Single-Nucleus RNA Sequencing Identifies New Classes of Proximal Tubular Epithelial Cells in Kidney Fibrosis

Yueh-An Lu[1,2,3], Chia-Te Liao[1,4,5,6,7], Rachel Raybould[1,4,8], Bnar Talabani[1,2,4], Irina Grigorieva[1,2], Barbara Szomolay[1,4], Timothy Bowen[1,2], Robert Andrews[1,4], Philip R. Taylor[1,4,8], Donald Fraser[1,2*]

1.Division of Infection and Immunity, School of Medicine, Cardiff University, Cardiff, United Kingdom

2.Wales Kidney Research Unit, Cardiff University, Cardiff, United Kingdom

3.Division of Nephrology, Kidney Research Center, Linkou Chang Gung Memorial Hospital, Linkou, Taiwan

4.Systems Immunity University Research Institute, Cardiff University, Cardiff, United Kingdom

5.Division of Nephrology, Department of Internal Medicine, Taipei Medical University-Shuang Ho Hospital, Taiwan

6.Division of Nephrology, Department of Internal Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taiwan

7.TMU-Research Center of Urology and Kidney, Taipei Medical University, Taiwan

8.Dementia Research Institute, Cardiff University, Cardiff, United Kingdom

Correspondence

Professor Donald Fraser

Wales Kidney Research Unit, Cardiff University, School of Medicine, Heath Park, Cardiff CF14 4XN, United Kingdom

fraserdj@cardiff.ac.uk

Running Title: Proximal tubular single cell sequencing

Abstract word count: 235

Text word count: 4738

Key words: cell biology and structure, chronic kidney disease, epithelial, kidney tubule, mRNA, proximal tubule, renal epithelial cell, renal fibrosis, renal tubular epithelial cells.

Significance Statement (120 words)

Proximal tubular cells (PTCs) are central to normal kidney function, and to kidney regeneration versus organ fibrosis following injury. Single-nucleus RNA sequencing of fibrotic kidneys from mice exposed to aristolochic acid and naïve controls segregated PTCs to five abundant clusters, mapping to S1, S1-2, S2, S2-cortical S3, and medullary S3 segments. Four additional new PTC clusters were present, that showed increased abundance in fibrotic kidneys. The new PTC clusters exhibited gene expression signatures suggesting roles in renal injury responses and fibrosis progression. Pathway analysis revealed metabolic reprogramming, enrichment of cellular communication and cell motility, and immune activation. New PTC clusters also associated with fibrotic signalling to fibroblasts and inflammatory signaling to macrophages.

Abstract

Background: Proximal tubular cells (PTCs) are the most abundant cell type in the kidney. PTCs are central to normal kidney function, and to kidney regeneration versus organ fibrosis following kidney injury. This study determined PTC phenotype in renal fibrosis by single-nucleus RNA sequencing (snRNA-seq).

Method: Kidneys were harvested from naïve mice and mice with renal fibrosis induced by chronic aristolochic acid administration. Nuclei were isolated using Nuclei EZ Lysis buffer. Libraries were prepared on the 10X platform and snRNA-seq completed using the Illumina NextSeq 550 System. Genome mapping was carried out with high-performance computing.

Results: A total of 23,885 nuclei were analysed. PTC's were found in five abundant clusters, mapping to S1, S1-2, S2, S2-cortical S3, and medullary S3 segments. Additional cell clusters were present ("new PTC clusters") at low abundance in normal kidney and in increased number in kidneys undergoing regeneration/fibrosis following injury. These clusters exhibited clear molecular phenotypes, permitting labelling as, proliferating, New-PT1, New-PT2 and (present only following injury) New-PT3. Each of these clusters exhibited a unique gene expression signature, including multiple genes previously associated with renal injury response and

fibrosis progression. Comprehensive pathway analyses revealed metabolic reprogramming, enrichment of cellular communication and cell motility, and various immune activations in new PTC clusters. In ligand-receptor analysis, new PTC clusters promoted fibrotic signaling to fibroblasts and inflammatory activation to macrophages.

Conclusion: These data identify unappreciated PTC phenotype heterogeneity, and reveal novel PTCs associated with kidney fibrosis.

Introduction

The kidneys excrete waste and maintain homeostasis by blood filtration and subsequent selective reabsorption. The core functional unit of the kidney is the nephron, of which about one million are present in each human kidney. The nephron extends from the glomerulus, which is the site of blood filtration, to the collecting duct system through which nephrons converge to drain urine to the lower urinary tract. Ultrafiltrate is processed to urine via active transport, which underpins reabsorption of water, electrolytes, glucose, amino acids, and other constituents, and is balanced by secretion to achieve homeostasis. Highly specialized, sequential processing of urine occurs in discrete nephron segments, reflecting underlying cellular specialization(1).

The proximal tubules are the site for the majority of the resorptive activity, comprising active transport of an estimated 150 liters per day of solute-rich fluid in a healthy human(2). The scale of proximal tubular resorptive activity makes the kidney a highly metabolically active organ, second only in oxygen consumption to the brain(3). Proximal tubular cells (PTCs) are highly susceptible to ischemic and toxic injury, and their responses are central to progression and recovery in acute and chronic kidney injury of diverse etiology(4). PTCs play

a core role in dictating renal recovery and fibrosis responses following injury, and in chronic kidney disease (CKD)(5).

PTCs are a pre-eminent contributor to the overall cellular composition of the kidney(6). Single-cell RNA sequencing (scRNAseq) is a powerful technique in discriminating cellular characteristics in the kidney and elsewhere. Surprisingly little advance in knowledge about PTC phenotypes has so far come from scRNAseq, where these cells have typically been reported as a single cluster, often numerically much larger than other clusters analyzed in detail(6, 7).

We hypothesized that there may be underappreciated heterogeneity in PTCs, and used a clinically relevant mouse model of chronic aristolochic acid nephropathy (AAN) to investigate the responses of PTCs during kidney fibrosis. We have delineated PTC subtypes found in normal mouse kidney, and further describe novel PTC phenotypes associated with kidney fibrosis.

Methods

Mouse model of chronic aristolochic acid nephropathy(AAN)

Eight 8-9 weeks old C57BL/6 male mice bred by Charles River Laboratories were involved in the experiment. Animals were randomly allocated to control vs. AA injection. Chronic AAN was induced in four mice by intraperitoneal injection of aristolochic acid (2.5 mg per kg body weight, Sigma A5512) twice weekly for two weeks (Fig. 1A). Inflammation and tissue fibrosis developed after repetitive injuries, followed by tissue remodeling and fibrosis in the subsequent two weeks. Another four mice that did not receive intervention were used as naïve controls. Mice were housed with free access to chow and tap water on a 12-hour day/night cycle in a specific pathogen free environment. All mice were euthanized 4 weeks after the first inject of the AAN group. We perfused chilled PBS (1x) via left ventricle before kidney harvest. Both kidneys were harvested and processed, one for single-nucleus RNA sequencing (snRNA-seq) and one for histopathology. Serum creatinine measurement was performed using enzymatic mouse creatinine assay kit (Crystal Chem). Masson's trichrome stain for the formalin-fixed paraffin-embedded sections was performed for each mouse to confirm healthy or fibrotic status. Experiments were performed in line with institutional and

UK Home Office guidelines under the authority of an appropriate project licence.

Isolation of nuclei

One quarter of a harvested kidney from each mouse was minced into <2 mm pieces and transferred to a Dounce tissue grinder containing 2 ml of lysis buffer (Nuclei EZ Lysis buffer, Sigma NUC101) supplemented with protease inhibitor (Sigma 5892970001) and RNase inhibitors (Promega N2615 and Life Technologies AM2696)). Kidneys were homogenized and transferred into 50 ml tubes containing 2 ml of lysis buffer, incubated for 7 min on ice, filtered through a 40 am cell strainer, and centrifuged at 500*g* for 5 min at 4ºC. The pellets were resuspended in 4 ml lysis buffer, incubated for another 7 min on ice, then centrifuged again at 500*g* for 5 min at 4ºC. The pellets were then resuspended in 4 ml wash & resuspension buffer (1xPBS, 1.0% BSA, and RNase Inhibitor (Sigma 3335399001)) and filtered through a 20 am cell strainer. Samples were then processed immediately using the 10x Genomics single-cell library preparation protocol.

Library preparation and RNA sequencing

Library preparation was performed using Chromium Single Cell 3′ Reagent Kits v3 (10x Genomics) and cDNA quality was evaluated by fragment analysis (5200 Fragment Analyzer

System, Agilent). RNA sequencing was carried out using the Illumina NextSeq 550 System.

Bioinformatics analysis

The sequencing data were processed using the zUMIs pipeline (version 2.3.0)(8). The pipeline was used to first discard reads with low-quality barcodes and UMIs, and then to map reads to the mouse reference assembly (Mus_musculus.GRCm38.95). The barcode-gene matrix generated by zUMIs was analyzed using the R package, Seurat (version 3.1.3)(9, 10).

In Seurat, cells for individual samples were retained if they contained ≧ 400 genes and genes identified in ≧ 3 nuclei. After merging the 4 naïve kidneys, and 4 AAN kidneys, cells were filtered again to remove nuclei expressing ≦ 400 genes or ≧ 7500 genes, or with mitochondrial gene expression ≧ 10%. The feature counts were normalized with scale.factor = 10000. The top 2000 variable genes were identified and scaled. The principal component analysis (PCA) result of the scaled data was obtained. The number of principal components (PCs) included in the downstream analysis was determined by identifying the knee point of the elbowplot generated after running *JackStraw* procedure. *FindNeighbors* and *FindClusters* (resolution 0.8 as default) functions were applied based on previously identified PCs to identify clusters of nuclei. T-Distributed Stochastic Neighbor Embedding (t-SNE) plot was used to

visualize clustering results of the naïve and the AAN dataset. DoubletFinder (version 2.0.2) was then used to exclude potential doublets (11). After doublet removal, the naïve and AAN datasets were integrated (dims = 1:50) . The integrated dataset was then scaled and processed with PCA, *FindNeighbors and FindClusters* (resolution = 3.0). Final clustering results were visualized using Uniform Manifold Approximation and Projection (UMAP). For differential gene expression (DGE) analysis, we used *FindMarker* command (wilcox method as the default). Significance was defined as a gene with an adjusted p value < 0.05, a ≧ 0.25 average log fold difference between the two groups of cells, and with presence detected in at least 10% of cells in either of the two populations. P-value adjustment was performed using bonferroni correction based on the total number of genes in the dataset. Harmony (version 1.0), which is another batch integration method, was used to validate the pipeline of cell clustering and showed consistent results with Seurat (not shown). (12, 13)

The cell cycle analysis used *CellCycleScoring* to identify cells in the G2/M and S status. Cells with G2M.Score > 0.15 and G2M.Score > S.Score were assigned as G2M status. Cells with S.Score > 0.15 and S.Score > G2M.Score were assigned as G2M status. Cell with G2M Score <0.15 and S.Score < 0.15 were assigned as G1/G0 phase.

For trajectory analysis, we used the R package, Monocle3 (version 0.2.2.0)(14). The metadata from the integrated Seurat object and the top 2000 variable genes from the integrated assay were loaded to monocle. The analysis involved all PTC clusters and the proliferative PTCs identified in proliferative cells using default parameters and removing clusters where low gene numbers were detected (clusters 0 and 1, Supplementary Fig.1). The PTCs were re-clustered by monocle and then the trajectory analysis was performed using *learn_graph* function. The RNA velocity of each cell was calculated by velocyto.py and velocyto.R (version 0.6) using the spliced and unspliced RNA counts provided by 10X cellranger package (15). Pseudotime analysis was performed on the basis of the RNA velocity result.

We performed combined analysis of single-nucleus RNA sequencing results using this AAN induced CKD dataset and the murine ischemic reperfusion injury induced AKI dataset published by Kirita et al (16). The two datasets were merged and the top 2000 variable genes from the combined dataset were obtained for PCA. We used harmony as the batch integration method. The clustering result of all cells from both datasets was shown by UMAP. For PTC analysis, we analyzed PTCs from both datasets using Seurat integrate data function.

Pathway analysis

We conducted gene set enrichment analysis (GESA) to understand pathways of the new-PT clusters by using an R package, WebGestaltR (version 0.4.3)(17). We evaluated the pathway enrichment in four major functional databases: KEGG, Panther, Reactome and WikiPathways. The recommended False Discovery Rate (FDR) cutoff of 0.25 was used (https://www.gsea-msigdb.org/gsea/index.jsp).

Ligand-Receptor analysis

The analysis was based on a well-reviewed dataset with 2557 ligand-receptor pairs from a published study(18). PT S1-S3 were combined to normal PT and fibroblast-1, 2 were combined to fibroblast in ligand-receptor analysis. Ligands and receptors with an average fold-change ≥ 0.25 in the DGE analysis were paired. Pairing results of ligand genes of New-PT1, New-PT2, New-PT3 and receptor genes of fibroblast, immune cells and normal PT were shown using the cross products of ligand/receptor gene expression on heatmap. We also show individual ligand-receptor pairs with the ligands from the New-PT clusters and receptors from fibroblast, immune cell and normal PTCs using a circular visualization tool, circlize (version 0.4.8), in R(19).

Immunohistochemistry(IHC) and immunofluorescence staining(IF)

Formalin-fixed paraffin-embedded kidney sections from the AAN and the control group were used for IHC/IF stain. Antigen retrieval at 120^{o}C for 20 minutes was carried out for rehydrated kidney reactions. For IHC stains, UltraVision LP HRP Polymer enhancer system was applied (Thermo Scientific, 12624007) for rabbit/mouse primary antibodies. Sections were blocked with 3% hydrogen peroxide then mouse on mouse block (Vector, MKB-2213-1) where needed, and Ultra V block. The sections were then stained with primary antibodies for Ki67 (Abcam, ab15580) and alpha-smooth muscle actin (Invitrogen, MA5-11547) followed by primary antibody enhancer, HRP Polymer, DAB and hematoxylin counter stain. For Havcr1 IHC stain, sections were blocked with 3% hydrogen peroxide, avidin/biotin block (Vector, SP-2001) and 10% donkey serum, and stained with goat anti-HAVCR1 primary antibody (R&D, AF3689). The sections were then stained with biotinylated donkey anti-goat secondary antibody (Abcam, ab6884), VECTASTAIN ABC-HRP enhancer kit (Vector, PK-4000), DAB and hematoxylin counter stain. For IF stains, sections were incubated with mouse on mouse block where needed and with 10% goat serum. Primary antibodies included anti-SLC4A4 (Invitrogen, PA5-57344), anti-VCAM1(Invitrogen, MA5-11447), anti FODX1(LifeSpan BioSciences, LS-B9155-LSP), anti-Akap12 (Abcam, ab49849), anti-WT1 (Sigma, MAB4234),

anti-NCAM1 (Abcam, ab220360), anti-Tenascin C (Abcam, ab108930) anti-P21 (novusbio, NBP2-29463) and anti-HAVCR1. For secondary antibodies we used goat anti-mouse or goat anti-rabbit Alexa Fluor 488 or 594 conjugated antibodies (Invitrogen). After primary and secondary antibody staining, the sections were stained with Hoechst 33342. TrueVIEW autofluorescence quenching kit (Vector, SP-8400-15) was used to diminish autofluorescence. Immunostained tissue slides were visualized and digitized using a confocal laser scanning microscope (LSM800, Carl Zeiss).

Quantitative image analysis

Immunostained tissue slides were visualized and digitized using Olympus DP27 5MP color camera attached to a Leica DMLA microscope or by a confocal laser scanning microscope (LSM800, Carl Zeiss). Images were analyzed with the ZEN2012 software (Zeiss) and quantification was performed with Qupath software(20). Pixel classifier was used to detect the collagen (cyan) stain and thresholds were determined to express the positive stain as a percentage of total area (in >10 fields of view per animal, n=3 per group). Quantification of the alpha-smooth muscle actin DAB signal was used to confirm fibrosis in the AAN model.

Nuclear Ki67 DAB signal was used to quantify proliferating cells as a percentage of all hematoxylin stained cells using QuPath's cell detection algorithm.

Results

Sample processing and quality control

We isolated kidney nuclei immediately following terminal anesthesia and tissue harvesting. Rapid processing of tissue to nuclei was employed, to minimize artefacts arising during processing that may have limited discrimination of PTC sub clusters in previous reports of single cell sequencing from kidney (6, 21). The AAN group had significantly higher serum creatinine (Fig 1B). Histological evidence confirmed the presence of renal fibrosis in mice exposed to aristolochic acid (AA) and its absence in control animals (Fig. 1C-F). Compared to healthy kidneys, AAN kidneys demonstrated a significant increase in collagen deposition (cyan signal in Masson's Trichrome stain) and alpha-smooth muscle actin DAB signal, indicating renal fibrosis (Fig. 1G, 1H). The AAN kidneys also had a higher percentage of Ki-67 positive cells (Fig 1I). Nuclei of kidneys acquired from mice exposed to AA to induce kidney fibrosis (AAN mice)) and ndeaïve controls were processed using the 10x platform. R packages, Seurat and DoubletFinder were used for data integration, quality control and doublet removal (Supplementary Data 1)(9-11). Downstream analysis included a total of 23,885 nuclei.

Nuclei clustering and cell-type identification

The clustering result was evaluated using UMAP (Fig. 2A, Supplementary Fig. 1). The number of nuclei comprising each cluster and the distribution on the UMAP of nuclei from each individual mouse are shown in Supplementary Data 2 and Supplementary Fig. 2. Nuclei from all mice contributed to every cluster, except for that subsequently labelled "New-PT3", which was seen in four of four AAN-treated mice and none of four naïve controls.

We next used canonical markers of kidney cell populations to identify major cell types in the kidney: podocyte *(Nphs1),* endothelial cells *(Flt1),* mesangial cells *(Igfbp5),* juxtaglomerular cells (JG cells) *(Ren1),* PTCs *(Slc34a1),* proliferative cells *(Top2a* and *Mki67),* descending thin limb *(Aqp1),* ascending thing limb *(Clcnka),* thick ascending limb (TAL) *(Slc12a1 and Umod),* distal convoluted tubule (DCT) 1 *(Slc12a3)* and 2 *(Slc12a3* and *Slc8a1)*, connecting tubule *(Slc8a1)*, principal cell (PC)-outer medullary collecting duct (OMCD) and inner medullary collecting duct (IMCD) *(Aqp2),* intercalated cells type A *(Atp6v1b1* and *Slc4a1)* and type B *(Atp6v1b1* and *Slc26a4),* transitional epithelium *(Upk1b),* immune cells *(Ptprc)* and fibroblasts *(Pdgfrb and Cfh).* Fig. 2B shows the expression level and the percentage of expression of the canonical genes per cluster. Regional-specific genes of mouse kidney identified by Ransick *et al.* were used to localize clusters in cortical-medullary and outer-inner

medullary regions.(22) Fig 2C showed regional expression of *Cyp2e1* in cortical PTCs and *Cyp7b1* in medullary PTCs. The expression of canonical genes for PC-OMCD and PC-IMCD shows the PCs from the two different regions were well-clustered (Fig 2D).

Clarifying the Pdgfrb+ clusters

Some functionally distinct renal cell types share lineage and have many common features in their gene expression profiles, which has complicated their recovery in discrete clusters in recent landmark studies(21, 23). These include mesangial cells, JG cells and fibroblasts, which differentiate from *Foxd1*+ cortical stromal cells and acquire similar genetic features after profibrotic stimulation(24-26). Four clusters closely proximal to one another on UMAP showed expression of the shared markers for mesangial cells, fibroblasts and JG cells, *Cfh, Fhl2* and *Pdgfrb* (Supplementary Fig. 3A, 3B). Mesangial cells were identified from expression of the mesangial-restricted *Igfbp5* and *Itga8* (http://www.proteinatlas.org)(27, 28) while JG cells were the only cluster expressing *Ren1*. Two renal fibroblast clusters were identified, of which the Fibroblast-1 cell number increased >200% in fibrotic kidney. We provisionally identified Fibroblast-1 as myofibroblast-containing, on the basis of *Meis1* expression(29). Our dataset is limited in this regard by low *Acta2* detection (Supplementary Fig. 3B) and further

characterization of renal fibroblast populations may benefit from supplementary technical approaches.

Analysis of proliferative cells and cell cycle

One cluster expressed markers of proliferative cells not seen in other clusters, notably including *Top2a* and *Mki67*. Uniquely, this cluster comprised four distinct and separated sub clusters of nuclei. These four sub clusters were located in close proximity to the endothelial, PTC, TAL and fibroblast clusters respectively (Supplementary Fig. 4A). Consistent with this cluster comprising proliferating cells from the adjacent clusters, the four sub clusters also expressed canonical markers for the cell type (e.g. *Flt1*-endothelial, *Slc34a1*-PTC, *Slc12a1*-TAL, *Pdgfrb*-fibroblast, Supplementary Fig. 4B). Cell cycle analysis identified cells in G2/M and S phase, localized mainly to the proliferative and immune cell clusters (Supplementary Fig. 4C). G2/M arrest of PTCs contributes to fibrogenesis after kidney injury(30), and cell number of G2M- and S-Phase PTCs increased in the AAN kidneys.

Analysis of PTC sub clusters

The proximal tubule is divided into segments S1, S2 and S3 on the basis of microscopic characteristics. The S1 segment is the longest, and comprises PTCs with extensive apical

microvilli, basolateral infoldings, cytoplasmic complexities, numerous long mitochondria, and a prominent endocytic region(31). These features are present but less evident in the S2 segment, which demonstrates a gradual transition from the S1 segments and additionally displays more numerous peroxisomes and larger secondary lysozymes. The S3 segment is more distinct, comprising simple cuboidal cells without the above features.

Five clusters were identified corresponding to proximal tubular segments S1-S3: “S1”, “S1-2”, “S2”, “S2-cortical S3” and “medullary S3”. These PTC clusters were labelled according to their expression of genes localized to PTC segments in previous studies. *Slc34a1* encodes Na^+-P_i cotransporter 2a (NaPi-2a) which is localized in fully differentiated PTCs along all segments except for medullary S3(22, 32, 33). *Slc5a2* encodes a low affinity/high capacity Na^+/glucose cotransporter, sodium/glucose cotransporter 2 (Sglt2), which is responsible for 80-90% of the glucose reabsorption in S1(34). *Slc22a6* encodes organic anion transporter 1 (Oat1) in S2, which plays a critical role in drug and xenobiotic elimination, and has been linked to AAN(33, 35-38). *Slc13a3* encodes sodium-dependent dicarboxylate transporter (NaDC3) in S2, which is responsible for transport of succinate and other Krebs cycle intermediates(39). *Slc7a13*, also known as Agt1, is an amino acid transporter localized to the apical membrane of

the S3 segment and is considered an S3-canonical gene(33, 40). The expression levels of segment-specific solute transporter related genes were reviewed to validate established PTC categories. Relevant genes were identified in the first instance from the study from the Knepper lab, in which bulk deep sequencing was applied to microdissected renal tubules, to identify nephron segment-specific transcriptomes(1). Additional studies are indicated for validation of key anchor genes below. These included markers for S1 (*Slc5a2, Slc5a12*), S2 (*Slc22a6, Ca4, Slc13a3*), S3 (*Slc5a10*, *Slc7a13, Atp11a*), medullary S3 (*Cyp7b1*, *Slc6a13, Slc34a1(-)*), and pan-PTC markers (*Slc34a1, Lrp2, Slc4a4*) (Fig. 3A, Supplementary Fig. 5A)(1, 22, 41-43). The differences in genes enriched in each PT cluster supported their attribution to specific PT segments (Supplementary Fig 6).

Identification of new PTC clusters

Four additional clusters were located in close proximity to clusters S1, S1-2, S2, S2-cortical S3, medullary S3 on UMAP. First, the largest component of the distributed proliferative cluster. This part of the proliferative cluster, labelled proliferative PTCs, expressed proximal tubular markers and proliferative marker genes, including *Ki67* and *Cdca3*, that Park *et al.* highlighted as identifying a novel cell type in normal mouse kidney, and that Wu *et al.*

identified as proliferative-PT in the unilateral ureter obstruction model of renal fibrosis(6, 21). The other three clusters showed strong signals for genes upregulated in tubules following kidney injury in previous bulk sequencing and other historical approaches (*Vcam1, Havcr1 and Akap12*), demonstrated expression of canonical PTC genes, and were labelled as New-PT clusters. Clusters New-PT1 and 2 were increased in abundance in kidneys undergoing fibrosis following injury, in which circumstances the "New-PT3" were also found. The discrete gene expression signatures of New-PT1-3 clusters included genes associated with renal injury response and fibrosis progression (Fig. 3A, Supplementary Fig. 5A). DGE analysis comparing each individual cluster to all other clusters of the whole dataset detected 786 differentially expressed genes in New-PT1, 637 in New-PT2 and 1318 in New-PT3 (Supplementary Data 3). Analysis of adjusted *p*-value and average log-fold change of differentially expressed genes in New-PT 1-3 clusters demonstrated that each had a discrete identifying signature (Fig. 3B-D).

Characteristics of the new classes of PTCs

Trajectory analysis was performed to infer paths of cell state transitions within the PTC clusters. Trajectory analysis showed a continuous change in RNA profile in normal PTC's, ordered along the anatomical axis from S1 to S3 tubular segments (Fig. 4A). RNA velocity

analysis was then employed, in which the RNA processing activity evident in the transcriptome of each cell is used to evaluate transcriptional reprogramming and to predict the future state of cells. Represented graphically by an arrow, the velocity of each cell indicates its rate and direction of transcriptional change. RNA velocity analysis showed that the New PT 1-3 and the proliferative PTCs exhibited strong directional change towards other states, while the normal PT clusters exhibited stable transcriptional profiles, concordant with mature cell phenotypes (Fig. 4B). The data further indicated that NewPT1 is an intermediate cell type, from which cells may differentiate in two directions, “NewPT1 - NewPT2 - proliferative PT - normal PTCs”, and, “NewPT1 - NewPT3” (Fig. 4B). Pseudotemporal ordering was also performed (Fig. 4C.) with time zero set at NewPT1, based on the results of the RNA velocity analysis. An ordered progression of cell states in pseudotime was seen, from New-PT states through anatomically distinct tubular segments (Fig. 4C.).

Genetic profiles suggested that the cells comprising these new PTC classes were dedifferentiated. Expression of *Kim-1, Vcam1, CD44, Anax3, Akap12* and *Ncam1* has been reported in dedifferentiated PTCs in the literature(21, 44, 45). Current evidence suggests that, following kidney injury, surviving differentiated PTCs can transform to dedifferentiated PTCs

then undergo proliferation and re-differentiation to restore normal proximal tubular morphology and function(32). We identified clusters labelled New-PT1 and New-PT2 as well as proliferative PTCs in normal kidney, suggesting that this dedifferentiation-proliferation-differentiation process occurs in normal circumstances.

Trajectory and RNA velocity analysis suggested that the New-PT1 represented an intermediate cell type bridging normal PTCs, New-PT2, New-PT3 and proliferative-PT. New-PT1 expressed injury markers, including *Vcam1, Pdgfb, Pdgfd, Bmp6, Il34, Itgb6,* and *Itgav*. New-PT1 also partially preserved typical PTC markers, *Slc4a4* and *Slc5a10*. Some cells in New-PT1 also expressed *Havcr1*, this subset of *Havcr1*-positive New-PT1 cells were only identified in AAN mice (Supplementary Fig. 5B). The New-PT1 cluster was labelled as "dedifferentiated-intermediate PTCs". New-PT2 cells expressed genes characteristic of the developing kidney and reactivated during tubular regeneration including *Ncam1*, *Tnc*, *Tgfbr3, Foxd1* and *Wt1*, and was labelled as "dedifferentiated-regenerating PTC". New-PT3 cells were detected only in AAN mice, indicating a PTC phenotype specific to progressive kidney fibrosis. *Havcr1*, also known as *Kim1*, was a prominent marker of New-PT3, which also had prominent expression of *Chd2*. Apart from genes related to kidney injury and fibrosis, this cluster

expressed *Ckd6* and *Cdkn1a* (also known as *p21*), associated with cell cycle and also with cell death. The senescence-associated secretory phenotype (SASP) pertains to senescent cells with cell-cycle arrest that remain metabolically active and release senescence-associated proteins(46, 47). SASP-related genes, including *Cdkn1a (P21), Cdkn2b (P15), Tp53, Tgfb1, Serpine1 (Pai1), Ccl2 (Mcp1), Cxc1,* and *Ccn2*, were enriched in New-PT3 (Supplementary Fig. 7). This cluster was identified as "dedifferentiated-senescent PTC".

Microscopic validation for New-PT markers was performed (Fig. 5). Occasional PTC's exhibiting VCAM-1 and SLC4A4 positive staining were identified as New-PT1 and were present in normal and AAN kidney (Figure 5A, B). HAVCR1 and P21 were selected as NewPT3 markers. Consistent with the snRNAseq dataset, these were not detected in normal kidney, but were seen in a proportion of PTCs in AAN kidneys (Fig 5 C, D), in keeping with the presence of the New-PT3 cluster only following injury. Microscopic evaluation of genes selected as New-PT2 markers identified not only PTC but also glomerular parietal epithelial cell staining. Therefore an extended set of markers, comprising FOXD1, AKAP12, WT1, TNC and NCAM1, were used to validate the New PT-2 cluster. Tubular staining for each marker was evident, with co-expression in rare tubular cells discernible through merged signal in normal

mouse kidney, and in increased numbers of cells in kidneys from the AAN group (Fig 5 E, G, I, K). Co-expression of markers was also evident through merged signal in parietal epithelial cells, consistent with the presence also of parietal epithelial cells in the New-PT2 cluster (Fig 5 F, H, J, L). As expected from its known expression pattern, and consistent with the snRNAseq data, WT1 was also detected in podocytes (Fig 5 J, L). Similarly, TNC was seen in mesangial cells, as well as demonstrating strong interstitial staining in the fibrotic kidney (Supplementary Figure 8).

For further validation, we performed combined analysis of single-nucleus RNA sequencing results using this AAN-induced CKD dataset and the murine ischemia reperfusion injury-induced AKI dataset published by Kirita et el (16). Results of cell clustering and cell-type identification were comparable (Supplementary Fig. 9A) with significantly less injured and severely injured PTCs in CKD (Supplementary Fig. 9B). UMAP of PTCs also addresses the relationship of new PT clusters identified in the two studies. The New PT clusters identified in the current paper cluster in the combined analysis with the failed repair PT's identified in the Kirita dataset, while the proliferative PT's cluster with the repairing PT's. The severely

injured PT cluster identified by Kirita et al in their ischemic AKI model was not found in our dataset.

Pathway analysis of the new classes of PTCs

Pathway enrichment analysis was carried out for a further understanding of the molecular interaction network of New-PT1-3. Functional and signaling pathway analysis in KEGG showed a general reduction of metabolic pathways among the three New-PT clusters (Fig. 6) (https://www.kegg.jp/kegg/). Environmental information processing was enriched, including signal transduction (KEGG category 3.2) and signaling molecules and interaction (KEGG category 3.3) pathways, except for the AMPK signal pathway, which acts as a sensor of cellular energy status. Consistent with categorization as "dedifferentiated-senescent PTCs", New-PT3 showed enrichment of cell growth and death (KEGG category 4.2), including cell cycle, apoptosis, p53 signaling pathway and cellular senescence. The cellular community (KEGG category 4.3) and cell motility (KEGG category 4.3) pathways were highly evident in the three new classes of PTCs, corresponding to their dedifferentiated phenotype. Immune system-activation pathways were also enriched in each new-PT cluster. A document summarizing the

pathway analysis results based on KEGG, Panther, Reactome, and WikiPathways databases is provided in Supplementary Data 4.

Intercellular cross-talk between the new classes of PTCs and other cell types

Ligands and receptors detected in clusters under analysis with a positive average fold-enrichment were paired, to elucidate intercellular signal transduction networks. For the purposes of this analysis, PT S1-3 were combined to "normal PT", and fibroblast-1, -2 were combined to "fibroblast" (Fig 7A). We summarized ligand-receptor pairs for ligands from the New-PT clusters and receptors from fibroblast, immune cell and normal PT clusters (Fig. 7B-D, Supplementary Fig. 10, Supplementary Data 5). The New-PT clusters expressed multiple profibrotic signals to fibroblasts, and analysis of the interaction of the New-PTs with immune cells showed several ligand-receptor pairs associated with macrophage activation. New-PTs also expressed signals to EGF receptors, expressed abundantly by normal PTCs. In comparison, normal PTs demonstrated low expression of ligands, when compared to New-PTs. The strongest ligand signal from normal PTs was for netrin, a secreted laminin-related protein that is linked to suppression of inflammatory and injury signals in the kidney. The expression level of ligands and receptors in each cluster were shown in Supplementary Fig.11.

Discussion

Here, we have characterized the cellular composition of adult mouse kidney, comparing healthy animals to those recovering from toxic proximal tubule injury caused by aristolochic acid. Our experimental approach benefitted from refinements introduced by other investigators, demonstrating the benefits of rapidly processing unsorted whole kidney nuclear preparations prior to transcriptomic profiling. We were thus able to delineate major populations of cells that have proved challenging to resolve in previous studies, including mesangial cells, fibroblasts and juxtaglomerular cells. Furthermore, we identified proliferating cells of distinct lineages.

PTCs make a predominant contribution to the wet weight of normal kidney, are highly metabolically active, and play central roles in kidney recovery versus fibrosis following injury. Existing single cell analyses from kidney tissue have grouped PTCs in a single large cluster, often numerically dwarfing all other clusters presented. Here, we have identified five clusters of PTCs abundant in normal kidney tissue, each mapping to proximal tubule segments on the basis of canonical marker expression. These comprised clusters with expression profiles consistent only with a single tubular segment (S1, S2, medullary S3), as well as those enriched in genes found across neighboring segments (S1-2 and S2-cortical S3). These data uncover

significant complexity in PTC phenotype, and provide an expression map of abundantly expressed genes within the major cell phenotypes at single cell resolution.

Kidneys from animals treated with recurrent doses of AA to induce renal injury and subsequent fibrosis contained an increased representation of proliferating cells, immune cells and fibroblasts(48). We further identified three new PTC clusters more prominent in kidney undergoing fibrosis. The first of these, New-PT1, displayed an expression profile intermediate between canonical PTC clusters and clusters New-PT2 and New-PT3, and trajectory analysis further suggested that this cluster may represent PTCs in transition between canonical and these rarer phenotypes. Intriguingly, clusters New-PT2 and New-PT3 demonstrated enriched expression of a panel of genes expressed in proximal tubules following injury. Rather than the diffuse expression in PTC suggested by prior bulk analyses, however, our data reveal restricted expression of specific markers by cluster. Cluster New-PT2 expressed multiple genes associated with tubular regeneration following injury and was labelled dedifferentiating-regenerating PTCs on this basis. Within these regeneration-associated genes, *Ncam1* is an early nephron progenitor marker that is also seen in proximal tubules after injury, and may contribute to recovery of PTC function(45, 49). *Tnc* protects against kidney injury and promotes tubular

regeneration(50). *Tgfbr3* attenuates TGF-beta signaling through processes including glycosaminoglycan modifications of the type I and type II TGF-beta receptors(51). New-PT2 also demonstrated enriched expression of *Foxd1* and *Wt1*, genes reactivated during tubular regeneration processes. Cluster New-PT3 exhibited unique enrichment for *Havcr1*, a transcript that is nearly undetectable in normal kidney but occurs promptly after acute kidney injury. *Havcr1* expression may be upregulated chronically after kidney injury, and its persistent expression leads to renal fibrosis(52, 53). New-PT3 cells also expressed other genes linked to fibrotic responses in tubular cells, namely *Cdh2 (N-Cadherin)*, which is associated with fibroblast growth factor signaling and cell invasiveness, and several genes linked to the SASP, including *Cdkn1a (P21), Cdkn2b (P15), Tp53, Tgfb1, Serpine1 (Pai1), Ccl2 (Mcp1), Cxc1,* and *Ccn2.* New-PT3 were accordingly labelled dedifferentiated-senescent PTC's.

At microscopy, combined expression of selected genes was confirmed in rare PTCs in normal kidney for NewPT1 and New PT2 markers, and at increased number for NewPT1-3, providing validation of identified expression phenotypes. NewPT2 markers were also identified in glomerular parietal epithelial cells, consistent with the presence also of parietal epithelial cells in the NewPT2 cluster. Shared expression of identifying markers has previously

been noted for scattered tubular cells and for a proportion of parietal epithelial cells (44) and as a consequence, the potential for a proportion of parietal epithelial cells to represent tubule-committed progenitor cells has been suggested.(54) Combined analysis with the data of Kirita *et al* provided further support for discrete, minority PT phenotypes, most notably post-injury. Alignment of the New-PT3 subset of cells identified in the current work with the failed-repair PT's identified in the work of Kirita et al supports their identification of this as a distinct and important phenotype in fibrosis following renal injury. Definitive establishment of the contribution of this and the other identified PTC phenotypes in renal recovery versus fibrosis after injury will be an important future direction for research.

In summary, these data identify principle cellular phenotypes existing in the proximal tubule of the kidney. They further uncover PTC clusters with discrete, fibrosis-associated phenotypes delineated by unique expression profiles of disease-associated markers.

Data availability

RNA sequencing data and the annotated barcode-gene matrices are available from the ArrayExpress database, accession code E-MTAB-9390.

Acknowledgements

P.R.T is funded by the Wellcome Trust Investigator Award (107964/Z/15/Z) and the UK Dementia Research Institute. C.T.L is funded by Kidney Research UK/MedImmune Post-Doctoral Fellowship (PDF_006_20151127). We thank the staff of our animal facilities for the care of the animals. All animal experiments were approved after institutional and UK Home Office review. We thank Genome Research Hub at School of Biosciences, Cardiff University for the next generation sequencing work. We acknowledge compute support provided by Advanced Research Computing Team at Cardiff University, Supercomputing Wales and Wales Gene Park.

Author disclosures

The authors have nothing to disclose.

Supplementary Data Table of Contents

Supplementary data 1. Summary of genome mapping results. Table shows details of nuclei and UMIs detection of each sample before data integration.

Supplementary data 2. Nuclei clustering results. Table shows the number of nuclei comprising each cluster from each individual mouse and the nuclei number left in each filtering step.

Supplementary data 3. Results of differentially expressed gene analysis (DGE). Results of DGE of each cell type by comparing the RNA expression profiles of one type of cell to all other cells of the dataset.

Supplementary data 4. Pathway analysis. Result of pathway analysis based on KEGG, Panther, Reactome, and WikiPathways databases.

Supplementary data 5. Ligand-receptor interaction. Result of ligand-receptor analysis show pairs of the new-PTs ligands and fibroblast, immune cell and normal PTC (all normal PTC clusters) receptors.

Reference

1. Lee JW, Chou CL, Knepper MA: Deep Sequencing in Microdissected Renal Tubules Identifies Nephron Segment-Specific Transcriptomes. *Journal of the American Society of Nephrology : JASN,* 26**:** 2669-2677, 2015 10.1681/ASN.2014111067
2. Meinild AK, Loo DD, Pajor AM, Zeuthen T, Wright EM: Water transport by the renal Na(+)-dicarboxylate cotransporter. *Am J Physiol Renal Physiol,* 278**:** F777-783, 2000 10.1152/ajprenal.2000.278.5.F777
3. Berg JM TJ, Stryer L.: *Biochemistry. 5th edition. New York: W H Freeman; 2002. Section 30.2, Each Organ Has a Unique Metabolic Profile. Available from:* https://www.ncbi.nlm.nih.gov/books/NBK22436/, 2002
4. Fattah H, Vallon V: Tubular Recovery after Acute Kidney Injury. *Nephron,* 140**:** 140-143, 2018 10.1159/000490007
5. Chevalier RL: The proximal tubule is the primary target of injury and progression of kidney disease: role of the glomerulotubular junction. *Am J Physiol Renal Physiol,* 311**:** F145-161, 2016 10.1152/ajprenal.00164.2016
6. Park J, Shrestha R, Qiu C, Kondo A, Huang S, Werth M, et al.: Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science,* 360**:** 758-763, 2018 10.1126/science.aar2131
7. Lake BB, Chen S, Hoshi M, Plongthongkum N, Salamon D, Knoten A, et al.: A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat Commun,* 10**:** 2832, 2019 10.1038/s41467-019-10861-2
8. Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I: zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *Gigascience,* 7, 2018 10.1093/gigascience/giy059
9. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, et al.: Comprehensive Integration of Single-Cell Data. *Cell,* 177**:** 1888-1902 e1821, 2019 10.1016/j.cell.2019.05.031
10. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R: Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol,* 36**:** 411-420, 2018 10.1038/nbt.4096
11. McGinnis CS, Murrow LM, Gartner ZJ: DoubletFinder: Doublet Detection in Single-Cell

RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst,* 8**:** 329-337 e324, 2019 10.1016/j.cels.2019.03.003

12. Tran HTN, Ang KS, Chevrier M, Zhang X, Lee NYS, Goh M, et al.: A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol,* 21**:** 12, 2020 10.1186/s13059-019-1850-9
13. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al.: Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods,* 16**:** 1289-1296, 2019 10.1038/s41592-019-0619-0
14. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al.: The single-cell transcriptional landscape of mammalian organogenesis. *Nature,* 566**:** 496-502, 2019 10.1038/s41586-019-0969-x
15. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, et al.: RNA velocity of single cells. *Nature,* 560**:** 494-498, 2018 10.1038/s41586-018-0414-6
16. 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
17. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B: WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res,* 47**:** W199-W205, 2019 10.1093/nar/gkz401
18. Ramilowski JA, Goldberg T, Harshbarger J, Kloppmann E, Lizio M, Satagopam VP, et al.: A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat Commun,* 6**:** 7866, 2015 10.1038/ncomms8866
19. Gu Z, Gu L, Eils R, Schlesner M, Brors B: circlize Implements and enhances circular visualization in R. *Bioinformatics,* 30**:** 2811-2812, 2014 10.1093/bioinformatics/btu393
20. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD, et al.: QuPath: Open source software for digital pathology image analysis. *Sci Rep,* 7**:** 16878, 2017 10.1038/s41598-017-17204-5
21. Wu H, Kirita Y, Donnelly EL, Humphreys BD: Advantages of Single-Nucleus over Single-Cell RNA Sequencing of Adult Kidney: Rare Cell Types and Novel Cell States Revealed in Fibrosis. *Journal of the American Society of Nephrology : JASN,* 30**:** 23-32, 2019 10.1681/ASN.2018090912
22. Ransick A, Lindstrom NO, Liu J, Zhu Q, Guo JJ, Alvarado GF, et al.: Single-Cell Profiling

Reveals Sex, Lineage, and Regional Diversity in the Mouse Kidney. *Dev Cell,* 51**:** 399-413 e397, 2019 10.1016/j.devcel.2019.10.005

23. Stewart BJ, Ferdinand JR, Young MD, Mitchell TJ, Loudon KW, Riding AM, et al.: Spatiotemporal immune zonation of the human kidney. *Science,* 365**:** 1461-1466, 2019 10.1126/science.aat5031
24. Kobayashi A, Mugford JW, Krautzberger AM, Naiman N, Liao J, McMahon AP: Identification of a multipotent self-renewing stromal progenitor population during mammalian kidney organogenesis. *Stem Cell Reports,* 3**:** 650-662, 2014 10.1016/j.stemcr.2014.08.008
25. Karaiskos N, Rahmatollahi M, Boltengagen A, Liu H, Hoehne M, Rinschen M, et al.: A Single-Cell Transcriptome Atlas of the Mouse Glomerulus. *Journal of the American Society of Nephrology : JASN,* 29**:** 2060-2068, 2018 10.1681/ASN.2018030238
26. Johnson RJ, Floege J, Yoshimura A, Iida H, Couser WG, Alpers CE: The activated mesangial cell: a glomerular "myofibroblast"? *Journal of the American Society of Nephrology : JASN,* 2**:** S190-197, 1992
27. Lu Y, Ye Y, Yang Q, Shi S: Single-cell RNA-sequence analysis of mouse glomerular mesangial cells uncovers mesangial cell essential genes. *Kidney international,* 92**:** 504-513, 2017 10.1016/j.kint.2017.01.016
28. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al.: Proteomics. Tissue-based map of the human proteome. *Science,* 347**:** 1260419, 2015 10.1126/science.1260419
29. Chang-Panesso M, Kadyrov FF, Machado FG, Kumar A, Humphreys BD: Meis1 is specifically upregulated in kidney myofibroblasts during aging and injury but is not required for kidney homeostasis or fibrotic response. *Am J Physiol Renal Physiol,* 315**:** F275-F290, 2018 10.1152/ajprenal.00030.2018
30. Yang L, Besschetnova TY, Brooks CR, Shah JV, Bonventre JV: Epithelial cell cycle arrest in G2/M mediates kidney fibrosis after injury. *Nat Med,* 16**:** 535-543, 531p following 143, 2010 10.1038/nm.2144
31. Jacobson HR: Functional segmentation of the mammalian nephron. *Am J Physiol,* 241**:** F203-218, 1981 10.1152/ajprenal.1981.241.3.F203
32. Kusaba T, Lalli M, Kramann R, Kobayashi A, Humphreys BD: Differentiated kidney epithelial cells repair injured proximal tubule. *Proc Natl Acad Sci U S A,* 111**:** 1527-1532, 2014 10.1073/pnas.1310653110

33. Chen L, Clark JZ, Nelson JW, Kaissling B, Ellison DH, Knepper MA: Renal-Tubule Epithelial Cell Nomenclature for Single-Cell RNA-Sequencing Studies. *Journal of the American Society of Nephrology : JASN,* 30**:** 1358-1364, 2019 10.1681/ASN.2019040415

34. Kanai Y, Lee WS, You G, Brown D, Hediger MA: The human kidney low affinity Na+/glucose cotransporter SGLT2. Delineation of the major renal reabsorptive mechanism for D-glucose. *J Clin Invest,* 93**:** 397-404, 1994 10.1172/JCI116972

35. Li C, Wang X, Bi Y, Yu H, Wei J, Zhang Y, et al.: Potent inhibitors of organic anion transporter 1 and 3 from natural compounds and their protective effect on aristolochic acid nephropathy. *Toxicol Sci*, 2020 10.1093/toxsci/kfaa033

36. Dantzler WH, Wright SH: The molecular and cellular physiology of basolateral organic anion transport in mammalian renal tubules. *Biochim Biophys Acta,* 1618**:** 185-193, 2003 10.1016/j.bbamem.2003.08.015

37. Hwang JS, Park EY, Kim WY, Yang CW, Kim J: Expression of OAT1 and OAT3 in differentiating proximal tubules of the mouse kidney. *Histol Histopathol,* 25**:** 33-44, 2010 10.14670/HH-25.33

38. Breljak D, Ljubojevic M, Hagos Y, Micek V, Balen Eror D, Vrhovac Madunic I, et al.: Distribution of organic anion transporters NaDC3 and OAT1-3 along the human nephron. *Am J Physiol Renal Physiol,* 311**:** F227-238, 2016 10.1152/ajprenal.00113.2016

39. Mouse Genome Informatics Gene Expression Data (http://www.informatics.jax.org/assay/MGI:5542047).

40. Nagamori S, Wiriyasermkul P, Guarch ME, Okuyama H, Nakagomi S, Tadagaki K, et al.: Novel cystine transporter in renal proximal tubule identified as a missing partner of cystinuria-related plasma membrane protein rBAT/SLC3A1. *Proc Natl Acad Sci U S A,* 113**:** 775-780, 2016 10.1073/pnas.1519959113

41. Cao J, Cusanovich DA, Ramani V, Aghamirzaie D, Pliner HA, Hill AJ, et al.: Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science,* 361**:** 1380-1385, 2018 10.1126/science.aau0730

42. Kamiyama M, Garner MK, Farragut KM, Kobori H: The establishment of a primary culture system of proximal tubule segments using specific markers from normal mouse kidneys. *Int J Mol Sci,* 13**:** 5098-5111, 2012 10.3390/ijms13045098

43. Thiagarajan RD, Georgas KM, Rumballe BA, Lesieur E, Chiu HS, Taylor D, et al.:

Identification of anchor genes during kidney development defines ontological relationships, molecular subcompartments and regulatory pathways. *PloS one,* 6**:** e17286, 2011 10.1371/journal.pone.0017286

44. Smeets B, Boor P, Dijkman H, Sharma SV, Jirak P, Mooren F, et al.: Proximal tubular cells contain a phenotypically distinct, scattered cell population involved in tubular regeneration. *J Pathol,* 229**:** 645-659, 2013 10.1002/path.4125
45. Abbate M, Brown D, Bonventre JV: Expression of NCAM recapitulates tubulogenic development in kidneys recovering from acute ischemia. *Am J Physiol,* 277**:** F454-463, 1999 10.1152/ajprenal.1999.277.3.F454
46. Sturmlechner I, Durik M, Sieben CJ, Baker DJ, van Deursen JM: Cellular senescence in renal ageing and disease. *Nat Rev Nephrol,* 13**:** 77-89, 2017 10.1038/nrneph.2016.183
47. Docherty MH, O'Sullivan ED, Bonventre JV, Ferenbach DA: Cellular Senescence in the Kidney. *Journal of the American Society of Nephrology : JASN,* 30**:** 726-736, 2019 10.1681/ASN.2018121251
48. Huang L, Scarpellini A, Funck M, Verderio EA, Johnson TS: Development of a chronic kidney disease model in C57BL/6 mice with relevance to human pathology. *Nephron Extra,* 3**:** 12-29, 2013 10.1159/000346180
49. Buzhor E, Omer D, Harari-Steinberg O, Dotan Z, Vax E, Pri-Chen S, et al.: Reactivation of NCAM1 defines a subpopulation of human adult kidney epithelial cells with clonogenic and stem/progenitor properties. *Am J Pathol,* 183**:** 1621-1633, 2013 10.1016/j.ajpath.2013.07.034
50. Chen S, Fu H, Wu S, Zhu W, Liao J, Hong X, et al.: Tenascin-C protects against acute kidney injury by recruiting Wnt ligands. *Kidney international,* 95**:** 62-74, 2019 10.1016/j.kint.2018.08.029
51. Eickelberg O, Centrella M, Reiss M, Kashgarian M, Wells RG: Betaglycan inhibits TGF-beta signaling by preventing type I-type II receptor complex formation. Glycosaminoglycan modifications alter betaglycan function. *J Biol Chem,* 277**:** 823-829, 2002 10.1074/jbc.M105110200
52. Kirk R: Renal fibrosis: KIM-1 expression links kidney injury with CKD in mice. *Nat Rev Nephrol,* 9**:** 627, 2013 10.1038/nrneph.2013.194
53. Humphreys BD, Xu F, Sabbisetti V, Grgic I, Movahedi Naini S, Wang N, et al.: Chronic epithelial kidney injury molecule-1 expression causes murine kidney fibrosis. *J Clin Invest,* 123**:** 4023-4035, 2013 10.1172/JCI45361

54. Shankland SJ, Smeets B, Pippin JW, Moeller MJ: The emergence of the glomerular parietal epithelial cell. *Nat Rev Nephrol,* 10**:** 158-173, 2014 10.1038/nrneph.2014.1