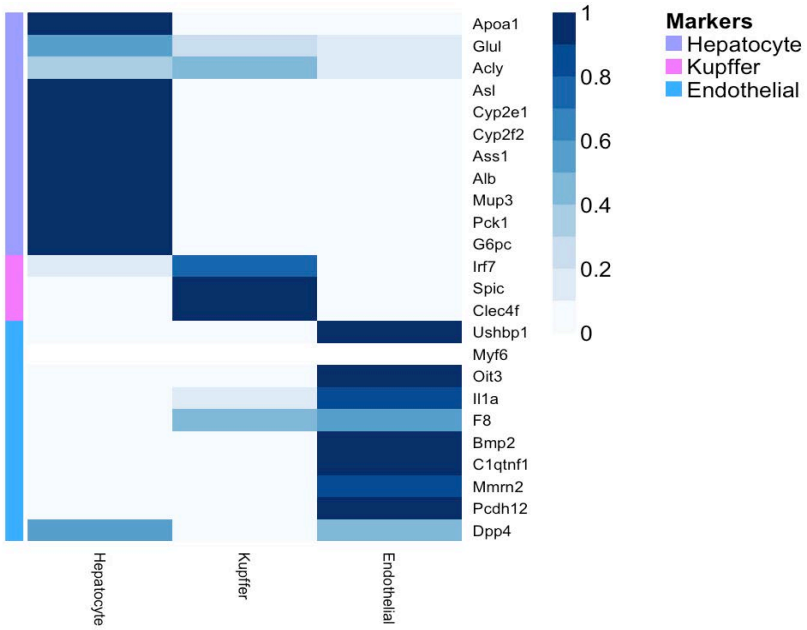
Figure S1

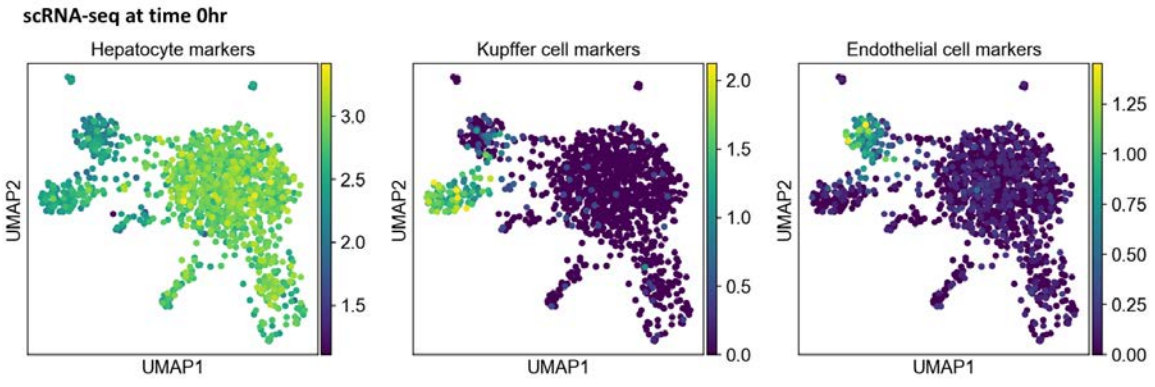
A



B

Tissue	Condition	Gender	Age	Hepatocytes	Kupffer cells	Endothelial cells
liver	PHx48h	male	adult	0.9589439	0.01895047	0.022105632
liver	PHx48h	male	adult	0.9559730	0.01865062	0.025376421
liver	PHx48h	male	adult	0.9559757	0.01865023	0.025374104
liver	PHx0h	male	adult	0.9821886	0.01389723	0.003914131
liver	PHx0h	male	adult	0.9719966	0.01701637	0.010987019
liver	PHx0h	male	adult	0.9785515	0.00000000	0.021448470

C



D

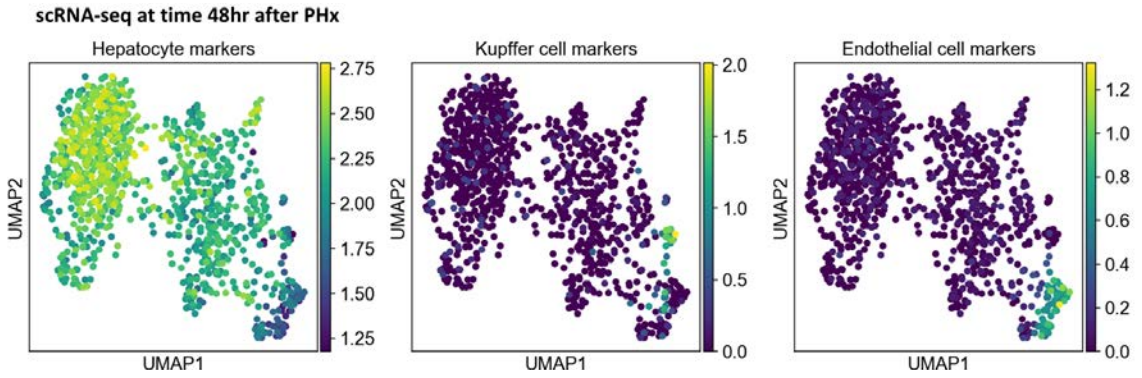


Figure S1. Deconvolution of single cell RNA-seq data demonstrates little contamination of the hepatocyte preparation by other cell types.

A. Cell types signatures including hepatocytes, Kupffer cells and Endothelial cells derived from previously published single cell RNA-seq data (ref.9).
B. Proportions of each cell type in healthy livers (PHx0h, n=3) and 48h PHx livers (PHx48h, n=3).
C-D. UMAP plot of distinct cell types signatures at PHx0h (C) and PHx48h (D).

Figure S2

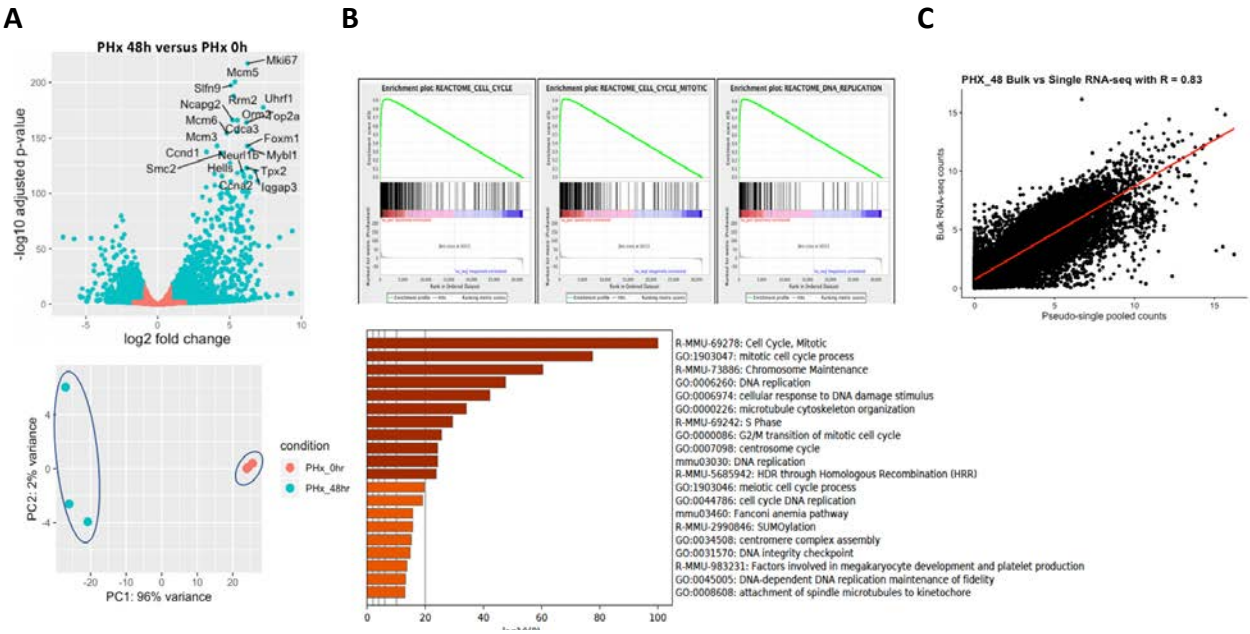
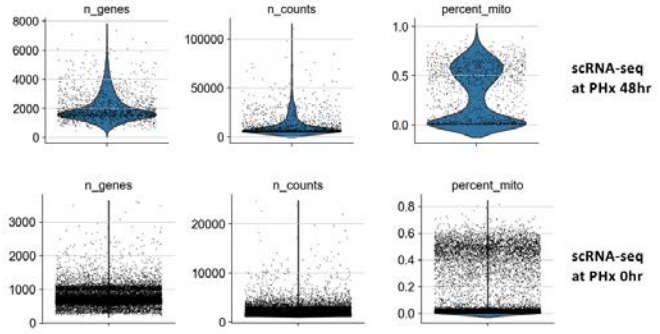


Figure S2. Bulk RNA-seq analysis of hepatocytes before and at 48h post-PHx.

A. Volcano plot showing comparative bulk RNA-seq analysis of UD hepatocytes versus hepatocytes at PHx 48h (top). Result of principle component analysis (bottom).
B. GSEA results (top) and GO terms (bottom) of enriched pathways in PHx 48h condition.
C. Correlation plot showing the overall similarity between bulk RNA-seq and pseudo-bulk scRNA-seq datasets at PHx 48h.

Figure S3

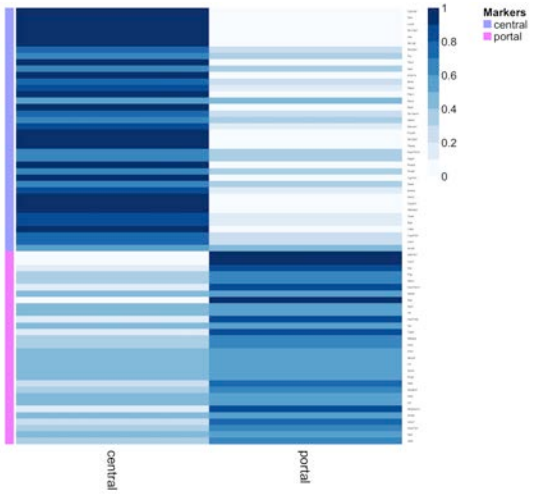
A



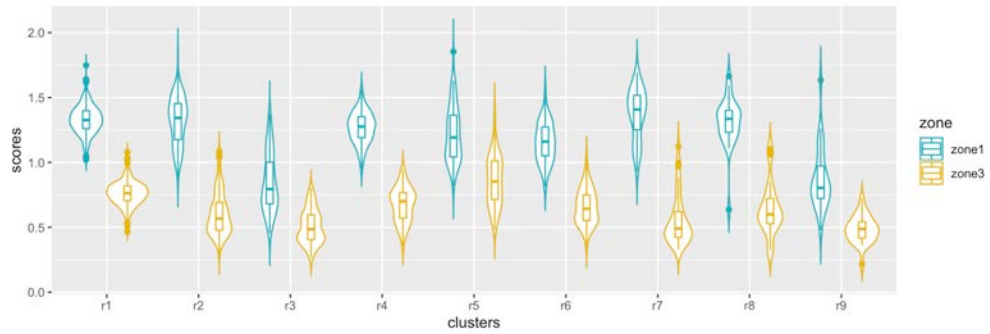
B

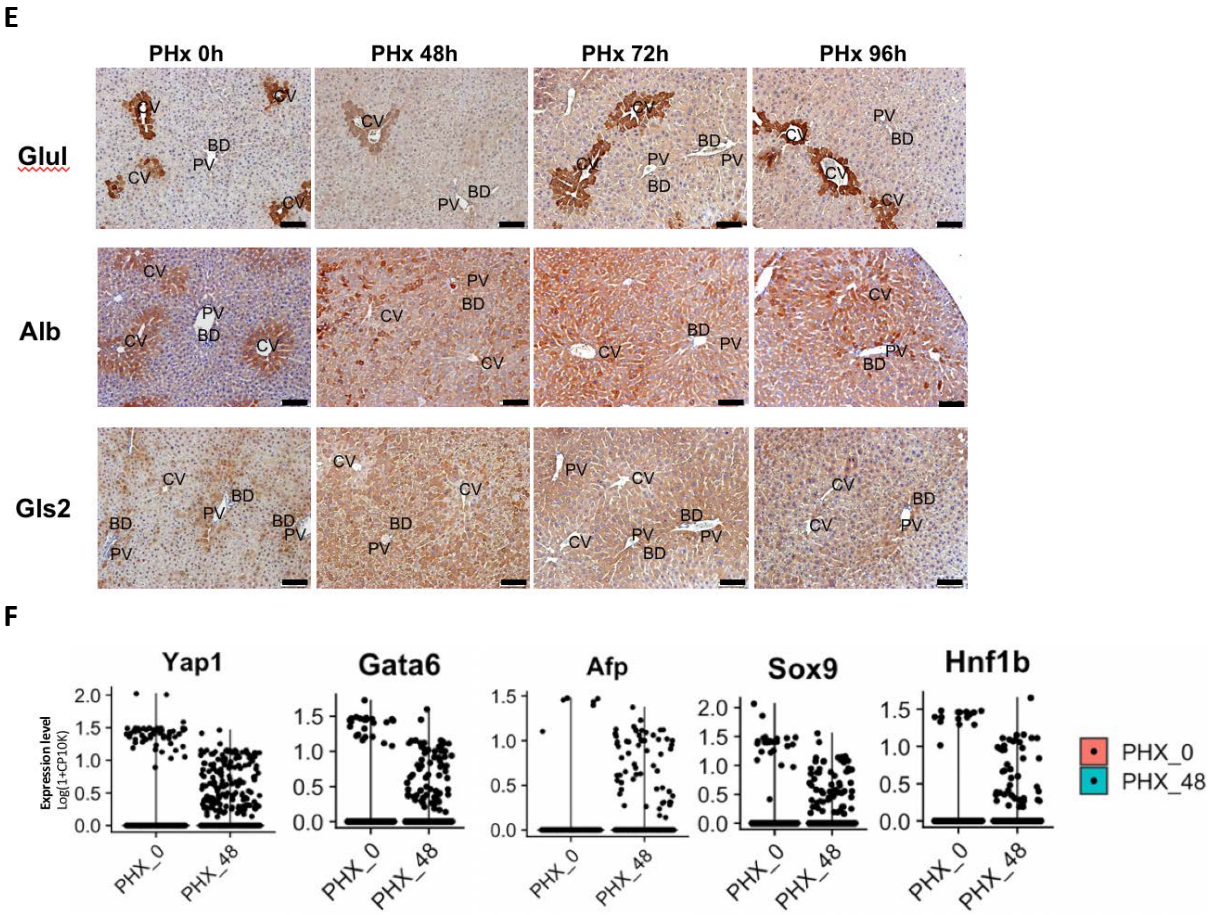
clusters	number_of_cells_per_cluster
r1	438
r2	205
r3	92
r4	90
r5	67
r6	62
r7	40
r8	28
r9	27

C



D




E
PHx 0h
PHx 48h
PHx 72h
PHx 96h
Glul
Alb
Gls2
CV
PV
BD
F
Yap1
Gata6
Afp
Sox9
Hnf1b
Expression level
Log[1+CP10K]
PHX_0
PHX_48
G
Cebpa
Esrp2
Gls2
Alb
Hnf4a
H
GO:0006790: sulfur compound metabolic process
mmu01200: Carbon metabolism
GO:0007005: mitochondrion organization
mmu00280: Valine, leucine and isoleucine degradation
GO:0006575: cellular modified amino acid metabolic process
R-MMU-199977: ER to Golgi Anterograde Transport
GO:0044257: cellular protein catabolic process
mmu04141: Protein processing in endoplasmic reticulum
mmu00020: Citrate cycle (TCA cycle)
GO:0044283: small molecule biosynthetic process
GO:0032787: monocarboxylic acid metabolic process
GO:0008202: steroid metabolic process
GO:0044282: small molecule catabolic process
GO:0051186: cofactor metabolic process
mmu04146: Peroxisome
GO:0017144: drug metabolic process
mmu04610: Complement and coagulation cascades
R-MMU-5389840: Mitochondrial translation elongation
mmu03050: Proteasome
R-MMU-72172: mRNA Splicing
r6 r1 r4 r2 r5
I
R-MMU-9006934: Signaling by Receptor Tyrosine Kinases
GO:0008380: RNA splicing
R-MMU-5653656: Vesicle-mediated transport
GO:0030036: actin cytoskeleton organization
R-MMU-109582: Hemostasis
R-MMU-195258: RHO GTPase Effectors
R-MMU-1280218: Adaptive Immune System
mmu04144: Endocytosis
R-MMU-6798695: Neutrophil degranulation
GO:0033365: protein localization to organelle
mmu04145: Phagosome
GO:0006413: translational initiation
GO:1903827: regulation of cellular protein localization
mmu04142: Lysosome
R-MMU-422475: Axon guidance
GO:0006897: endocytosis
mmu04510: Focal adhesion
GO:0034329: cell junction assembly
GO:0022604: regulation of cell morphogenesis
GO:0048514: blood vessel morphogenesis
r3 r9

J

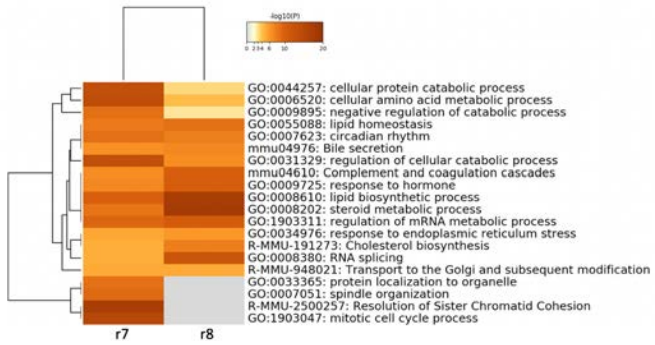


Figure S3. Single cell RNA-seq analysis reveals heterogeneity of hepatocytes in regenerating liver.

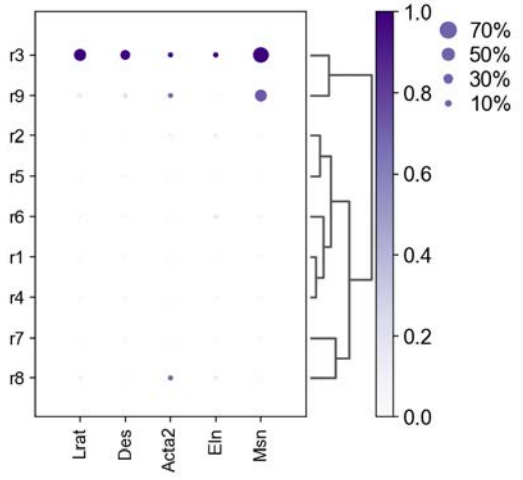
A. Quality metrics of single cell RNA-seq data of hepatocytes at 48h PHx (top) and 0h PHx (bottom).
B. Table showing the respective cell number for each cluster (r1-9).
C. Central and portal zone signatures derived from previously published single cell RNA-seq data (ref. 9).
D. Violin plot showing distribution of zonal signatures in each cluster (r1-9).
E. IHC staining of various zonal markers before and after PHx. Scale bar = 100uM

F-G. Violin plots showing the expression of representative markers for stem/progenitor cells (F) and mature hepatocytes (G).

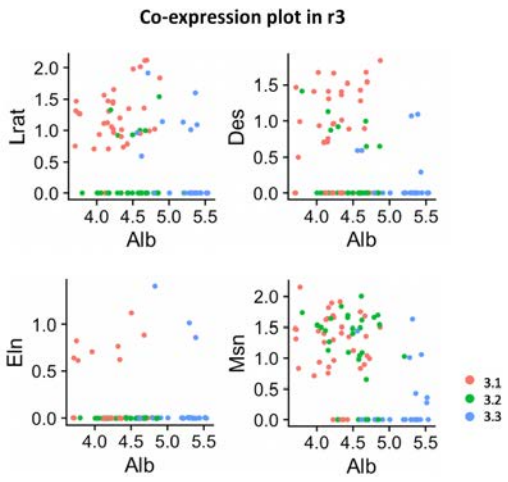
H. Enriched GO terms for group A including r1, r2, r4, r5 and r6.
I. Enriched GO terms for group B including r3 and r9.
J. Enriched GO terms for group C including r7 and r8.

Figure S4

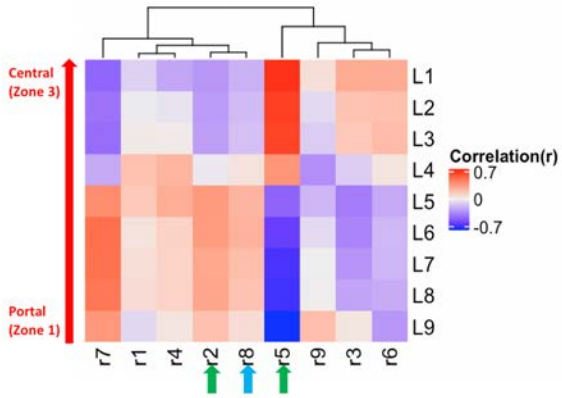
A



B



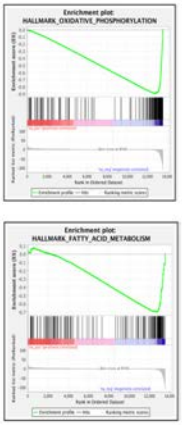
C



D

GSEA upregulated pathways in clusters r1 and r4

Enriched pathways	p-val
Oxidative phosphorylation	0
MYC_targets_v1	0
Adipogenesis	0
Reactive oxygen species pathway	0.004
Interferon alpha response	0.009
Fatty acid metabolism	0.017
Allograft rejection	0.024
Glycolysis	0.044
DNA repair	0.075



E

GSEA upregulated pathways in clusters r2, r5 and r6

Enriched pathways	p-val
UV_response_down	0.007
KRAS signaling_down	0.015
Hedgehog signaling	0.021
Estrogen response early	0.019
IL6_JAK_STAT3 signaling	0.031
Angiogenesis	0.035
TNFa singaling	0.018
Inflammatory response	0.057
Epithelial_mesenchymal_	0.059

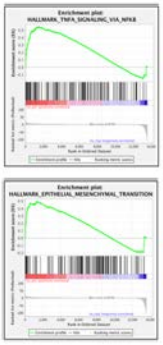


Figure S4. Refined analysis of regenerative clusters demonstrates functional heterogeneity.

A. Dot plot showing expression of representative hepatic stellate cells (HSCs) genes in each cluster (r1-9).

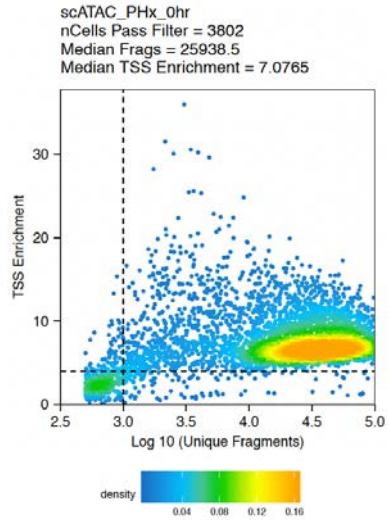
B. Co-expression plot showing various markers in r3.

C. Heatmap showing zonal enrichment for each cluster (r1-9).

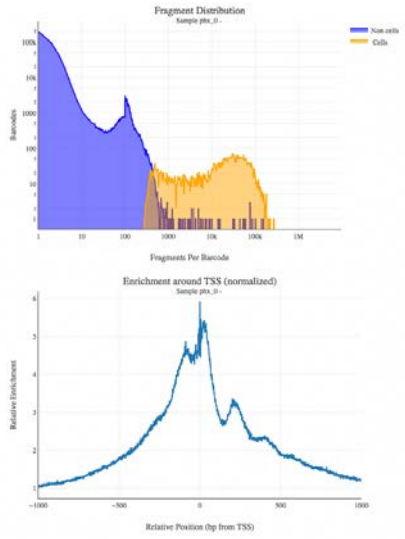
D-E. GSEA results showed pathways enriched in clusters r1 and r4 (D) and enriched pathways in clusters r2, r5 and r6 (E).

Figure S5

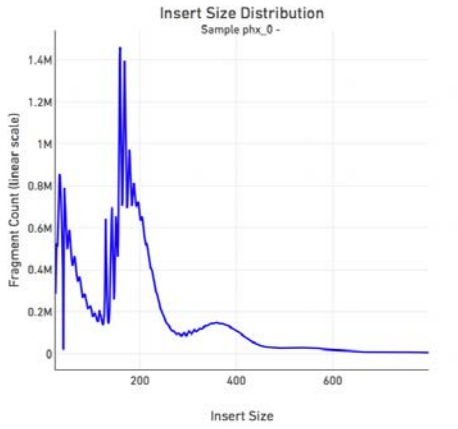
A

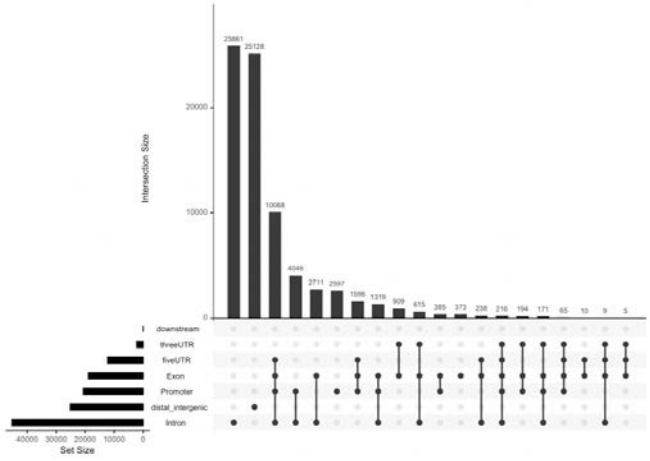


B

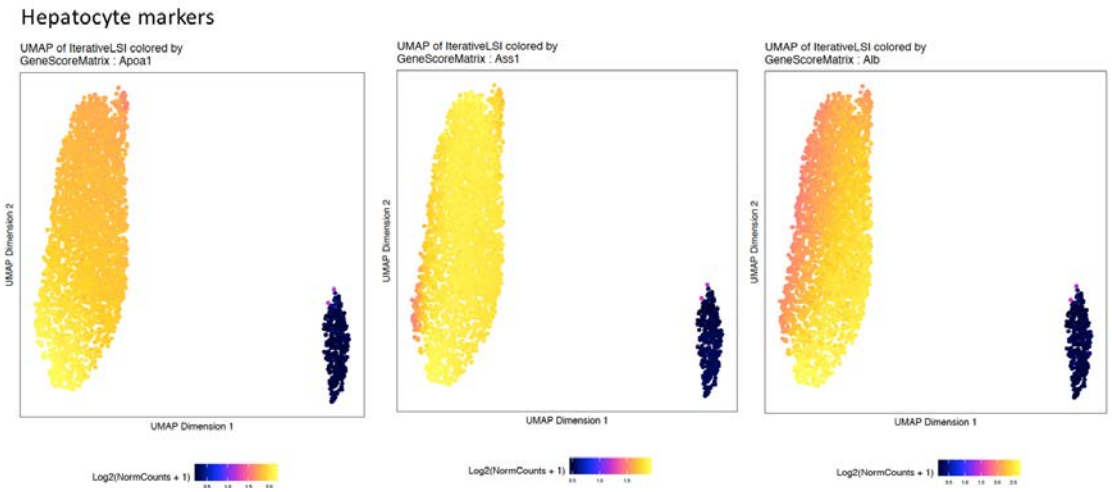


C

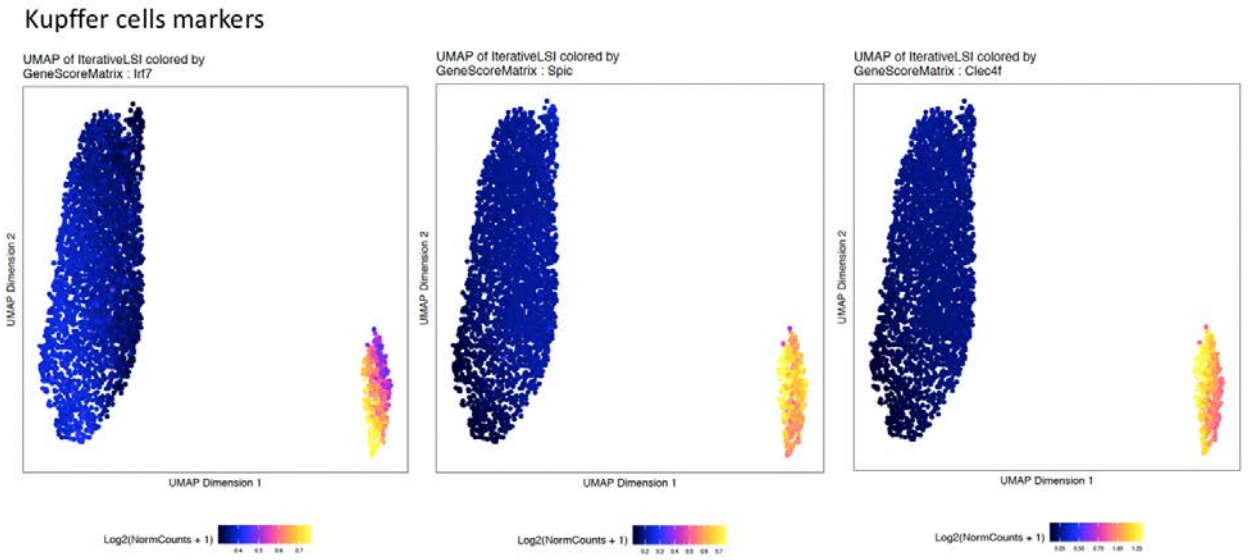




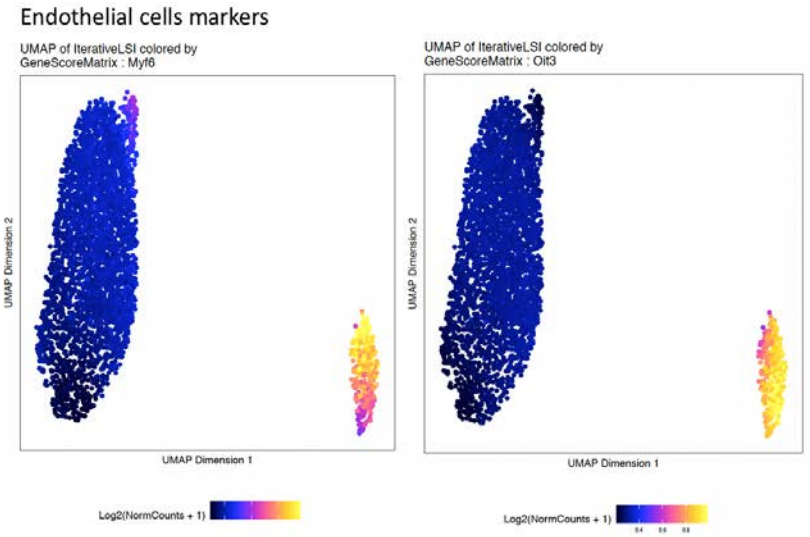
D



E



F



G

	C2	C4	C5	C3
L1	0.2	-0.04	-0.11	-0.4***
L2	0.16	0.06	0.04	-0.28**
L3	0.18	0.08	0.11	-0.23*
L4	0.17	0.19	0.35**	0.08
L5	-0.05	0.1	0.26*	0.46***
L6	-0.14	0.07	0.13	0.43***
L7	-0.17	0.02	0.06	0.38***
L8	-0.18	-0.02	0.01	0.33**
L9	-0.27**	-0.15	-0.2	0.11

***P<0.001, **P<0.01, *P<0.05

Figure S5. Quality metrics of single cell ATAC-seq data at time 0hr.

A-B. Quality metrics of transcription starting sites (TSS) scores versus unique fragments per nucleus in 0hPHx (undamaged liver, UD) hepatocytes.

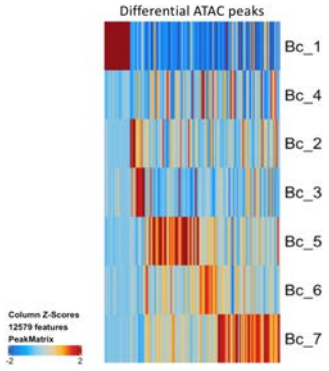
C. ScATAC-seq library fragment size distribution and their genome locations.

D-F. scATAC-seq UMAP projections of distinct cell types signatures at time 0h including hepatocytes (D), Kupffer cells (E) and Endothelial cells (F).

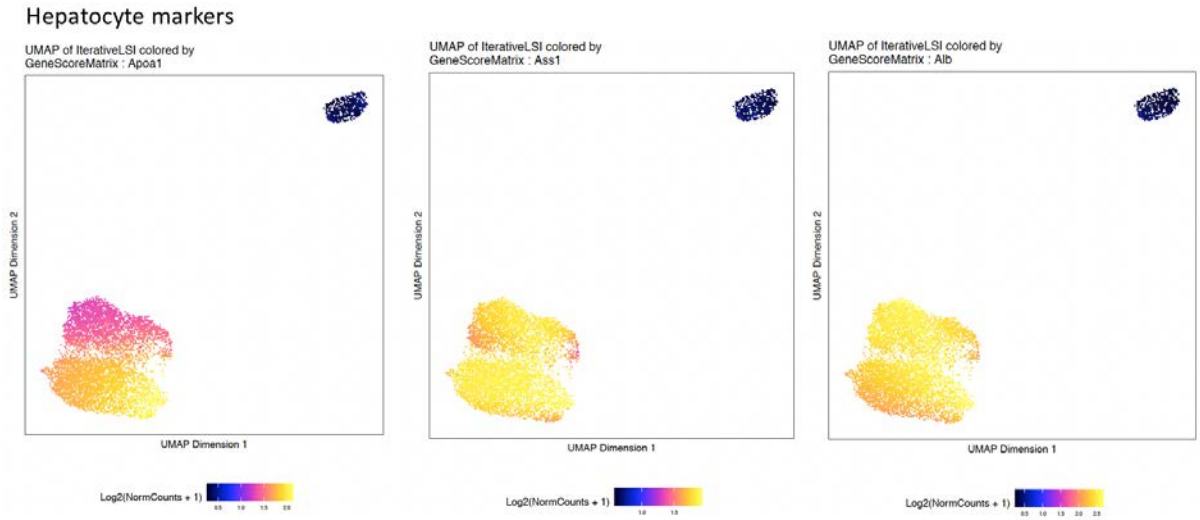
G. Table showing Pearson correlation coefficient calculated for hepatocyte clusters (C2-5) with the pre-defined (layer L1-9) zonal markers.

Figure S6

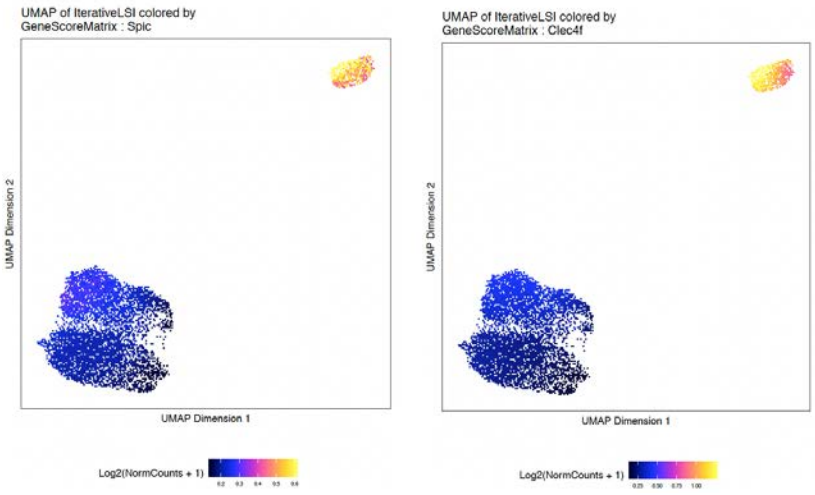
A



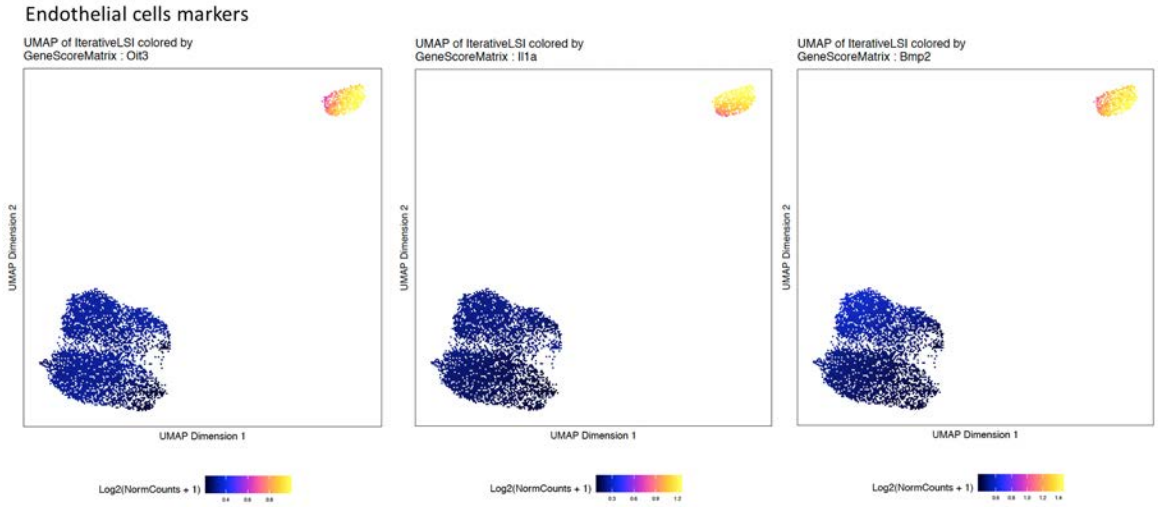
B



C



D



E

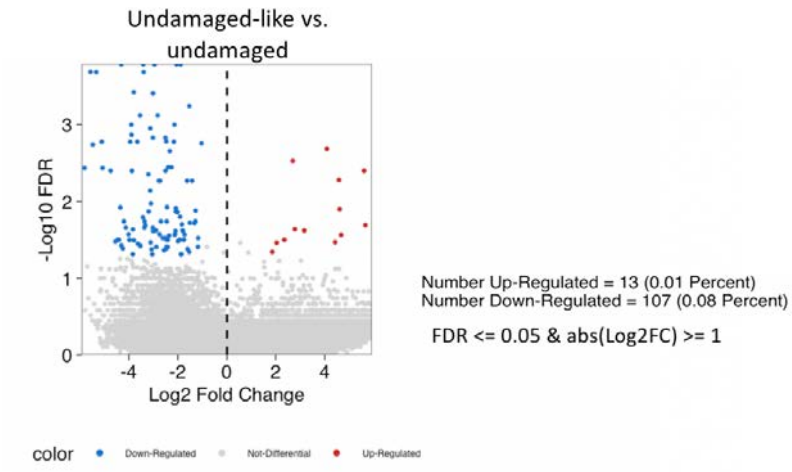


Figure S6. Quality metrics of combined single cell ATAC-seq data at time 0hr and 48hr PHx.

A. Heatmap showing differential peaks identified for each cluster (Bc_1-7).

B-D. scATAC-seq UMAP projections of distinct cell types signatures including hepatocytes (B), Kupffer cells (C) and Endothelial cells (D).

E. Volcano plot showing distribution of chromatin regions with increased (red) or decreased (blue) accessibility in undamaged-like hepatocytes at 48h PHx.

Figure S7

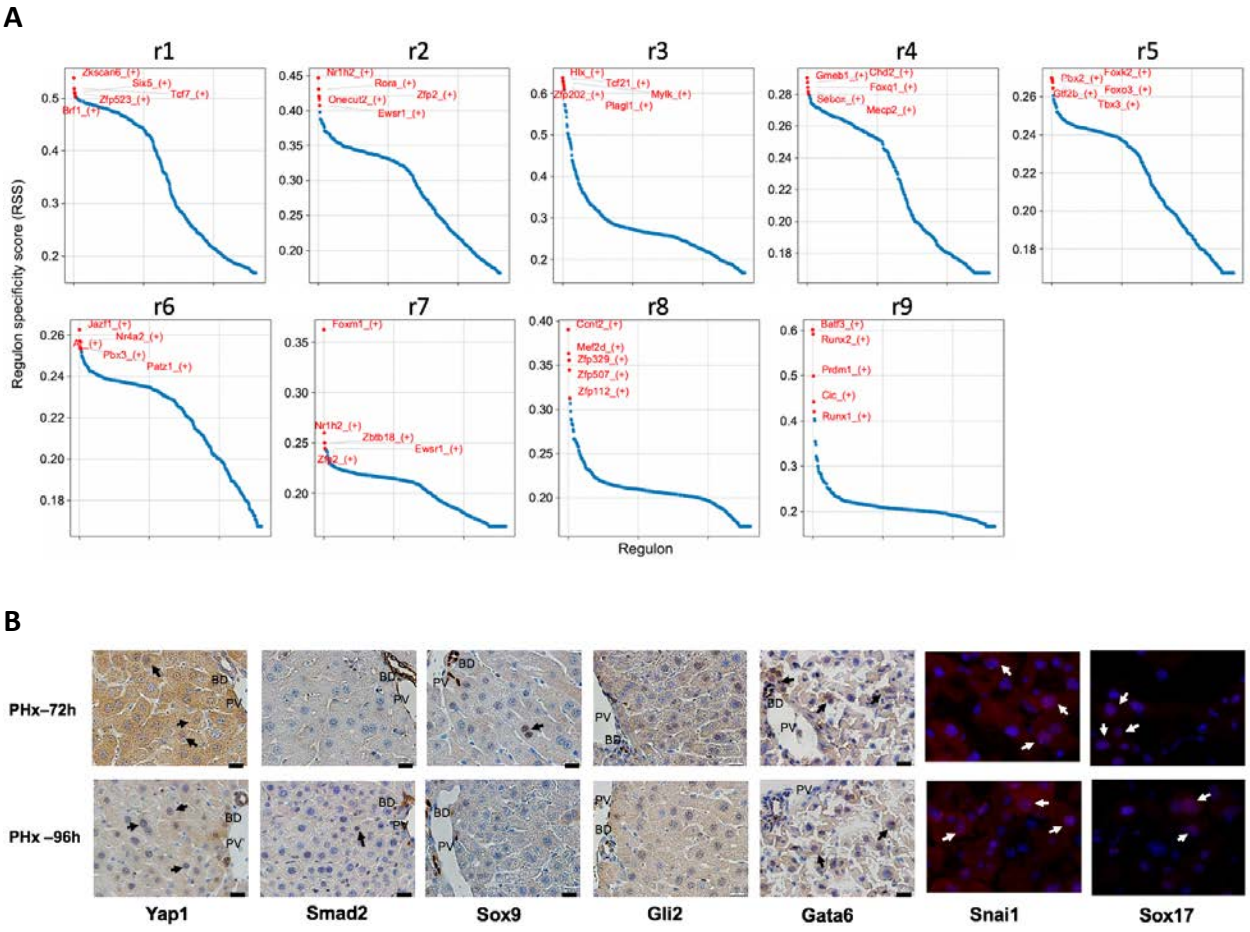


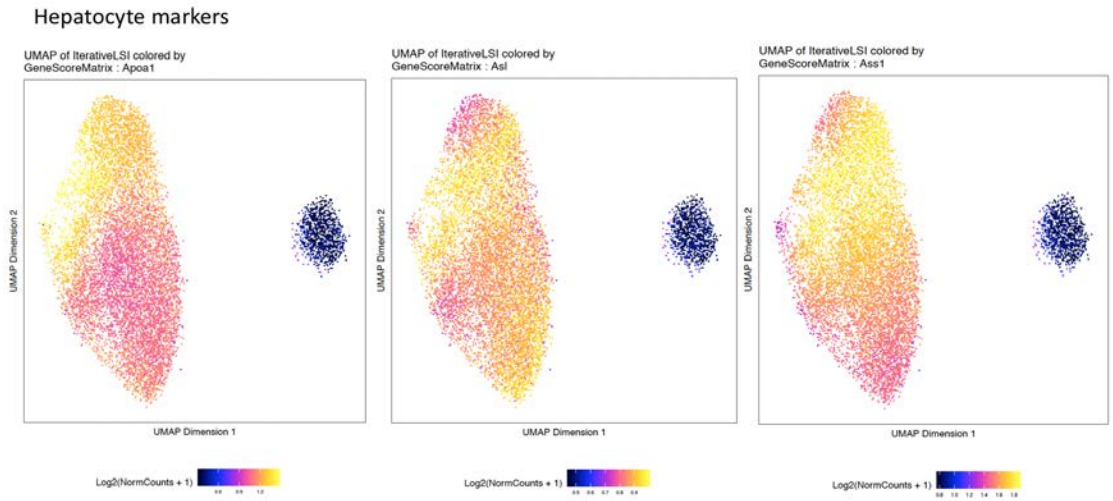
Figure S7. Top ranked regulons in each regenerative cluster and immunohistochemical confirmation of regulon-related transcription factors in fetal-like cluster r8.

A. Top ranked regulons identified for each cluster r1-r9 of scRNA-seq data at 48h PHx.

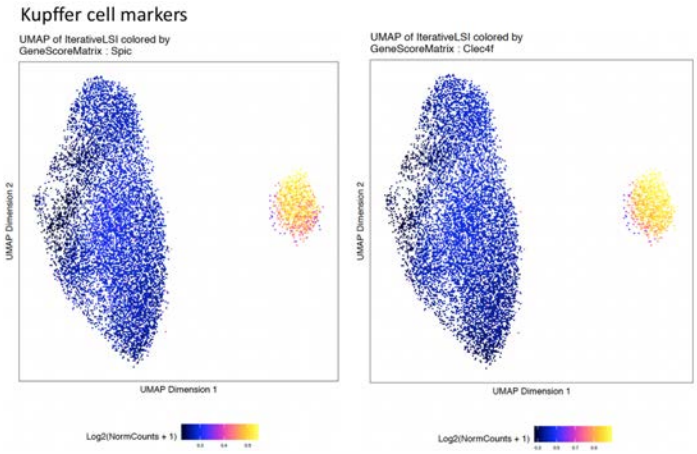
B. High magnification view of hepatocyte expression of transcription factors varies during the regenerative process. Immunohistochemical (IHC) and Immunofluorescence (IF) staining for Yap1, Smad2, Sox9, Gli2, Gata6 (brown), Snai1 and Sox17 (red) in representative sections from livers after partial hepatectomy (PH) counterstained with Hematoxylin or nuclear DAPI (blue). Black arrow in IHC stained results (Yap1, Smad2, Sox9, Gli2 and Gata6) indicates positive signal in hepatocyte nuclei. White arrows in IF stained results (Snai1 and Sox17) indicates positive signal in hepatocyte nuclei. Size bar: 50μm.

Figure S8

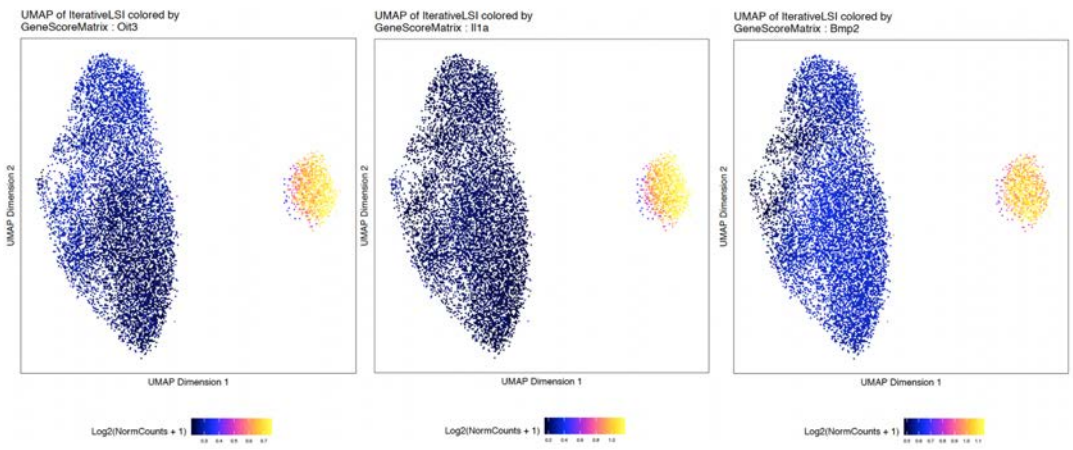
A



B



C



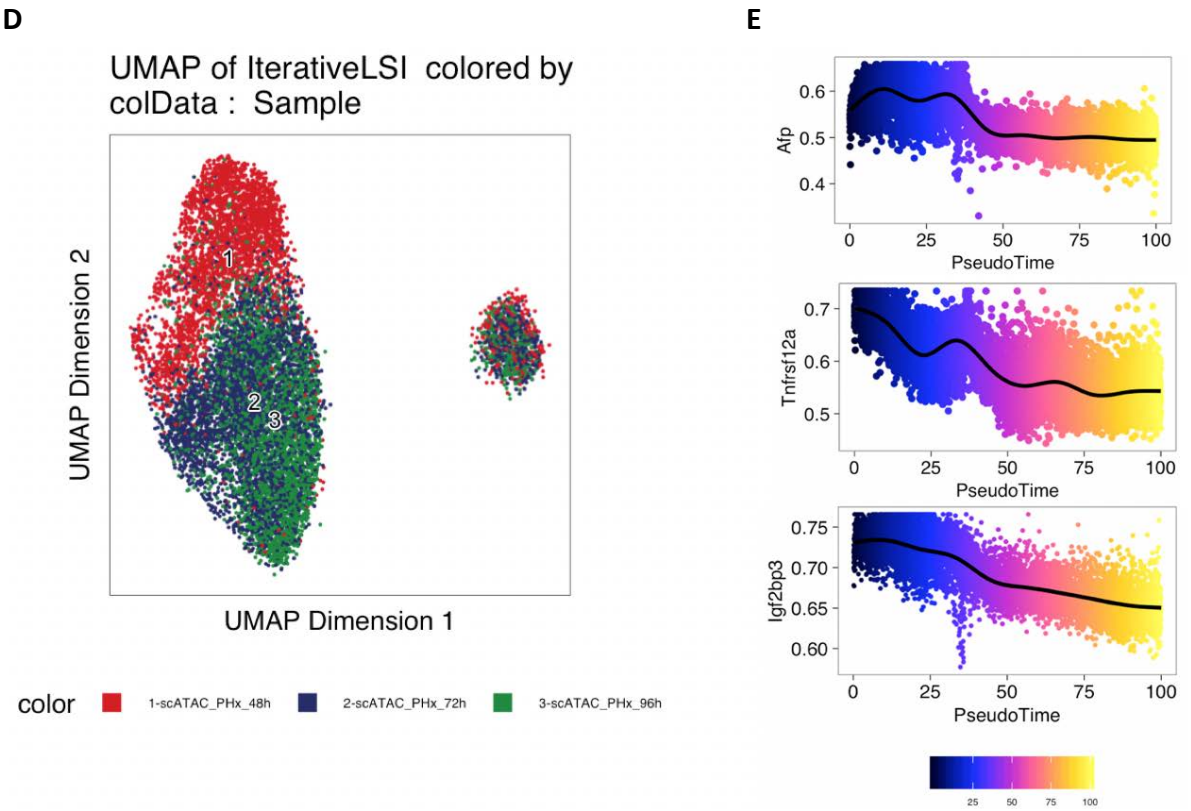


Figure S8. Quality metrics of combined single cell ATAC-seq data at 48h, 72h and 96h PHx.

A-C. scATAC-seq UMAP projections of distinct cell types signatures including hepatocytes (A), Kupffer cells (B) and Endothelial cells (C).

D. UMAP projection of combined data sets. Data are colored according to different time points.

E. Expression of representative fetal hepatocyte markers including Afp, Tnfrsf12a and Igf2bp3 across inferred pseudotime trajectory.

Supplemental materials and methods

IHC and IF. Liver tissues were frozen in OCT or fixed in formalin, sectioned and immunohistochemistry was performed with the following antibodies: Yap1 (14074, Cell Signaling), Smad2 (sc-101153, Santa Cruz Biotechnology), Sox9 (AB5535, EMD Millipore), Sox17 (09-038-I, EMD Millipore; PA5-72815, Invitrogen), Gli2 (18-732-292462, Genway), Snai1 (ab180714, Abcam), Gata6 (ab22600, Abcam). Secondary antibodies were HRP-conjugated anti-rabbit (K4003, Dako) or anti-mouse (K4003, Dako) antibodies. Blocking and chromogenic detection was performed using the DAKO Envision System with DAB substrate according to the manufacturer's protocol. For detection of Snai1 and Sox17, Alexa Fluor 594 donkey anti-rabbit IgG (A10042, Carlsbad) were used. Nuclei were identified by DAPI positivity (H-1200, Vector Labs). Tissue sections were counterstained with Aqua Hematoxylin-INNOVEX (Innovex Biosciences).

***Bulk RNA-seq and analysis*.** RNAs were extracted from fresh hepatocyte isolates using the TRIzol RNA Purification kit and cDNA libraries were made with NEBNext Ultra RNA Library Prep Kit from Illumina. Sample qualities were assessed by Agilent RNA Bioanalyzer chip traces before pooled and sequenced with an Illumina HiSeq 2500. Sequencing reads were trimmed using TrimGalore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and aligned to the mm10 reference genome using STAR.[1] Gene expression counts were generated using HTSeq [2] and DEseq2 [3] was used to perform differential expression analyses (absolute log2FC >= 1 and q <0.05). Cell type specific marker genes, including hepatocytes, Kupffer cells, and endothelial cells were generated by Halpern *et al.* 2007. Deconvolution was performed in bseqsc [4] using the gene expression matrix with values converted to TPM.

1. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.
2. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166-9.
3. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
4. M. Baron, A. Veres, S.L. Wolock, A. L. Faust, R. Gaujoux, A. Vetere, J. Hyoje Ryu, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Systems. 2016 Oct 26