

Supplemental Information

Comprehensive single-cell transcriptional profiling defines shared and unique epithelial injury responses during kidney fibrosis

Haikuo Li¹, Eryn E. Dixon¹, Haojia Wu¹ and Benjamin D. Humphreys^{1,2*}

¹Division of Nephrology, Department of Medicine, Washington University in St. Louis, St. Louis, MO, USA

²Department of Developmental Biology, Washington University in St. Louis, St. Louis, MO, USA

*Correspondence to B.D.H (humphreysbd@wustl.edu)

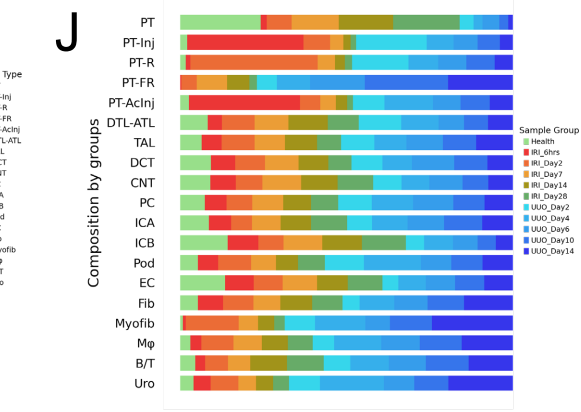
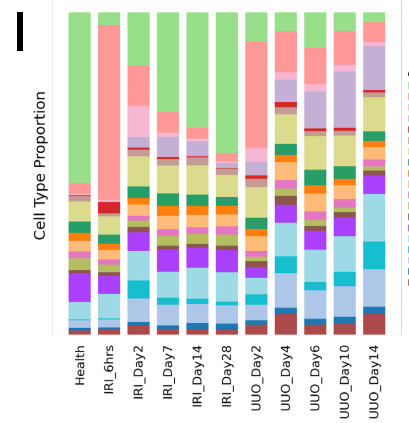
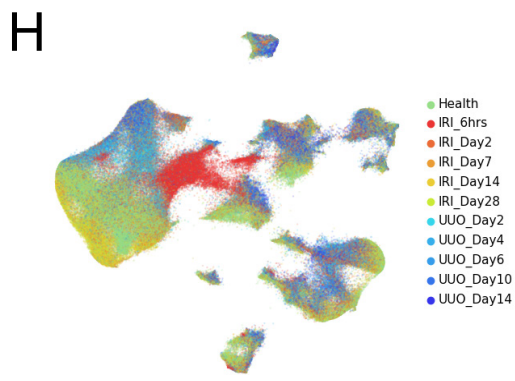
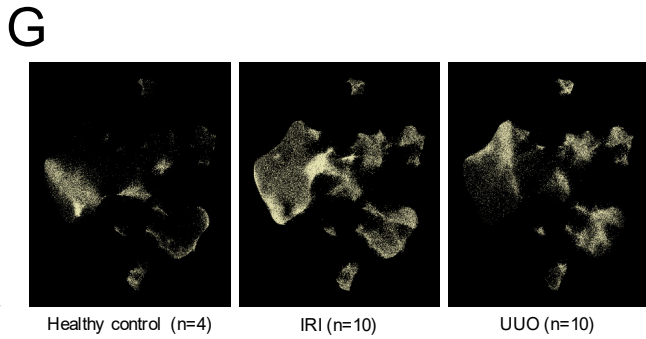
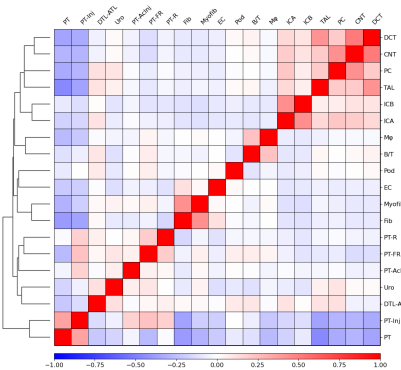
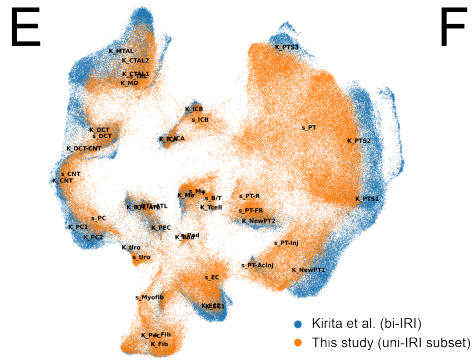
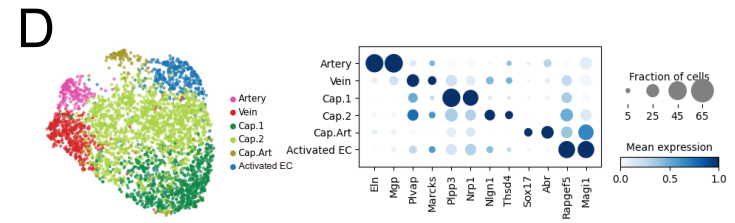
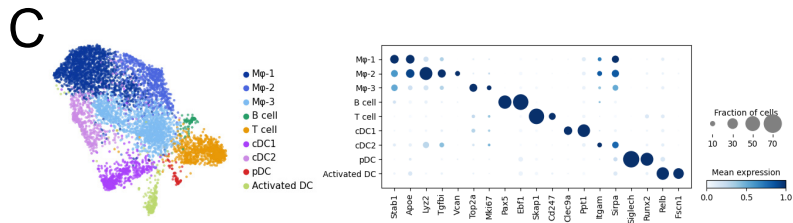
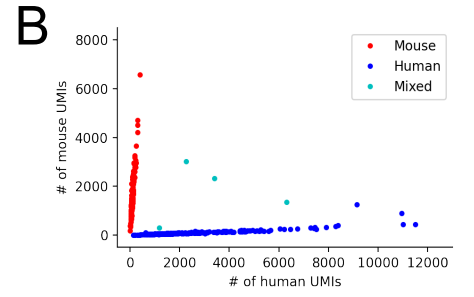
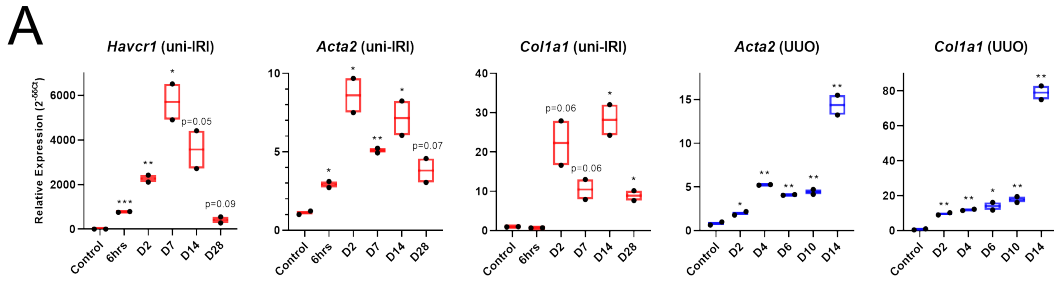


Figure S1. A single-cell transcriptomics landscape of mouse kidney fibrogenesis profiled with sci-RNA-seq3 (related to Figure 1).

- (A) Relative expression of marker genes measured by qPCR. Each dot represents a tissue sample. Data are shown as the line at mean. Comparisons were made between each condition and healthy controls. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$ by Student's t test.
- (B) Cell collision rate determined by the species-mixing experiment. Cells with over 85% of UMIs assigned to one species are annotated as species-specific cells and the other cells were annotated as cell collisions.
- (C) UMAP plot of immune cell subtypes identified in subclustering analysis and dot plot showing expression of cluster-specific marker genes. cDC1, Type 1 conventional dendritic cell; cDC2, Type 2 conventional dendritic cell; pDC, plasmacytoid dendritic cell.
- (D) UMAP plot of endothelial cell subtypes identified in subclustering analysis and dot plot showing expression of cluster-specific marker genes. Cap, capillary endothelium; Cap.Art, capillaries connecting the artery; EC, endothelial cell.
- (E) Integration of our uni-IRI subset with a previous scRNA-seq on renal bi-IRI samples (by 10X Genomics) with batch effects removed, using BBKNN.
- (F) Correlation plot of all major cell types identified in Figure 1D. High correlation score indicates high similarity of cell type transcriptomes.
- (G) Condition map presenting unique cell distribution in different experimental groups.
- (H) Projection of cells based on their group condition on the global UMAP.
- (I) Bar plots showing the abundance of each cell type in a group condition.
- (J) Bar plots showing the abundance of each group condition in a cell population.

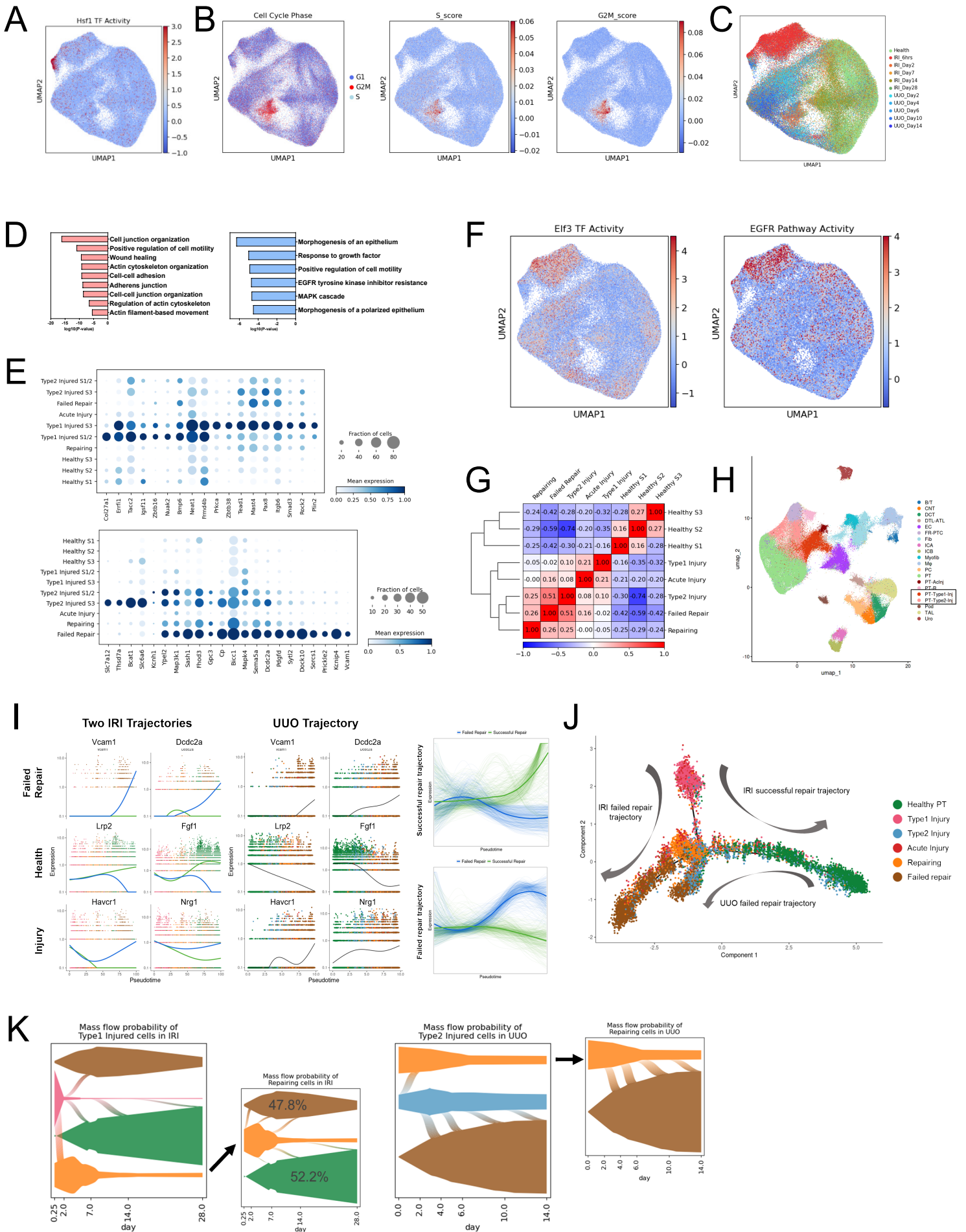


Figure S2. Diverse cell states of injured proximal tubule (related to Figure 2).

- (A) Transcription factor activity of Hsf1 in PT cells determined by gene regulatory network analysis.
- (B) Cell phase prediction of PT cells with cell cycle scoring analysis.
- (C) Projection of cells based on their group condition on the PT subclustering UMAP.
- (D) Top significant terms of gene ontology enrichment analysis on DEGs of Type1 (left panel) and Type2 (right panel) injured PT.
- (E) Dot plots showing expression pattern of DEGs of Type1 (upper panel) and Type2 injured PT cells and FR-PTC (lower panel).
- (F) Transcription factor activity of Elf3 in PT cells determined by gene regulatory network analysis (left panel) and activity of EGFR pathway determined by single-cell pathway analysis (right panel).
- (G) Correlation plot of all PT subtypes identified. High correlation score indicates high similarity of cell type transcriptomes.
- (H) UMAP showing results of mapping Type1 and Type2 injured PT cells back to the global clustering.
- (I) Left two panels: Expression dynamics of marker genes of failed repair (*Vcam1*, *Dcdc2a*), healthy (*Lrp2*, *Fgf1*) and injured PT (*Havcr1*, *Nrg1*) in two uni-IRI trajectories identified in the left panel of Figure 2F and the UUO trajectory identified in the right panel of Figure 2F across pseudotime. For the two uni-IRI trajectories, successful repair trajectory is colored by green and failed repair trajectory is colored by blue. Right panel: Expression dynamics of DEGs of the two uni-IRI trajectories identified with BEAM. Thick lines indicate the average gene expression patterns in each catalog.
- (J) Pseudotemporal ordering of cells derived from data combining both uni-IRI and UUO samples colored by cluster identity.
- (K) Single-cell fate mapping of PT-R cells differentiated from Type1 (left panel) and Type2 injured PT (right panel).

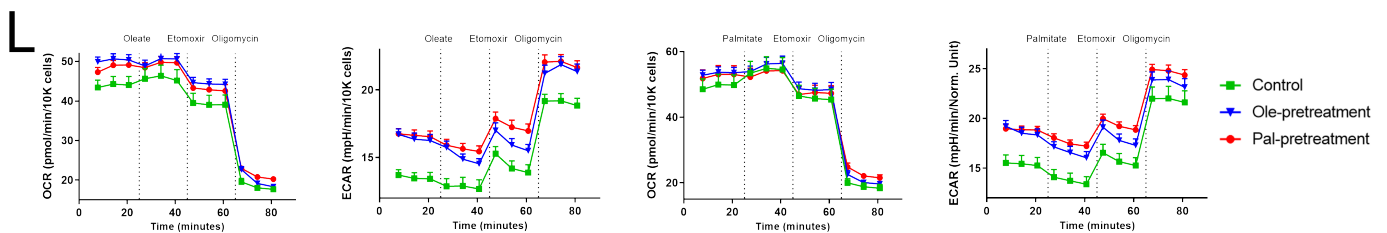
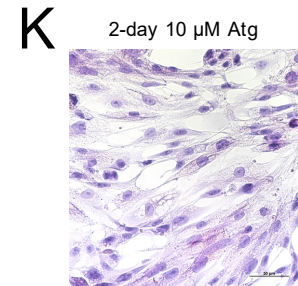
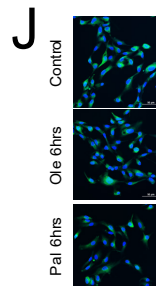
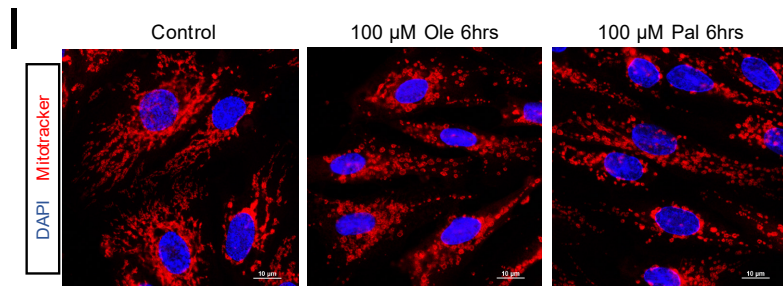
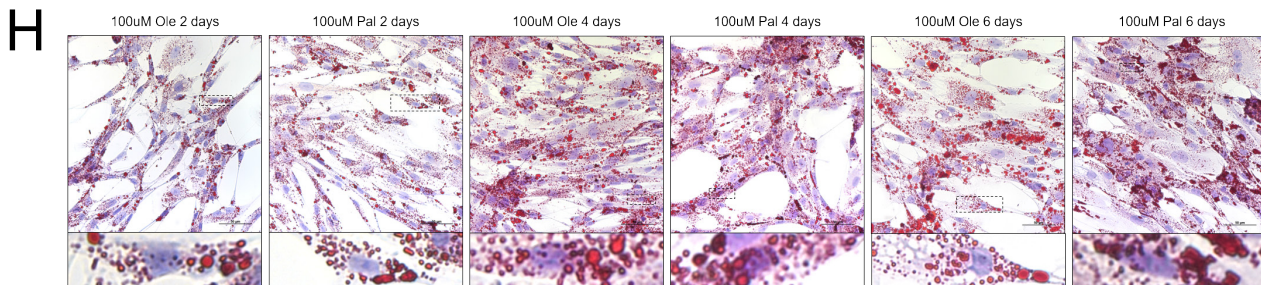
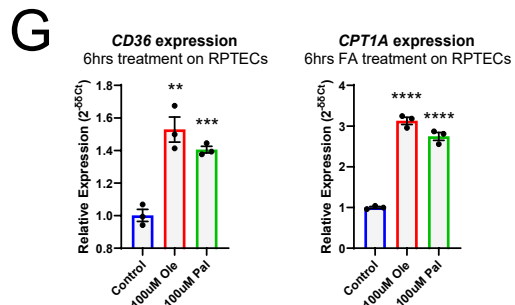
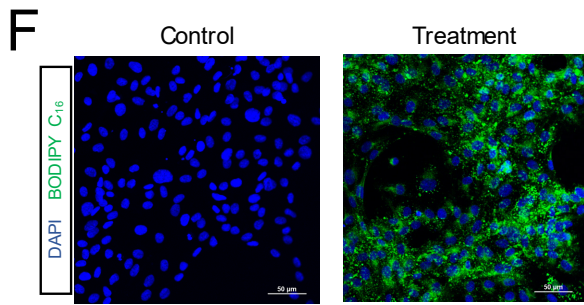
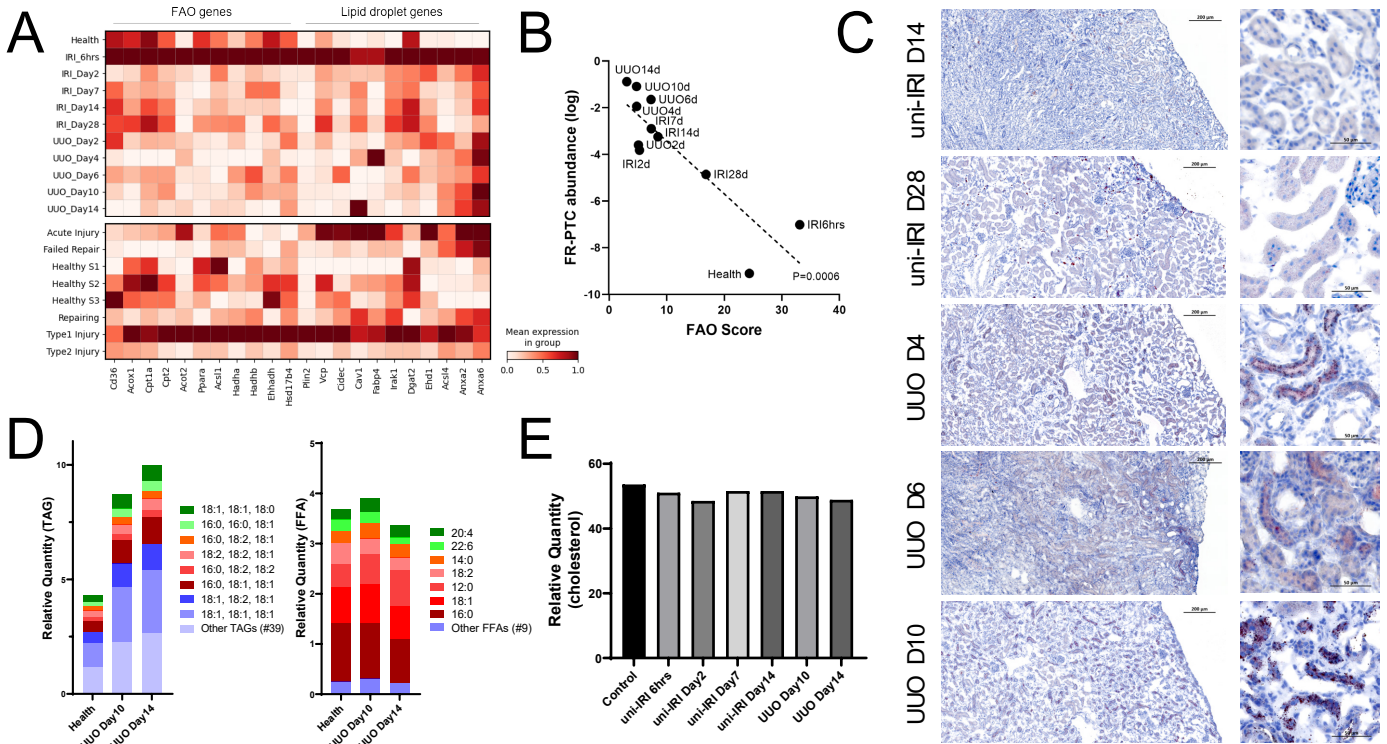


Figure S3. Dysregulated lipid metabolism in proximal tubule cells during fibrogenesis and activated fatty acid oxidation after short-term lipid deposition (related to Figure 3).

- (A) Heat maps displaying expression of genes involved in FAO and lipid droplets in all group conditions (upper panel) and PT subtypes (lower panel)
- (B) Proportion of FR-PTC of each group (log-transformed) is negatively correlated with FAO score. Linear regression is used after data transformation.
- (C) Oil Red O staining on multiple disease conditions, related to Figure 3D. Red color indicates lipids and blue indicates cell nucleus. Scale bar: 200 μm for large-area scanning and 50 μm for the right panel.
- (D) Relative quantity of triglyceride (TAG) and free fatty acid (FFA) species in mouse kidney tissues of healthy, UUO D10 and UUO D14, related to Figure 3E-F.
- (E) Relative quantity of cholesterol in mouse kidney tissues measured by mass spectrometry.
- (F) Internalization assay counterstained with DAPI showing RPTECs could internalize BODIPY C16 in 6-hour treatment.
- (G) Relative expression of CD36 and CPT1A measured by qPCR. Data are shown as the mean \pm SEM. Comparisons were made between each treatment group and control. $**p < 0.01$, $***p < 0.001$ and $****p < 0.0001$ by Student's t test.
- (H) Oil Red O staining on RPTECs treated with oleate or palmitate fatty acids for 2/4/6 days. Scale bars: 50 μm . Zoom-in figures of a single cell are also presented.
- (I) Mitochondria staining on RPTECs, counterstained with DAPI, exposed to fatty acids for 6 hours showing increased mitochondrial fragmentation after treatment. Scale bars: 10 μm .
- (J) Staining of reactive oxygen species (green) and DAPI showing no obvious changes after 6-hour fatty acid exposure. Scale bars: 50 μm .
- (K) Oil Red O staining on RPTECs treated with Atglistatin for 2 days. Scale bars: 50 μm .
- (L) OCR and ECAR measurement on RPTECs with or without 6-hour oleate or palmitate fatty acid pretreatment, related to Figure 3I. Left two panels: oleic acid is injected during real-time measurement. Right two panels: palmitic acid is injected during real-time measurement. Data are shown as the mean \pm SEM. Drug concentrations are described in Methods.

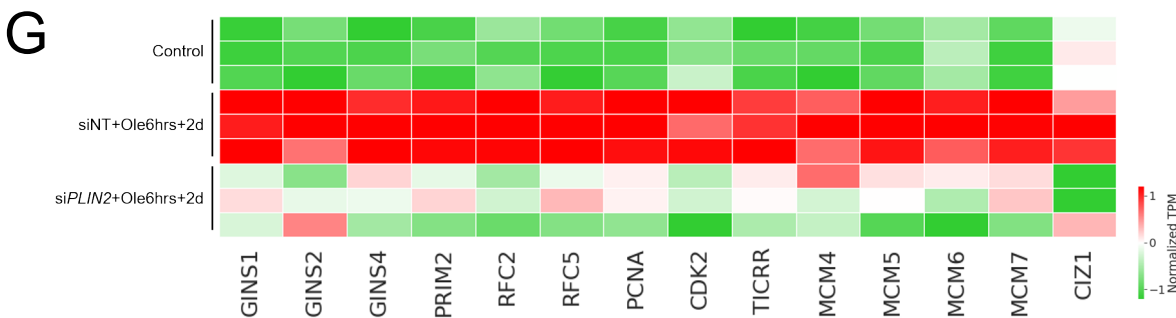
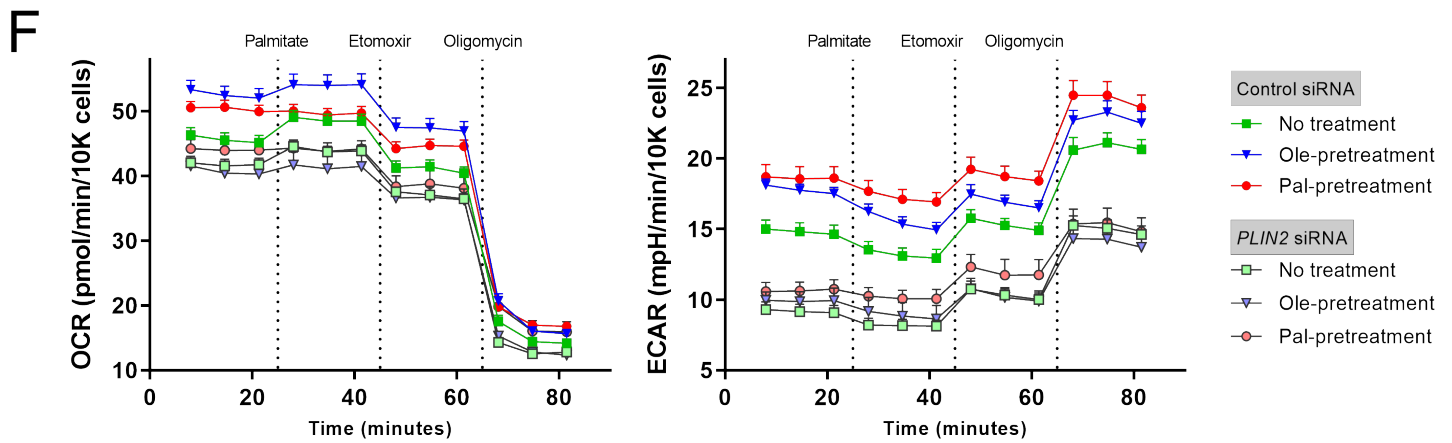
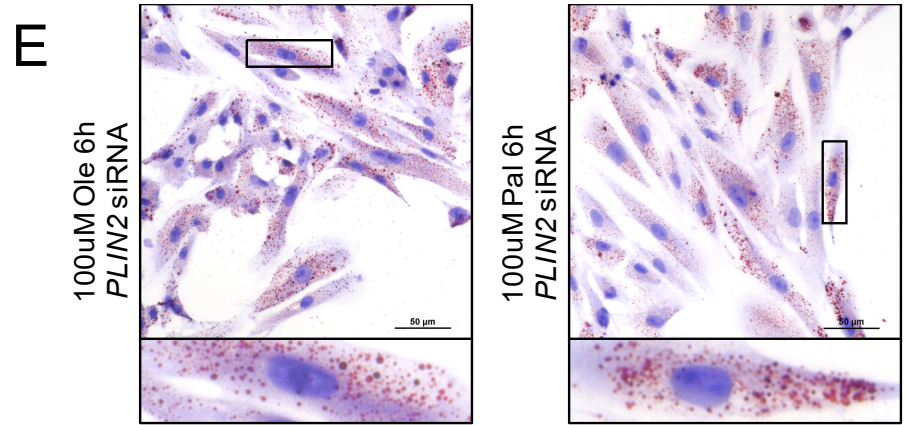
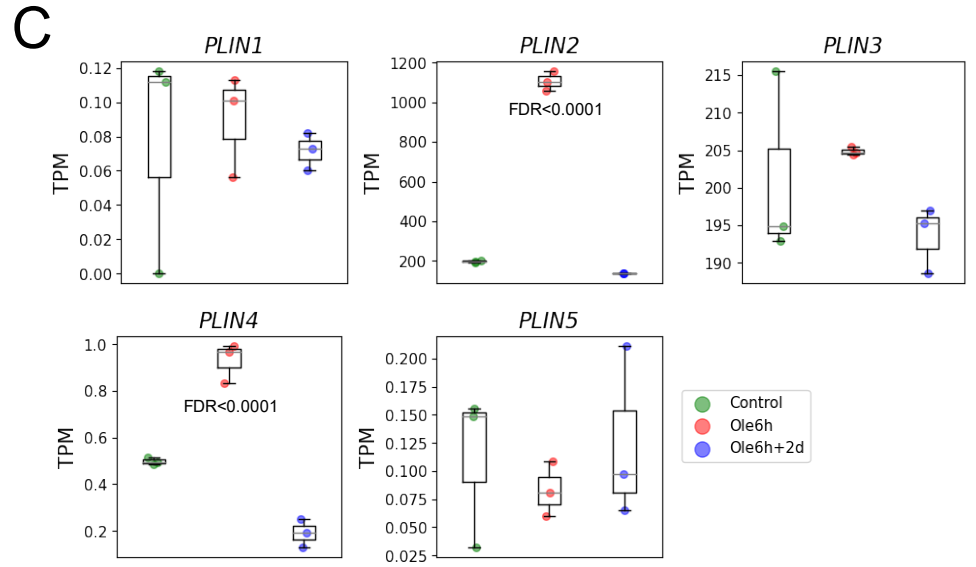
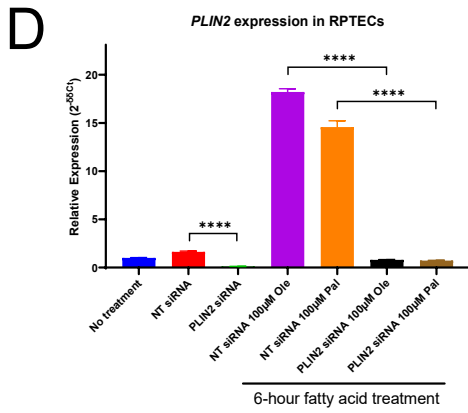
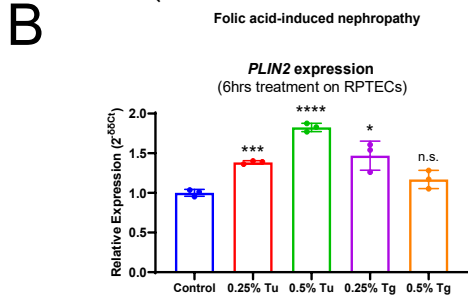
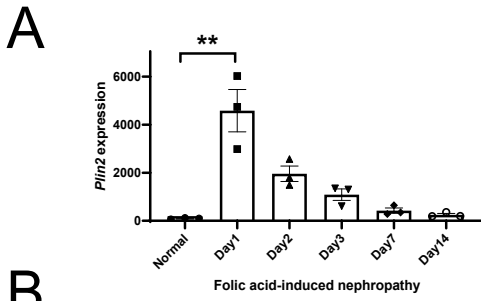


Figure S4. PLIN2 marks lipid droplets in Type1 injured proximal tubule and maintains cell energy state (related to Figure 4).

- (A) Bar plots showing normalized expression of Plin2 in a previous RNA-seq dataset on folic acid-induced mouse nephropathy at multiple time points.
- (B) Relative expression of PLIN2 in RPTECs after treatment ER stress inducers, tunicamycin (Tu) or thapsigargin (Tg), measured by qPCR. Data are shown as the mean \pm SEM. Comparisons were made between each treatment group and control. See Methods for drug stock concentrations. * $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$ and n.s. (not significant) by Student's t test.
- (C) Expression of genes, shown in TPM, encoding PLIN family proteins on control RPTECs and cells treated with 6-hour oleic acids (Ole6h) and harvested at 2 days after the 6-hour treatment (Ole6h+2d), measured by RNA-seq.
- (D) Relative expression of PLIN2 in RPTECs treated with control siRNA (NT siRNA) or PLIN2 siRNA, with or without 6-hour fatty acid stimulus measured by qPCR. Data are shown as the mean \pm SEM. **** $p < 0.0001$ by Student's t test.
- (E) Oil Red O staining on RPTECs treated with PLIN2 siRNA and 6-hour oleate/palmitate fatty acids. Scale bars: 50 μm . Zoom-in figures of a single cell are also presented.
- (F) OCR (left panel) and ECAR (right panel) measurement on RPTECs treated with control siRNA or PLIN2 siRNA, with or without 6-hour oleate or palmitate fatty acid pretreatment, related to Figure 4I. Data are shown as the mean \pm SEM. Drug concentrations are described in Methods.
- (G) Heat map showing expression of genes involved in cell proliferation in RPTECs treated with control siRNA (control), cells harvested at 2 days after control siRNA and 6-hour 100 μM oleic acid treatment (siNT+Ole6hrs+2d) and cells harvested at 2 days after PLIN2 siRNA and 6-hour 100 μM oleic acid treatment (siPLIN2+Ole6hrs+2d). Each group has three biological replicates. TPM expression is normalized for visualization.

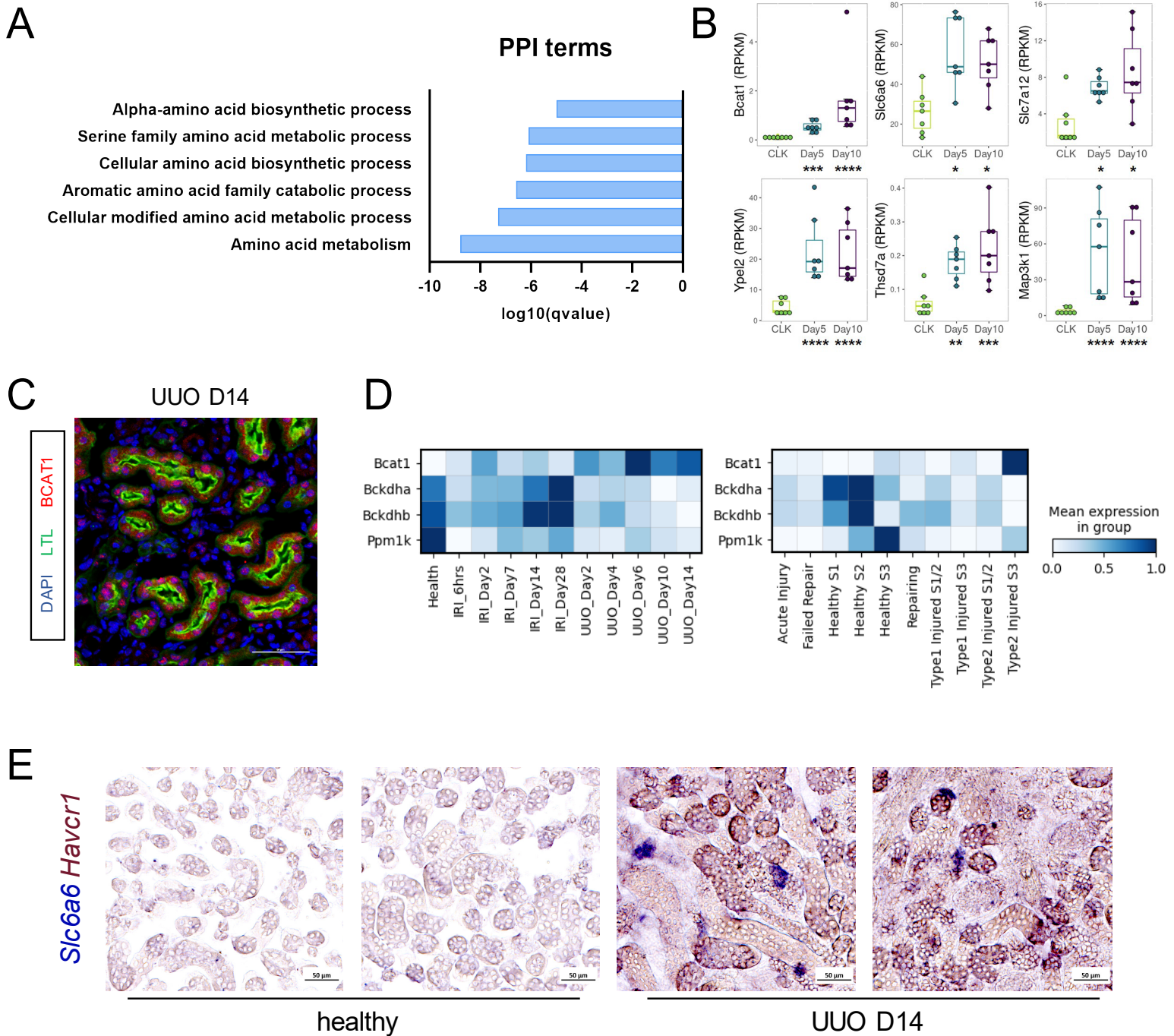


Figure S5. Dysregulation of genes involved in amino acid metabolisms in Type2 injured PT cells (related to Figure 5).

- (A) Terms related to amino acid metabolisms are downregulated in Type2 injured PT compared to healthy PT in PPI GO analysis.
- (B) Bar plots showing significant upregulation of DEGs of Type2 injured PT by surveying a previous bulk RNA-seq study on PT-enrichment transcripts collected from UO mice. CLK, contralateral kidney.
- (C) Immunostaining of BCAT1 (red), LTL (green) and DAPI (blue) on the UO D14 tissue section. Outer medulla regions are presented. Scale bars: 50 μ m.
- (D) Heat maps displaying expression of genes involved in BCAA and BCKA metabolisms.
- (E) Double-probe RNA ISH staining of Havcr1 and Slc6a6 on healthy and UO D14 tissue sections showing Slc6a6 is expressed in Havcr1-expressing injured PT.

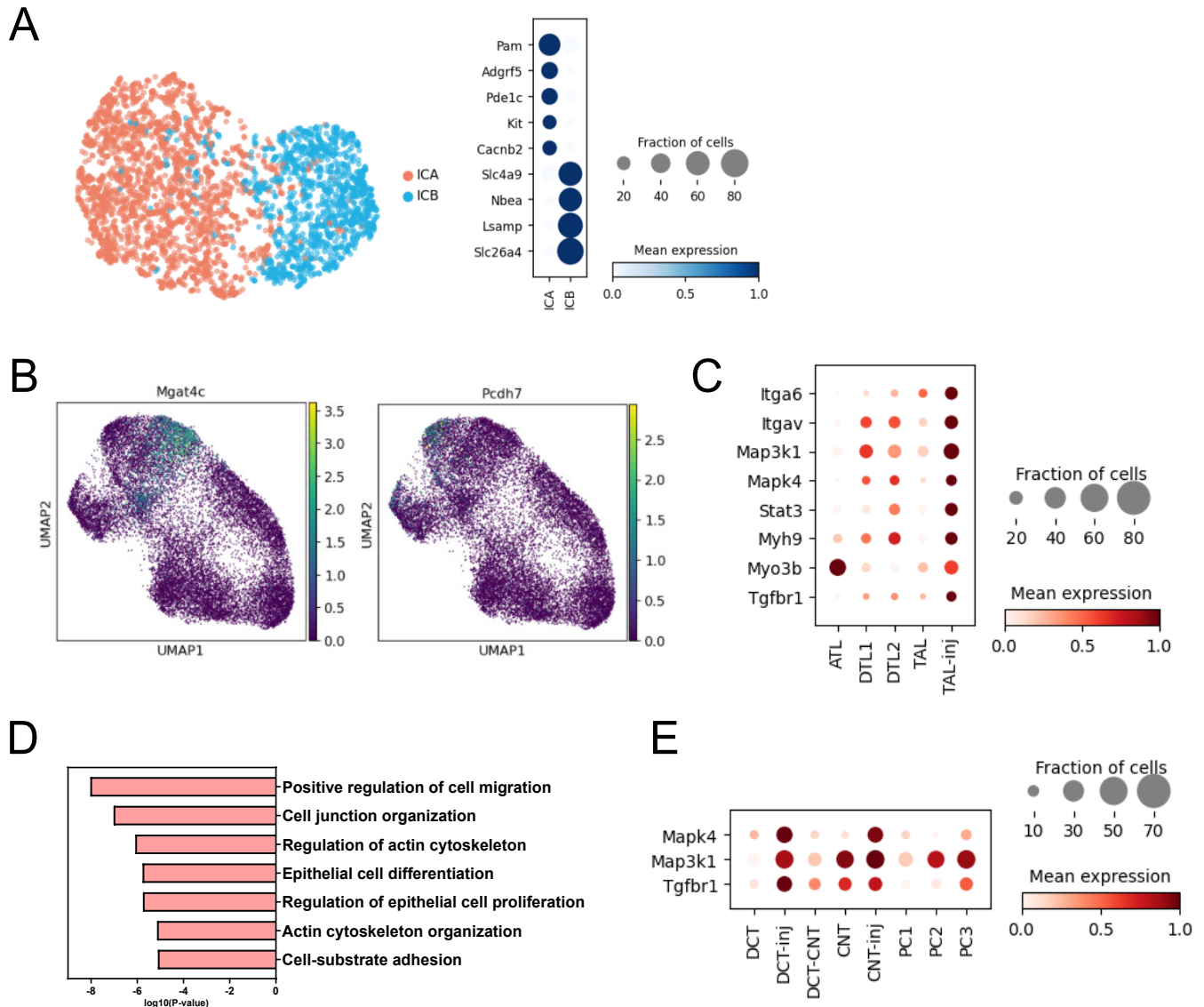


Figure S6. Shared and unique injury responses of renal tubular epithelial cells (related to Figure 6).

- (A) Subclustering results of intercalated cells of collecting duct including a UMAP plot (left) and a dot plot of expression of marker genes of type A cells (ICA) and type B cells (ICB) (right).
- (B) Expression of two regional markers of principal cells of collecting duct, *Mgat4c* (cortical) and *Pcdh7* (medullary).
- (C) Dot plot displaying expression pattern of genes involved in fibrotic responses in LoH cell subtypes.
- (D) Top significant terms of gene ontology enrichment analysis on DEGs of injured TAL (TAL-inj).
- (E) Dot plot showing expression pattern of genes involved in fibrotic responses in distal nephron cell subtypes.

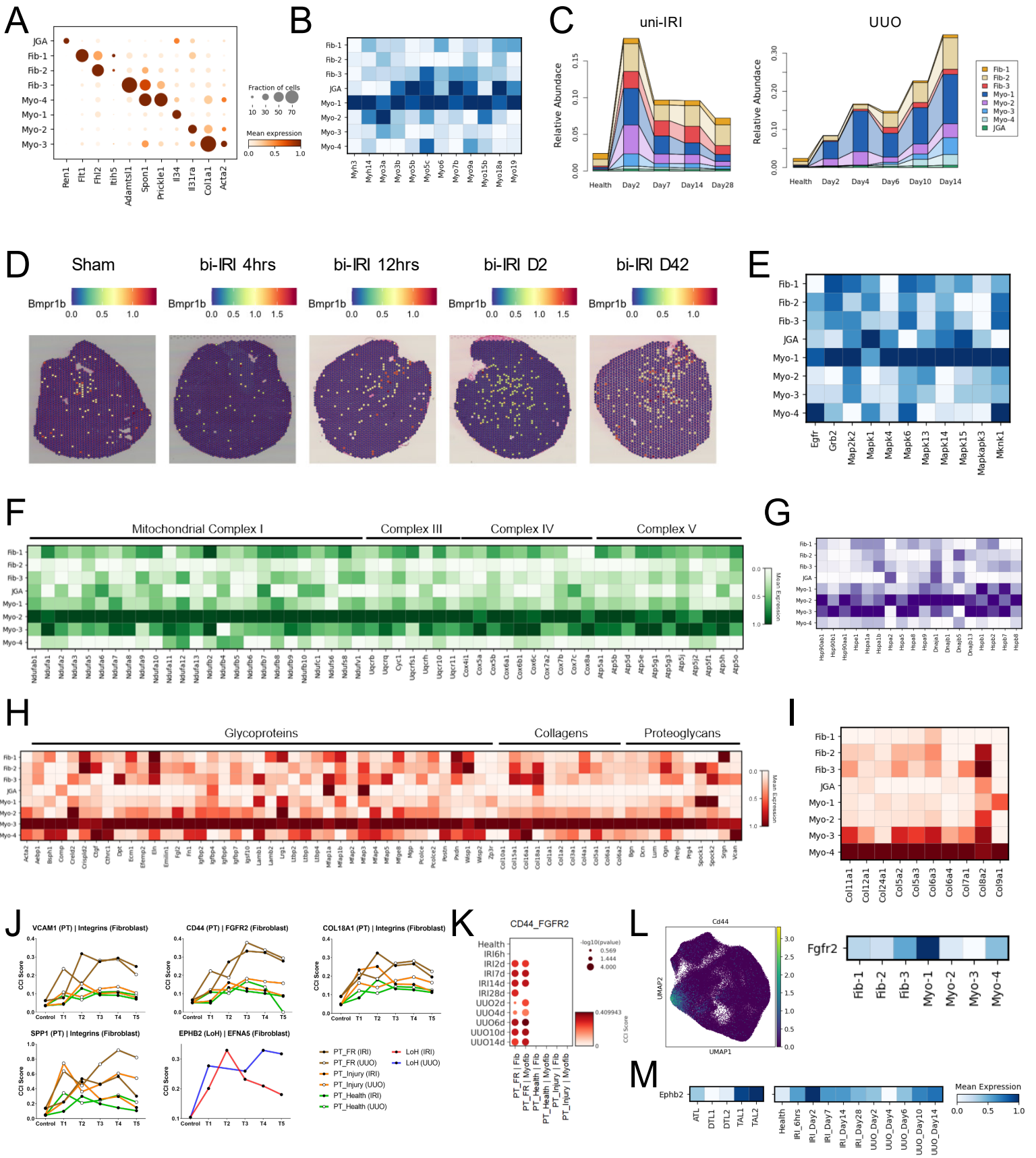


Figure S7. Heterogeneity of kidney stromal cells and cell-cell communications in kidney fibrogenesis (related to Figure 7).

- (A) Dot plot showing expression pattern of cluster-specific genes of stromal cells, related to Figure 7A.
- (B) Heat map showing expression of essential myosin genes in stromal cell types.
- (C) Connected bar plots displaying the abundance of each stromal cell type in each disease condition. Relative abundance is defined as the proportion of the number of cells to the total cell number.
- (D) Expression of *Bmpr1b*, a medullary fibroblast marker, in a spatial transcriptomics dataset on female bi-IRI kidneys. Each spot of a tissue section is colored by gene expression.
- (E) Heat map showing expression of genes involved in ERK/MAPK signaling in stromal cell types.
- (F) Heat map showing expression of genes that encode subunits of mitochondrial respiratory chain complexes enriched in Myo-2 cells.
- (G) Heat map showing expression of genes that encode major HSPs in stromal cell types.
- (H) Heat map showing expression of ECM components enriched in Myo-3 cells.
- (I) Heat map showing expression of genes that encode rare collagens in stromal cell types.
- (J) Activity of featured CCI pairs across the time courses of uni-IRI and UUO. Classification of PT subtypes is same as in Figure 7H.
- (K) Dot plot displaying dynamics of CD44-FGFR2 interaction activity in communications between PT and fibroblasts.
- (L) Expression of *Cd44* in PT single-cell clustering and heat map showing expression of *Fgfr2* in stromal cell types.
- (M) Heat maps showing expression of *Ephb2* in cell types of LoH (left panel) and across the time courses of uni-IRI and UUO in LoH (right panel). The color legend is shared with other subplots of Figure S7.

Table S1. sci-RNA-seq3 protocol optimization and dataset overview (related to Figure 1).

Protocol Optimization		
Steps	Original procedures (Cao et al. 2019)	Modified procedures of this work
Tissue homogenization	With the rubber tip of a syringe plunger	With a dounce homogenizer
Lysis buffer	Homemade buffer supplemented with RNase inhibitor	Nuclei EZ Lysis Buffer supplemented with both RNase and protease inhibitor
Nuclei filtration	40um cell strainer	First 200um and then 40um cell strainers
Nuclei fixation	4% PFA; 15 minutes	2.4% PFA; 10 minutes
Cell concentration when snap-frozen	Not mentioned	~5 million/mL
Cell permeabilization	100ul material; 3 minutes	300ul material; 5 minutes
Sonication	Bioruptor® Plus sonication device: low power mode 12s	Bioruptor® Pico sonication device: regular mode 10s
Tagmentation	1:400 diluted N7-Tn5	We recommend researchers first perform a Tn5 titration test on a sublibrary plate to determine the best Tn5 dilution ratio.
Size-select purification	Column-based purification	Bead-based purification (higher efficiency and yield)
Dataset Overview		
Samples multiplexed in the library	24 mouse kidney samples and 1 human/mouse species mixing sample; the sample-baracode lookup table is described in https://github.com/TheHumphreysLab/sci-RNA-seq-kidney	
Number of cells (UMI number>200)	413681	
Number of cells (post QC)	309666	
Number of cells (gene number>200)	260028	
Mean number of UMIs/cell	1165	
Mean number of genes/cell	545	
Fraction of intronic reads	60.4%	
Ratio of reads mapped to protein-coding genes	81%	
Mean mitochondrial read fraction	1.33%	

Table S2. Summary of all cell subtypes and cell states identified (related to Figure 1).

Cell type/state	Brief description	Figure in manuscript
Proximal Tubule		
Healthy S1	S1 segment of PT (Slc34a1/Slc5a12 high)	Figure 2
Healthy S2	S2 segment of PT (Slc34a1/Slc22a30 high, Slc5a12 low)	
Healthy S3	S2 segment of PT (Slc22a30/Slc7a13 high)	
PT Acute injury	High heat shock protein expression; low healthy marker expression	
PT Repairing	High expression of proliferation markers; low healthy marker expression	
PT Type1 injured S1/2	uni-IRI specific; expressing Plin2/Col27a1; low healthy marker expression	
PT Type1 injured S3		
PT Type2 injured S1/2	UUO specific; expressing Beat1/Slc6a6; low healthy marker expression	
PT Type2 injured S3		
PT Failed repair (FR-PTC)	Vcam1-expressing; Highly inflammatory and fibrotic; low healthy marker expression	
Loop of Henle		
TAL	TAL in health (expressing Slc12a1/Umod). It could be further stratified by cortical/medullary markers.	Figure 6
TAL-inj	TAL in injury (Slc12a1/Umod low; Kctd1/Lcn2 high)	
ATL	Sgc2/Pax2 high thin ascending limb	
DTL1	Slc4a11/Cdh13 high descending limb	
DTL2	Slc14a2-expressing descending limb	
DCT/CNT/PC/IC		
DCT	DCT in health (Slc12a3/Egf high)	Figure 6
DCT-inj	DCT in injury (Slc12a3/Egf low; Psd3 high)	
DCT-CNT	Slc12a3/Slc8a1 high	
CNT	CNT in health (Slc8a1/Tbck high)	
CNT-inj	CNT in injury (Slc8a1/Tbck low; Trpv5 high)	
PC1	Lcn2 is mainly expressed by PC1.	
PC2	Medullary PC	
PC3	Cortical PC	
ICA	Pam/Kit high	
ICB	Slc4a9/Slc26a4 high	
Fibroblast		
Fib-1	Cortical fibroblasts (Bmpr1b/Itih5 high); dominant fibroblast population in health	Figure 7
Fib-2		
Fib-3		
Myo-1		
Myo-2		
Myo-3		
Myo-4		
Immune		
Mφ-1	Stab1/Apoe high macrophages	Figure S1C
Mφ-2	Macrophages expression Tgfb1, Vcan and Lyz2	
Mφ-3	Macrophages expression proliferation markers	
B cell	Pax5/Ebf1-expressing	
T cell	Skap1/Cd247-expressing	
cDC1	Clec9a/Ppt1-expressing	
cDC2	Dendritic cells expressing Itgam and Sirpa	
pDC	Siglech/Runx2-expressing	
Activated DC	Relb/Fscn1 high	
Endothelia		
Artery	See Figure S1D	
Vein		
Cap.1		
Cap.2		
Cap.Art		
Activated EC		
Others		
Podocyte	Podxl/Magi2-expressing	Figure 1D
JGA	Ren1-expressing	Figure 7A
Urothelium	Upk1b/Krt19-expressing	Figure 1D