Supplemental Information

Unified mouse and human kidney single cell atlas reveal commonalities and differences in disease states

Jianfu Zhou, Amin Abedini, Michael Balzer, Rojesh Shrestra, Poonam Dhillon, Hongbo Liu, Hailong Hu, and Katalin Susztak

Renal, Electrolyte, and Hypertension Division, Department of Medicine, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104, USA

Institute for Diabetes, Obesity, and Metabolism, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104, USA

Department of Genetics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104, USA

Kidney Innovation Center, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104, USA

Correspondence:

Katalin Susztak, MD, PhD

University of Pennsylvania, School of Medicine

3400 Civic Center Blvd,

Smilow Translational building 12-123,

Philadelphia, PA, 19104

Phone: (215)898-2009

ksusztak@pennmedicine.upenn.edu

Supplemental Material List

Supplemental Figure 1. Quality control of the mouse scRNA-seq data.
Supplemental Figure 2. Integration of the mouse scRNA-seq samples.
Supplemental Figure 3. Cell type correlation.
Supplemental Figure 4. Cell type-specific DEG conservation.
Supplemental Figure 5. Loadings matrices produced by tensor decomposition.
Supplemental Figure 6. Integration of PT cells from the mouse scRNA-seq data.
Supplemental Figure 7. Monocle 2 trajectory analysis of PT cells from the mouse scRNA-seq data.
Supplemental Figure 8. RNA velocity analysis of PT cells from the mouse scRNA-seq data.
Supplemental Figure 9. Conservation of genes regulated along Monocle 2 trajectories of PT cells from the mouse scRNA-seq data.
Supplemental Figure 10. Gene set enrichment analysis of Monocle 2 trajectories of PT cells from the mouse scRNA-seq data.
Supplemental Figure 11. WGCNA of PT cells from the mouse scRNA-seq data.
Supplemental Figure 12. Quality control of the human and mouse snRNA-seq data.
Supplemental Figure 13. Integration of the human snRNA-seq samples.
Supplemental Figure 14. Integration of the mouse snRNA-seq samples.
Supplemental Figure 15. Integration of the human and mouse snRNA-seq samples.
Supplemental Figure 16. Human-mouse cell type-specific DEG conservation.
Supplemental Figure 17. Integration of PT nuclei from the human and mouse snRNA-seq data.
Supplemental Figure 18. Gene set enrichment analysis of Monocle 2 trajectory of PT nuclei from the human and mouse snRNA-seq data.
Supplemental Figure 19. WGCNA of PT nuclei from the human and mouse snRNA-seq data.
Supplemental Figure 20. Gene set enrichment analysis of WGCNA gene modules of PT nuclei from the human and mouse snRNA-seq data.
Supplemental Dataset 1: Information of mouse and human samples.
Supplemental Dataset 2: DEGs of each cell type against all the other cell types in the mouse kidney scRNA-seq data.
Supplemental Dataset 3: DEGs of each disease model against the control in each identified cell type of the mouse kidney scRNA-seq data.
Supplemental Dataset 4: DEGs of each disease model against the control in the mouse kidney bulk RNA-seq data.
Supplemental Dataset 5: Cell fraction-adjusted DEGs of each disease model against the control in the mouse kidney bulk RNA-seq data.
Supplemental Dataset 6: DEGs of each PT cell subtype against all the other PT cell subtypes in the mouse kidney scRNA-seq data.

Supplemental Dataset 7: Genes regulated along each Monocle2 trajectory of PT cells in each mouse model of the mouse kidney scRNA-seq data.

Supplemental Dataset 8: Enriched KEGG pathways and GO BP terms along each Monocle2 trajectory of PT cells in each mouse model of the mouse kidney scRNA-seq data.

Supplemental Dataset 9: Gene modules identified by WGCNA of PT cells in the mouse kidney scRNA-seq data.

Supplemental Dataset 10: KEGG pathways and GO BP terms enriched in each identified WGCNA gene module of PT cells in the mouse kidney scRNA-seq data.

Supplemental Dataset 11: DEGs of each cell type against all the other cell types in the human DKD snRNA-seq data.

Supplemental Dataset 12: DEGs of each cell type against all the other cell types in the mouse DKD snRNA-seq data.

Supplemental Dataset 13: DEGs of DKD samples against healthy samples in each identified cell type of the human DKD snRNA-seq data.

Supplemental Dataset 14: DEGs of DKD samples against control samples in each identified cell type of the mouse DKD snRNA-seq data.

Supplemental Dataset 15: Genes regulated along the Monocle2 trajectory of PT nulcei in the human and mouse DKD snRNA-seq data.

Supplemental Dataset 16: Enriched KEGG pathways and GO BP terms along the Monocle2 trajectory of PT nulcei in the human and mouse DKD snRNA-seq data.

Supplemental Dataset 17: Gene modules identified by WGCNA of PT nuclei in the human DKD snRNA-seq data.

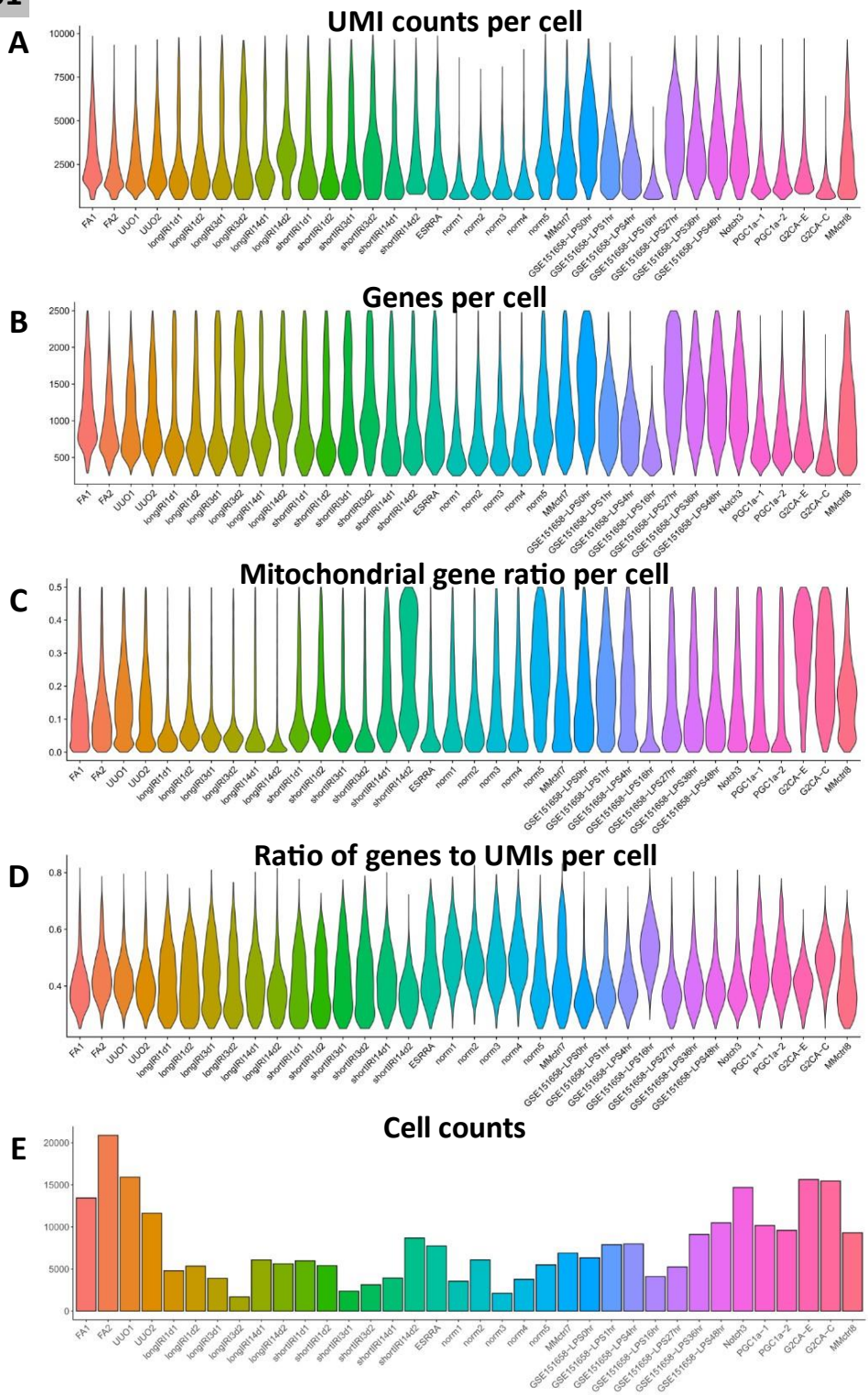
Supplemental Dataset 18: Gene modules identified by WGCNA of PT nuclei in the mouse DKD snRNA-seq data.

Supplemental Dataset 19: KEGG pathways and GO BP terms enriched in each identified WGCNA gene module of PT nuclei in the human DKD snRNA-seq data.

Supplemental Dataset 20: KEGG pathways and GO BP terms enriched in each identified WGCNA gene module of PT nuclei in the mouse DKD snRNA-seq data.

Supplemental Figures

Figure S1



Supplemental Figure 1. Quality control of the mouse scRNA-seq data.

(A) Violin plot showing the number of UMIs per single cell after QC, split by mouse samples.

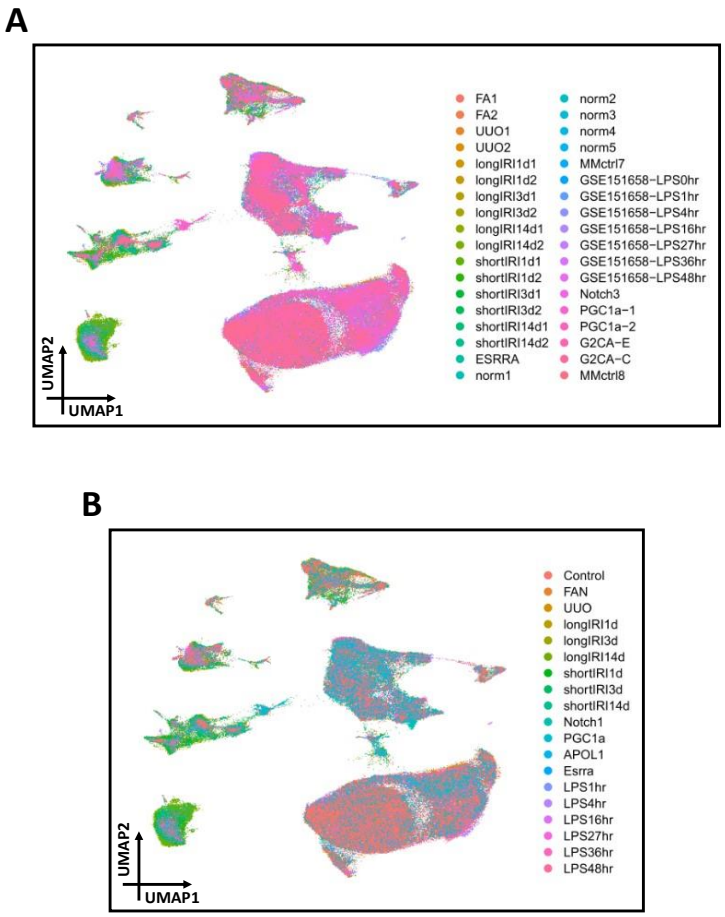
(B) Violin plot showing the number of detected genes per single cell after QC, split by mouse samples.

(C) Violin plot showing the percentage of mitochondrially encoded gene reads per single cell after QC, split by mouse samples.

(D) Violin plot showing the ratio of detected genes to UMIs per single cell after QC, split by mouse samples.

(E) Bar plot showing the cell count of each mouse scRNA-seq sample after QC.

Figure S2

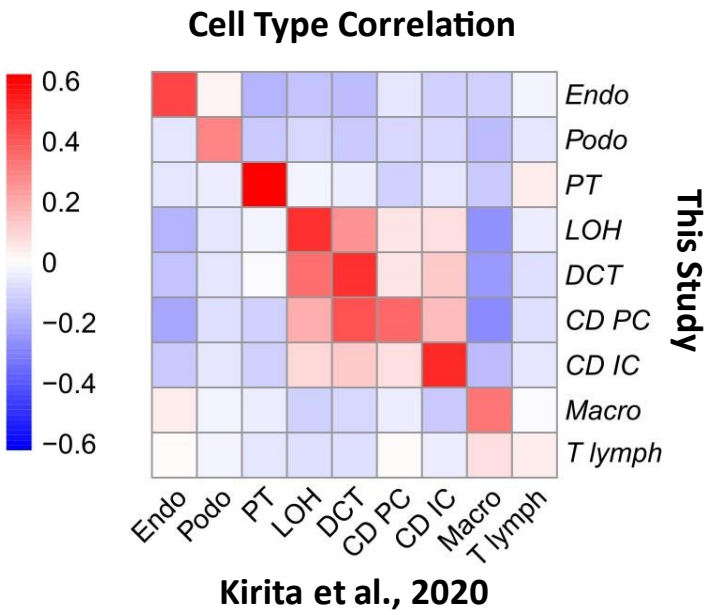


Supplemental Figure 2. Integration of the mouse scRNA-seq samples.

(A) UMAP of 280,521 mouse kidney single cells, colored by samples.

(B) UMAP of 280,521 mouse kidney single cells, colored by mouse models.

Figure S3

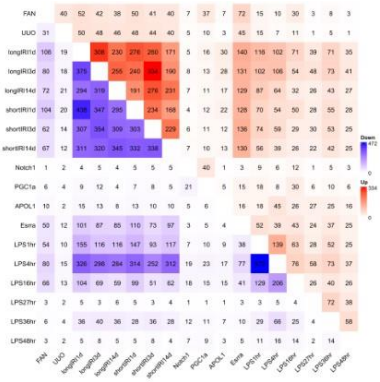


Supplemental Figure 3. Cell type correlation.

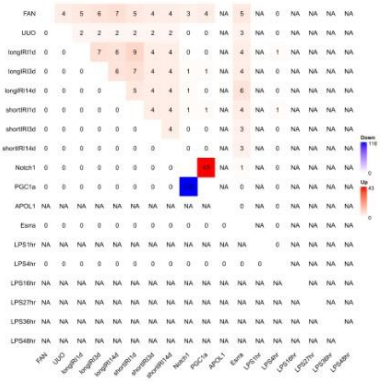
Heatmap shows Pearson's correlation coefficients of averaged cell type gene expression between mouse kidney scRNA-seq atlas generated in this study and a published mouse snRNA-seq dataset with IRI and sham kidney samples[1].

Figure S4

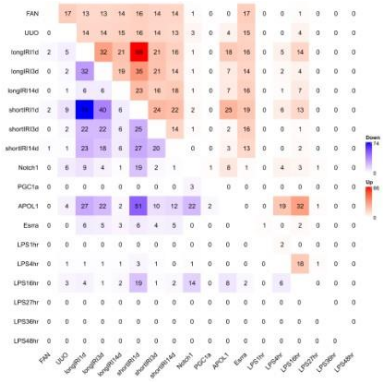
Endo



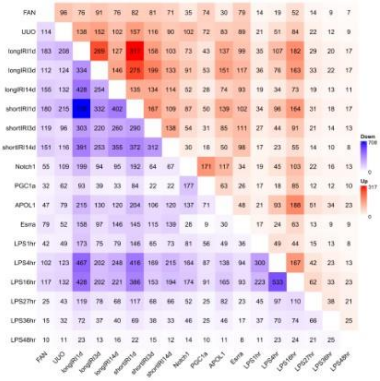
Podo



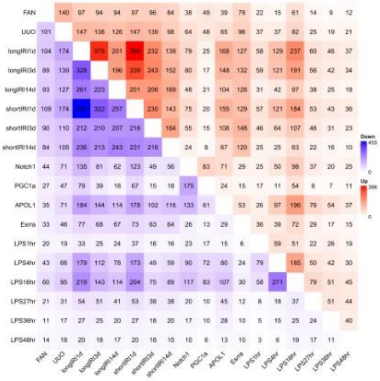
DLOH



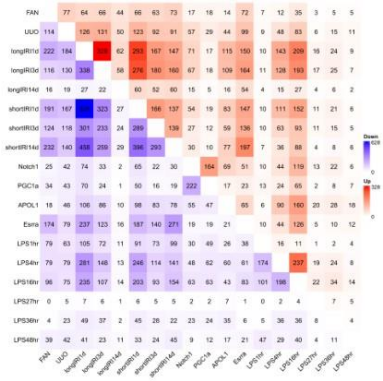
ALOH



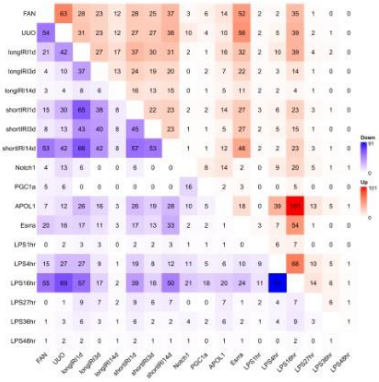
DCT



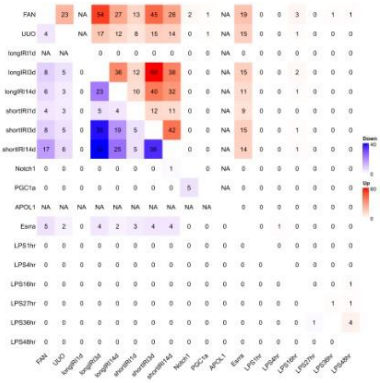
CD PC



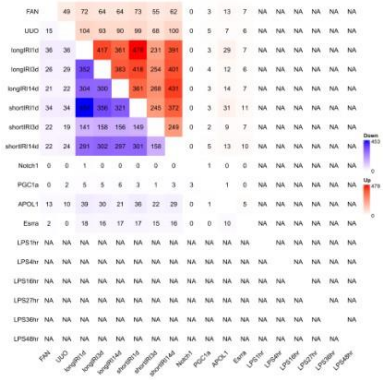
CD IC

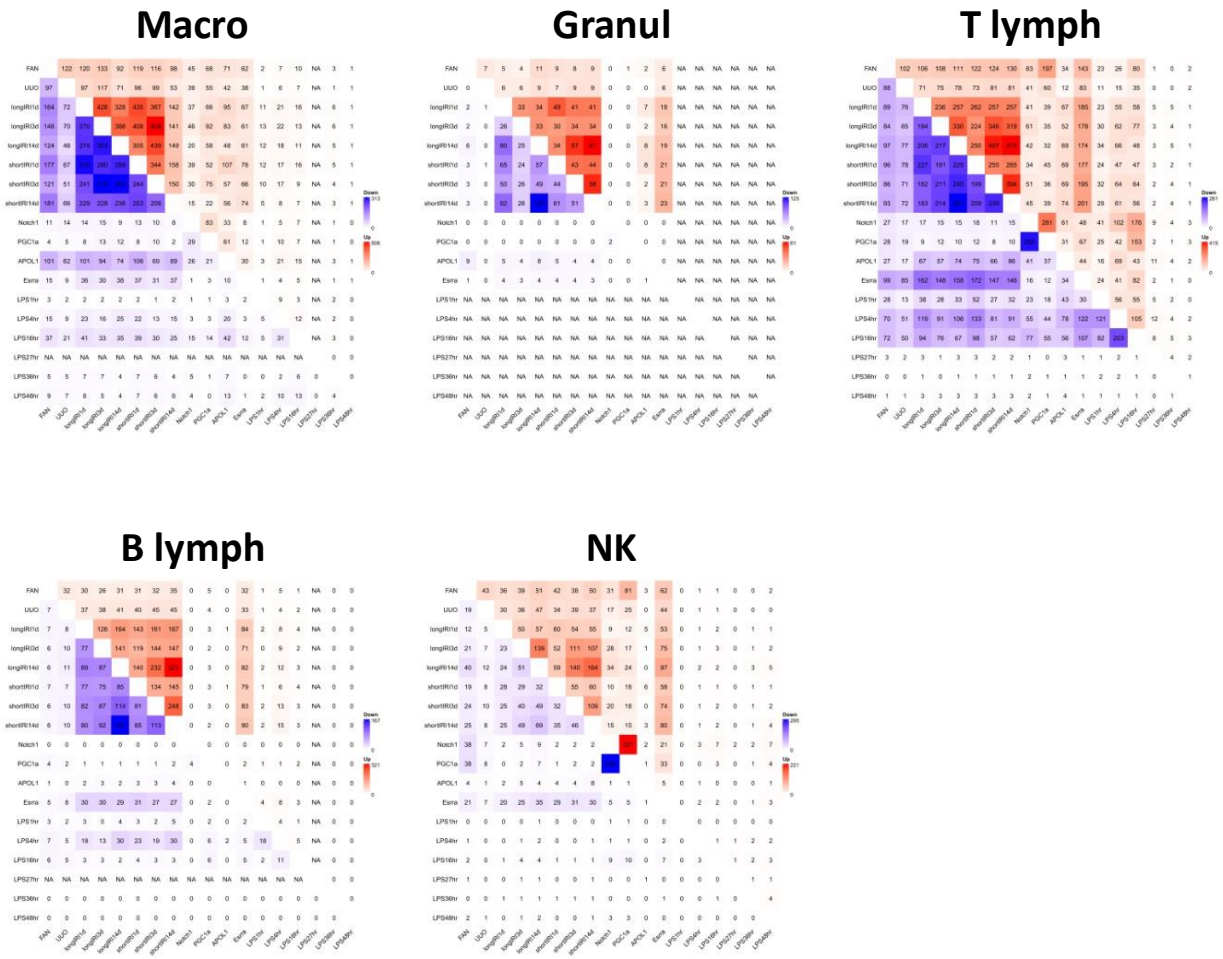


Proliferating



Mono

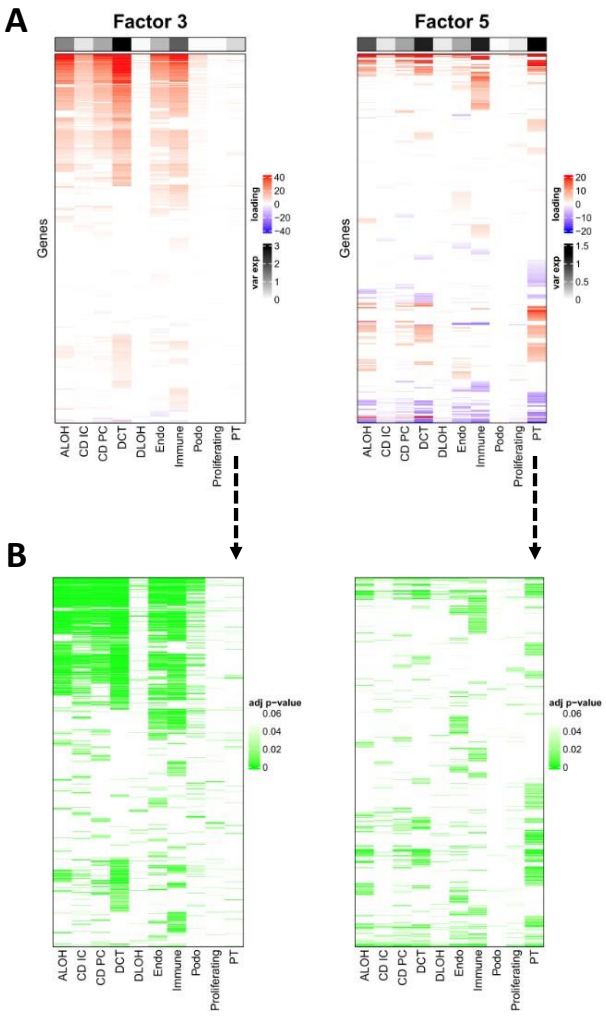




Supplemental Figure 4. Cell type-specific DEG conservation.

Heatmaps show the numbers of up- (upper triangle) and down-regulated (lower triangle) cell type-specific DEGs (against the control) conserved between any two studied mouse kidney disease models in identified cell types (one heatmap per cell type). NA (i.e., not applicable) means not enough cells in either disease or control groups for DEG identification.

Figure S5

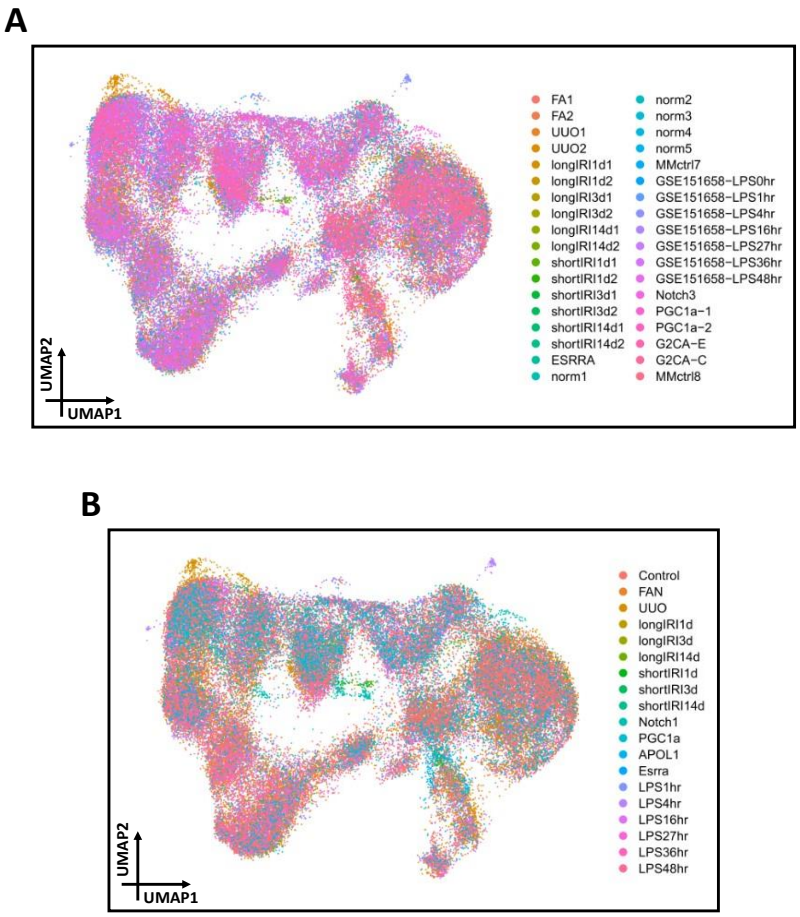


Supplemental Figure 5. Loadings matrices produced by tensor decomposition.

(A) Loadings matrices for factors 3 and 5 limited to significant genes. The top annotation shows the percentage of overall explained variance for each cell type of the factor. Rows are hierarchically clustered.

(B) The same matrices for factors 3 and 5 as those in (A) except that each entry shows the association significance p-value of each gene in each cell type of the factor .

Figure S6

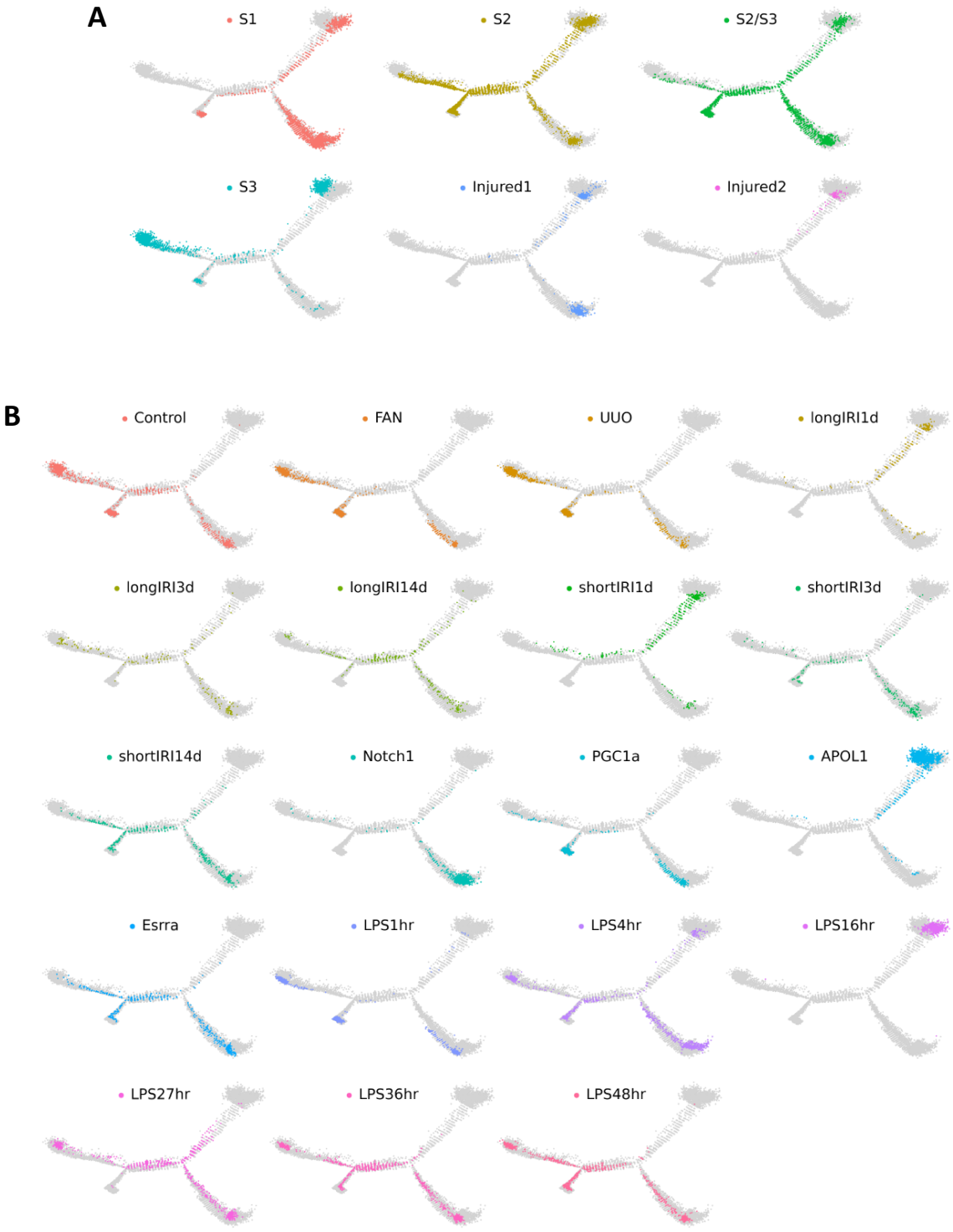


Supplemental Figure 6. Integration of PT cells from the mouse scRNA-seq data.

(A) UMAP of 70,501 PT cells from the mouse scRNA-seq data, colored by samples.

(B) UMAP of 70,501 PT cells from the mouse scRNA-seq data, colored by mouse models.

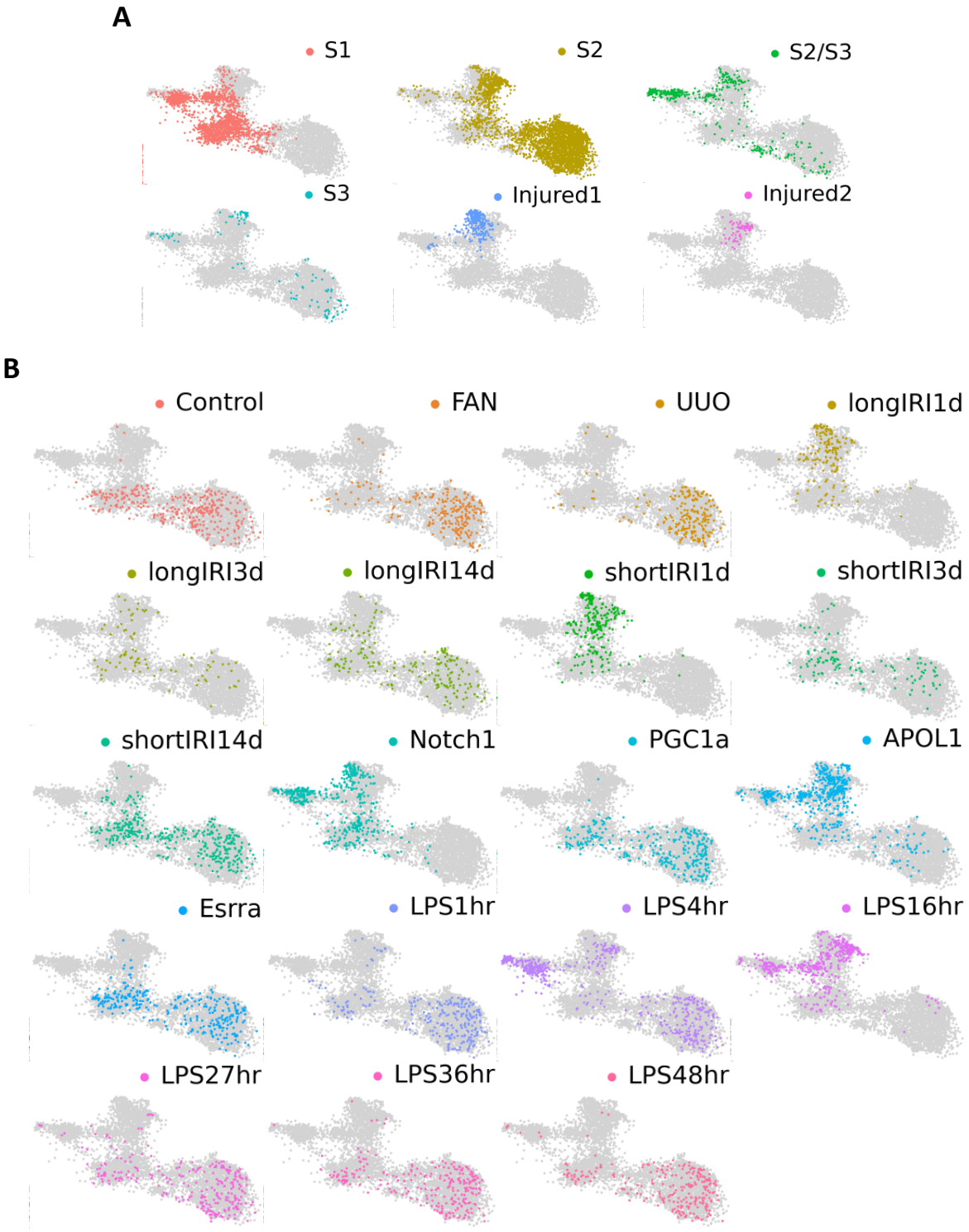
Figure S7



Supplemental Figure 7. Monocle 2 trajectory analysis of PT cells from the mouse scRNA-seq data.

(A) All six panels show the same trajectories. Each panel shows the location of the cells along the trajectories for one PT cell subtype.
(B) All 19 panels show the same trajectories. Each panel shows the location of the cells along the trajectories for one studied mouse model.

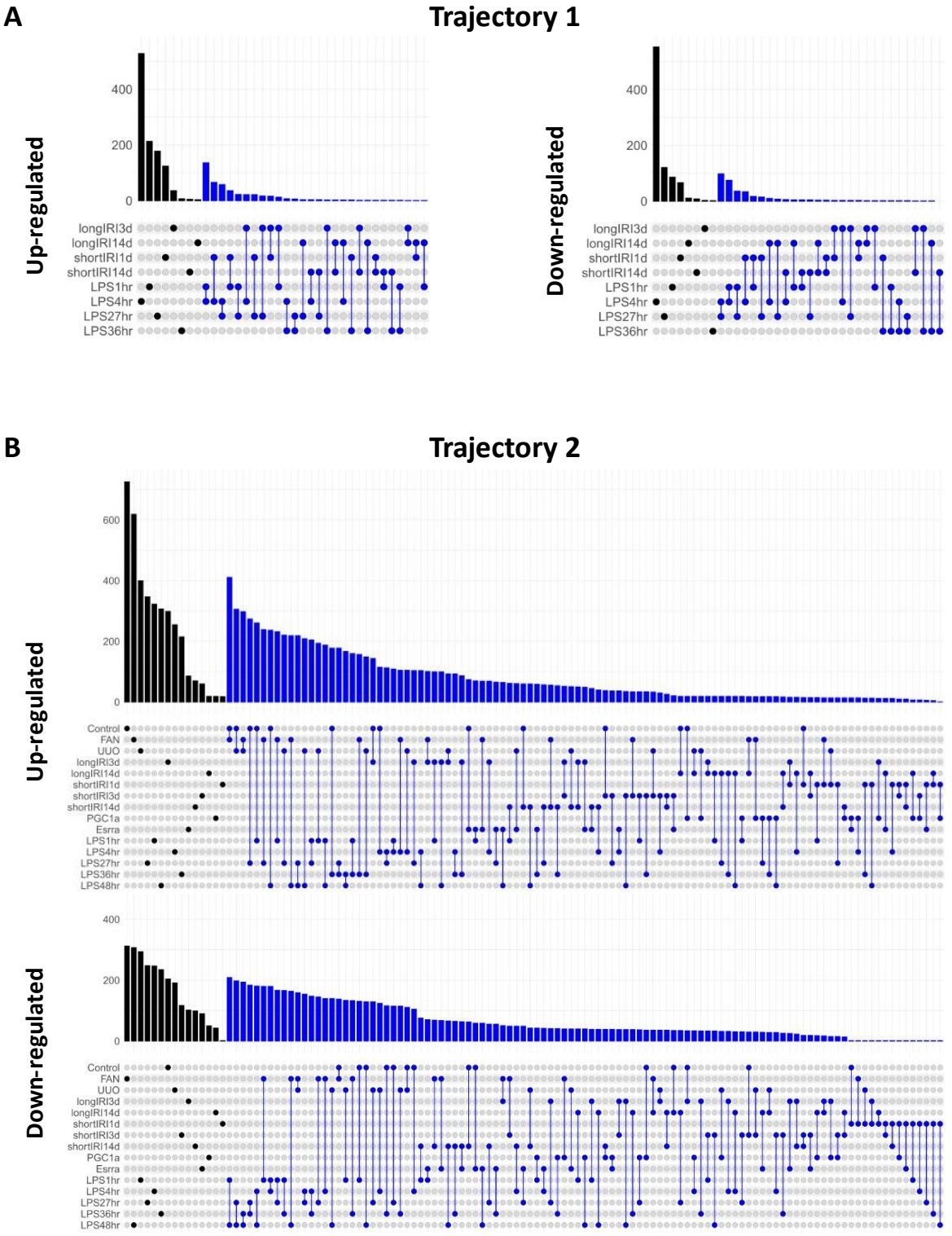
Figure S8



Supplemental Figure 8. RNA velocity analysis of PT cells from the mouse scRNA-seq data.

(A) All six panels show the same RNA velocity UMAP. Each panel shows the location of the cells on the UMAP for one PT cell subtype.
(B) All 19 panels show the same RNA velocity UMAP. Each panel shows the location of the cells on the UMAP for one studied mouse model.

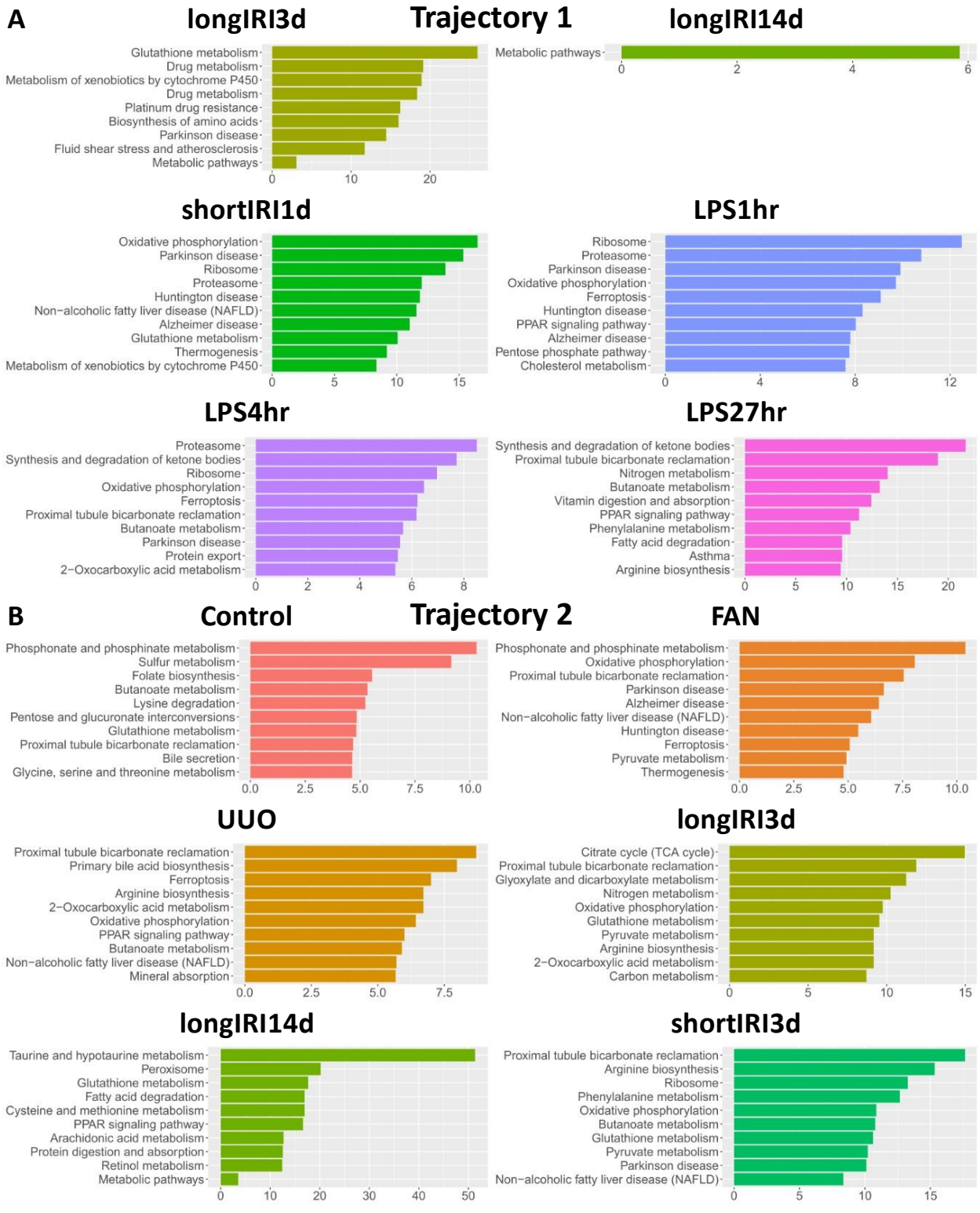
Figure S9

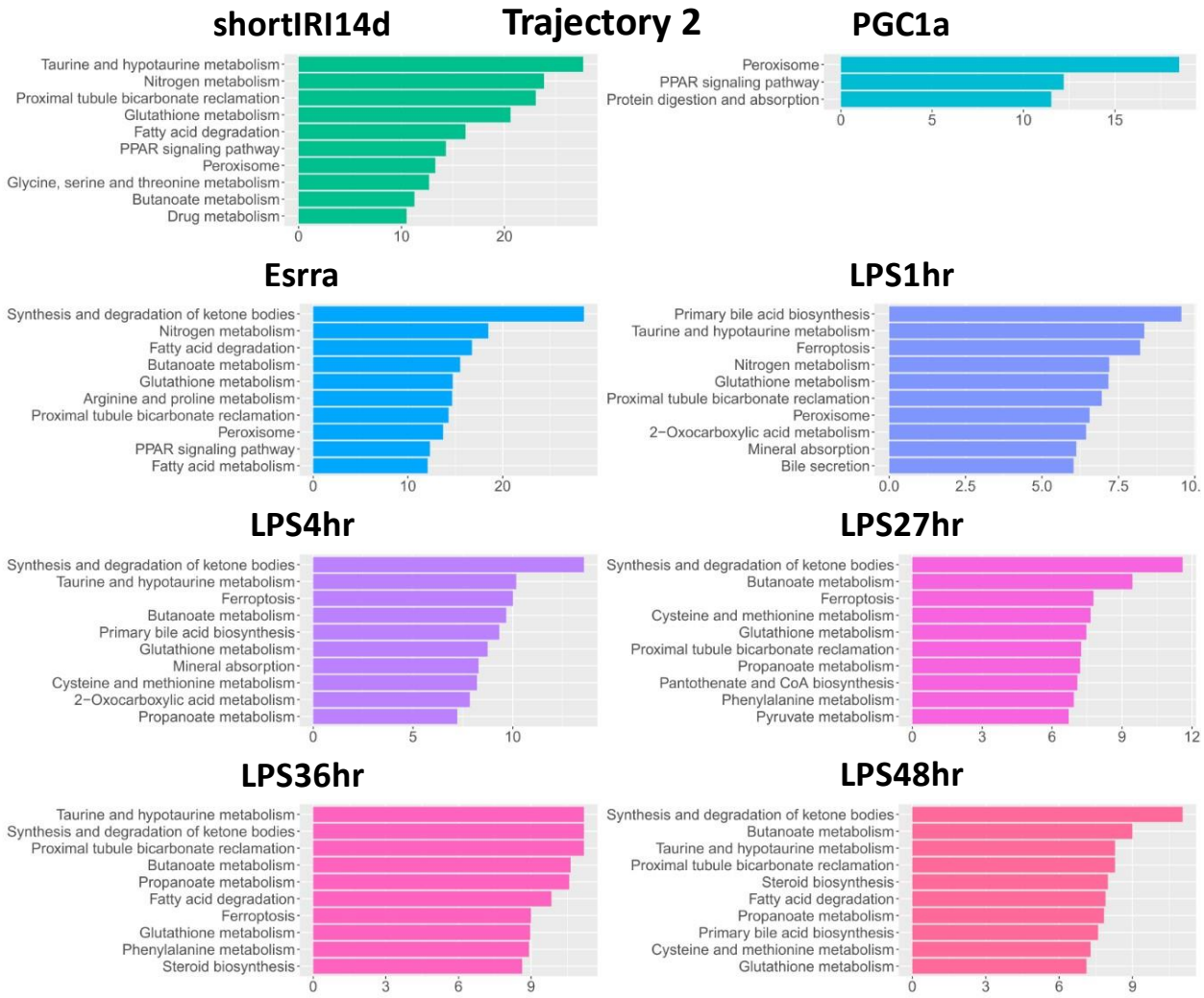


Supplemental Figure 9. Conservation of genes regulated along Monocle 2 trajectories of PT cells from the mouse scRNA-seq data.

(A) Upset plot showing the numbers of up- (left panel) and down-regulated (right panel) genes along Trajectory 1 that were conserved among studied mouse models.
(B) Upset plot showing the numbers of up- (top panel) and down-regulated (bottom panel) genes along Trajectory 2 that were conserved among studied mouse models.

Figure S10


A
longIRI3d
Trajectory 1
longIRI14d
Glutathione metabolism
Drug metabolism
Metabolism of xenobiotics by cytochrome P450
Drug metabolism
Platinum drug resistance
Biosynthesis of amino acids
Parkinson disease
Fluid shear stress and atherosclerosis
Metabolic pathways
0
10
20
Metabolic pathways
0
2
4
6
shortIRI1d
Oxidative phosphorylation
Parkinson disease
Ribosome
Proteasome
Huntington disease
Non−alcoholic fatty liver disease (NAFLD)
Alzheimer disease
Glutathione metabolism
Thermogenesis
Metabolism of xenobiotics by cytochrome P450
0
5
10
15
LPS1hr
Ribosome
Proteasome
Parkinson disease
Oxidative phosphorylation
Ferroptosis
Huntington disease
PPAR signaling pathway
Alzheimer disease
Pentose phosphate pathway
Cholesterol metabolism
0
4
8
12
LPS4hr
Proteasome
Synthesis and degradation of ketone bodies
Ribosome
Oxidative phosphorylation
Ferroptosis
Proximal tubule bicarbonate reclamation
Butanoate metabolism
Parkinson disease
Protein export
2−Oxocarboxylic acid metabolism
0
2
4
6
8
LPS27hr
Synthesis and degradation of ketone bodies
Proximal tubule bicarbonate reclamation
Nitrogen metabolism
Butanoate metabolism
Vitamin digestion and absorption
PPAR signaling pathway
Phenylalanine metabolism
Fatty acid degradation
Asthma
Arginine biosynthesis
0
5
10
15
20
B
Control
Trajectory 2
FAN
Phosphonate and phosphinate metabolism
Sulfur metabolism
Folate biosynthesis
Butanoate metabolism
Lysine degradation
Pentose and glucuronate interconversions
Glutathione metabolism
Proximal tubule bicarbonate reclamation
Bile secretion
Glycine, serine and threonine metabolism
0.0
2.5
5.0
7.5
10.0
Phosphonate and phosphinate metabolism
Oxidative phosphorylation
Proximal tubule bicarbonate reclamation
Parkinson disease
Alzheimer disease
Non−alcoholic fatty liver disease (NAFLD)
Huntington disease
Ferroptosis
Pyruvate metabolism
Thermogenesis
0.0
2.5
5.0
7.5
10.0
UUO
longIRI3d
Proximal tubule bicarbonate reclamation
Primary bile acid biosynthesis
Ferroptosis
Arginine biosynthesis
2−Oxocarboxylic acid metabolism
Oxidative phosphorylation
PPAR signaling pathway
Butanoate metabolism
Non−alcoholic fatty liver disease (NAFLD)
Mineral absorption
0.0
2.5
5.0
7.5
Citrate cycle (TCA cycle)
Proximal tubule bicarbonate reclamation
Glyoxylate and dicarboxylate metabolism
Nitrogen metabolism
Oxidative phosphorylation
Glutathione metabolism
Pyruvate metabolism
Arginine biosynthesis
2−Oxocarboxylic acid metabolism
Carbon metabolism
0
5
10
15
longIRI14d
shortIRI3d
Taurine and hypotaurine metabolism
Peroxisome
Glutathione metabolism
Fatty acid degradation
Cysteine and methionine metabolism
PPAR signaling pathway
Arachidonic acid metabolism
Protein digestion and absorption
Retinol metabolism
Metabolic pathways
0
10
20
30
40
50
Proximal tubule bicarbonate reclamation
Arginine biosynthesis
Ribosome
Phenylalanine metabolism
Oxidative phosphorylation
Butanoate metabolism
Glutathione metabolism
Pyruvate metabolism
Parkinson disease
Non−alcoholic fatty liver disease (NAFLD)
0
5
10
15

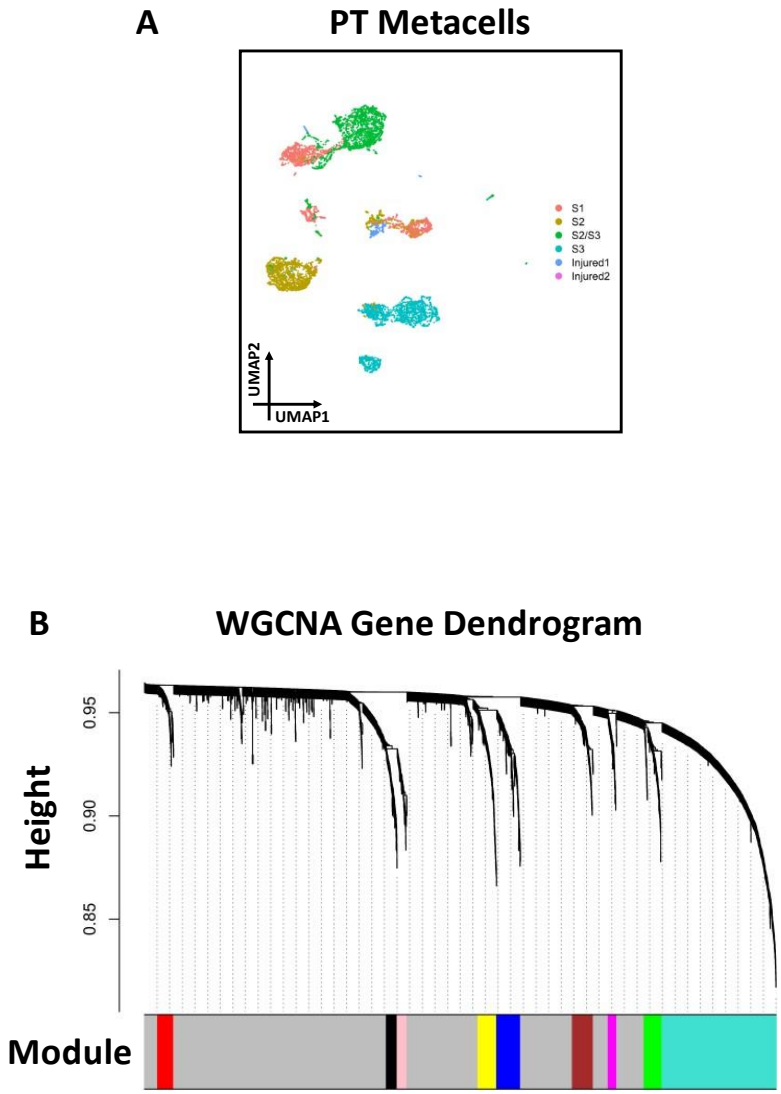


Supplemental Figure 10. Gene set enrichment analysis of Monocle 2 trajectories of PT cells from the mouse scRNA-seq data.

(A) Bar plots showing the top KEGG pathways enriched in studied mouse models along Trajectory 1 (one bar plot per mouse model).

(B) Bar plots showing the top KEGG pathways enriched in studied mouse models along Trajectory 2 (one bar plot per mouse model).

Figure S11

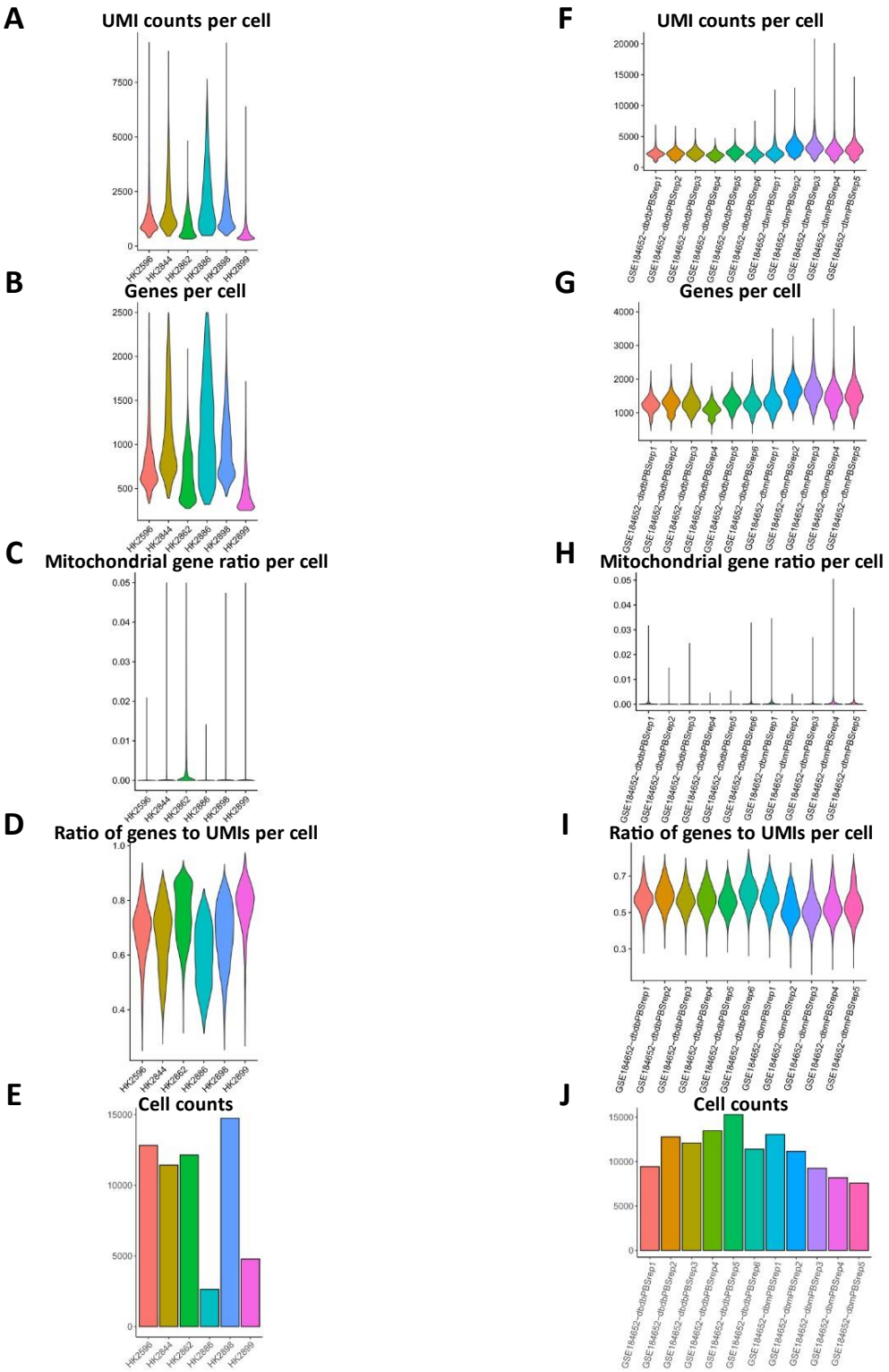


Supplemental Figure 11. WGCNA of PT cells from the mouse scRNA-seq data.

(A) UMAP of 7,432 PT metacells from the mouse scRNA-seq data, colored by cell types.

(B) Hierarchical cluster tree showing gene co-expression modules identified by WGCNA of PT cells from the mouse scRNA-seq data.

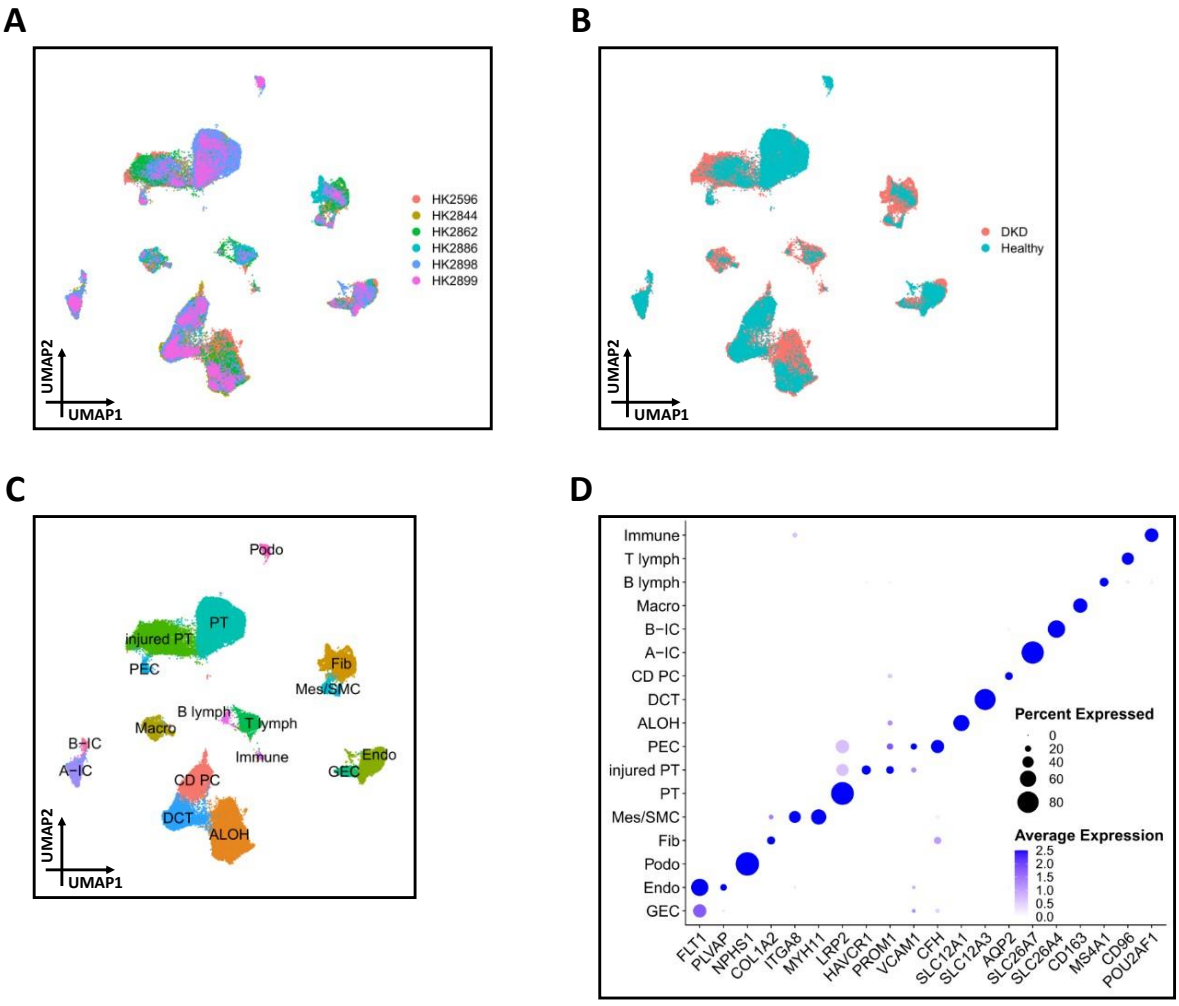
Figure S12



Supplemental Figure 12. Quality control of the human and mouse snRNA-seq data.

(A) Violin plot showing the number of UMIs per single nucleus after QC, split by human snRNA-seq samples.
(B) Violin plot showing the number of detected genes per single nucleus after QC, split by human snRNA-seq samples.
(C) Violin plot showing the percentage of mitochondrially encoded gene reads per single nucleus after QC, split by human snRNA-seq samples.
(D) Violin plot showing the ratio of detected genes to UMIs per single nucleus after QC, split by human snRNA-seq samples.
(E) Bar plot showing the nucleus count of each human snRNA-seq sample after QC.
(F) Violin plot showing the number of UMIs per single nucleus after QC, split by mouse snRNA-seq samples.
(G) Violin plot showing the number of detected genes per single nucleus after QC, split by mouse snRNA-seq samples.
(H) Violin plot showing the percentage of mitochondrially encoded gene reads per single nucleus after QC, split by mouse snRNA-seq samples.
(I) Violin plot showing the ratio of detected genes to UMIs per single nucleus after QC, split by mouse snRNA-seq samples.
(J) Bar plot showing the nucleus count of each mouse snRNA-seq sample after QC.

Figure S13



Supplemental Figure 13. Integration of the human snRNA-seq samples.

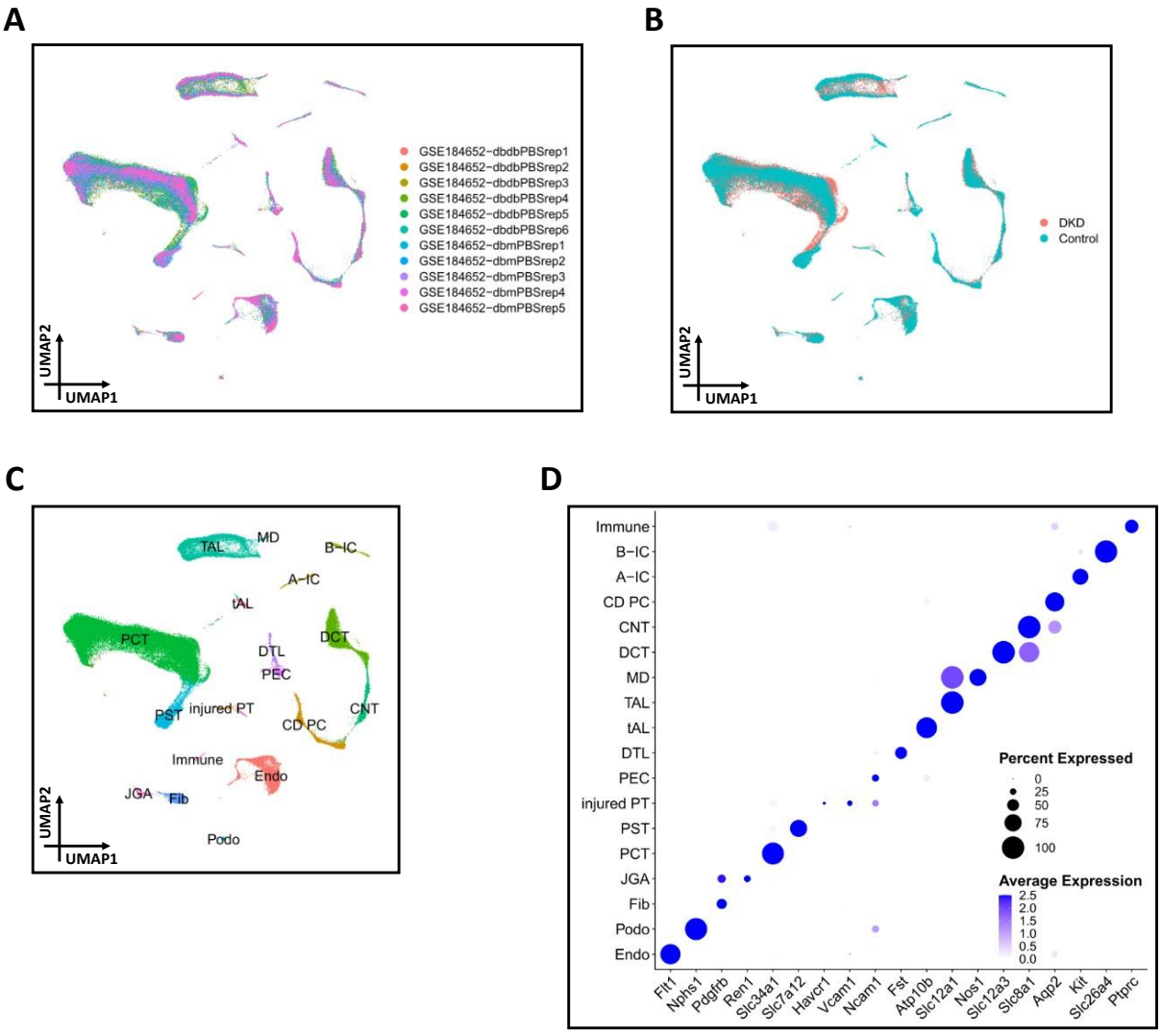
(A) UMAP of 54,945 human kidney single nuclei, colored by samples.

(B) UMAP of 54,945 human kidney single nuclei, colored by conditions.

(C) UMAP of 54,945 human kidney single nuclei, colored by cell types. 18 cell types were identified: GEC, glomerular endothelial cell; Endo, peritubular endothelial cell; Podo, podocyte; Fib, fibroblast; Mes, mesangial cell; SMC, smooth muscle cell; PT, proximal tubule; injured PT, injured proximal tubule; PEC, parietal epithelial cell; ALOH, ascending loop of Henle; DCT, distal convoluted tubule; CD PC, collecting duct principal cell; A-IC, alpha intercalated cell; B-IC, beta intercalated cell; Macro, macrophage; B lymph, B lymphocyte; T lymph, T lymphocyte; Immune, immune cell.

(D) Dot plot of cell type-specific marker genes (dot size denotes percentage of nuclei expressing the marker, and color scale represents average gene expression values).

Figure S14



Supplemental Figure 14. Integration of the mouse snRNA-seq samples.

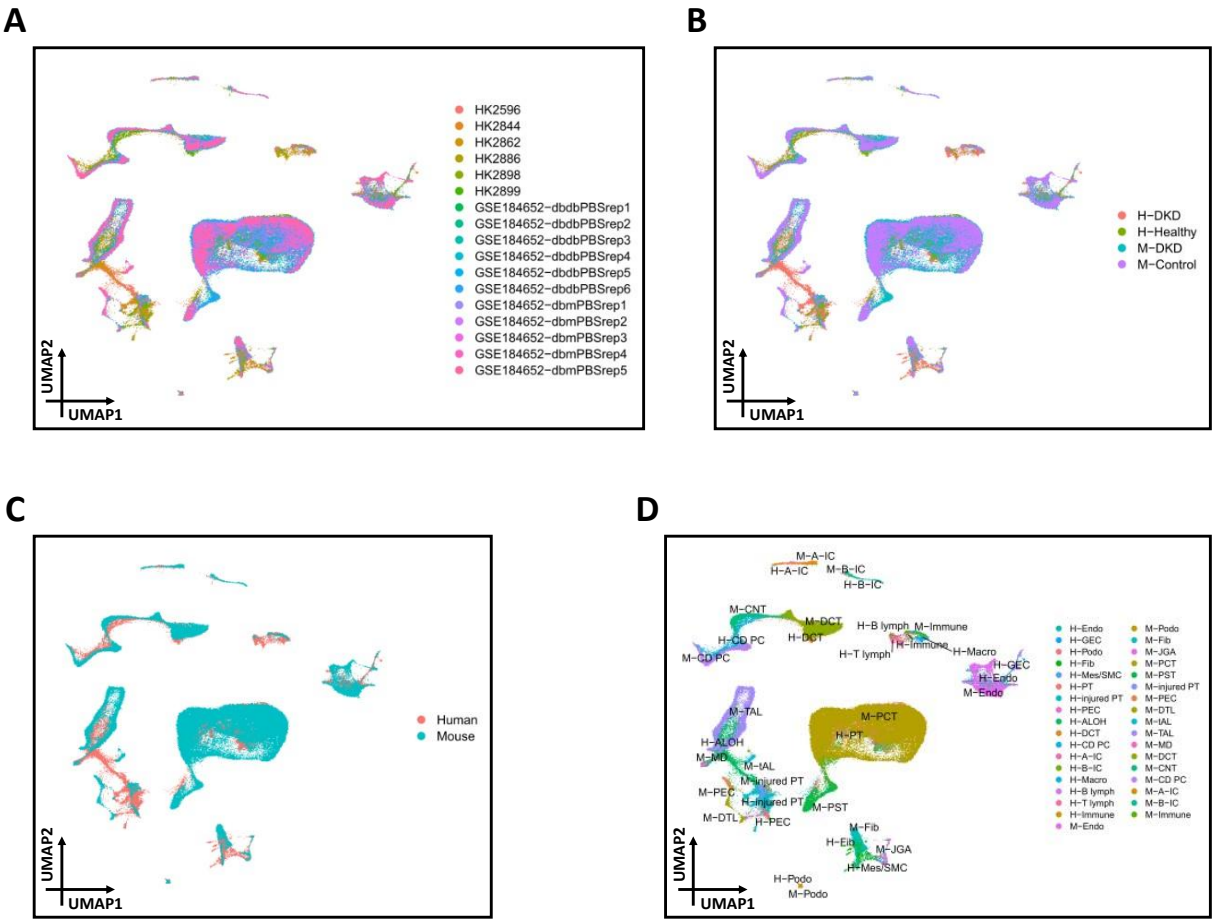
(A) UMAP of 123,704 mouse kidney single nuclei, colored by samples.

(B) UMAP of 123,704 mouse kidney single nuclei, colored by conditions.

(C) UMAP of 123,704 mouse kidney single nuclei, colored by cell types. 18 cell types were identified: Endo, endothelial cell; Podo, podocyte; Fib, fibroblast; JGA, juxtaglomerular apparatus; PCT, proximal convoluted tubule; PST, proximal straight tubule; injured PT, injured proximal tubule; PEC, parietal epithelial cell; DTL, descending thin limb of Henle's loop; tAL, thin ascending limb of Henle's loop; TAL, thick ascending limb of Henle's loop; MD, macular densa; DCT, distal convoluted tubule; CNT, connecting tubule; CD PC, collecting duct principal cell; A-IC, alpha intercalated cell; B-IC, beta intercalated cell; Immune, immune cell.

(D) Dot plot of cell type-specific marker genes (dot size denotes percentage of nuclei expressing the marker, and color scale represents average gene expression values).

Figure S15



Supplemental Figure 15. Integration of the human and mouse snRNA-seq samples.

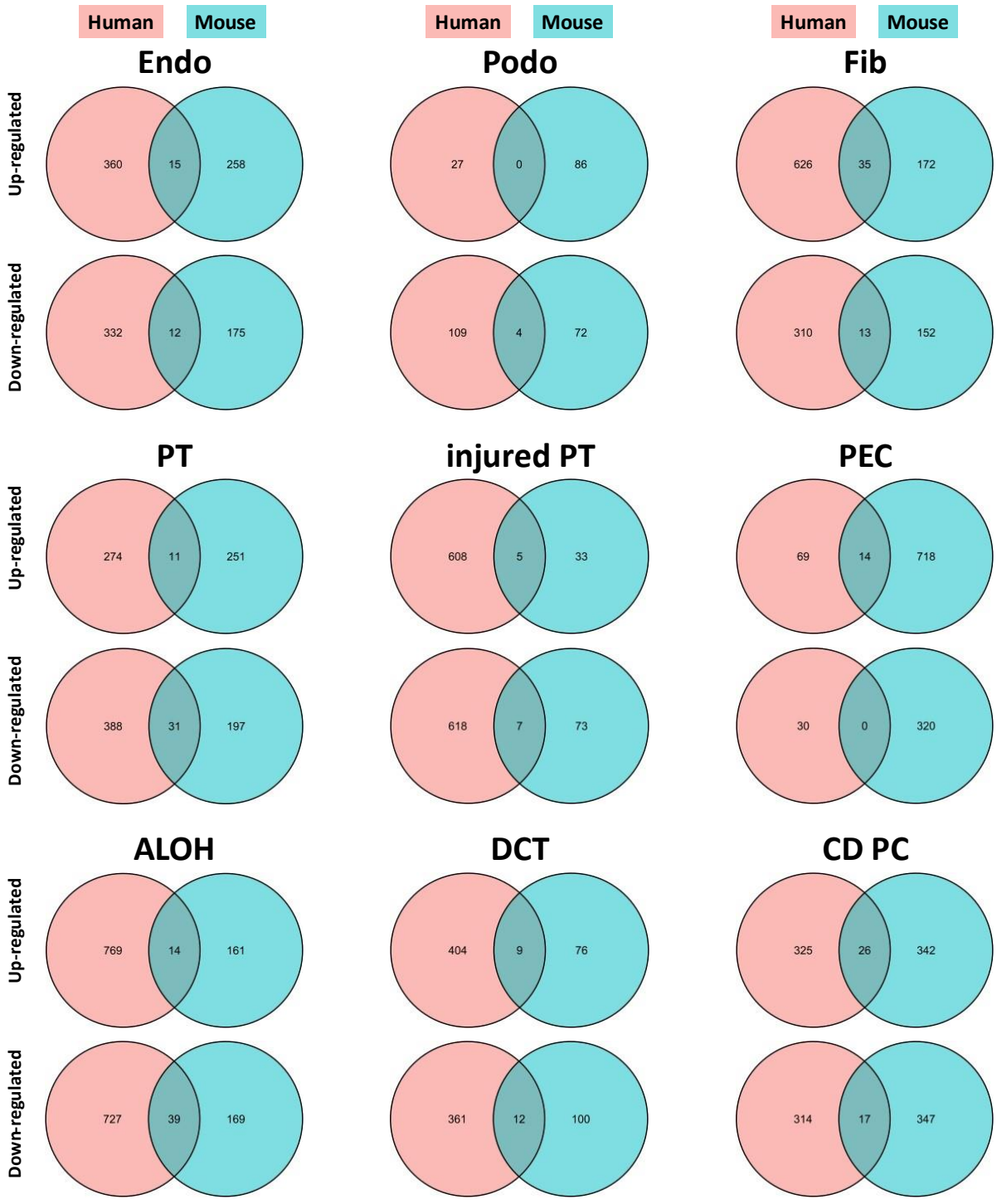
(A) UMAP of 54,945 human and 123,704 mouse kidney single nuclei, colored by samples.

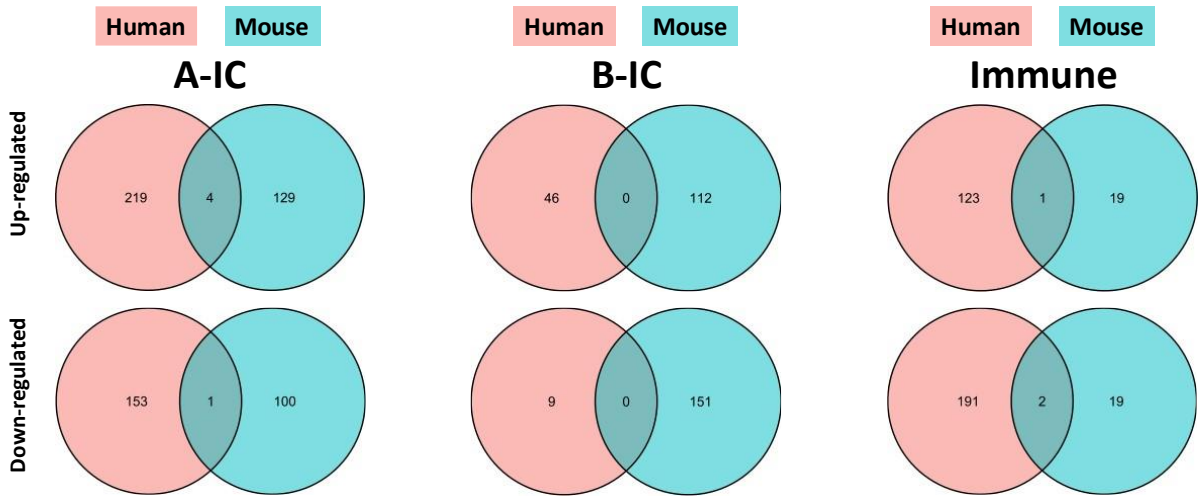
(B) UMAP of 54,945 human and 123,704 mouse kidney single nuclei, colored by conditions.

(C) UMAP of 54,945 human and 123,704 mouse kidney single nuclei, colored by species.

(D) UMAP of 54,945 human and 123,704 mouse kidney single nuclei, colored by cell types. The cell-type annotations were directly taken from Figures S13D and S14D. A prefix was added to each annotation to indicate whether it is a human (i.e., “H-”) or mouse (i.e., “M-”) cell type.

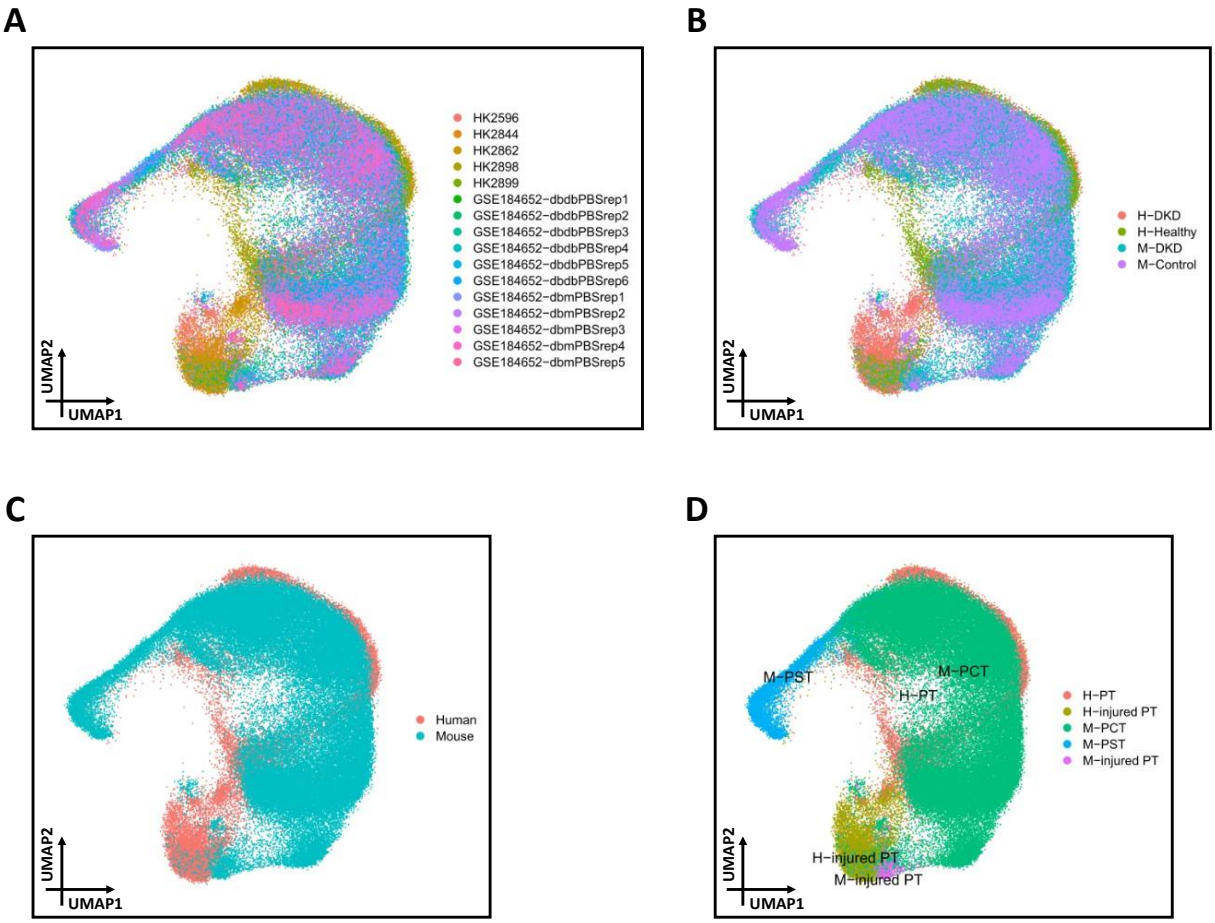
Figure S16


Human
Mouse
Endo
Podo
Fib
PT
injured PT
PEC
ALOH
DCT
CD PC
Up-regulated
Down-regulated
360
15
258
27
0
86
626
35
172
332
12
175
109
4
72
310
13
152
274
11
251
608
5
33
69
14
718
388
31
197
618
7
73
30
0
320
769
14
161
404
9
76
325
26
342
727
39
169
361
12
100
314
17
347



Supplemental Figure 16. Human-mouse cell type-specific DEG conservation.
Venn diagrams show the numbers of up- (top panel) and down-regulated (bottom panel) cell type-specific DEGs (against the control) conserved between human and mouse DKD in identified cell types (two Venn diagrams per cell type).

Figure S17



Supplemental Figure 17. Integration of PT nuclei from the human and mouse snRNA-seq data.

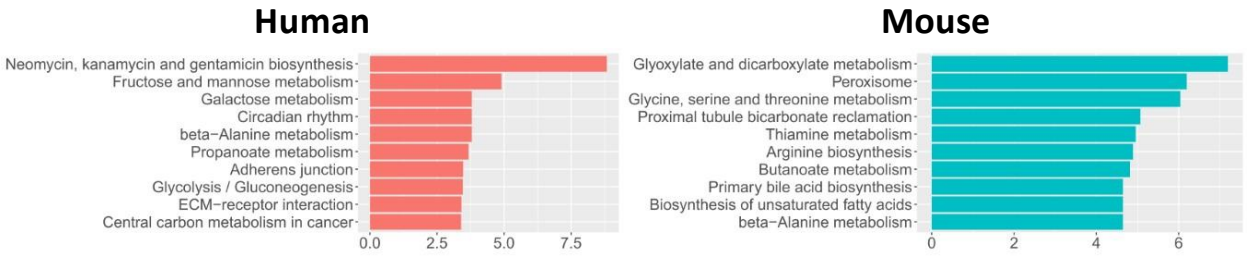
(A) UMAP of 19,319 human and 70,125 mouse PT nuclei, colored by samples.

(B) UMAP of 19,319 human and 70,125 mouse PT nuclei, colored by conditions.

(C) UMAP of 19,319 human and 70,125 mouse PT nuclei, colored by species.

(D) UMAP of 19,319 human and 70,125 mouse PT nuclei, colored by cell types. The cell-type annotations were directly taken from Figure S15D.

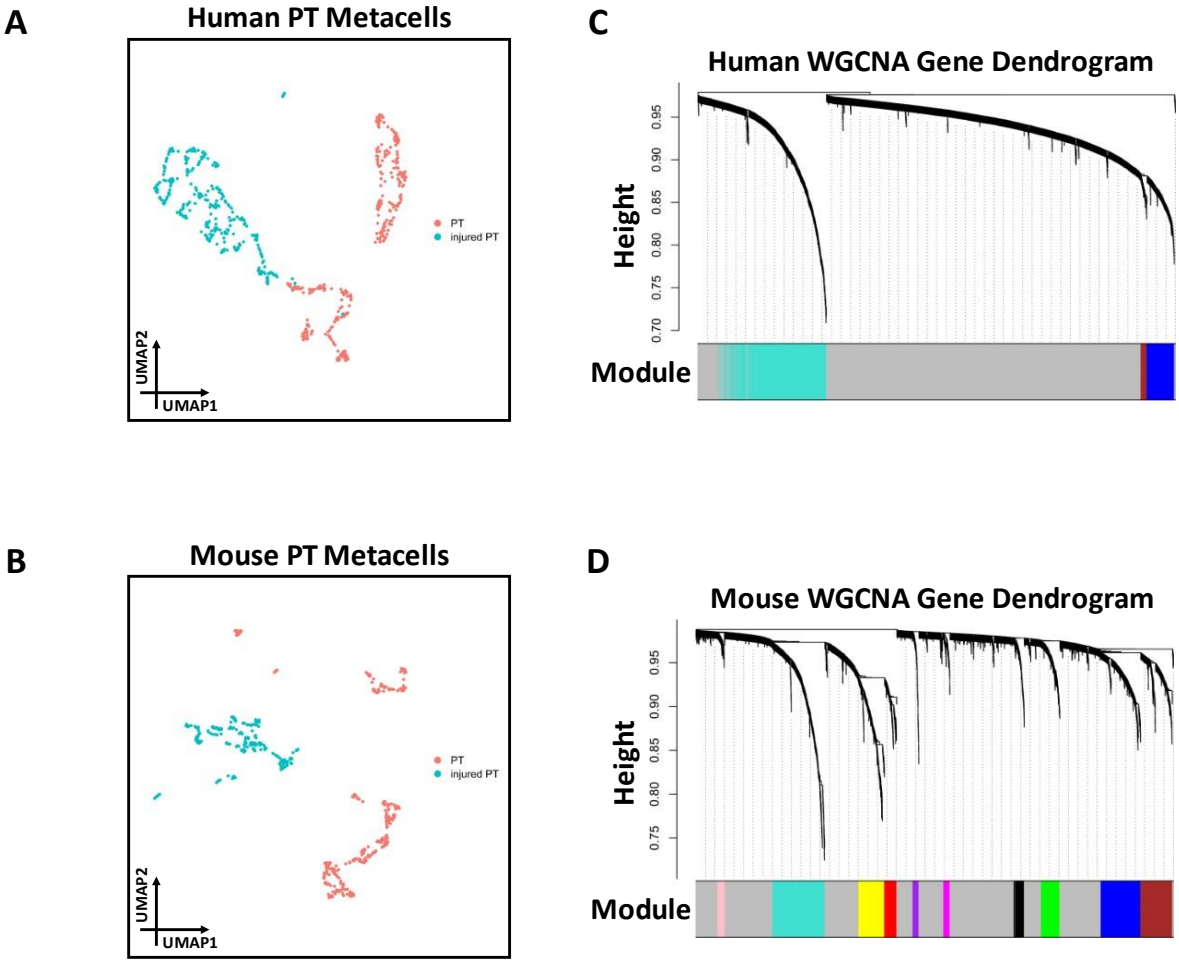
Figure S18



Supplemental Figure 18. Gene set enrichment analysis of Monocle 2 trajectory of PT nuclei from the human and mouse snRNA-seq data.

Bar plots showing the top KEGG pathways enriched in the PT trajectory of the human (left panel) and mouse (right panel) snRNA-seq data.

Figure S19



Supplemental Figure 19. WGCNA of PT nuclei from the human and mouse snRNA-seq data.

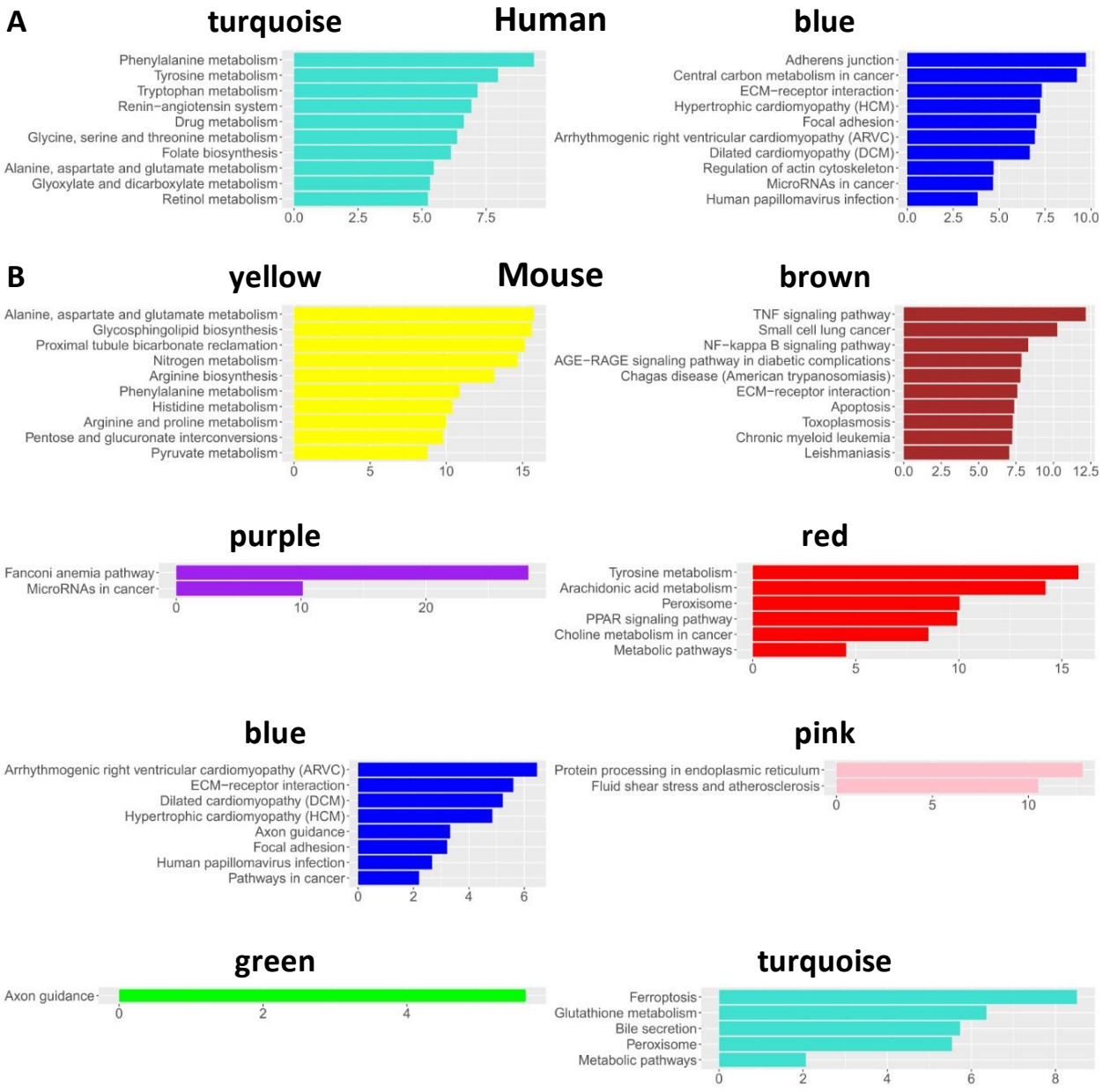
(A) UMAP of 616 PT metacells from the human snRNA-seq data, colored by cell types.

(B) UMAP of 509 PT metacells from the mouse snRNA-seq data, colored by cell types.

(C) Hierarchical cluster tree showing gene co-expression modules identified by WGCNA of PT nuclei from the human snRNA-seq data.

(D) Hierarchical cluster tree showing gene co-expression modules identified by WGCNA of PT nuclei from the mouse snRNA-seq data.

Figure S20



Supplemental Figure 20. Gene set enrichment analysis of WGCNA gene modules of PT nuclei from the human and mouse snRNA-seq data.

(A) Bar plots showing the top KEGG pathways enriched in WGCNA gene modules of PT nuclei from the human snRNA-seq data (one bar plot per gene module).

(B) Bar plots showing the top KEGG pathways enriched in WGCNA gene modules of PT nulcei from the mouse snRNA-seq data (one bar plot per gene module).

Supplemental Reference

1. Kirita Y, Wu H, Uchimura K, Wilson PC, Humphreys BD: Cell profiling of mouse acute kidney injury reveals conserved cellular responses to injury. *Proc Natl Acad Sci U S A,* 117**:** 15874-15883, 2020 10.1073/pnas.2005477117