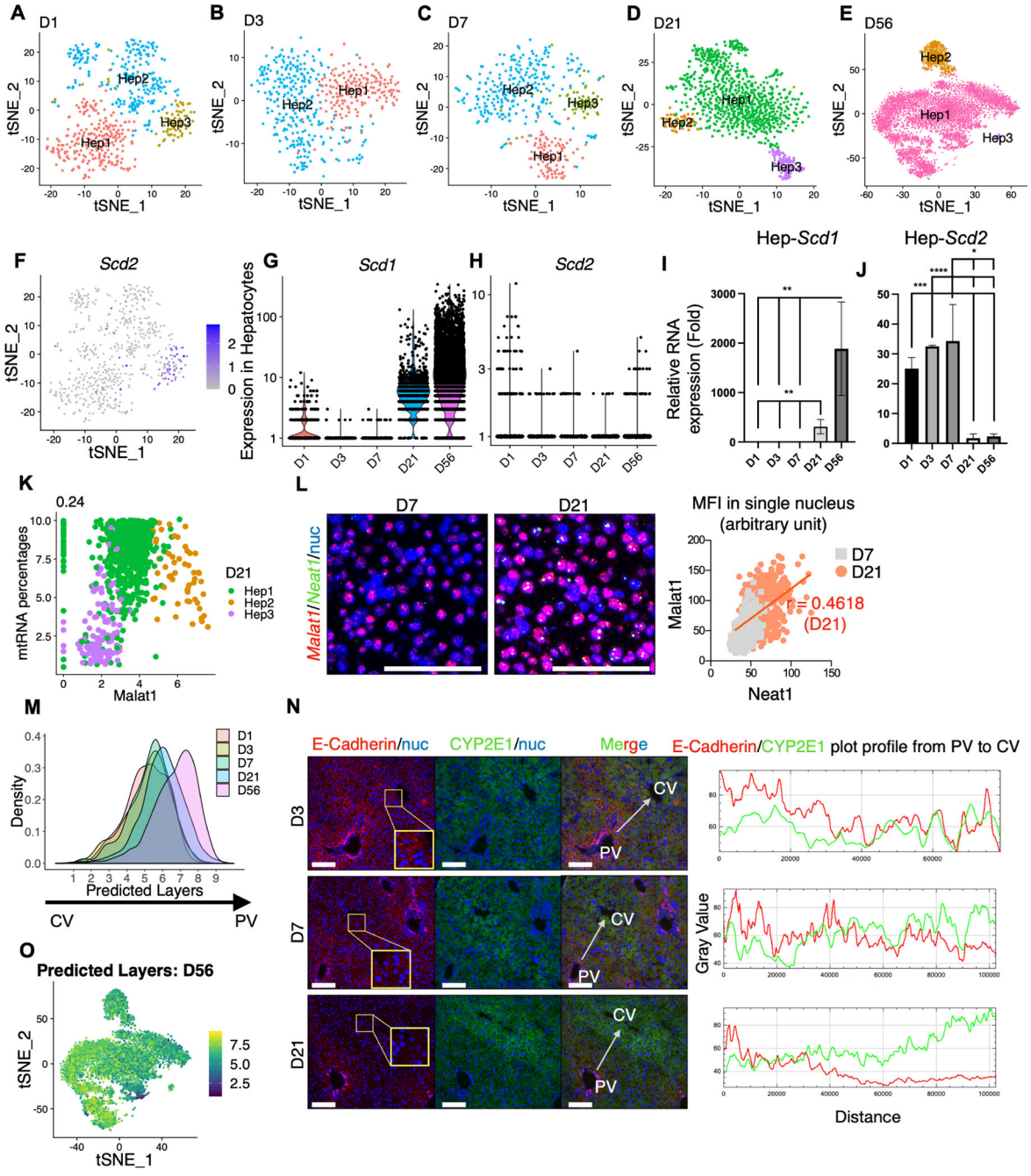


Figure 2. scRNA-seq identifies distinct transcriptome profiles in hepatocytes at each time point

(A-E). tSNE map of hepatocytes from D1 (A), D3 (B), D7 (C), D21 (D) and D56 (E). Colors indicate subpopulations identified.

(F). tSNE map displaying *Scd2* expression in D1 hepatocytes from Fig.2A.

(G-H). Violin plots displaying *Scd1* (G) and *Scd2* (H) expression in hepatocytes. Y-axis indicates log-normalized expression.

(I-J). qRT-PCR analysis of *Scd1* (I) and *Scd2* (J) expression in isolated hepatocytes. Statistical analysis used student's T-test. Values are presented as means ± SD. (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$). See also Figure S1L for cell type marker expression.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(K). Scatter plot displaying correlation between *Malat1* expression (x-axis) and mtRNA percentage (y-axis). Pearson correlation coefficient (0.24) is shown above the plot. *Malat1* expression is log normalized value. mtRNA percentage is the proportion of transcripts mapped to mitochondrial genes. Colors indicate D21 subpopulations in Fig.2D.
(L). (Left) smFISH of *Malat1* and *Neat1* at D7 and D21. (Right) The quantitative results showing fluorescence signal intensity (MFI) and correlation coefficient between *Malat1* and *Neat1*. Scale bar, 100 μm.
(M). Density plot displaying distribution of predicted layers of hepatocytes. The predicted layer for each hepatocyte was normalized from layer 1 (pericentral) to layer 9 (periportal). See also Table S3 for gene signatures used for zonation prediction.
(N). (Left) Immunostaining of E-cadherin and CYP2E1 in D3, D7 and D21 livers. Scale bar, 100 μm. The insets show loss of E-Cad around CV over time, establishing a clear zonation. (Right) Quantitative results showing the changing expression levels from PV to CV as the arrow indicates.
(O). tSNE map displaying predicted layers for D56 hepatocytes from 2E. Color bar indicates predicted layers from Layer 1 (CV area) to Layer 9 (PV area).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

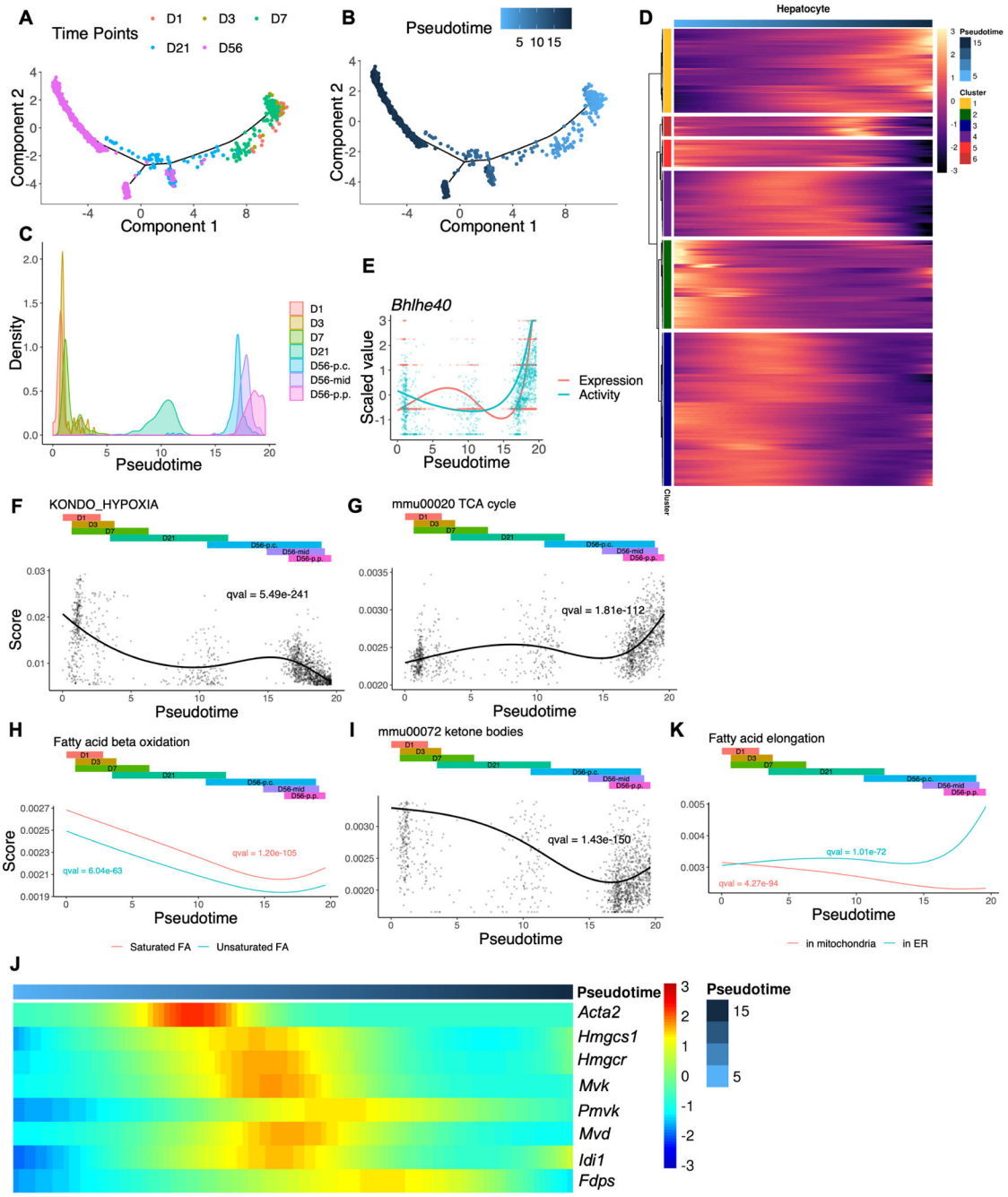


Figure 3. Dynamic changes of transcription factor activities and metabolic functions in hepatocytes

(A). Pseudotime analysis of hepatocyte development from D1 to D56 with Monocle 2.

(B). Pseudotime analysis of data in 3A. Color indicates inferred pseudotime used for 3C-3K.

(C). Density plot displaying distribution of inferred pseudotime (x-axis).

(D). Heatmap representing trends of differentially expressed genes as a function of inferred pseudotime. Genes in row are grouped into 6 clusters based on expression patterns. See also Figure S10 and Table S4 for the gene list and enriched pathways.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(E). Scatter and fitted plots of *Bhlhe40* scaled expression and activity values along pseudotime. The inferred pseudotime was stretched from 0 to 100.
(F-I). Scatter plot and fitted plots of enrichment scores for indicated pathways along inferred pseudotime. Only fitted plots are shown for 3H and 3K.
(J). Heatmap displaying genes from the mevalonate pathway with differential expression along inferred pseudotime.
(K). Fitted plots of enrichment scores for indicated pathway along inferred pseudotime.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

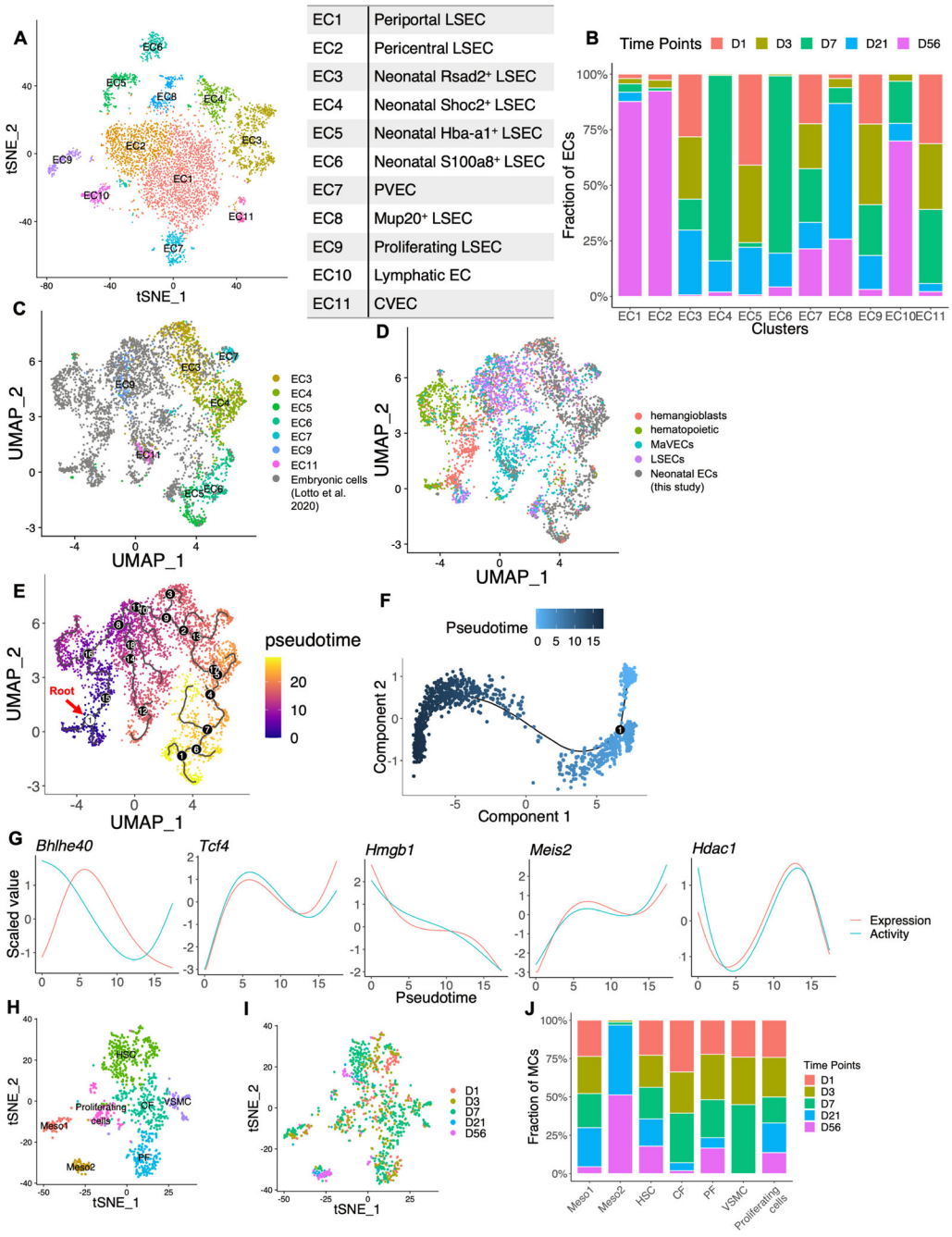


Figure 4. Development of liver endothelial and mesenchymal cells

(A). tSNE map of endothelial cells from D1 to D56. Cells in EC from figure 1A were segregated and re-analyzed.

(B). Time point compositions of each EC subpopulation labelled in 4A.

(C-D). UMAP visualization of cells from D1, D3, D7 and embryonic cells (including hemangioblasts, hematopoietic cells, HSEC and endothelium) from published data (Lotto *et al.*, 2020). Colors indicate endothelial cell subpopulation labelled in figure 4A (C), or embryonic cell types (D).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(E). Pseudotime analysis of cells from 4C with Monocle 3. Root cells are labelled in white circle 1; branch points are labelled in black circles with numbers.
(F). Pseudotime analysis of ECs from D1 to D56 with Monocle 2. Color indicates inferred pseudotime.
(G). Fitted plot of scaled expression and activity values along pseudotime of indicated genes. The inferred pseudotime from 4F was stretched from 0 to 100.
(H). tSNE map of mesenchymal cells from D1 to D56. Cells in meso, HSC and fibroblast from figure 1A were segregated and re-analyzed. Colors indicate cell types assigned.
(I). tSNE map of mesenchymal cells from 4H.
(J). Time point compositions of each cell type labelled in 4H.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

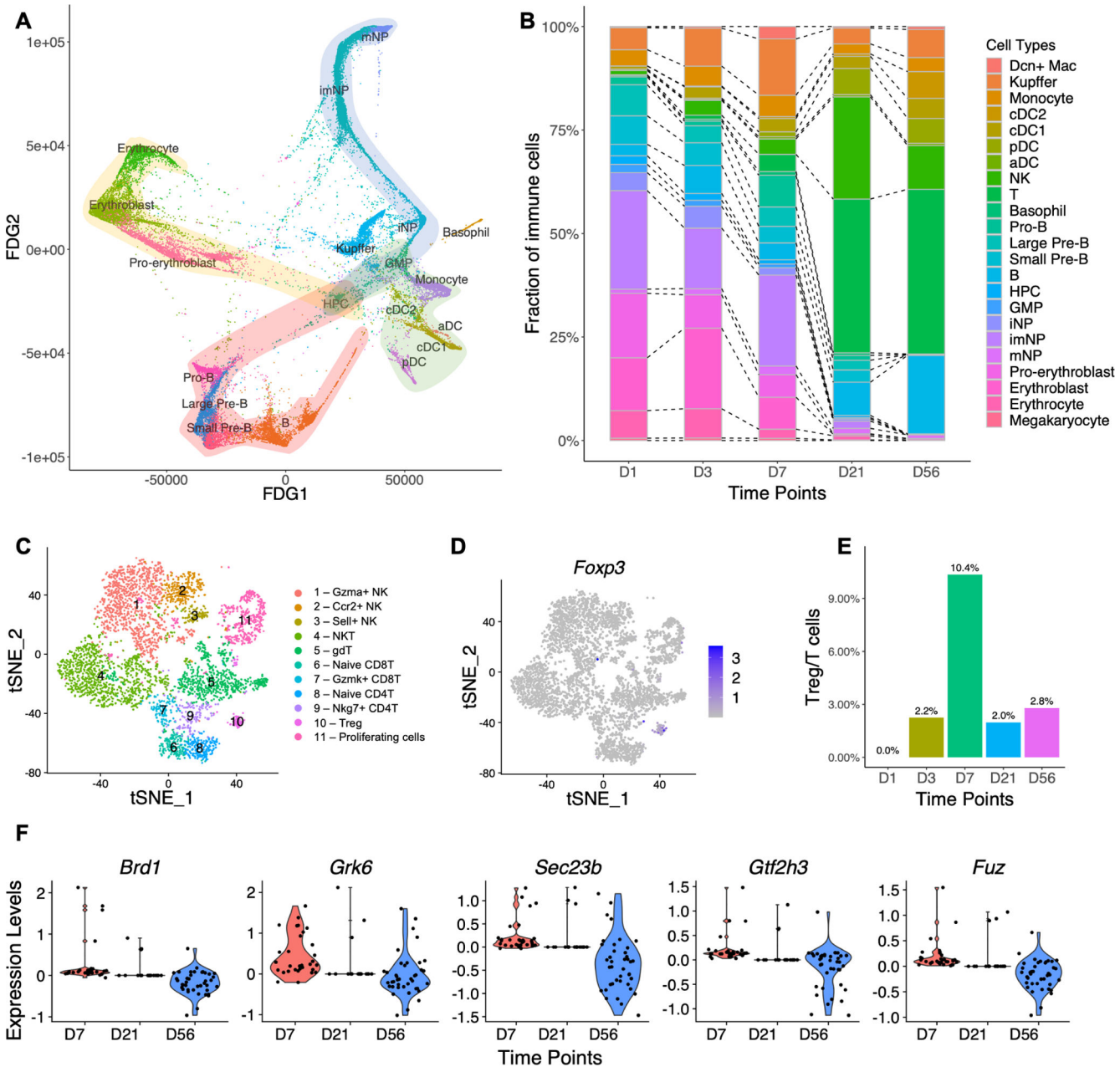


Figure 5. Dynamic changes of hematopoietic and immune cell populations

(A). Force-directed graph (FDG) of HPCs, GMPs, erythroid cells, neutrophils, B cells, DCs, basophils, monocytes and Kupffer cells.

(B). Cell type compositions of immune cells at each time point.

(C). tSNE map of T and NK cells from D1 to D56. Cells in T and NK clusters from figure 1A were segregated and re-analyzed.

(D). tSNE map displaying *Foxp3* expression in cells from 5C.

(E). Percentages of Treg cells out of all T cells at indicated time point.

(F). Violin plots displaying differentially expressed genes in Treg cells at D7, D21 and D56.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

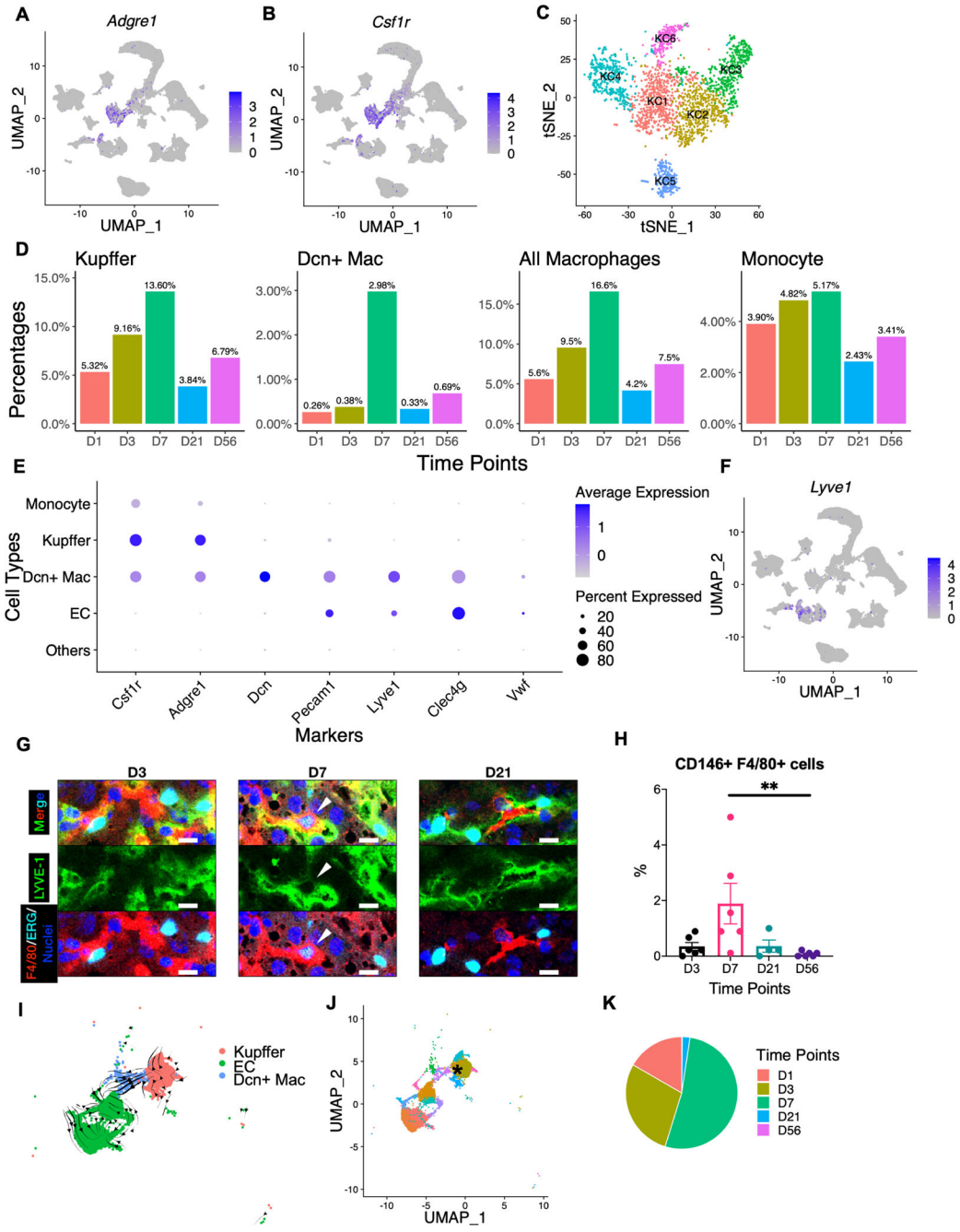


Figure 6. A subtype of macrophages emerges transiently around postnatal day 7

(A-B). UMAP visualization of *Adgre1* (A) and *Csf1r* (B) expression levels in all cells included (Figure 1A).

(C). tSNE map of Kupffer cells from D1 to D56. Cells in Kupffer clusters from figure 1A were segregated and re-analyzed.

(D). Percentages of indicated cell types out of total immune cell population.

(E). Dot plots displaying expression levels of selected markers for indicated cell types. Dot size corresponds to the ratio of cells expressing the gene in the cell type. The color corresponds to the scaled average expression level.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(F). UMAP visualization of *Lyve1* expression levels in all cells included in this study (Figure 1A).

(G). Immunostaining of F4/80, ERG and LYVE-1. Representative image taken under confocal microscope. *Dcn*$^{+}$ Mac cells were indicated by white arrowhead. Scale bar, 10 μm.

(H). Quantified FACS data of CD146^{+} F4/80^{+} cells in isolated NPCs, showing an enrichment at D7. Statistical analysis was done with Kruskal-Wallis test, followed by Dunn's test. (** $p < 0.01$).

(I). RNA velocities of Kupffer cell, EC and *Dcn*$^{+}$ Mac from D1 to D56, visualized on UMAP.

(J). Clustering analysis of cells from 6I. The Kupffer cell cluster showing differentiation potential to *Dcn*$^{+}$ Mac was labelled by asterisk (transitioning KC).

(K). Time point composition of asterisk labelled cluster (transitioning KC) from 6J.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

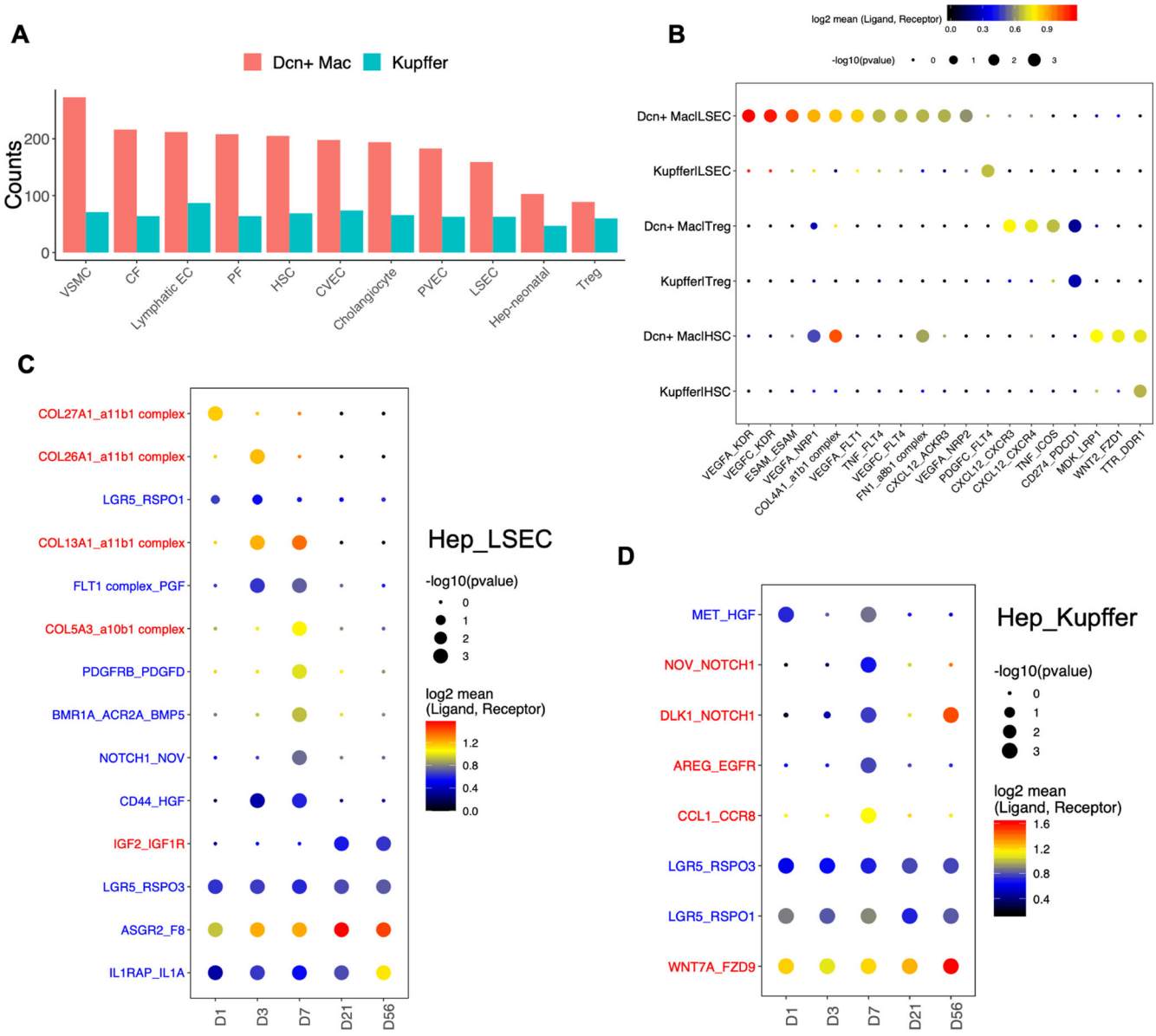


Figure 7. Predicted hepatic cell-cell interactions

(A). Bar plot comparing numbers of significant ligand-receptor interactions between indicated cell types and *Dcn*$^+$ Mac or Kupffer cells.

(B). Dot plot of selected ligand-receptor interactions between Dcn$^+$ Mac and indicated cell types (*Dcn*$^+$ Mac|LSEC, *Dcn*$^+$ Mac|Treg, *Dcn*$^+$ Mac|HSC) or between Kupffer and LSEC (Kupffer|LSEC, Kupffer|Treg, Kupffer|HSC). For example, 'moleculeA_moleculeB in cellC|cellD' indicates the putative interaction between molecule A expressed by cell type C and molecule B expressed by cell type D. P values are indicated by circle sizes, and the means of the average expression levels of interacting ligand and receptor are indicated by color.

(C-D). Dot plot of selected ligand-receptor interactions between hepatocytes and LSECs (C) or Kupffer cells (D). Interacting pairs in red indicate ligands expressed by hepatocytes; interacting pairs in blue indicate ligands expressed by LSECs or Kupffer cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

KEY RESOURCES TABLES

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Immunostaining Antibodies		
Rabbit monoclonal anti-E-Cadherin	Santa Cruz	Cat # sc-7870
Rabbit polyclonal Anti-CYP2E1	Santa Cruz	Cat # sc-133491
Rat monoclonal anti-F4/80	eBioscience	Cat # 14-4801-82
Rabbit monoclonal anti-ERG	Abeam	Cat # ab92513
Goat polyclonal anti-LYVE-1	R&D Systems	Cat # AF2125
FACS Antibodies		
LIVE/DEAD Fixable Aqua Stain	ThermoFisher	Cat # L34957
Rat monoclonal anti-CD16/CD32	eBioscience	Cat # 14-0161-85
Rat monoclonal anti-CD45, PerCP-Cy5.5	BioLegend	Cat # 147706
Rat monoclonal anti-B220, BV711	BioLegend	Cat # 103255
Rat monoclonal anti-CD146, PE-Cy7	BioLegend	Cat # 134714
Rat monoclonal anti-F4/80, Alexa Fluor 594	BioLegend	Cat # 123140
Chemicals, Peptides, and Recombinant Proteins		
Hanks' Balanced Salt solution (HBSS)	gibco	Cat # 14185-052; 14065-056
EGTA	Sigma-Aldrich	Cat # E3889
HEPES	HyClone	Cat # SH30237.01
DNase I	Roche	Cat # 10104159001
Collagenase H	Roche	Cat # 11074032001
ACK Lysing Buffer	gibco	Cat # A10492-01
Percoll	Cytiva	Cat # 17-0891-02
Dulbecco's Modified Eagle's Medium (DMEM)	gibco	Cat # 10313-021
Fetal bovine serum (FBS)	HyClone	Cat # SH30070.03
Tissue-Tek O.C.T. Compound	Sakura	Cat # 4583
Paraformaldehyde (PFA)	Sigma-Aldrich	Cat # P6148
Triton X-100	Sigma-Aldrich	Cat # X100
Bovine Serum Albumin	Spectrum	Cat # A3611
VECTASHIELD with DAPI	Avantor	Cat # H-1200
Anti-Fade Fluorescence Mounting Medium - Aqueous, Fluoroshield	Abeam	Cat # abl04135
Critical Commercial Assays		
gentleMACS Dissociator	Miltenyi	Cat # 130-093-235
gentleMACS C Tubes	Miltenyi	Cat # 130-093-237
MidiMACS separator	Miltenyi	Cat # 130-042-302
LS columns	Miltenyi	Cat # 130-042-401
Dead cell removal kit	Miltenyi	Cat # 130-090-101
Chromium Single Cell 3′ Reagent Kits (v2 Chemistry)	Miltenyi	Cat # CG00052

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Scd1 – qPCR FWD	Eton Bioscience	GCGATACACTCTGGTGCTCA
Scd1 – qPCR REV	Eton Bioscience	CCCAGGGAAACCAGGATATT
Scd2 – qPCR FWD	Eton Bioscience	GCTCTCGGGAGAACATCTTG
Scd2 – qPCR REV	Eton Bioscience	CAGCCCTGGACACTCTCTTC
F4/80 – qPCR FWD	Eton Bioscience	CTTTGGCTATGGGCTTCCAGTC
F4/80 – qPCR REV	Eton Bioscience	GCAAGGAGGACAGAGTTTATCGTG
Clec4f – qPCR FWD	Eton Bioscience	CTTCGGGGAAGCAACAACTC
Clec4f – qPCR REV	Eton Bioscience	CAAGCAACTGCACCAGAGAAC
Lyve1 – qPCR FWD	Eton Bioscience	GAAAACTCTGTTGCGGGTGT
Lyve1 – qPCR REV	Eton Bioscience	CCTCCAGCCAAAAGTTCAAA
Alb – qPCR FWD	Eton Bioscience	GCAGATGACAGGGCGGAACTTG
Alb – qPCR REV	Eton Bioscience	CAGCAGCAATGGCAGGCAGAT
Gapdh – qPCR FWD	Eton Bioscience	CGACTTCAACAGCAACTCCCACTCTTCC
Gapdh – qPCR REV	Eton Bioscience	TGGGTGGTCCAGGGTTTCTTACTCCTT
Deposited Data		
Adult hepatocytes (processed scRNA-seq data)	(Halpern *et al.,* 2017)	Supplementary tables
E7.5, E8.75, E9.5 and E10.5 endothelial cell (processed scRNA-seq data)	Single Cell Portal (Lotto *et al,* 2020)	SCP1022
Software and Algorithms		
FlowJo v10.6.2	BD	https://www.flowio.com
GraphPad Prism 9	GraphPad	https://www.graphpad.com/scientific-software/prism/
ImageJ/FIJI	ImageJ	https://imagei.net/Fiji
Cellranger v3.0.2	10x Genomics	https://www.10xgenomics.com/
R	The R Foundation	https://www.r-project.org/
Seurat v3.0.2	(Butler et al., 2018; Stuart et al., 2019)	https://github.com/satiialab/seurat
Monocle	(Cao et al., 2019; Qiu *et al.*, 2017; Trapnell *et al.*, 2014)	http://cole-trapnell-lab.github.io/monocle-release/
SCENIC v1.1.2.2	(Aibar *et al.*, 2017)	https://github.com/aertslab/SCE NIC
RNA Velocity	(La Manno et al., 2018)	http://velocyto.org/velocyto.py/index.html
scVelo	(Bergen et al., 2020)	https://scvelo.readthedocs.io/
CellphoneDB	(Efremova *et al.*, 2020)	https://github.com/Teichlab/cellphonedb
Other		
HCR FISH probe: *Neat1*	Molecular Instruments	NR_131212.1
HCR FISH probe: *Malat1*	Molecular Instruments	NR_002847.2