

Supplemental Methods

Sample procurement

Kidney samples were obtained from surgical nephrectomies. Nephrectomies were de-identified, and the corresponding clinical information was collected through an honest broker; therefore, no consent was obtained from the subjects. Collected tissue was immersed in RNAlater (Ambion#AM7020) and manually microdissected for glomerular and tubular compartments. In general, 60-150 glomeruli that readily released from the capsule were collected and placed into RNeasy RNA Tissue Lysis Buffer Solution from Qiagen RNeasy kit (Qiagen#74106). We refer to the remaining compartment as tubule throughout the article.

Part of the tissue core was formalin fixed and paraffin embedded. These samples were later sectioned and stained with Hematoxylin eosin or periodic acid Schiff. Our local renal pathologist performed an unbiased review of the tissue section by scoring multiple parameters.

DNA was isolated using the Qiagen DNAeasy or MagAttract High Molecular Weight DNA Kit (Qiagen#67563) according to the manufacturer's instructions. DNA was quantified by the Invitrogen Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen#P11496).

Genotyping and RNASeq

Genomic DNA isolated from whole kidney or tubule tissue was used for genotyping using Affymetrix Axiom Biobank genotyping array. Quality control steps were performed using PLINK 1.9 (1). To quantify population structure, Principal Component Analysis (PCA) implemented in EIGENSTRAT (2) was conducted on these final 408 samples, with additional genotype data from 395 HapMap2 samples (112 CEU, 84 CHB, 86 JPT, 113 YRI) and 2504 samples from the 1,000 Genomes Project Phase 1 (release v3, 661 AFR, 347 AMR, 504 EAS, 503 EUR, 489 SAS), respectively. Subsequently, genotypes

of the 408 samples were phased with SHAPEIT2 (3) and imputed with IMPUTE2, using multi-ethnic panel reference from 1,000 Genomes Phase 1 v3.

RNA was isolated from glomeruli and tubule compartments separately. RNA quality was assessed with the Agilent Bioanalyzer 2100. Only samples with RNA integrity number (RIN) scores above 7 and a minimum total RNA of 100 ng were used for cDNA production. Non-strand specific, polyA+ selected RNA-seq libraries were generated using the Illumina TruSeq protocol. Libraries were sequenced to a median depth of 35 million 100-bp single-end reads. Sequence quality was assessed using FastQC and then the adaptor and lower-quality bases were trimmed using Trim-galore. Samples with less than 15 million mapped reads were excluded.

Only samples with European ancestry and absence of significant kidney structural changes (tubular fibrosis < 10%, glomerular sclerosis < 10%) were used. Variants were excluded from analysis if they had: (1) call rate < 95%; (2) Minor allele frequency (MAF) < 5%; (3) deviated from Hardy-Weinberg Equilibrium ($P < 10^{-6}$); and (4) imputation info score < 0.4. Finally, 121 samples were used for tubule TWAS analysis.

RNA-seq reads were aligned to the human genome (hg19/GRCh37) using STAR (v2.4.1d) in the 2-pass mode based on GENCODE v19 annotations. Low mapping quality reads (MAPQ < 10) were removed using samtools (v1.1) .

Meta-analysis of multiple-tissue eQTL mapping

The eQTL summary results of 44 other human tissues were downloaded from GTEx (v6p). METASOFT, a meta-analysis method, was performed on all variant-gene pairs that were significant (FDR < 5%) in at least one of the 46 tissues (2 kidney compartments and 44 GTEx tissues). A random effects model in METASOFT (called RE2) was used and the posterior probability that an eQTL effect in a given tissue (called m) was calculated for each SNP-gene pair and tissue tested. A significance cutoff of $m > 0.9$ was used to discover high-confidence eQTLs.

Human kidney snATACseq

Fresh human kidneys were collected after surgical nephrectomies, transferred from hospital to lab in RPMI on ice. Kidneys were minced and lysed in 5 mL lysis buffer for 15 min. Tissue lysis reactions were then blocked by adding 10 mL 1x PBS into each tube, and solution was passed through 40 μ m cell strainers. Cell debris and cytoplasmic contaminants were removed by Nuclei PURE Prep Nuclei Isolation Kit (Sigma, NUC-201) after centrifugation at 13,000 RPM for 45 min. Nuclei concentration was calculated by Countess AutoCounter (Invitrogen, C10227). Diluted nuclei suspension was loaded and incubated in transposition mix from Chromium Single Cell ATAC Library & Gel Bead Kit (10X Genomics, PN-1000110) by targeting 10,000 nuclei recovery. GEMs were then captured on the Chromium Chip E (10x Genomics, PN-1000082) in the Chromium Controller according to the manufacturer's protocol (10X Genomics, CG000168). Libraries were generated using the Chromium Single Cell ATAC Library & Gel Bead Kit and Chromium i7 Multiplex Kit N (10X Genomics, PN-1000084) according to the manufacturer's manual. Quality control for constructed library was performed by Agilent Bioanalyzer High Sensitivity DNA kit. The library was sequenced on an Illumina HiSeq 2x50 paired-end kits using the following read length: 50 bp Read1 for DNA fragments, 8 bp i7 index for sample index, 16 bp i5 index for cell barcodes and 50 bp Read2 for DNA fragments.

Mouse kidney snATAC-seq and CICERO analysis

Mouse kidney snATAC-seq data were obtained from (4). This dataset consisted of mouse kidney samples from P0 and adult, and cells were assigned to cell types before peak-calling. The open chromatin landscapes were visualized using IGV. CICERO (5) analysis was conducted on the P0 samples to infer the peak-peak co-accessibilities. A heuristic cutoff 0.25 was implemented to filter the co-accessible peaks.

Human kidney tubule microarray samples

95 tubule compartments of human kidney samples were microdissected from nephrectomies as previously described (ArrayExpress: E-MTAB-5929, E-MTAB-2502)(6). Correlation analysis between eGFR and gene expression were performed using a linear regression model adjusted for gender, age, and race.

Antibodies

The following antibodies were used; rabbit polyclonal antibody against STMN1 (#3352, CST), Dach1 (10914-1-AP, Proteintech), rabbit monoclonal antibody against Ki67 (clone D3B5, #12202; CST), GAPDH (clone 14C10, #2118, CST), mouse polyclonal antibody against UMOD (SAB1400296, Sigma), mouse monoclonal antibody against aquaporin 1 (clone 7D11, ab11205, Abcam 500:1), CDH16 (clone E-7, sc-393153, Santa Cruz), Aquaporin2 (clone E-2, sc-515770, Santa Cruz), Calbindin (clone 401025, MA5-24135, Thermo Fisher), rat monoclonal antibody against F4/80 (clone A3-1, ab6640, Abcam).

Histology

Kidney tissue was fixed in 10% formalin. 5- μ m thick Paraffin-embedded sections were used. Periodic acid Schiff (PAS) staining was used to evaluate kidney injury. Sirius red staining (Polysciences) were performed according to manufacturer's protocol to evaluate kidney fibrosis.

For immunohistochemical and immunofluorescence staining, sections were deparaffinized using xylene, 100% alcohol, 95 alcohol, 70% alcohol, and distilled water. Epitope retrieval was performed using citrate buffer pH6.0 at 95 degree for 10 minutes or proteinase K (20 μ g/ml) at RT for 10 minutes. Sections were immersed in 3% H₂O₂ for blocking endogenous peroxidase for 10 minutes. After blocking nonspecific binding with blocking buffer, the following primary antibodies diluted in blocking buffer were applied

at 4 degree overnight; anti-Stmn1 (#3352, CST, 1:50), anti-Ki67 (#9027, CST, 1:500), anti-Dach1 (Proteintech, 10914-1-AP, 1:500), anti-AQP1 (Abcam, ab11025, 1:50). Anti-AQP2 (Santa Cruz, sc-515770, 1:50), anti-calbindin (Thermo fisher, MA524135, 1:500), anti-UMOD (Sigma, SAB1400296, 1:500), anti-CDH16 (Santa Cruz, sc-393153, 1:50), anti-F4/80 (Abcam, ab6640, 1:100), FITC-conjugated LTL (vector, FL-1321-2, 1:20). For immunohistochemical staining, VectaStain ABC Standard Kit and ImmPACT DAB Peroxidase (HRP) Substrate (Vector Laboratories) were used to visualize antibodies. Counterstaining were performed with Mayer modified Hematoxylin. For immunofluorescence staining, sections were incubated with the following secondary antibodies; Chicken anti-Mouse Alexa fluor 488 (Thermo fisher, 1:500), Donkey anti-Rabbit Alexa fluor 555 (Thermo fisher, 1:500), Goat anti-Rat Alexa fluor 488 (Thermo fisher, 1:500). Sections were mounted using ProLong Gold Antifade Mounting with DAPI (Life Sciences). The Ki67 and Stmn1 positive cells were counted at $\times 200$ magnification. 5 independent fields in kidney tissue were analyzed and the percentage of Sirius Red and F4/80 positive area was quantified using Image J (National Institutes of Health, NIH).

Immunoblot

Proteins from primary tubular epithelial cell (PTEC) were collected using RIPA buffer (#9806, CST). Samples were incubated on ice for 30 minutes with regular vortexing and briefly sonicating. For SDS-PAGE preparation, 4x Laemmli Sample Buffer (#1610747 Bio-Rad) containing β -mercaptoethanol were added, and samples were boiled at 95 degree for 5 minutes. Separated proteins by SDS-PAGE were transferred onto 0.2 μm pore size PVDF membrane. After blocking with 5% non-fat dry milk, membranes were incubated with anti-Dach1 antibody (Proteintech, 1:1000), anti-GAPDH antibody (CST, 1:1000) at 4 degree overnight. Membranes were incubated with appropriate secondary

antibodies conjugated with HRP, and signals were detected using ECL Western Blotting Substrate (Thermo Fisher).

Real time Quantitative PCR

Mouse kidneys were homogenized and total RNA were extracted using Trizol, according to manufacturer's protocol (Thermo fisher). Quality and concentration of extracted RNA were examined using nanodrop. cDNA was generated using Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's protocol. Real time quantitative PCR were performed using Cyber green Master Mix Reagents (Thermo Fisher) with ViiA 7 System (Life Technologies) instrument. Primer sequence can be found in **Supplemental table 8**.

MTT assay

The MTT assay was performed according to manufacturer's instructions (Ab211091, abcam). Primary renal tubule cells from *Dach1^{flox/flox}* mice were seeded in 96 well plate at the density of 1×10^5 cells/well. The next day, adenovirus-GFP or adenovirus-Cre-GFP were added. To evaluate cell numbers, renal tubular cells were incubated with MTT reagents at 37 degree for 3hours and analyzed by measurement absorbance at 590 nm at; 0-hour, 24-hour and 48-hour after adenoviral infection.

Chemotaxis assay

RAW 264.7 were cultured with Dulbecco's Modified Eagle's Medium containing 10% FBS and 1% penicillin/streptomycin. Cells were maintained at 37 degree in 5% CO₂ incubator.

Cell migration of Raw 264.7 cell was evaluated by Chemotaxis assay (polycarbonate membrane, 5µm pore size) according to manufacturer's protocol (Cell Biolabs). Supernatant of the primary renal tubules isolated from 3- to 4-week-old *KSP^{Cre}*

Dach1^{Flox/Flox} mice 2 days after adeno-GFP or adeno-Cre-GFP infection were collected and added to the feeder tray. Cell suspension containing 1×10^5 Raw 264.7 cells in serum free medium were added to the upper chamber. After incubation for 2 or 24 hours, migrated cells were dissociated from the polycarbonate membrane by detachment buffer, and lysed with lysis buffer containing fluorescent dye. Fluorescence was measured in the fluorescence plate reader at 480nm/520nm. Relative fluorescence units (RFU) of samples were calculated by subtracting blank value.

Functional analysis

Gene ontology analysis for genes correlated with *DACH1* was performed using DAVID (7). *DACH1* binding peaks identified by ChIP-seq in K562 cell were download from the ENCODE portal (8) with the identifier ENCFF526HLP. GREAT software was used to analyze functional enrichment of *DACH1* binding peaks in pathways from Molecular Signatures Database (MsigDB). Full list of enriched MsigDB pathways is in **supplemental table 11**.

Web resource

Mouse kidney snATAC-seq database; <http://susztaklab.com/igv/>

Human kidney eQTL atlas; <http://susztaklab.com/eqtl/>

TWAS-FUSION software; <http://gusevlab.org/projects/fusion/>

DAVID; <https://david.ncifcrf.gov/>

GREAT; <http://great.stanford.edu/public/html/>, version 3.0.0

ENCODE; <https://www.encodeproject.org/>

Seurat 3; <https://satijalab.org/seurat/>

Cell Ranger 1.3; <http://10xgenomics.com>

PLINK1.9; <http://zzz.bwh.harvard.edu/plink/download.shtml>

EIGENSTRAT; <https://data.broadinstitute.org/alkesgroup/EIGENSOFT/>

SHAPEIT2; https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html

IMPUTE2; https://mathgen.stats.ox.ac.uk/impute/impute_v2.html

Trim Galore; https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

STAR (v2.4.1d); <https://github.com/alexdobin/STAR>

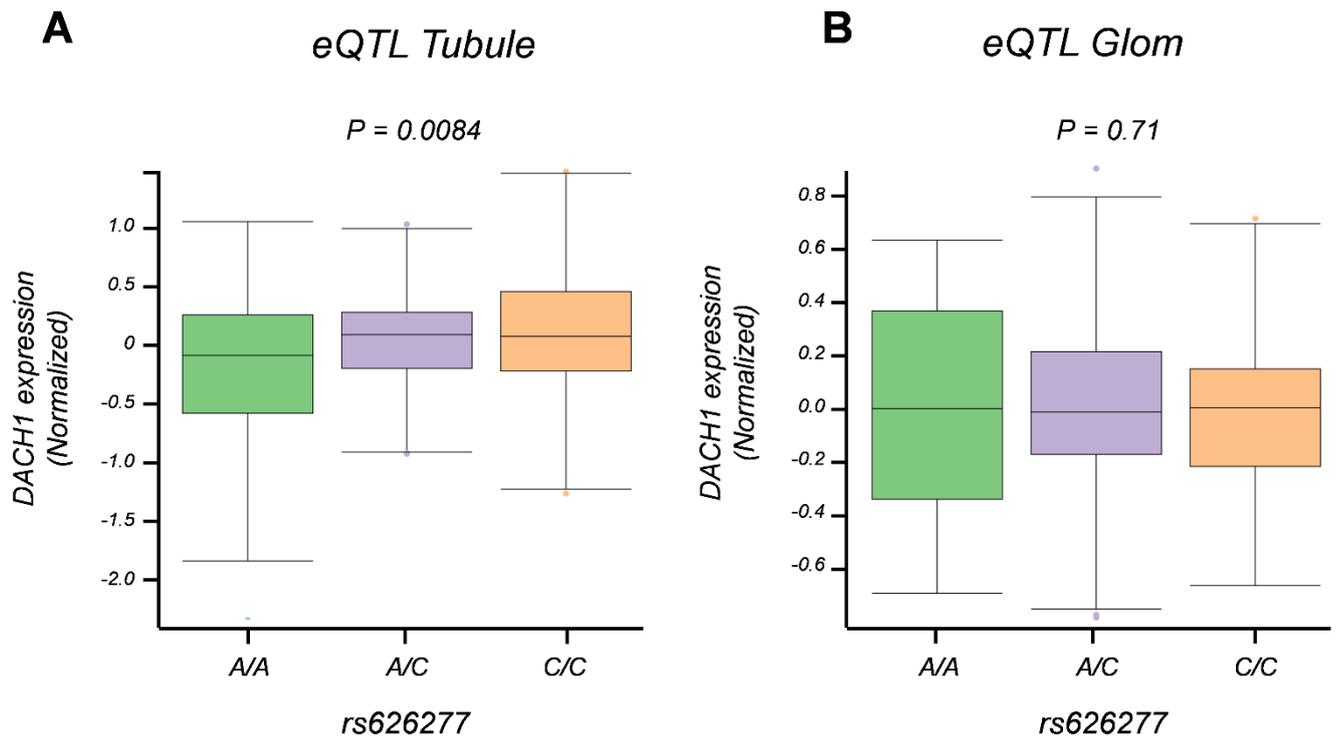
Samtools; <http://www.htslib.org/>

Gtex; <https://www.gtexportal.org/home/>

CRISPOR; <http://crispor.tefor.net/>

1. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*. 2007;81(3):559-75.
2. Sankararaman S, Sridhar S, Kimmel G, and Halperin E. Estimating local ancestry in admixed populations. *Am J Hum Genet*. 2008;82(2):290-303.
3. Delaneau O, Marchini J, and Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods*. 2011;9(2):179-81.
4. Miao Z, Balzer MS, Ma Z, Liu H, Wu J, Shrestha R, et al. Single cell resolution regulatory landscape of the mouse kidney highlights cellular differentiation programs and renal disease targets. *bioRxiv*. 2020:2020.05.24.113910.
5. Pliner HA, Packer JS, McFaline-Figueroa JL, Cusanovich DA, Daza RM, Aghamirzaie D, et al. Cicero Predicts cis-Regulatory DNA Interactions from Single-Cell Chromatin Accessibility Data. *Mol Cell*. 2018;71(5):858-71 e8.
6. Beckerman P, Qiu C, Park J, Ledo N, Ko YA, Park AD, et al. Human Kidney Tubule-Specific Gene Expression Based Dissection of Chronic Kidney Disease Traits. *EBioMedicine*. 2017;24:267-76.
7. Huang da W, Sherman BT, and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2009;4(1):44-57.
8. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic acids research*. 2018;46(D1):D794-d801.

Supplemental Figure 1

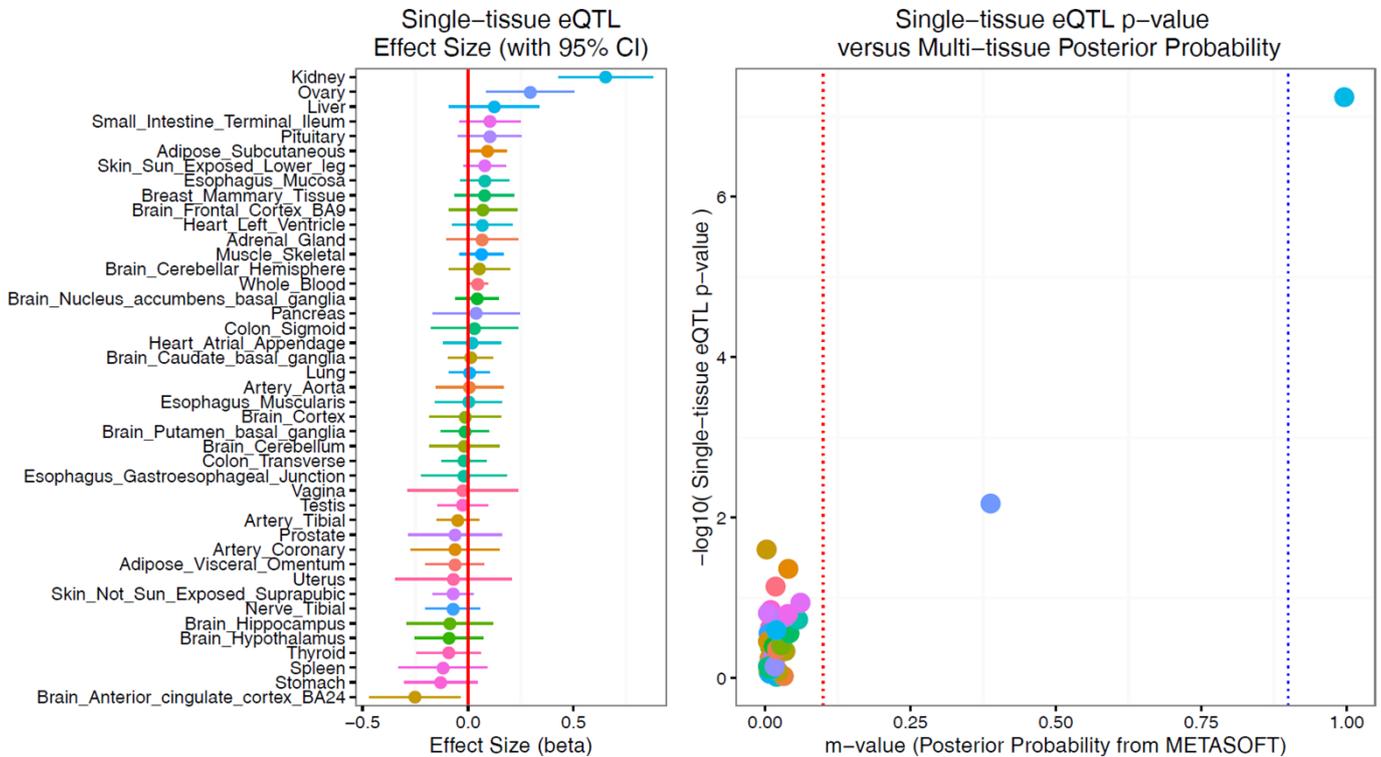


Supplemental figure 1. Cis-eQTL effect of the glomerular and tubulointerstitial tissues of the kidney in human nephrotic syndrome (NEPTUNE cohort).

(A) The box plot of cis-eQTL effect of index SNP (rs626277) in human kidney tubulointerstitial samples, showing correlation between genotype (A/A, A/C, C/C) and normalized *DACH1* expression ($p = 0.0084$). (B) The box plot of cis-eQTL effect of index SNP (rs626277) in human kidney glomerular samples, showing lack of correlation between genotype (A/A, A/C, C/C) and normalized *DACH1* expression ($p = 0.71$). Tubule: tubulointerstitial, Glom; glomerular, eQTL; expression quantitative trait loci.

Supplemental Figure 2

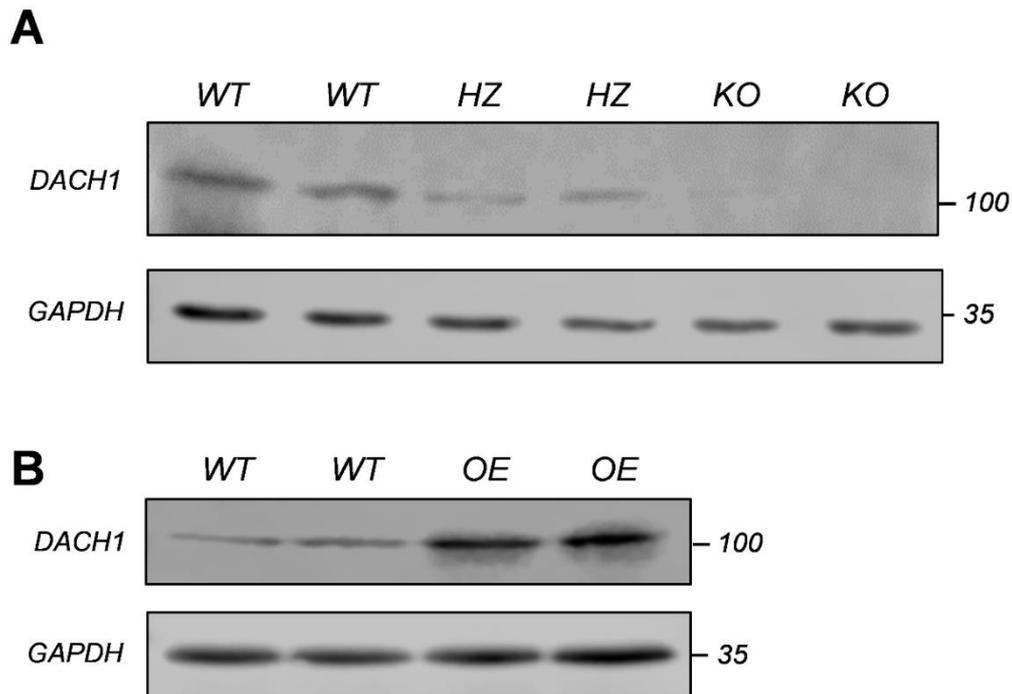
rs626277 vs DACH1



Supplemental figure 2. Multi-tissue eQTL comparison for *DACH1* expression (rs626277).

(Left) The effect size of single tissue eQTL with 95% confidence interval. The organs were arranged according to the effect size. (Right) Single-tissue eQTL p-value vs. multi-tissue posterior probability. The Y axis indicates $-\log_{10}$ single-tissue eQTL p-value. X axis indicates m-value (0-1), which indicates posterior probability of eQTL effect in each tissue. M-value < 0.1 indicates no significant eQTL effect in tissue, while m-value > 0.9 indicates significant eQTL effect in tissue.

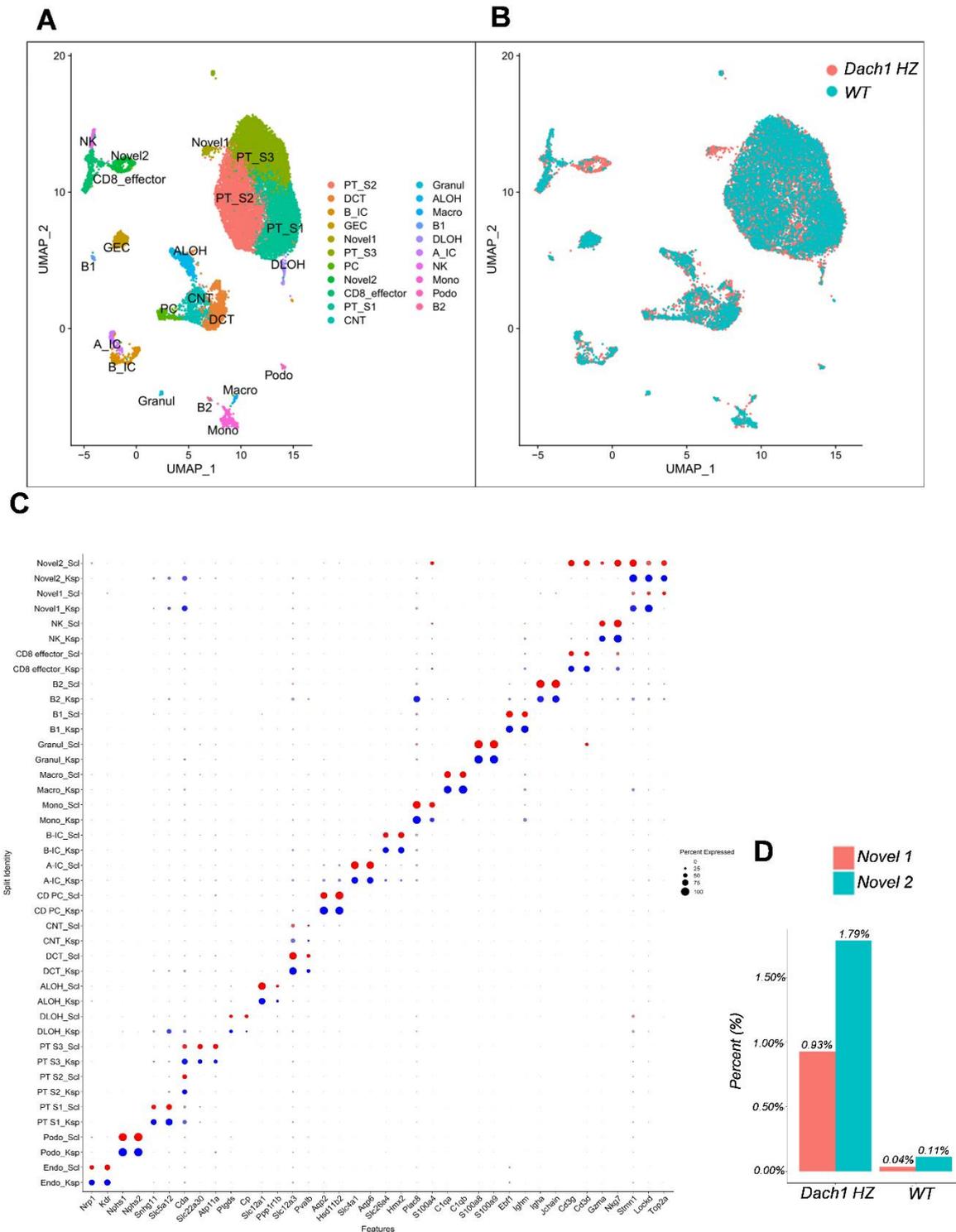
Supplemental Figure 3



Supplemental figure 3. DACH1 protein expression in kidneys of mice with tubule-specific *Dach1* deletion and overexpression.

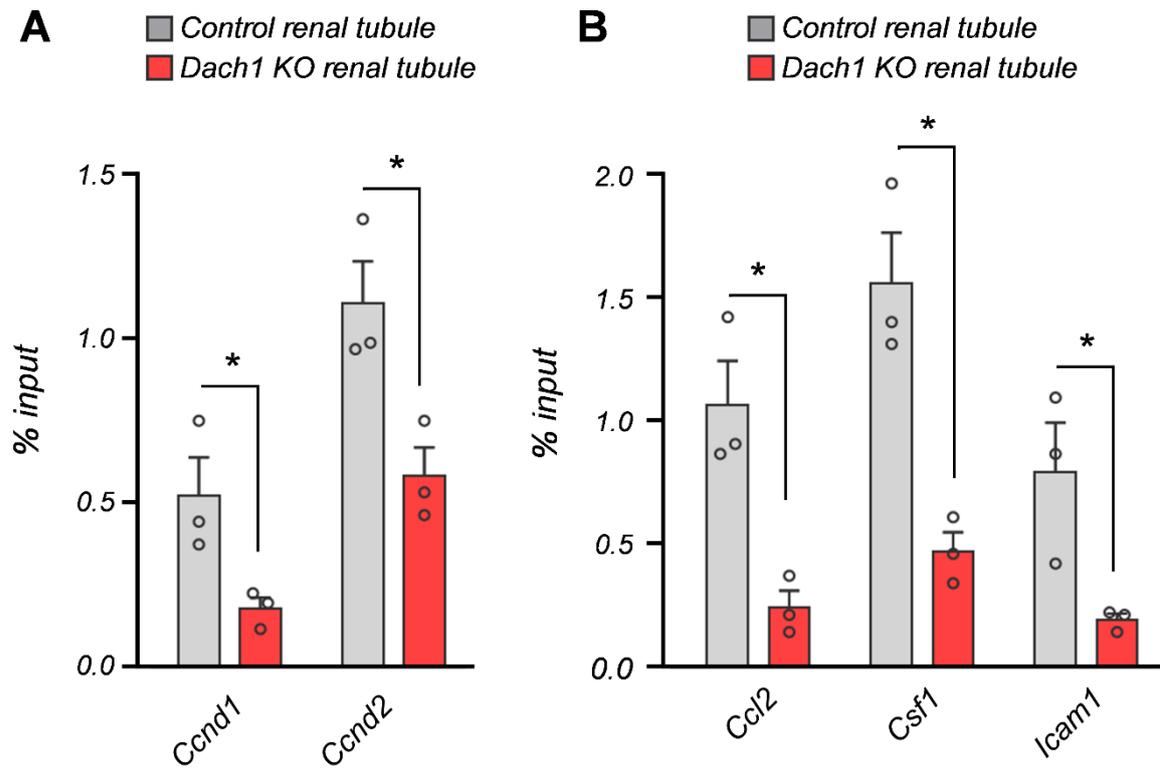
(A) The representative image of DACH1 protein expression in whole kidney lysate from wild type (WT), *Ksp^{Cre}/Dach1^{Flox/WT}* (*Dach1* HZ) and *Ksp^{Cre}/Dach1^{Flox/Flox}* (*Dach1* KO) mice. GAPDH was used as loading control. (B) The representative image of DACH1 protein expression in whole kidney lysate from wild type (WT), *Pax8rtTA/TRE-Dach1* (*Dach1* OE) mice. GAPDH was used as loading control.

Supplemental Figure 4



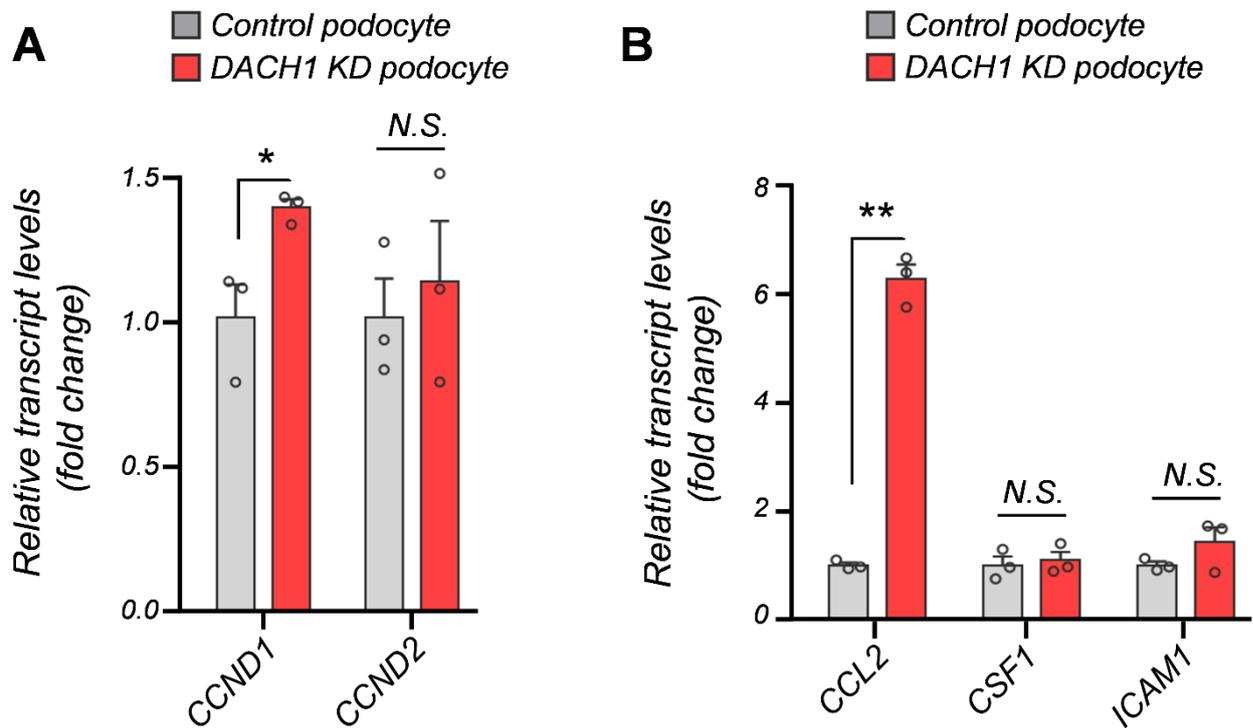
Supplemental figure 4. Single cell analysis of wild type (WT) and $Ksp^{Cre}/Dach1^{Flox/WT}$ (*Dach1* HZ) mice. (A) Dimension reduction (UMAP) clustering. The following cell clusters were identified. Endo: endothelial cells, Podo; podocytes, PT-S1, S2, S3: proximal tubule S1, S2 and S3 segments. DLOH, ALOH: descending and ascending loop of Henle, DCT: distal convoluted tubules, CNT: connecting tubule, PC: principal cells of the collecting tubule, A-IC, B-IC: A-type and B-type intercalated cells, Mono: monocytes, Macro; macrophages, granul: granulocytes, B1,B2: B-lymphocytes, CD8 effector: CD8-effector T-cells, NK: natural killer cells, Novel1,2: novel cell types. (B) Batch effect of wild type (WT) and $Ksp^{Cre}/Dach1^{Flox/WT}$ (*Dach1* HZ) mice. Light blue dots indicate the cells from WT mice, and light red dots indicate the cells from *Dach1* HZ mice. (C) Bubble plot for wild type (WT, red circle) and $Ksp^{Cre}/Dach1^{Flox/WT}$ (*Dach1* HZ, blue circle) mice. Y axis indicates cluster of specific cell type. Scl; WT, Ksp; *Dach1* HZ. X axis indicates representative cell marker of each cell cluster. (D) The fraction of Novel 1 (light red bars) and Novel 2 (light blue bars) is shown as percentage (%) in $Ksp^{Cre}/Dach1^{Flox/WT}$ (*Dach1* HZ) mice and wild type (WT) mice.

Supplemental Figure 5



Supplemental figure 5. DACH1-ChIP-PCR results in *Dach1* deficient mouse kidney tubule cells. (A, B) DACH1-ChIP-PCR of promoter region of cell cycle genes; *Ccnd1*, *Ccnd2* (A), cytokine genes; *Ccl2*, *Csf1* and *Icam1* (B) in control (gray bars) and *Dach1* deficient (red bars) cultured mouse kidney tubule cells. The y axis is presented as percent input (n= 3). N.S. not significant. *P < 0.05.

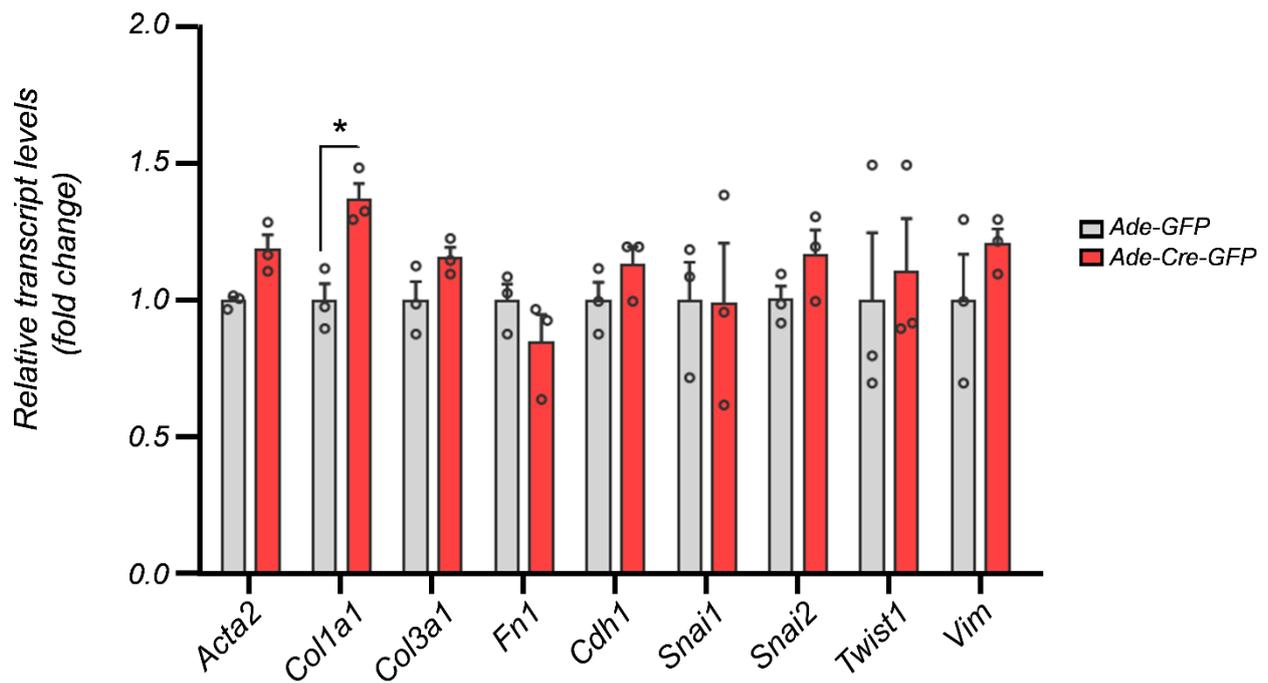
Supplemental Figure 6



Supplemental figure 6. Cell cycle and inflammatory gene expression change in cultured *DACH1*-null human podocytes.

(A) The relative gene expression of *CCND1*, *CCND2* in control (gray bars) or *DACH1* knock down (KD) (red bars) human podocytes (n=3, independent experiments). *GAPDH* was used as internal control. N.S.; not significant, *P < 0.05. (B) The relative gene expression of *CCL2*, *CSF1* and *ICAM1* in control (gray bars) or *DACH1* KD (red bars) human podocyte (n=3, independent experiments). *GAPDH* was used as internal control. N.S.; not significant, **P < 0.01.

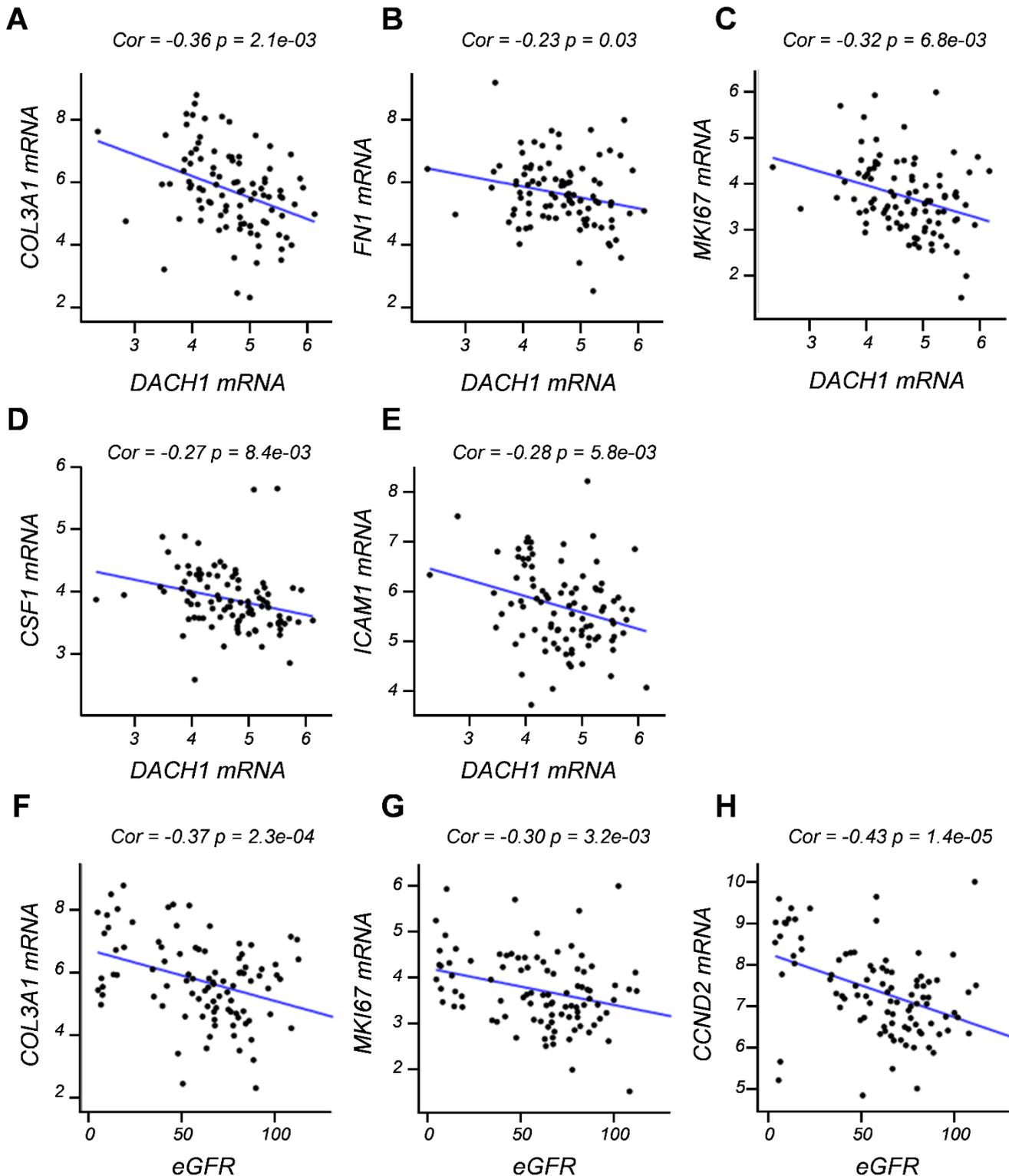
Supplemental Figure 7



Supplemental figure 7. Expression of genes associated with epithelial-to-mesenchymal transition (EMT)

The relative gene expression of *Acta2*, *Col1a1*, *Col3a1*, *Fn1*, *Cdh1*, *Snai1*, *Snai2*, *Twist1* and *Vim* in cultured kidney tubule cells from *Dach1^{Fllox/Fllox}* mice transfected with Ade-GFP (light gray bars) or Ade-Cre-GFP (red bars). Three independent experiments were performed. *Gapdh* was used as internal control. * $p < 0.05$. Ade; adenovirus.

Supplemental Figure 8



Supplemental figure 8. Correlation between *DACH1* expression and expression of markers of inflammation and proliferation in human kidney tubules.

(A-E) Correlation between normalized gene expression of *COL3A1* (A), *FN1* (B), *MKI67* (C), *CSF1* (D), *ICAM1* (E) and *DACH1* in human kidney tubule samples. (F-H) Correlation between normalized gene expression of *COL3A1* (F), *MKI67* (G) and *CCND2* (H) in human kidney tubule samples and eGFR (ml/min/1.73m²) levels. Cor; correlation coefficient. Correlation analysis was performed using a linear regression model adjusted for gender, age and race.