PNAS

Supporting Information for

A single cell multiomic analysis of kidney organoid differentiation

Yasuhiro Yoshimura, Yoshiharu Muto, Nicolas Ledru, Haojia Wu, Kohei Omachi, Jeffrey H. Miner and Benjamin D. Humphreys

Corresponding author: Benjamin D. Humphreys MD, PhD. Division of Nephrology, Department of Medicine, Washington University in St. Louis School of Medicine, 600 South Euclid Avenue, St. Louis, MO 63110

E-mail: humphreysbd@wustl.edu

This PDF file includes:

Supporting Materials and Methods
Figures S1 to S13
Tables S1 to S3
Legends for Dataset S1
SI References

Other supporting materials for this manuscript include the following:

Dataset S1

Supporting Materials and Methods

Human adult kidney multiome dataset data processing

We used our previously published multi-omics dataset of human adult kidneys generated using snATAC-seq paired with snRNA-seq (1). Cell types in the snATAC-seq dataset were predicted by label transfer using the cell types assigned in the snRNA-seq dataset. We used peaks identified by peak calling for each cell type separately using MACS2. Cis-coaccessibility networks were predicted using the snATAC-seq dataset and Cicero (v1.3.5) according to instructions provided on GitHub. Briefly, the snATAC-seq library was partitioned into individual cell types and converted to cell dataset (CDS) objects using the make_atac_cds function. The CDS object was processed using the detect_genes and estimate_size_factors functions with default parameters prior to dimensional reduction and conversion to a Cicero CDS object. Cicero connections were obtained using the run_cicero function with default parameters and visualized with plot_connections function as described previously (1).

Gene imputation using MAGIC

Given that the gene dropout events were high for the low expression genes in our kidney organoid differentiation dataset, we used a Markov affinity-based graph imputation algorithm, called MAGIC (2), to impute the expression values for the PIM. The parameters for running MAGIC are as follows: MAGIC(a=None, decay=1, k=None, knn=5, knn_dist='euclidean', knn_max=15, n_jobs=1, n_pca=100, random_state=None, solver='exact', t=3, verbose=1). After the imputation process was complete, we visualized the expression patterns of developmental genes on the original UMAP graph generated by Seurat.

Human adult kidney nuclei isolation

Frozen human adult kidney cortex was cut into < 1 mm pieces, homogenized using a dounce tissue homogenizer (DWK Life Sciences, 885302-0002), and incubated on ice for 3 min in chilled Nuclei EZ lysis buffer (Sigma-Aldrich, NUC-101) supplemented with protease inhibitor cocktail (Roche, 5892791001). The homogenate was filtered through a 40-μm cell strainer and centrifuged at 500 g for 5 minutes at 4°C. The pellet was resuspended and washed with chilled lysis buffer, and further incubated on ice for 3 minutes. Following centrifugation, the pellet was resuspended in 1% BSA in PBS(−), and filtered through a 5-μm cell strainer (pluriselect, 43-50005-13).

CUT&RUN

CUT&RUN assay libraries for kidney organoid-derived proximal tubule cells or human adult kidneys were generated with the CUTANA ChIC/CUT&RUN Kit (EpiCypher, 14-1048). Kidney organoid-derived proximal tubule cells were sorted as an LTL^+ population by flow cytometry, filtered through a 40-μm cell strainer, and fixed with 0.5% formalin for one minute at room temperature. After inactivating unreacted aldehyde by adding glycine, 1.0×10^5 cells were treated with chilled multiome lysis buffer on ice for 5 minutes to obtain nuclei. For human adult kidney, 5.0×10^5 isolated nuclei were fixed with 0.5% formalin for one minute at room temperature, followed by addition of glycine. Fixed nuclei were then incubated with Concanavalin A (ConA) conjugated paramagnetic beads and stained with anti-IgG (EpiCypher, 13-0042), anti-histone H3K4me3 (EpiCypher, 13-0041), or anti-histone H3K27Ac antibodies (EpiCypher, 13-0045). The remaining steps were performed according to manufacturer's instructions. Library preparation was performed using the NEBNext Multiplex Oligos for Illumina (New England Biolabs, E6440S) with manufacturer's instructions, including minor modifications indicated by CUTANA described above. CUT&RUN libraries were sequenced on a NovaSeq instrument (Illumina, 150 bp paired-end reads). FASTQ files were trimmed with Trim Galore (Cutadapt [v2.8]) and aligned with Bowtie2 (parameters: --local --very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700) using hg38. The bam files were then converted to BigWig format with

bigwigbamCoverage function in deepTools (v3.5.0), and the data were visualized with Interactive Genome Viewer (v2.6.3).

RNA purification and quantitative RT-PCR (qRT-PCR)
RNA was extracted with a RNeasy Plus Micro Kit (Qiagen, 74034) or a Direct-zol RNA Microprep Kit (Zymo Research, R2062), and reverse-transcribed using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, 4368814). Quantitative PCR was carried out in the BioRad CFX96 Real-Time System using iTaq Universal SYBR Green Supermix (Bio-Rad, 1725124). Expression levels were normalized to *GAPDH*, and the data were analyzed using the 2-ΔΔCt method. The following primers were used in this study: *GAPDH*, Fw 5´- GACAGTCAGCCGCATCTTCT -3´, Rv 5`- GCGCCCAATACGACCAAATC -3´; *HNF1B*, Fw 5´- ACCAAGCCGGTCTTCCATACT -3´, Rv 5`- GGTGTGTCATAGTCGTCGCC -3´.

Flow Cytometry
Organoids were washed with PBS(−) and dissociated with 0.25% trypsin-EDTA at 37℃ for 10 minutes. Cell number of each sample was counted after dissociation. After blocking with normal mouse serum (Thermo Fisher Scientific, 31881) for 10 minutes on ice, cell surface marker staining was carried out with biotinylated lotus tetragonolobus lectin (LTL) (Vector Laboratories, B-1325), followed by APC Streptavidin (BioLegend, 405207) in 1× HBSS (Thermo Fisher Scientific, 14065056) containing 1% bovine serum albumin (Sigma-Aldrich, 3116956001) and 0.035% $NaHCO_3$ (Sigma-Aldrich, S6014) on ice. Stained cells were analyzed using a FACS Canto II or sorted using a FACS Aria II (BD Biosciences). Data analyses were performed with FLOWJO software (BD Biosciences).

Immunofluorescence
For frozen sections, samples were fixed with 4% paraformaldehyde for 20 minutes, washed with PBS(−), cryoprotected with 10%, 20%, and 30% sucrose in PBS(−), embedded in optimal cutting temperature (OCT) compound, and cryosectioned at 7-μm thicknesses. For paraffin embedded sections, samples were fixed with 10% formalin overnight, embedded in paraffin, and cut at 5-μm thicknesses. Antigen retrieval in antigen unmasking solution (Vector Laboratories, H-3301-250) was performed before staining. The sections were blocked with 1% bovine serum albumin in PBS(−) for 60 minutes at room temperature, followed by incubation with primary antibodies against WT1 (abcam, ab89901), biotinylated LTL, CDH1 (BD Biosciences, 610182), CRABP1 (cell signaling, 13163), MLANA (Thermo Fisher Scientific, MA5-14168), SLC5A2/SGLT2 (Novus Biologicals, NBP1-92384), SLC22A12/URAT1 (Sigma-Aldrich, HPA024575), CAS9 (Novus Biologicals, NBP2-36440), and SLC12A1 (Proteintech, 18970-1-AP) at 4°C overnight. Next, the sections were incubated with secondary antibodies for 90 minutes at room temperature. For staining of HNF1B, sections were incubated with an anti-HNF1B antibody (Sigma-Aldrich, AMAB90733) at 1:10,000 dilution, followed by incubation with an anti-mouse ImmPRESS HRP polymer (Vector Laboratories; MP-7402). The sections were then incubated with Alexa Fluor 488-tyramide (Thermo Fisher Scientific, B40953) for 10 minutes at room temperature. We confirmed no cross-reactivities between anti-HNF1B and anti-CAS9 antibodies as previously reported (3). Nuclei were counterstained with DAPI (4,6-diamidino-2-phenylindole, Thermo Fisher Scientific, D1306) and mounted in Prolong Gold (Thermo Fisher Scientific, P36930). Fluorescence images were captured by Nikon C2+ Eclipse confocal microscopy.

CRISPR interference
Single guide RNAs (sgRNAs) targeting *HNF1B* promoter or putative enhancer regions were designed using CHOPCHOP (https://chopchop.cbu.uib.no/) (4). Sequences of the sgRNAs are as follows:
HNF1B_promoter-1, 5´- GTCCCGGAGGCTCCTCCGAA -3´; HNF1B_promoter-2, 5´- GCAGACACCTGTTACTCCCCG -3´; HNF1B_enhancer_1-1, 5´- GTTTGCAGAGACCACGACCA -3´; HNF1B_enhancer_1-2, 5´- GTCACAGGTAACTGACACAAG -3´; HNF1B_enhancer_2-1, 5´- GCCCTCGCCGAGTTGGAGACG -3´; HNF1B_enhancer_2-2, 5´- GCAAGCGCAGGTGTCAAACGC -3´; _HNF1B_enhancer_3-1, 5´- GAGTTCTCGCCTGTAGAGCAT -3´; HNF1B_enhancer_3-2, 5´-

GCCCAAGACGCCACTGGACAG -3´. The sgRNAs were inserted into downstream of the U6 promoter of the dCas9-KRAB repression plasmid (pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro, a gift from Charles Gersbach, Addgene, 71236). Briefly, sense and anti-sense oligonucleotides were annealed by cooling from 95°C to 25°C for 1.5 hours. The annealed oligonucleotides were then cloned into the dCas9-KRAB repression plasmid by Golden Gate Assembly with Esp3I restriction enzyme (New England Biolabs, R0734L) and T4 DNA ligase (New England Biolabs, M0202L) on a thermal cycler repeating 37 °C for 5 minutes and 16°C for 5 minutes for 60 cycles, followed by transformation to NEB 5-alpha Competent E. coli (New England Biolabs, C2987H) as manufacturer's instruction. The cloned lentiviral vectors were purified with high-speed mini plasmid kit (IBI Scientific, IB47102) or PureLink HiPure plasmid midiprep kit (Thermo Fisher Scientific, K210004), and sgRNA insertion was confirmed with Sanger sequencing by GENEWIZ.

To generate lentivirus, 293T cells were seeded at 3.0×10^6 cells on a 10-cm tissue culture dish one day before transfection. The cells were transfected with 9.0 µg of psPAX2 (a gift from Didier Trono, Addgene, 12260), 0.9 µg of pMD2.G (a gift from Didier Trono, Addgene, 12259), and 9.0 µg of dCas9-KRAB repression plasmid with sgRNAs by Lipofectamine 3000 transfection reagent (Thermo Fisher Scientific, L3000015) as the manufacturer's instructions. Culture media were changed to DMEM supplemented with 30% FBS at 24 hours after transfection. Lentivirus-containing supernatants were collected 24 and 48 hours later and filtered through 0.45 µm PVDF filters (CELLTREAT, 229745). The resultant supernatants were stored at -80℃ or concentrated using Lenti-X concentrator (Takara, 631231). The titer of the concentrated lentivirus was measured using qPCR Lentivirus Titer Kit (Applied Biological Materials, LV900).

For transduction to RPTECs, cells were seeded at 5.0×10^4 cells per well on 6-well tissue culture plates one day before transfection. Culture media was then changed to the lentiviral supernatants supplemented with 3 µg/mL polybrene (Santa Cruz Biotechnology, sc-134220) and transduced for 24 hours. After 48 hours from transduction, transduced cells were selected by 3 µg/ml of puromycin (InvivoGen, ant-pr-1) for 72 hours. For transduction to iPSCs, we transiently changed iPSC culture method to the single-cell passage using StemFit Basic04 (Ajinomoto, BASIC04CT) (5) to increase transduction efficiency. AN1.1 iPSCs were seeded at 2.0×10^4 cells per well on matrigel-coated 12-well tissue culture plates and MOI 5 of the concentrated lentivirus was added during seeding. 10 µM of Y27632 was also added to the media for 48 hours after passage. After 48 hours, lentivirus containing media was replaced to StemFit Basic04 supplemented with 0.5 µg/ml puromycin. The transduced cells were passaged at 2.6×10^4 cells per well on matrigel-coated 12-well tissue culture and treated with 0.5 µg/ml puromycin for a total of 8-10 days.

Bulk ATAC-seq library preparation and data processing

RPTECs after CRISPRi targeting HNF1B were scraped from culture dishes and centrifuged at 500 g for 5 minutes. 10,000 cells per each condition were resuspended in 50 µl of ice-cold lysis buffer containing 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 1% BSA, 0.1% Tween-20 (Sigma-Aldrich, P7949), 0.1% NP-40 (Thermo Fisher Scientific, 28324) and 0.01% Digitonin (Thermo Fisher Scientific, BN2006) (6). The suspension was incubated for 4 minutes on ice. Subsequently, 450 µl of ice-cold wash buffer containing 10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 1% BSA, 0.1% Tween-20 was added and centrifuged at 600 g for 6 minutes. The pellet was resuspended in 25 µl of ATAC-seq transposition mix (12.5 µl 2 × Illumina Tagment DNA (TD) buffer; 10.5 µl nuclease-free water; 2.0 µl Tn5 transposase (Illumina, FC-121-1030)) and incubated at 37 °C for 1 hour on a thermomixer. The transposed DNA was purified using MinElute PCR purification kit (Qiagen, 28004). DNA samples were then amplified with PCR [72 °C for 5 minutes and 98 °C for 30 seconds] followed by 9 cycles of [98 °C for 10 seconds; 63 °C for 30 seconds; 72 °C for 1 minute] using unique 10-bp dual indexes and NEBNext High-Fidelity 2 × PCR Master Kit (New England Biolabs, M0541L). Following the first amplification, DNA size selection was performed using solid-phase reversible immobilization (SPRI) beads, AMPure XP (Beckman Coulter, A63881), at a SPRI to DNA ratio of 0.5. The supernatant was further mixed with SPRI beads at a SPRI to DNA ratio of 1.2. The resulting supernatant was discarded, and the magnet-immobilized SPRI beads were washed twice with 80% ethanol. DNA was subsequently eluted in 20 µl of EB elution buffer (Qiagen). The size-selected DNA was amplified with an additional 9-cycle PCR. Subsequently, the amplified DNA was purified with AMPure XP at a SPRI

to DNA ratio of 1.7 and eluted with 25 µl of EB elution buffer. The resultant ATAC-seq libraries were sequenced on a NovaSeq instrument (Illumina, 150 bp paired-end reads).FASTQ files were trimmed with Trim Galore (Cutadapt [v2.8]) and aligned with Bowtie2 [v2.3.5.1] with --very-sensitive -X 2000 using hg38. PCR duplicates were removed with Picard's MarkDuplicates function.

Figure S1

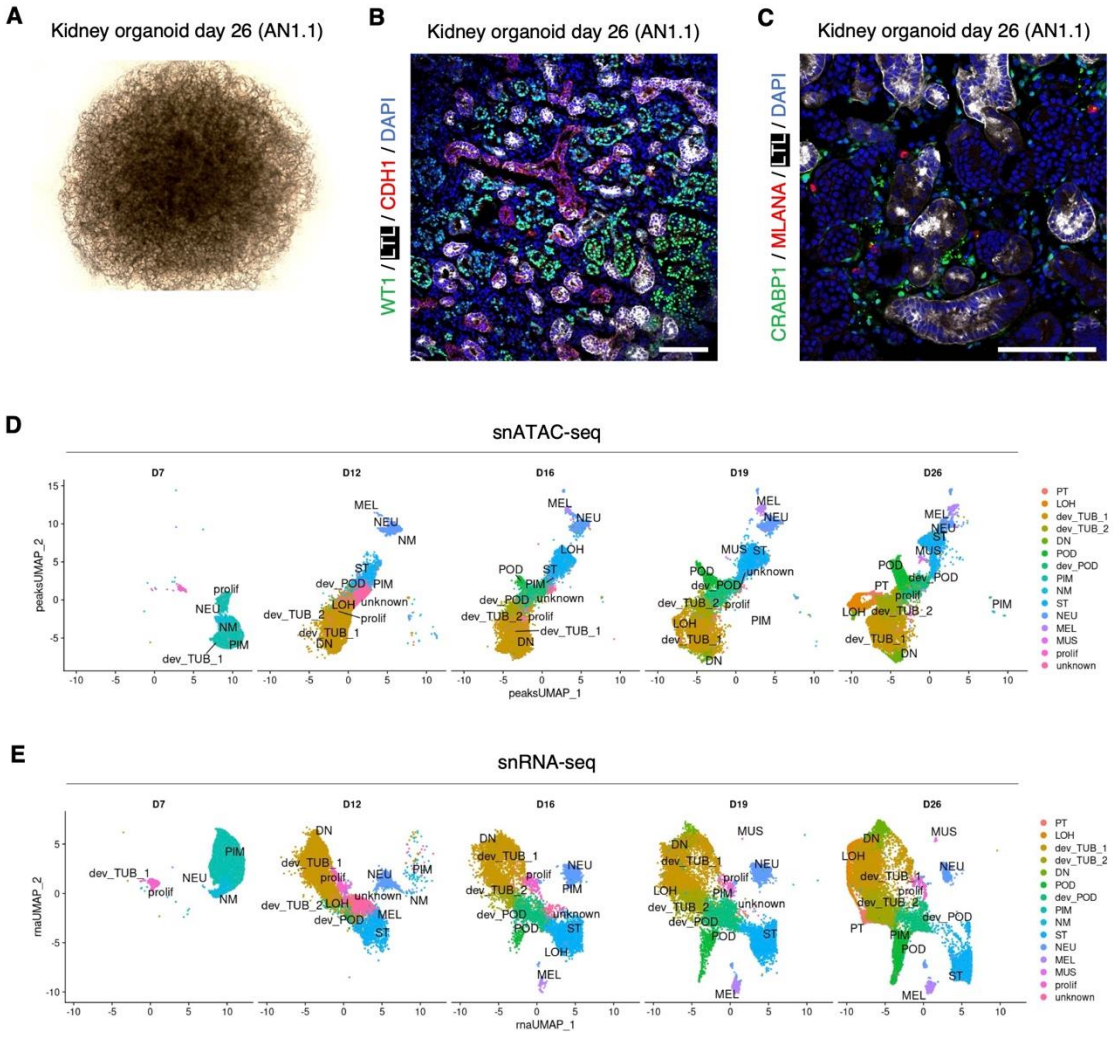

Figure S1. Characterization of human kidney organoid differentiation multiome dataset.
(A) Blight field image of kidney organoid (day 26) induced from AN1.1 iPSCs.
(B) Immunofluorescence images of WT1 (green, podocyte), LTL (white, proximal tubule), CDH1 (red, loop of Henle and distal nephron), and nuclear DAPI staining (blue) of kidney organoid (day 26) induced from AN1.1 iPSCs. Scale bar indicates 100 µm.
(C) Immunofluorescence images of CRABP1 (green, neural cell), MLANA (red, melanocyte), LTL (white, proximal tubule), and nuclear DAPI staining (blue) of kidney organoid (day 26) induced from AN1.1 iPSCs. Scale bar indicates 100 µm.
(D,E) UMAP plots of chromatin accessible peaks (D) and gene expression (E) at each time point (day7, day12, day16, day19, and day26) of the merged kidney organoid differentiation multiome dataset. PT, proximal tubule; LOH, Loop of Henle; dev_TUB, developing tubule; DN, distal nephron; POD, podocyte; dev_POD, developing podocyte; PIM, posterior intermediate mesoderm; NM, nascent mesoderm; ST, stromal cell; NEU, neural cell; MEL, melanocyte; MUS, muscle cell; prolif, proliferating cell; unknown, undefined type of cell.

Figure S2

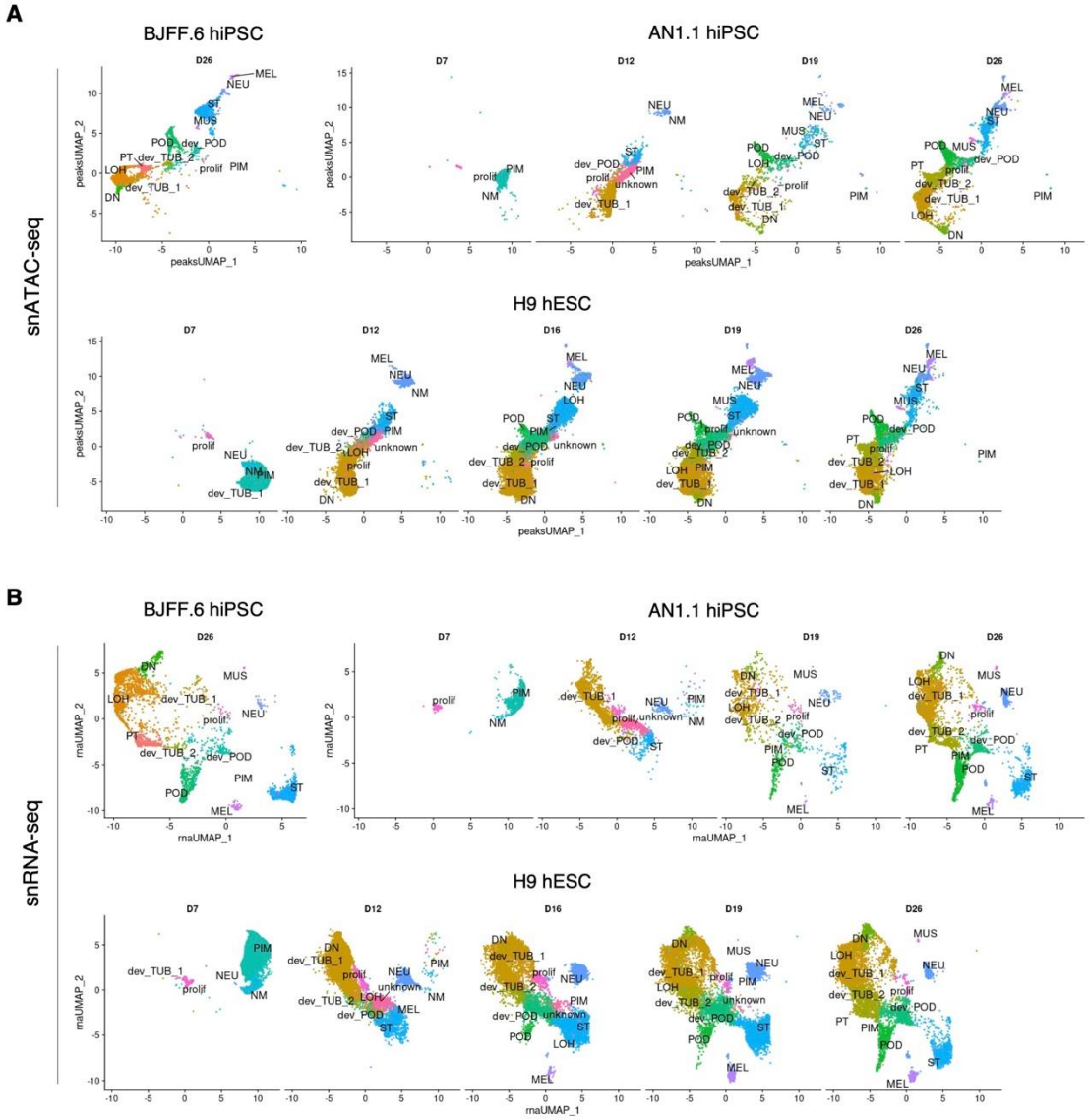

Figure S2. Time-course changes of cell types per cell line of the human kidney organoid differentiation multiome dataset.
UMAP plots of chromatin accessible peaks (A) and gene expression (B) at each time point (day7, day12, day16, day19, and day26) of each PSC clone in the kidney organoid differentiation multiome dataset. PT, proximal tubule; LOH, Loop of Henle; dev_TUB, developing tubule; DN, distal nephron; POD, podocyte; dev_POD, developing podocyte; PIM, posterior intermediate mesoderm; NM, nascent mesoderm; ST, stromal cell; NEU, neural cell; MEL, melanocyte; MUS, muscle cell; prolif, proliferating cell; unknown, undefined type of cell.

Figure S3

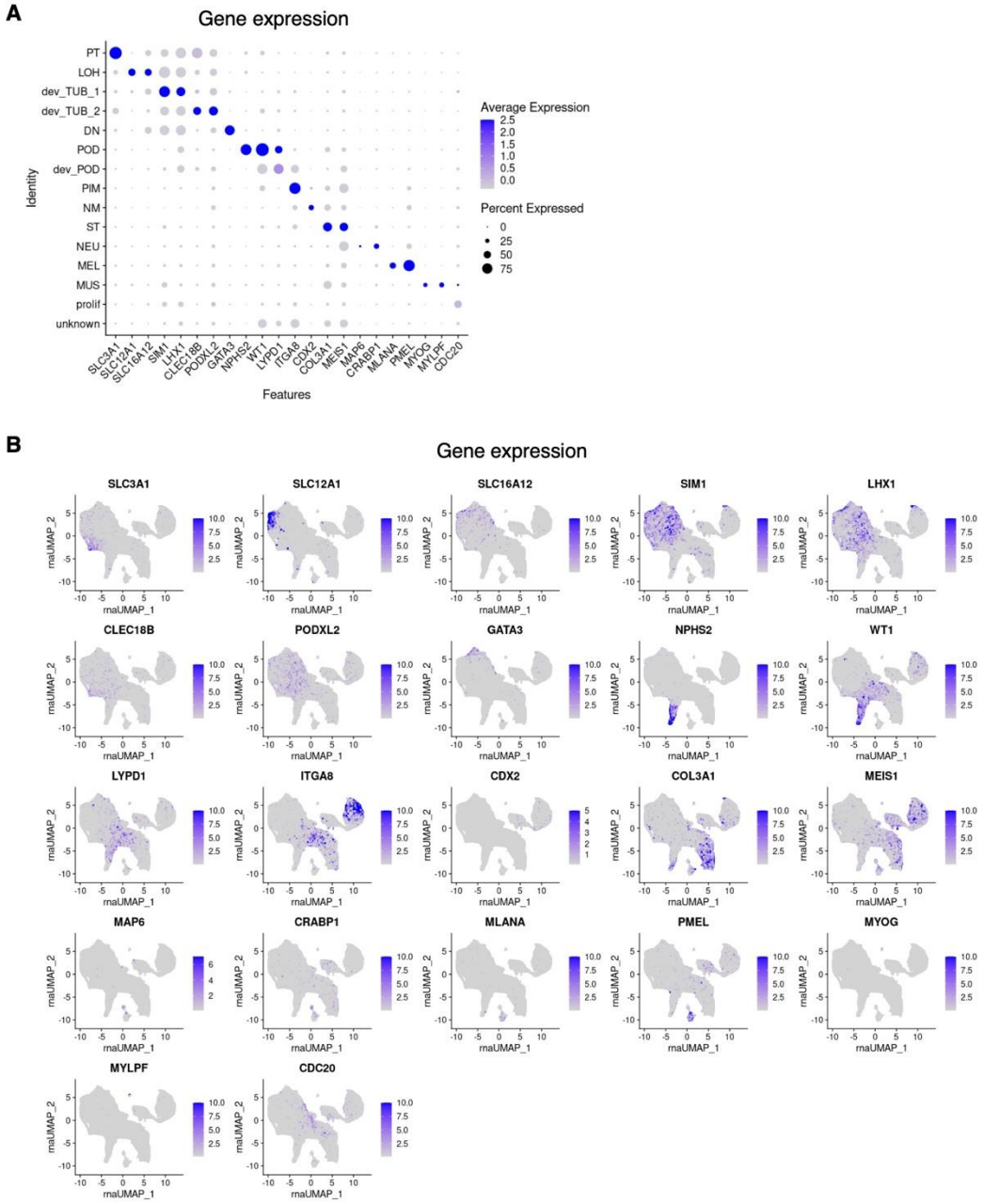

Figure S3. Gene expression of cell type-specific markers in the merged kidney organoid differentiation multiome dataset.
(A) Dot plot showing gene expression levels of cell type-specific markers in each cell type of the merged kidney organoid differentiation multiome dataset.
(B) Feature plots showing gene expression levels of cell type-specific markers in the merged kidney organoid differentiation multiome dataset.

Figure S4

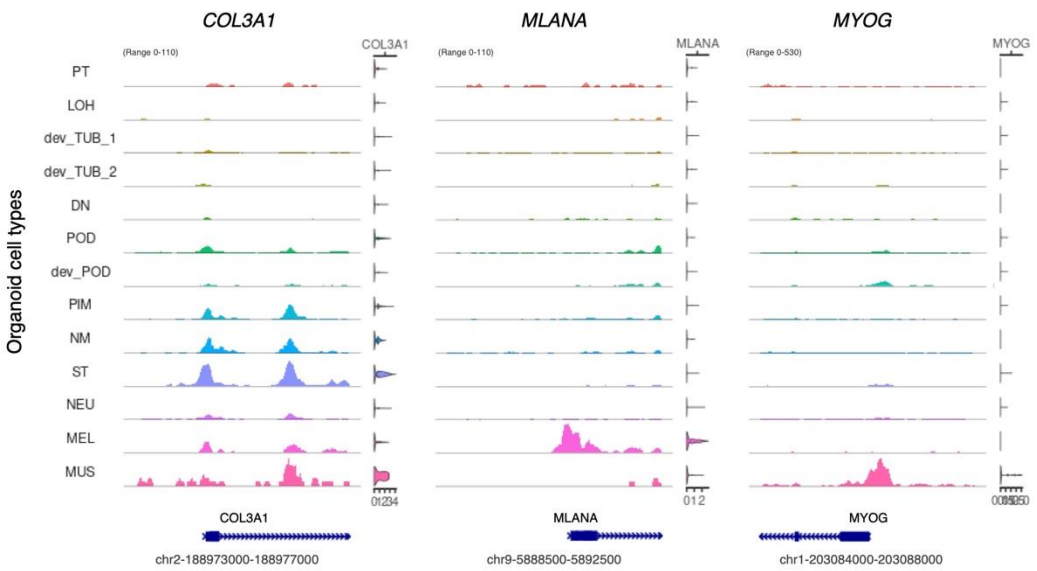

Figure S4. Open chromatin accessibility for stromal and off-target genes of the kidney organoid differentiation multiome dataset.
Coverage plots showing ATAC peaks around transcription start site and gene expression levels of *COL3A1* (stromal cell), *MLANA* (melanocyte), and *MYOG* (muscle cell) in each cluster of the merged kidney organoid differentiation multiome dataset.

Figure S5

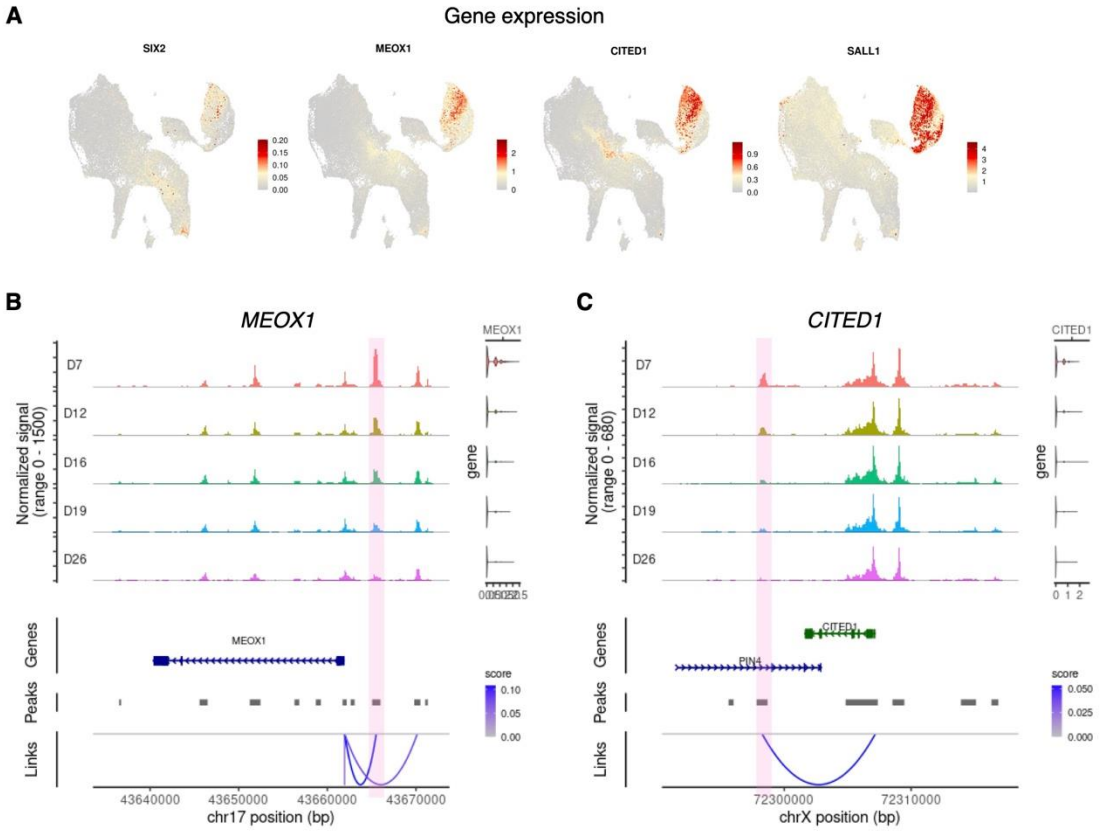

Figure S5. Gene expression and chromatin accessibility changes of nephron progenitor marker genes of the kidney organoid differentiation multiome dataset.
(A) UMAP plots showing gene expression levels of nephron progenitor marker genes using MAGIC gene imputation.
(B, C) Coverage plots showing changes in ATAC peaks and gene expression levels of *MEOX1* (B) and *CITED1* (C) during organoid differentiation time course. Pink boxes indicate putative enhancer regions.

Figure S6

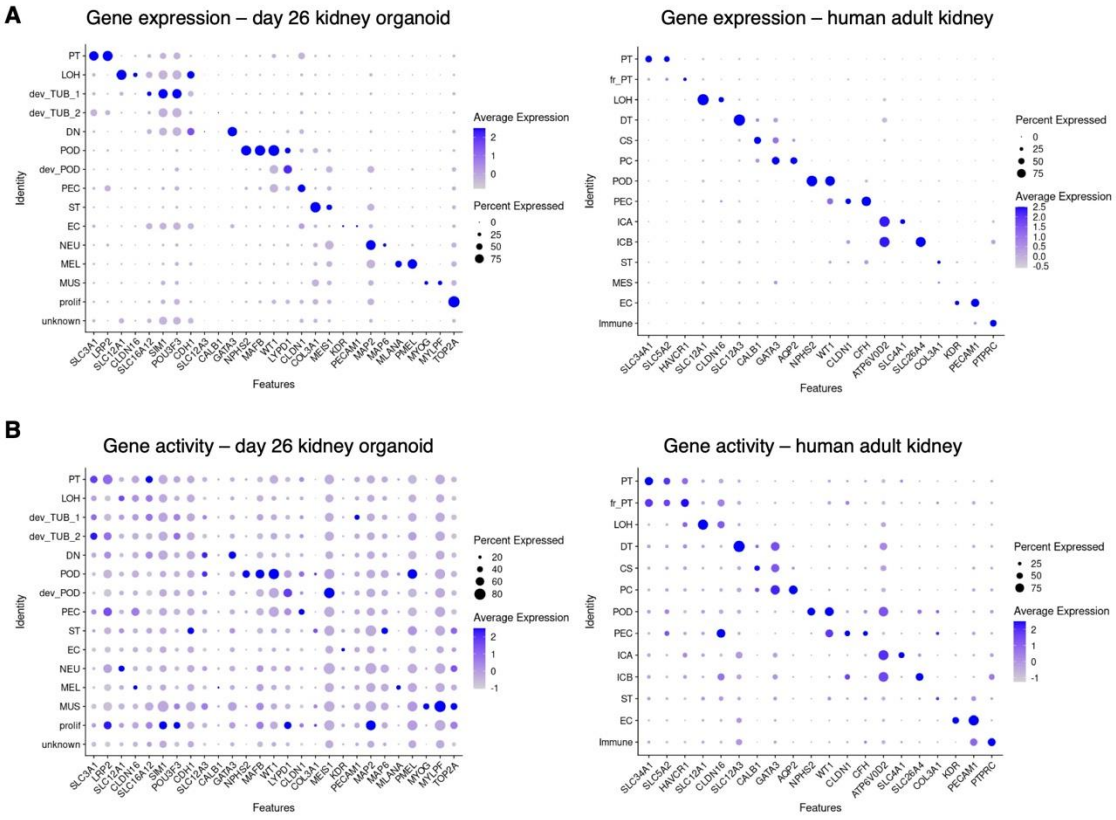

Figure S6. Comparison of kidney organoid and human adult kidney multiome datasets. (A, B) Dot plots showing gene expression levels (A) and gene activity scores (B) of cell type-specific markers in each cell type of the merged kidney organoid (day 26) (left) and human adult kidney (right) multiome datasets.

Figure S7

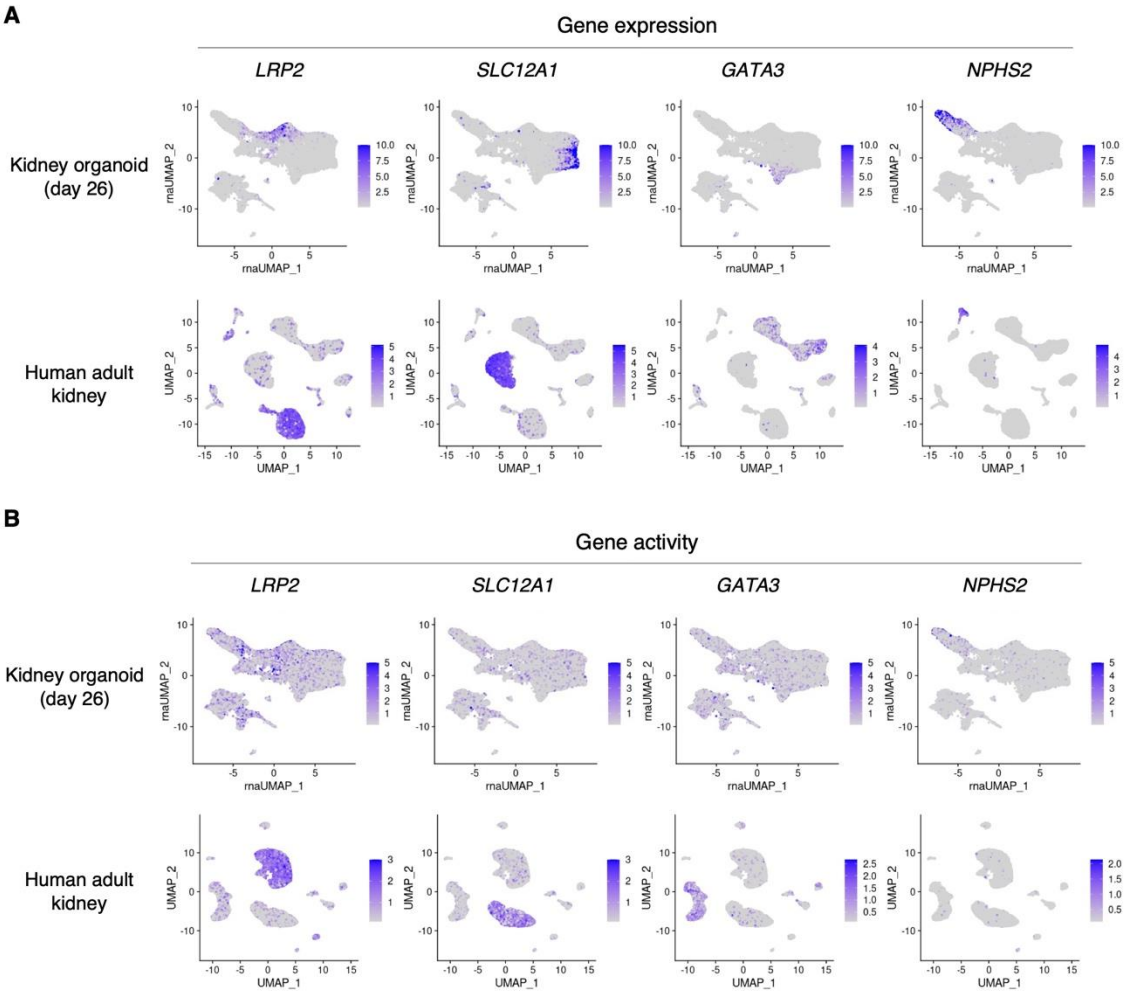

Figure S7. Kidney organoid showed less cell-specific gene activity compared to human adult kidney.
(A) UMAP plots showing gene expression levels of *LRP2* (proximal tubule), *SLC12A1* (loop of Henle), *GATA3* (distal nephron, distal tubule, and connecting segment), and *NPHS2* (podocyte) in the day 26 kidney organoid (upper) and human adult kidney (lower) multiome datasets.
(B) UMAP plots showing gene activity scores of *LRP2* (proximal tubule), *SLC12A1* (loop of Henle), *GATA3* (distal nephron, distal tubule, and connecting segment), and *NPHS2* (podocyte) in the day 26 kidney organoid (upper) and human adult kidney (lower) multiome datasets.

Figure S8

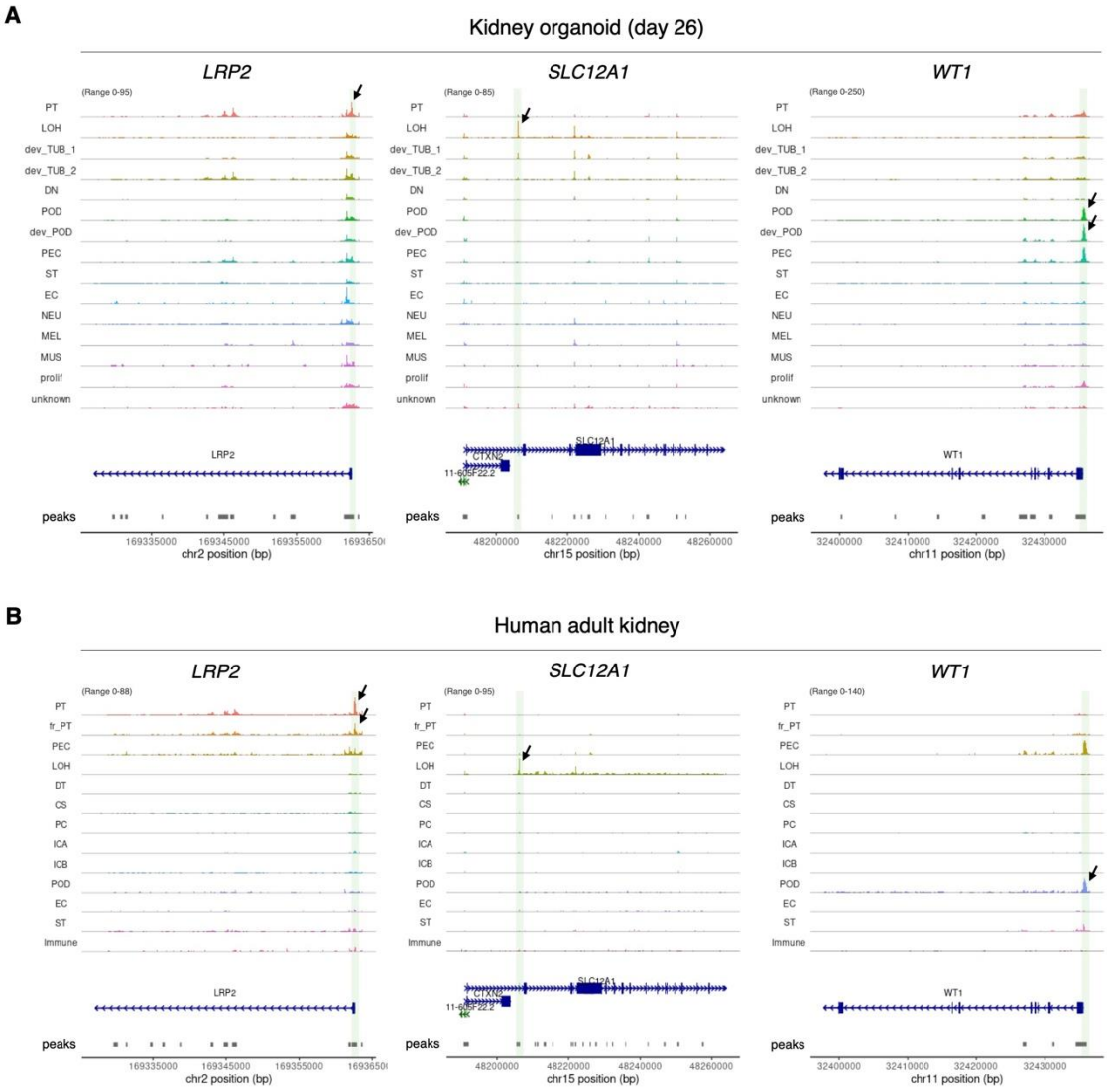

Figure S8. Chromatin accessible peaks in human adult kidney showed more cell type-specific patterns compared to those in kidney organoid.
(A, B) Coverage plots showing ATAC peaks in the gene body regions of *LRP2* (proximal tubule), *SLC12A1* (loop of Henle), and *WT1* (podocyte) genes in each cluster of the day 26 kidney organoid (A) and human adult kidney (B) multiome datasets. Green boxes indicate promoter regions. Black arrows indicate specific ATAC peaks of promoter regions in cell types expressing the marker genes.

Figure S9

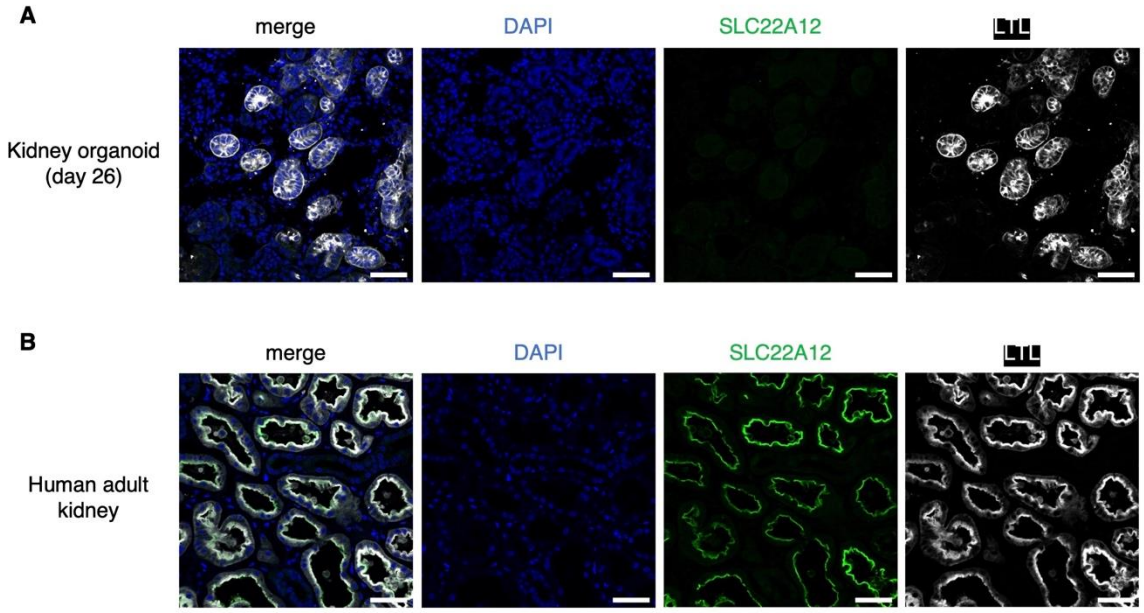

Figure S9. Proximal tubule in kidney organoids does not express SLC22A12 (URAT1).
(A, B) Immunofluorescence images of SLC22A12 (green, URAT1), LTL (white), and nuclear DAPI (blue) staining in the proximal tubule of kidney organoids (day 26) (A) and human adult kidney (B). Scale bars indicate 50 µm.

Figure S10

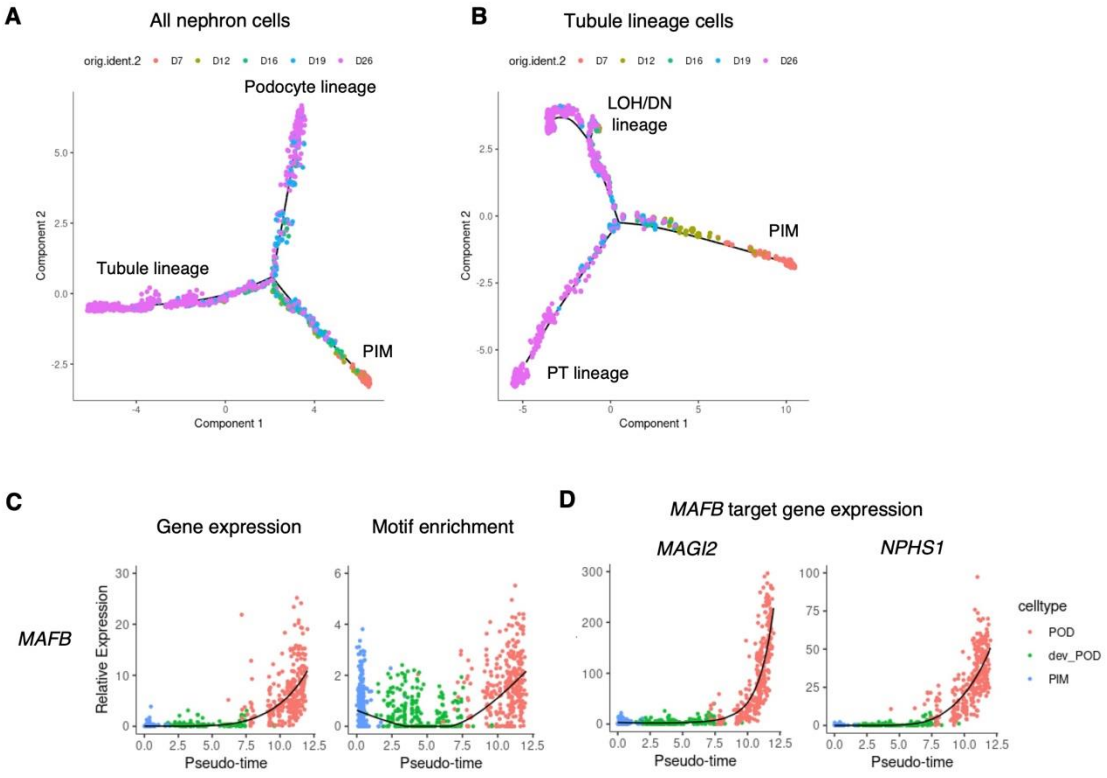

Figure S10. Changes in chromatin accessible landscape during kidney organoid differentiation.
(A) Pseudotime ordering plots of nephron lineage cells showing podocyte and tubule lineage branches. Colors are based on collecting time points during differentiation.
(B) Sub-ordering plots of tubule lineage cells along pseudotime showing PT and LOH/DN lineage branches. Colors are based on collecting time points during differentiation.
(C, D) Dynamics of gene expression (C, left) and motif enrichment (C, right) of the *MAFB* gene and their target gene expression changes (D) along pseudotime.

Figure S11

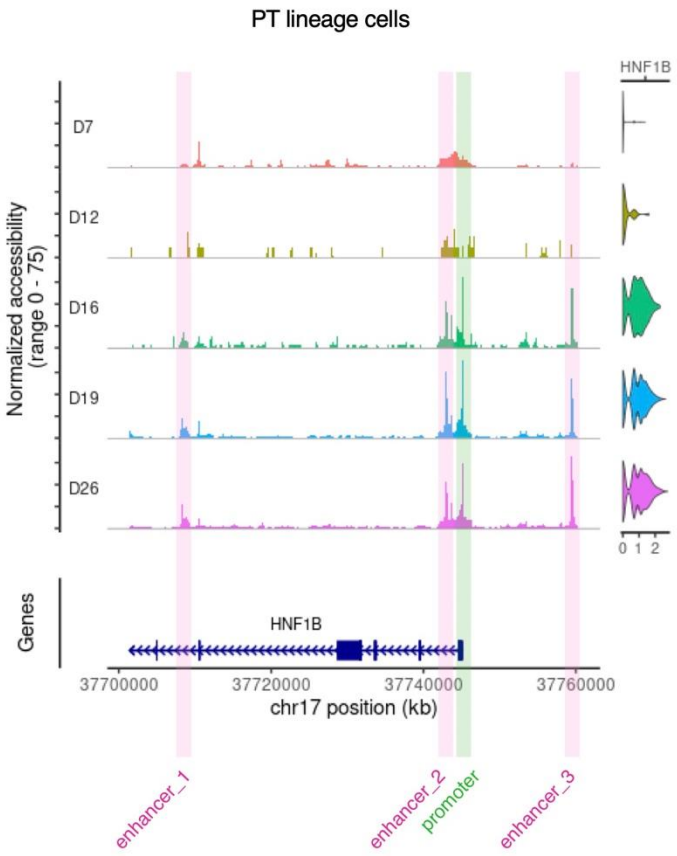

Figure S11. Changes in chromatin accessibility for the *HNF1B* gene during organoid proximal tubule differentiation.
Coverage plots showing changes in ATAC peaks and gene expression level of *HNF1B* gene in the proximal tubule lineage cells during organoid differentiation time course. Green box indicates promoter region; pink boxes indicate putative enhancer regions.

Figure S12

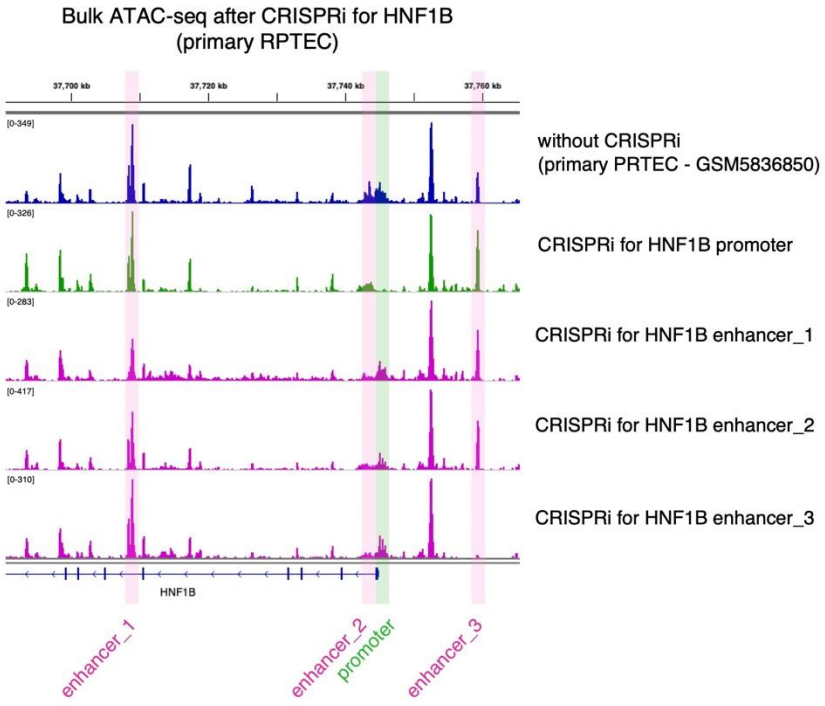

Figure S12. CRISPRi for *HNF1B* specifically reduced chromatin accessibility of the target region of sgRNA.
Integrative genomics viewer (IGV) images of bulk ATAC-seq after CRISPRi targeting *HNF1B* promoter and enhancers using primary RPTEC. Publicly available bulk ATAC-seq dataset of primary RPTEC (GEO: GSM5836850) was used for comparison (1st row, without CRISPRi). The gRNA promoter_1 was used for promoter CRISPRi; the gRNA enhancer_1-1 was used for enhancer_1 CRISPRi; the gRNA enhancer_2-2 was used for enhancer_2 CRISPRi; the gRNA enhancer_3-2 was used for enhancer_3 CRISPRi. Green box indicates promoter region; pink boxes indicate putative enhancer regions. Scales of Y-axis were determined using autoscale function of IGV.

Figure S13

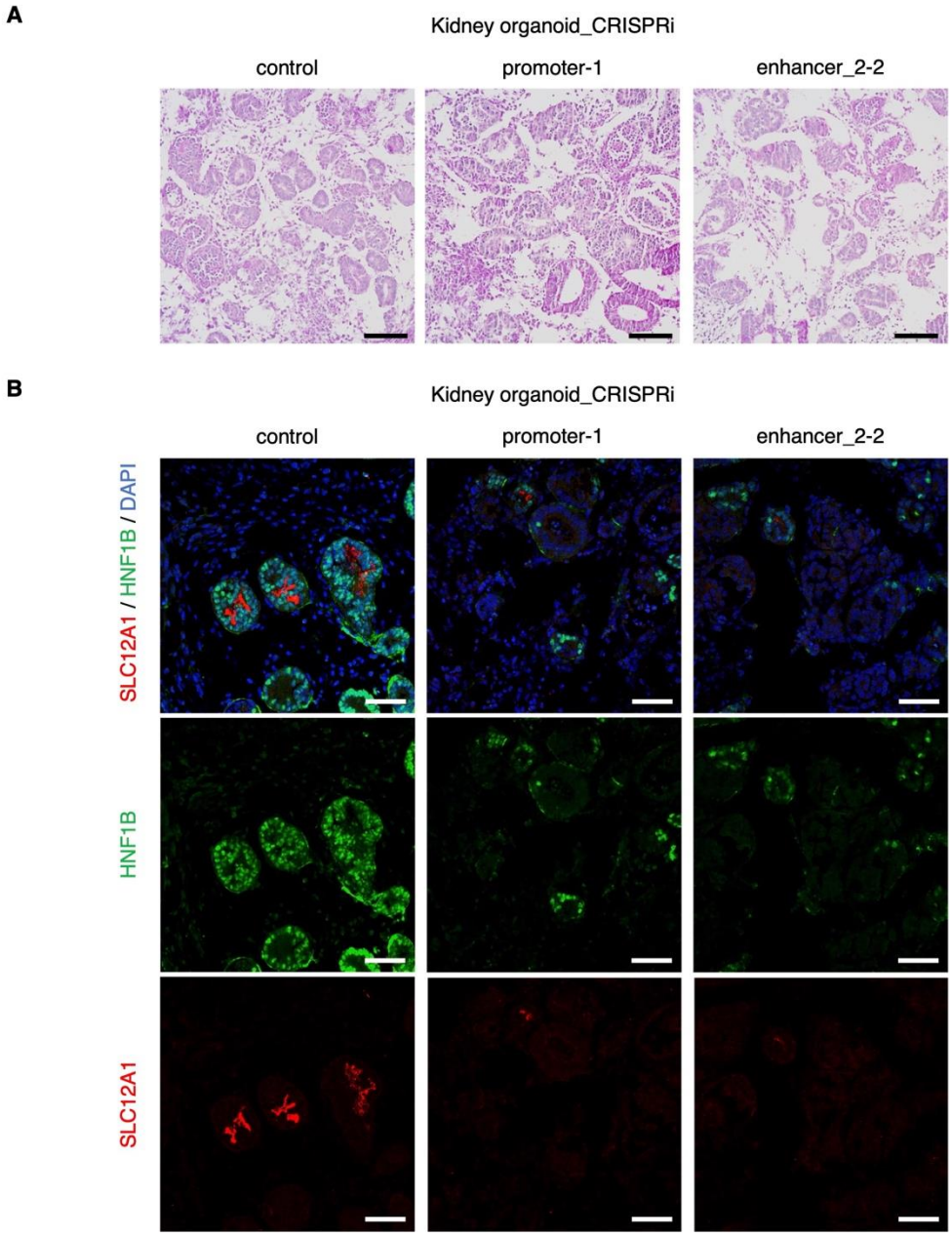

Figure S13. CRISPRi kidney organoids for *HNF1B* showed impaired PT and LOH differentiation

(A) PAS-stained images of kidney organoids (day 26) with CRISPRi for *HNF1B* using control vector (dCAS9-KRAB without sgRNA, left), sgRNA targeting promoter (middle), and sgRNA targeting enhancer_2 (right). Scale bars indicate 100 µm.

(B) Immunofluorescence images of SLC12A1 (red, loop of Henle), HNF1B (green), and nuclear DAPI (blue) staining in kidney organoids (day 26) with CRISPRi for HNF1B using control vector (dCAS9-KRAB without sgRNA, left), sgRNA targeting promoter (middle), and sgRNA targeting enhancer_2 (right). Scale bars indicate 50 µm.

Table S1. Quality control filtration of kidney organoid differentiation multiome (scATAC-seq and scRNA-seq) samples.

Sample ID	Cell line	Day	nCount_RNA	nCount_ATAC	percent.mt	nucleosome_signal	TSS.enrichment	filtered cell number	doublet cell number	filtered single cell number
BJFF.6_d26	BJFF.6 hiPSC	d26	<40000, >1000	<100000, >1000	<30	<2.0	>1.0	3967	204	3763
AN1.1_d7	AN1.1 hiPSC	d7	<25000, >1000	<125000, >1000	<20	<2.0	>1.0	1570	49	1521
AN1.1_d12	AN1.1 hiPSC	d12	<20000, >1000	<100000, >1000	<25	<2.0	>1.0	3925	151	3144
AN1.1_d19	AN1.1 hiPSC	d19	<30000, >1000	<150000, >1000	<30	<2.0	>1.0	1183	33	1150
AN1.1_d26	AN1.1 hiPSC	d26	<30000, >1000	<100000, >1000	<20	<2.0	>1.0	5129	313	4816
H9_d7	H9 hESC	d7	<30000, >1000	<50000, >1000	<25	<2.0	>1.0	7536	613	6923
H9_d12	H9 hESC	d12	<25000, >1000	<50000, >1000	<25	<2.0	>1.0	8869	820	8049
H9_d16	H9 hESC	d16	<25000, >1000	<40000, >1000	<20	<2.0	>1.0	10443	1104	9339
H9_d19	H9 hESC	d19	<25000, >1000	<40000, >1000	<30	<2.0	>1.0	11150	1245	9905
H9_d26	H9 hESC	d26	<25000, >1000	<40000, >1000	<40	<2.0	>1.0	9117	862	8255
										56865
(reference1*)										
pbmc_granulocyte_sorted_10k	hPBMC	10x Genomics	<25000, >1000	<70000, >5000	<20					
(reference2*)										
pbmc_granulocyte_sorted_10k	hPBMC	10x Genomics	<25000, >1000	<100000, >1000		<2.0	>1.0			
*reference1: https://satijalab.org/seurat/articles/weighted_nearest_neighbor_analysis.html#wnn-analysis-of-10x-multiome-rna-atac-1										
*reference2: https://satijalab.org/signac/articles/pbmc_multiomic.html										

Table S2. Cluster annotation of kidney organoid multiome (scRNA-seq and scATAC-seq) samples.

Kidney organoid differentiation multiome dataset (Related to Figure 1)				Kidney organoid day 26 multiome dataset (Related to Figure 2)		
Cluster ID	**cell type**	**Full name of cell type**		**Cluster ID**	**cell type**	**Full name of cell type**
c0	ST	Stromal cell/Interstitial cell		c0	dev_TUB_1	developing renal tubule_1
c1	dev_TUB_1	developing renal tubule_1		c1	dev_TUB_1	developing renal tubule_1
c2	dev_POD	developing podocyte		c2	ST	Stromal cell/Interstitial cell
c3	dev_TUB_1	developing renal tubule_1		c3	POD	Podocyte
c4	PIM	Posterior intermediate mesoderm		c4	PT	Proximal tubule
c5	dev_TUB_1	developing renal tubule_1		c5	LOH	Loop of Henle
c6	dev_TUB_2	developing renal tubule_2		c6	dev_POD	developing podocyte
c7	ST	Stromal cell/Interstitial cell		c7	NEU	Neuron
c8	dev_TUB_1	developing renal tubule_1		c8	prolif	Proliferating cell
c9	NEU	Neuron		c9	dev_TUB_1	developing renal tubule_1
c10	ST	Stromal cell/Interstitial cell		c10	PEC	Glomerular parietal epithelial cell
c11	POD	Podocyte		c11	ST	Stromal cell/Interstitial cell
c12	dev_TUB_1	developing renal tubule_1		c12	dev_TUB_2	developing renal tubule_2
c13	PIM	Posterior intermediate mesoderm		c13	DN	Distal nephron
c14	dev_TUB_1	developing renal tubule_1		c14	dev_TUB_1	developing renal tubule_1
c15	unknown	unknown type of cell		c15	POD	Podocyte
c16	PIM	Posterior intermediate mesoderm		c16	MEL	Melanocyte
c17	LOH	Loop of Henle		c17	unknown	unknown type of cell
c18	DN	Distal nephron		c18	NEU	Neuron
c19	NM	Nascent mesoderm		c19	EC	Endothelial cell
c20	prolif	Proliferating cell		c20	ST	Stromal cell/Interstitial cell
c21	MEL	Melanocyte		c21	MUS	Muscle cell
c22	NEU	Neuron				
c23	PT	Proximal tubule				
c24	NEU	Neuron				
c25	MUS	Muscle cell				

Table S3. Top 10 transcription factors in each nephron segment of the kidney organoid differentiation multiome (scATAC-seq and scRNA-seq) samples.

celltype	rank	gene	RNA.auc	RNA.pval	motif.feature	motif.auc	motif.pval	avg_auc
PT	1	HNF4G	0.812701116	4.90E-217	ENSG00000164749-LINE3040-HNF4G-D-N1	0.98062005	8.80E-200	0.896660583
PT	2	HNF4A	0.748397507	0	ENSG00000101076-LINE2936-HNF4A-D-N7	0.984738925	3.48E-203	0.866568216
PT	3	HNF1B	0.745134156	1.15E-77	ENSG00000108753-LINE2249-HNF1B-D-N3	0.970507945	1.50E-191	0.857821051
PT	4	MITF	0.751479587	6.80E-76	ENSG00000187098-LINE295-MITF-D	0.862748204	1.12E-114	0.807113895
PT	5	TFEC	0.740631253	0	ENSG00000105967-LINE129-TFEC-D	0.861635726	5.51E-114	0.801133489
PT	6	HNF1A	0.60074904	3.17E-202	ENSG00000135100-LINE2344-HNF1A-D-N4	0.973206836	9.91E-194	0.786977938
PT	7	JUND	0.696422768	2.05E-74	ENSG00000130522-LINE411-JUND-D-N6	0.874323281	5.54E-122	0.785373024
PT	8	JUNB	0.720027099	1.12E-74	ENSG00000171223-LINE487-JUNB-D-N3	0.840711868	2.14E-101	0.780369484
PT	9	JUN	0.693707407	7.04E-46	ENSG00000177606-LINE517-JUN-D-N5	0.855186087	5.03E-110	0.774446747
PT	10	CEBPD	0.772886725	1.18E-167	ENSG00000221869-LINE558-CEBPD-D-N2	0.741274887	8.98E-52	0.757080806
LOH	1	HNF1B	0.731241737	3.80E-252	ENSG00000108753-LINE2249-HNF1B-D-N3	0.905459415	0	0.818350576
LOH	2	HNF1A	0.513443194	6.04E-15	ENSG00000135100-LINE2344-HNF1A-D-N4	0.905528158	0	0.709485676
LOH	3	ESRRG	0.853411788	0	ENSG00000196482-LINE3129-ESRRG-D-N2	0.563605454	1.45E-14	0.708508621
LOH	4	MITF	0.78271488	0	ENSG00000187098-LINE295-MITF-D	0.63105342	1.44E-56	0.70688415
LOH	5	POU3F3	0.767849502	2.89E-275	ENSG00000198914-LINE2689-POU3F3-D-N1	0.644128223	4.93E-68	0.705988863
LOH	6	ID4	0.54128823	7.69E-10	ENSG00000172201-LINE271-ID4-D	0.860336197	0	0.700812213
LOH	7	JUN	0.755453737	5.56E-286	ENSG00000177606-LINE517-JUN-D-N5	0.61030834	1.36E-40	0.682881038
LOH	8	JUNB	0.735465355	0	ENSG00000171223-LINE487-JUNB-D-N3	0.614215288	2.16E-43	0.674840322
LOH	9	GRHL1	0.516739433	2.70E-10	ENSG00000134317-LINE2139-GRHL1-D-N2	0.82626885	0	0.671504141
LOH	10	HOXD8	0.59148051	2.60E-42	ENSG00000175879-LINE2498-HOXD8-D	0.725323729	1.73E-163	0.65840212
dev_TUB_1	1	HNF1B	0.707172379	0	ENSG00000108753-LINE2249-HNF1B-D-N3	0.846148983	0	0.776660681
dev_TUB_1	2	ID4	0.562566735	1.24E-198	ENSG00000172201-LINE271-ID4-D	0.868989422	0	0.715778078
dev_TUB_1	3	POU3F3	0.817593762	0	ENSG00000198914-LINE2689-POU3F3-D-N1	0.605189453	0	0.711391608
dev_TUB_1	4	PAX2	0.790091674	0	ENSG00000075891-LINE2612-PAX2-D-N1	0.597809431	0	0.693950553
dev_TUB_1	5	HOXD8	0.584881622	0	ENSG00000175879-LINE2498-HOXD8-D	0.78625466	0	0.685568141
dev_TUB_1	6	HNF1A	0.503495181	5.98E-11	ENSG00000135100-LINE2344-HNF1A-D-N4	0.846112584	0	0.674803882
dev_TUB_1	7	SIM1	0.804205701	0	ENSG00000112246-LINE139-SIM1-I	0.538387786	1.07E-50	0.671296743
dev_TUB_1	8	LMO2	0.502506788	1.00E-05	ENSG00000135363-LINE3654-LMO2-D-N1	0.801442396	0	0.651974592
dev_TUB_1	9	GRHL1	0.51065147	1.99E-38	ENSG00000134317-LINE2139-GRHL1-D-N2	0.774700684	0	0.642676077
dev_TUB_1	10	ZBTB3	0.505717125	9.44E-15	ENSG00000185670-LINE1324-ZBTB3-I	0.721537229	0	0.613627177
dev_TUB_2	1	HNF1B	0.748988288	0	ENSG00000108753-LINE2249-HNF1B-D-N3	0.886849757	0	0.817919022
dev_TUB_2	2	HNF1A	0.527141365	4.17E-132	ENSG00000135100-LINE2344-HNF1A-D-N4	0.886531935	0	0.70683665
dev_TUB_2	3	LHX1	0.681102542	0	ENSG00000132130-LINE2332-LHX1-I	0.62293005	7.55E-118	0.652016296
dev_TUB_2	4	ID4	0.623654955	6.63E-180	ENSG00000172201-LINE271-ID4-D	0.673658485	3.77E-233	0.64865672
dev_TUB_2	5	EMX2	0.734914949	0	ENSG00000170370-LINE2476-EMX2-D-N1	0.548433064	9.65E-20	0.641674006
dev_TUB_2	6	CEBPD	0.682203823	0	ENSG00000221869-LINE558-CEBPD-D-N2	0.594778429	7.91E-71	0.638491126
dev_TUB_2	7	PAX2	0.672974375	2.80E-258	ENSG00000075891-LINE2612-PAX2-D-N1	0.596635657	1.48E-73	0.634805016
dev_TUB_2	8	PAX8	0.657704377	1.00E-207	ENSG00000125618-LINE3159-PAX8-D	0.595320051	1.28E-71	0.626512214
dev_TUB_2	9	MITF	0.656119435	5.28E-257	ENSG00000187098-LINE295-MITF-D	0.593406234	7.60E-69	0.624762834
dev_TUB_2	10	HNF4G	0.639257282	0	ENSG00000164749-LINE3040-HNF4G-D-N1	0.55302965	2.38E-23	0.596143466
DN	1	GATA3	0.838721321	0	ENSG00000107485-LINE2094-GATA3-D-N6	0.588589009	4.89E-23	0.713655165
DN	2	HNF1B	0.662114196	1.15E-106	ENSG00000108753-LINE2249-HNF1B-D-N3	0.752657001	7.95E-175	0.707385599
DN	3	GRHL1	0.546548516	4.87E-59	ENSG00000134317-LINE2139-GRHL1-D-N2	0.83038942	1.89E-297	0.688468968
DN	4	TFAP2A	0.631773595	6.81E-203	ENSG00000137203-LINE29-TFAP2A-D-N3	0.690646041	2.13E-100	0.661209818
DN	5	PAX2	0.767926759	3.57E-219	ENSG00000075891-LINE2612-PAX2-D-N1	0.539840882	8.79E-06	0.65388382
DN	6	TFAP2B	0.557743609	4.03E-44	ENSG00000008196-LINE2-TFAP2B-D-N1	0.72648207	7.05E-141	0.64211284
DN	7	FOS	0.672695697	1.64E-119	ENSG00000170345-LINE476-FOS-D-N4	0.570245928	4.60E-15	0.621470813
DN	8	MGA	0.526714078	0.00149738	ENSG00000174197-LINE321-MGA-D-N1	0.689598872	2.55E-99	0.608156475
DN	9	JUNB	0.611649044	3.85E-61	ENSG00000171223-LINE487-JUNB-D-N3	0.600844892	2.28E-29	0.606246968
DN	10	ELF3	0.607042745	3.58E-111	ENSG00000163435-LINE1919-ELF3-D-N2	0.594506767	5.40E-26	0.600774756
POD	1	MAFB	0.941392574	0	ENSG00000204103-LINE553-MAFB-D	0.832346491	0	0.886869532
POD	2	WT1	0.949069215	0	ENSG00000184937-LINE1305-WT1-D	0.619411505	1.51E-91	0.78424036
POD	3	TCF21	0.802480627	0	ENSG00000118526-LINE152-TCF21-I	0.668088959	1.88E-179	0.735284793
POD	4	MAFG	0.54577067	6.22E-48	ENSG00000197063-LINE538-MAFG-D-N1	0.905173267	0	0.725471969
POD	5	TEAD4	0.541620775	1.14E-58	ENSG00000197905-LINE3584-TEAD4-D-N1	0.894528121	0	0.718074448
POD	6	TEAD3	0.572688141	2.31E-78	ENSG00000007866-LINE3572-TEAD3-D-N1	0.855736714	0	0.714212427
POD	7	MAFK	0.516912462	1.44E-42	ENSG00000198517-LINE543-MAFK-D-N5	0.896573979	0	0.706743221
POD	8	FOXP1	0.675897111	6.97E-200	ENSG00000114861-LINE1957-FOXP1-D-N1	0.661914843	1.17E-166	0.668905977
POD	9	FOXJ2	0.513163142	3.60E-09	ENSG00000065970-LINE1948-FOXJ2-D-N2	0.788730565	0	0.650946853
POD	10	MYF5	0.500656235	0.00042693	ENSG00000111049-LINE482-MYF5-I-N8	0.797932648	0	0.649294442
dev_POD	1	WT1	0.77552174	0	ENSG00000184937-LINE1305-WT1-D	0.6240368	1.65E-157	0.69977927
dev_POD	2	PAX8	0.758852527	0	ENSG00000125618-LINE3159-PAX8-D	0.593622026	1.40E-90	0.676237276
dev_POD	3	PAX2	0.591788319	3.68E-97	ENSG00000075891-LINE2612-PAX2-D-N1	0.622968153	7.69E-155	0.607378236
dev_POD	4	HMGA1	0.527670255	4.97E-11	ENSG00000137309-LINE57-HMGA1-D	0.669346743	8.68E-292	0.598508499
dev_POD	5	EGR3	0.504432229	8.29E-09	ENSG00000179388-LINE1241-EGR3-D-N2	0.675564454	0	0.589998341
dev_POD	6	ZNF219	0.516084053	8.43E-09	ENSG00000165804-LINE1060-ZNF219-D	0.654546199	2.26E-243	0.585315126
dev_POD	7	FOXB1	0.518483739	1.47E-99	ENSG00000171956-LINE2021-FOXB1-D-N2	0.647361301	1.84E-221	0.58292252
dev_POD	8	FOXM1	0.507873903	0.00107998	ENSG00000111206-LINE1956-FOXM1-D	0.653930202	1.87E-241	0.580902053
dev_POD	9	MAZ	0.539209213	4.17E-37	ENSG00000103495-LINE797-MAZ-D	0.607427209	1.19E-118	0.573318211
dev_POD	10	ID3	0.626539191	5.78E-179	ENSG00000117318-LINE151-ID3-I	0.517805487	0.00012384	0.572172339

Dataset S1 (separate file). Quality control metrics of kidney organoid differentiation multiome (scATAC-seq and scRNA-seq) samples.

SI References

1. Y. Muto, *et al.*, Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nat Commun* **12**, 2190 (2021).

2. D. van Dijk, *et al.*, Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716-729.e27 (2018).

3. Z. E. Tóth, É. Mezey, Simultaneous Visualization of Multiple Antigens with Tyramide Signal Amplification using Antibodies from the same Species. *J Histochem Cytochem.* **55**, 545–554 (2007).

4. K. Labun, *et al.*, CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research* **47**, W171–W174 (2019).

5. M. Nakagawa, *et al.*, A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. *Sci Rep* **4**, 3594 (2014).

6. M. R. Corces, *et al.*, An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* **14**, 959–962 (2017).