The Company of Biologists

Single-cell analysis of progenitor cell dynamics and lineage specification in the human fetal kidney

Rajasree Menon[1,‡], Edgar A. Otto[2,‡], Austin Kokoruda[3], Jian Zhou[4,5], Zidong Zhang[4,5], Euisik Yoon[6], Yu-Chih Chen[6], Olga Troyanskaya[4,7,8], Jason R. Spence[3,9,§], Matthias Kretzler[2,§] and Cristina Cebrián[3,*,§]

ABSTRACT

The mammalian kidney develops through reciprocal interactions between the ureteric bud and the metanephric mesenchyme to give rise to the entire collecting system and the nephrons. Most of our knowledge of the developmental regulators driving this process arises from the study of gene expression and functional genetics in mice and other animal models. In order to shed light on human kidney development, we have used single-cell transcriptomics to characterize gene expression in different cell populations, and to study individual cell dynamics and lineage trajectories during development. Single-cell transcriptome analyses of 6414 cells from five individual specimens identified 11 initial clusters of specific renal cell types as defined by their gene expression profile. Further subclustering identifies progenitors, and mature and intermediate stages of differentiation for several renal lineages. Other lineages identified include mesangium, stroma, endothelial and immune cells. Novel markers for these cell types were revealed in the analysis, as were components of key signaling pathways driving renal development in animal models. Altogether, we provide a comprehensive and dynamic gene expression profile of the developing human kidney at the single-cell level.

KEY WORDS: Branching, Fetal, Human, Kidney, Single cell

INTRODUCTION

The development of the embryonic kidney is a paradigm for the complex inductive and regulatory mechanisms driving organogenesis (Grobstein, 1953). In mammals, the epithelial ureteric bud (UB) tips undergo repetitive branching morphogenesis until birth and generate

[1]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA. [2]Department of Internal Medicine, Division of Nephrology, University of Michigan, Ann Arbor, MI 48109, USA. [3]Department of Internal Medicine, Division of Gastroenterology, University of Michigan, Ann Arbor, MI 48109, USA. [4]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA. [5]Graduate Program in Quantitative and Computational Biology, Princeton University, Princeton, NJ 08544, USA. [6]Department of Electrical Engineering and Computer Science, Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA. [7]Flatiron Institute, Simons Foundation, New York, NY 10010, USA. [8]Department of Computer Science, Princeton University, Princeton, NJ. [9]Department of Cell and Developmental Biology, Division of Gastroenterology, University of Michigan, Ann Arbor, MI 48109, USA.
*Present address: Department of Pediatrics, Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA.
‡These authors contributed equally to this work

§Authors for correspondence (spencejr@umich.edu; kretzler@umich.edu; cristina.cebrian@cchmc.org)

J.R.S., 0000-0001-7869-3992; M.K., 0000-0003-4064-0582; C.C., 0000-0002-1297-6591

the entire collecting system, including the ureter, calyces and collecting ducts, but not the bladder (Chi et al., 2009; Riccio et al., 2016; Shakya et al., 2005). Around the tips of the UB, nephron progenitors condense into cap mesenchyme, epithelialize and undergo complex morphogenesis to differentiate into the vast majority of cells in the nephron, including parietal epithelial cells, the podocytes, proximal and distal tubules, loops of Henle and the connecting tubule (Cebrian et al., 2014; Kobayashi et al., 2008; Mugford et al., 2008). Interspersed between the nephron progenitors, the stroma promotes survival and differentiation of the progenitors as well as branching of the UB and contribute to the mesangial and endothelial lineage (Das et al., 2013; Mugford et al., 2008).

Studies in mice have identified several signaling pathways involved in the reciprocal epithelial-mesenchymal interactions driving kidney development. For example, soluble ligands are expressed by the mesenchymal nephron progenitors and signal to receptors in the epithelial UB tip to activate downstream transcription factors. Reciprocally, the UB tip cells also release signals to the mesenchyme to maintain its self-renewal. These signaling cascades are crucial for UB outgrowth and branching, and involve a host of morphogens and signaling networks, including Gdnf-Ret, Wnt, Fgf and Notch signaling, to name a few (Cacalano et al., 1998; Lu et al., 2009; Pachnis et al., 1993; Pichel et al., 1996; Schuchardt et al., 1994). Collectively, these signaling events regulate the balance between self-renewal and differentiation of nephron progenitors (Carroll et al., 2005; Park et al., 2007), ensure proper positioning of the metanephric mesenchyme (MM) and UB outgrowth (Huang et al., 2014; Pietilä et al., 2016; Yun and Perantoni, 2017), and promote further differentiation into specific cell types (Cheng et al., 2007; Surendran et al., 2010).

Although the murine and the human kidney seem to share a common developmental pattern, they also present anatomical, physiological and pathophysiological differences, as well as significant differences in gene expression (O'Brien et al., 2016), suggesting that significant new information can be gained by directly studying human kidney development. Here, we have used single-cell RNA-sequencing to interrogate 6414 cells from five developing human kidneys and have identified distinct cell populations based on their gene expression profile. These populations include progenitors, immature and mature renal cell types and proliferating cells. We further identify new markers for specific cell types in the developing human kidney, and we use computational approaches to infer developmental trajectories and interrogate the complex network of signaling pathways and cellular transitions during development.

RESULTS

Single-cell sequencing identifies progenitors, and intermediate and mature cell types in the developing human kidney

Analysis of single-cell sequencing data of 6414 cells from five human embryonic kidneys revealed 11 clusters (Fig. 1A-C and Table S1).
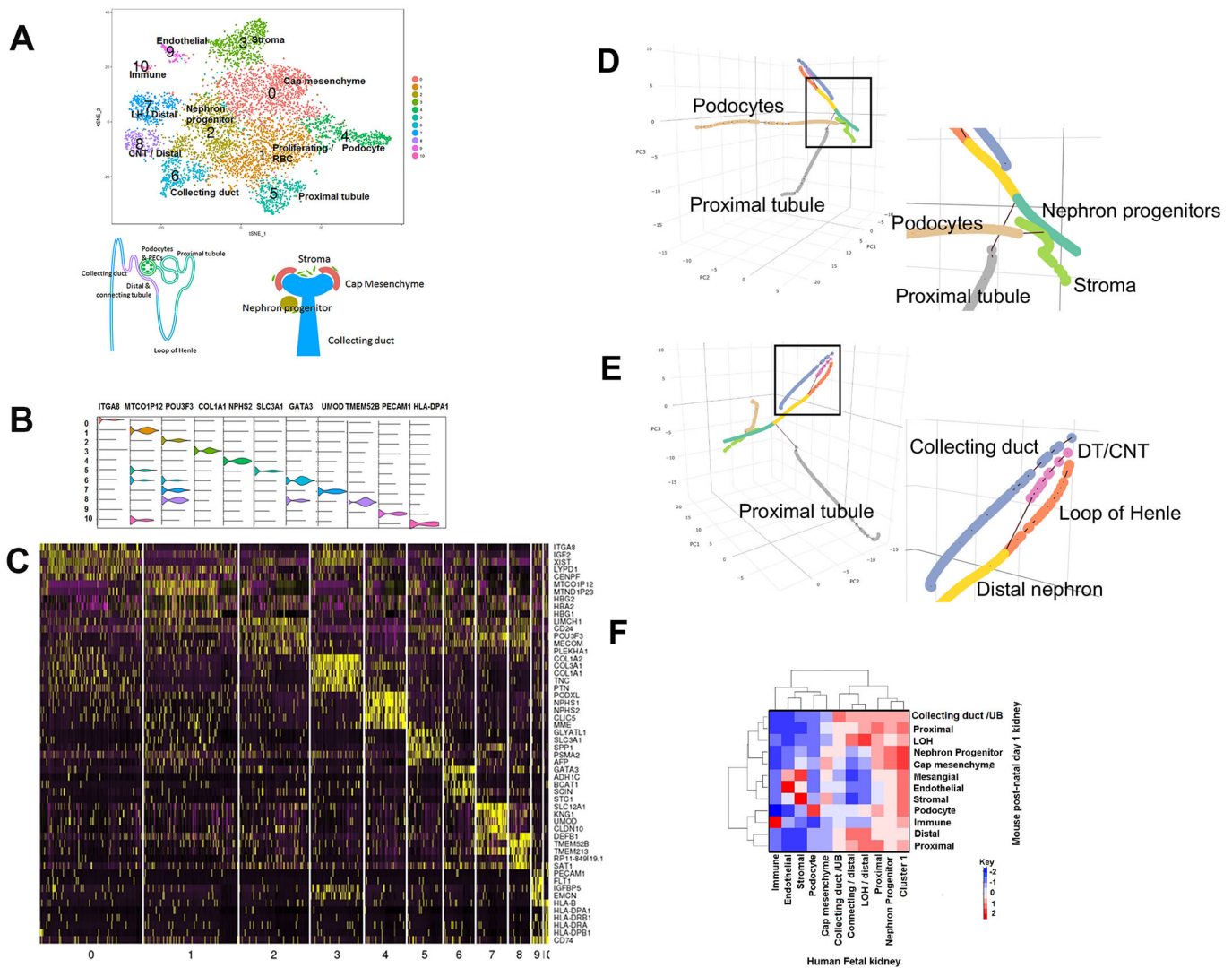
Fig. 1. Human fetal kidney single cell types. (A) tSNE plot showing the 11 cell clusters from the combined analysis of the five fetal kidney datasets. A total of 6414 cells with at least 500 genes per cell were identified. The diagram depicts developing and mature renal structures identified in the tSNE plot. The number of cells per cluster is as follows: 0 (1384), 1 (1203), 2 (895), 3 (662), 4 (568), 5 (430), 6 (406), 7 (395), 8 (270), 9 (135), 10 (66). (B) Violin plots showing the expression pattern of a top differentially expressed (cluster-specific) gene from each of the 11 clusters. (C) Heatmap with the expression pattern of the top five cluster-specific genes in 11 clusters. (D,E) Two different perspectives of the lineage trajectory analysis of single-cell RNAseq data from human embryonic kidney. Boxed area highlights the initial branch out of podocytes and proximal tubules from the nephron progenitor lineage. Boxed area in E highlights the branch out of distal tubules (DT) and connecting tubules (CNT) from the loop of Henle. (F) Heatmap showing the correlation between the average expression of all genes expressed in human fetal kidney cell clusters to mouse postnatal day 1 kidney clusters. The color intensity in the heatmap is based on the z-scores from the row-wise scaling of the correlation matrix.

The cellular identity of these clusters was assigned based on a list of representative genes most significantly expressed in these single cells (Table 1). A violin plot showing the expression of a representative gene and a heat map with the expression of the five most significantly differentially expressed genes for each cluster are shown in Fig. 1B and C, respectively. This analysis identifies most of the cell types expected in a developing kidney, from undifferentiated cap mesenchyme to differentiated epithelia (podocytes, proximal tubules, loops of Henle, collecting ducts and distal/connecting tubules), stroma, endothelial and immune cells as well as erythrocytes. In addition to these differentiated cell types, cluster 2 represents a population of cells with gene expression that is characteristic of the developing nephrons. The QC parameters indicated similarities between the five datasets (Fig. S1A,B). Fig. S1C,D shows clearly that the 11 clusters contained cells from all five datasets. It was interesting to observe

higher level of mitochondrial read content in cluster 7 (loop of Henle) compared with all other clusters (Fig. S1E).

tSNE (t-distributed stochastic neighbor embedding) dimensionality reduction identifies progenitor and differentiated cell types in 2D but fails to identify lineage relationships among different clusters. Therefore, to inquire whether single-cell data can be used to identify lineage relationships, we have performed pseudotemporal trajectory analysis, which uses an unbiased computational approach to infer developmental trajectories and distinguish between progenitor cell types in 3D (Fig. 1D,E and Movie 1). This approach identifies three independent developmental trajectories: the stroma, the collecting duct and the nephron lineage. For each trajectory, immature cell types are positioned towards the center of the plot whereas more mature cell types are positioned towards the periphery. Differentiated cell types branch out from the nephron trajectory (shown as black lines in Fig. 1D,E); interestingly, the

Table 1. Representative genes from clusters 0-10

Gene	Cluster	P	%*	%‡	Location	References
H19	0	7.39E-124	0.802	0.503	MM	GUDMAP; Harding et al., 2011; McMahon et al., 2008
ITGA8	0	9.11E-117	0.296	0.07	CM	Valerius et al., 2002
NCAM1	0	4.04E-52	0.289	0.121	CM, RV	Pode-Shakked et al., 2016
NTRK2	0	3.26E-43	0.287	0.133	hCM	Metsuyanim et al., 2009
MEIS2	0	5.78E-36	0.284	0.14	CM, ST	GUDMAP
MTCO1P12	1	0	0.674	0.156	n/a	
MTND1P23	1	1.19E-173	0.478	0.133	n/a	
RPL10P9	1	1.58E-157	0.618	0.261	n/a	
HBG2	1	3.99E-137	0.893	0.799	RBCs	Marks and Kovach, 1966
HBA2	1	7.50E-103	0.86	0.747	RBCs	Marks and Kovach, 1966
POU3F3	2	6.36E-81	0.433	0.165	SSB, HL, DT, CD	GUDMAP
MECOM	2	1.60E-40	0.406	0.207	CNT, DT, LH	GUDMAP
PLEKHA1	2	1.99E-40	0.354	0.175	CD, SSB, CNT	GUDMAP
EMX2	2	2.09E-26	0.431	0.267	PTA, RV, SSB	GUDMAP
LHX1	2	1.39E-22	0.376	0.232	PTA, RV, SSB	GUDMAP
COL1A2	3	0	0.735	0.14	ST	GUDMAP
COL3A1	3	2.94E-306	0.664	0.129	ST	GUDMAP
TNC	3	9.77E-225	0.444	0.064	ST	Aufderheide et al., 1987
COL6A3	3	4.83E-211	0.415	0.058	ST	GUDMAP
LGALS1	3	7.19E-200	0.557	0.129	ST, MES	GUDMAP
PODXL	4	0	0.779	0.128	POD	GUDMAP
NPHS1	4	0	0.749	0.066	POD	GUDMAP
NPHS2	4	0	0.623	0.063	POD	GUDMAP
CLIC5	4	0	0.664	0.073	POD	GUDMAP
MAFB	4	0	0.728	0.09	CM, POD, ST	GUDMAP
SLC3A1	5	1.96E-214	0.352	0.026	PT, LH	GUDMAP
CUBN	5	2.10E-147	0.269	0.023	PT	Assémat et al., 2005
PDZK1	5	1.12E-139	0.281	0.027	PT	Valerius and McMahon, 2008
LRP2	5	9.57E-133	0.332	0.043	PT	GUDMAP
GATA3	6	6.38E-300	0.711	0.091	CD	GUDMAP
ELF5	6	1.33E-263	0.281	0.007	CD	Grassmeyer et al., 2017
STC1	6	5.30E-130	0.257	0.023	CD	GUDMAP
TBX3	6	1.33E-121	0.346	0.047	CD	GUDMAP
CLDN4	6	3.78E-74	0.368	0.088	CD	GENEPAINT
SLC12A1	7	0	0.963	0.126	DT, LH	GUDMAP
KCNJ1	7	0	0.512	0.035	CD, LH	GUDMAP
UMOD	7	7.44E-308	0.585	0.055	DT, LH	GUDMAP
CLDN10	7	5.73E-274	0.607	0.069	LH	GUDMAP
CLCNKB	8	4.57E-142	0.419	0.046	DT, LH	GUDMAP
CA2	8	2.34E-137	0.437	0.053	CD, PT, DT	GUDMAP
ATP6V1B1	8	1.51E-109	0.264	0.022	CD, CNT	GUDMAP
ATP1B1	8	7.20E-104	0.841	0.325	CD, PT, LH, DT	GUDMAP
FXYD2	8	1.17E-88	0.657	0.191	PT, LH, DT, CNT	GUDMAP
ADGRL4	9	2.59E-228	0.284	0.004	EC	Wallgard et al., 2008
ADGRF5	9	2.65E-212	0.351	0.009	EC	Wallgard et al., 2008
NOSTRIN	9	7.13E-206	0.313	0.007	EC	Zimmermann et al., 2002
PECAM1	9	2.08E-200	0.396	0.014	EC	GUDMAP
CD34	9	8.64E-152	0.396	0.02	EC	GUDMAP
LST1	10	1.11E-230	0.439	0.007	HC	Holzinger et al., 1995
HLA-DPA1	10	3.24E-215	0.47	0.01	HC	Harada et al., 1992
SRGN	10	8.11E-206	0.333	0.004	HC	Kolset and Gallagher, 1990
MNDA	10	6.74E-189	0.258	0.002	HC	Hudson et al., 1988
CORO1A	10	3.63E-183	0.258	0.003	HC	Astier et al., 2003

*Percentage of cells in the cluster of interest.

‡Percentage of cells in all other clusters.

CD, collecting duct; CM, cap mesenchyme; CNT, connecting tubule; DT, distal tubule; EC, endothelial cell; HC, hematopoietic cell; hCM, human cap mesenchyme; LH, loop of Henle; MES, mesangial cells; POD, podocytes; PT, proximal tubule; PTA, pre-tubular aggregate; RBCs, red blood cells; RV, renal vesicle; SSB, S-shaped bodies; ST, stroma.

Unless stated otherwise, these expression patterns correspond to previously reported expression in mouse tissues.

podocytes are the first ones to emerge from the nephron progenitors (see insert box in Fig. 1D) followed by the proximal and distal segments. The distal nephron further differentiates into loop of Henle and distal/connecting tubules (insert box in Fig. 1E). Altogether, our approach identifies cell populations and developmental trajectories that are consistent with the major renal mature cell types, their progenitors and intermediate stages of differentiation.

To further validate our analysis, we have clustered single-cell sequencing data from postnatal day 1 mice (Adam et al., 2017) and compared those clusters (Fig. S1F) with the ones generated from human single-cell sequencing. As shown in Fig. 1F, mouse and

human clusters show significant correlation. Therefore, despite differential expression of specific genes, the overall gene expression and cellular identity between the mouse and human developing kidney is highly conserved. Interestingly, geneontology enrichment analysis of the significantly expressed genes in each cluster compared with all other clusters showed relevant biological processes for the identified cell types (Fig. S1G). For example, cell-cell adhesion and extracellular matrix organization were among the top terms for podocytes.

Sub-clustering of undifferentiated and nephron progenitor cells reveals several maturation stages

The initial clustering analysis shown in Fig. 1A identified discrete differentiated cell populations consistent with the existence of specific mature renal cell types in the human embryonic kidney (clusters 3-9). In addition, a large population of cells with the expression pattern of renal progenitors was also identified (clusters 0-2). On the other hand, the intermediate stages that are anatomically described as renal vesicles, pre-tubular aggregates, and comma- and s-shaped bodies were not distinctly identified in this initial clustering. To gain further insight into the gene expression signature of these intermediate stages, we have performed sub-clustering of clusters 0-2, as shown in Fig. 2. A total of eight clusters were identified and shown in a tSNE plot (Fig. 2A and Table S2). A violin plot showing the expression of a representative gene and a heat map with the expression of the five most significantly differentially expressed genes for each cluster are shown in Fig. 2B and C, respectively. By analyzing cell specificity of the most significantly differentially expressed genes in these clusters (Table 2), we can begin to unravel the cellular identity of these clusters. Fig. S2A,B shows the distribution of cells from the five datasets in these eight sub-clusters.

Sub-cluster 0 contains cells that express ubiquitous markers, erythrocyte markers, nephron progenitor markers (see *LHX1* in Fig. 2B violin plot) as well as various proliferation markers, including *CCND1*, *UBE2C*, *TOP2A*, *MKI67*, *CDK4* and *CDK6* (Fig. S2C). In addition, cells in sub-cluster 5 present elevated expression of H1 linker histones; the mRNA level of these histones is greatly increased as cells progress from G1 to S phase, indicating that cells in sub-cluster 5 are undergoing mitosis (Harris et al., 1991). In addition, cells in sub-cluster 7 express genes associated with developing and mature erythrocytes, consistent with them being embryonic red blood cells. Sub-cluster 1 is defined by the expression of *ITGA8* and *EYA1*, reported markers of the nephrogenic mesenchyme in mice, the undifferentiated murine nephron progenitors (Kalatzis et al., 1998; Valerius et al., 2002). The expression pattern in cells in sub-cluster 2 is consistent with cells that have already initiated the differentiation process, with the expression of *JAG1* and *LHX1*; both these genes are markers of the pre-tubular aggregate in mice and humans (Crosnier et al., 2000; Fujii et al., 1994; McCright et al., 2001). Sub-cluster 3 contains cells with a more distal profile, with the expression of genes such as *MECOM*, *POU3F3* and *GATA3*; similarly, sub-cluster 6 contains distal nephron cells with an expression profile consistent with a loop of Henle identity. Finally, cells in sub-cluster 4 express *SIM2* and *EPHA4*, genes that, in mouse, are expressed by the s-shaped bodies (Brunskill et al., 2008; Peuckert et al., 2016).

To gain further insight into the dynamic gene expression of early nephron progenitors as they differentiate, we have plotted the expression of relevant genes in cells from sub-clusters 1 (cap mesenchyme), 2 (pre-tubular aggregates) and 4 (s-shaped bodies). The resulting heat map is shown in Fig. 2D. This analysis reveals

distinct populations of cells based on the expression of these marker genes. The expression of *CITED1*, *TMEM100*, *SFRP1*, *SIX1*, *SIX2*, *WT1*, *ITGA8*, *HNF1A* and *GDNF* is consistent with these cells being cap mesenchyme. Most of these genes are also expressed, but at lower levels, in a second population of cells characterized by the expression of *HNF4A*, *EYA1* and *PAX2*; this expression profile is consistent with these cells being part of the proximal renal vesicle. A third population is characterized by the expression of *FGF8* and *SFRP2*, which in mouse are expressed in the distal segments of the renal vesicles and comma-shaped bodies (Grieshammer et al., 2005; Heliot et al., 2013; Leimeister et al., 1998). Finally, two more populations are identified: one corresponding to the distal segment of the s-shaped body expressing *POU3F3* (Lindstrom et al., 2018b) and one corresponding to the medial segment of the s-shaped body expressing *FOXC2* and *LHX1* (Lindstrom et al., 2018b). We have plotted the expression of *POU3F3* and *LHX1* on 1, 2 and 4 sub-clusters (cap mesenchyme, proximal and distal nephron) of the original clusters 0, 1 and 2 (Fig. 2E). The cells with higher *POU3F3* expression cluster away from those of high *LHX1* expression, further confirming that they belong to different segments (distal and medial) of the developing nephron.

Parietal epithelial cells and immature and mature podocytes present distinct gene expression profiles in the developing human kidney

We have also performed sub-clustering analysis of the initial cluster 4, which is characterized by the specific expression of *PODXL*, *NPHS1*, *NPHS2* and *CLIC5*, and we first identified these cells as podocytes (Table 1). Fig. 3A-C shows the tSNE plot of this sub-clustering and heatmap of expression of representative genes; the full list of genes is provided in Table S3. The tSNE plot showing the distribution of cells from each embryonic sample demonstrates contribution from all samples to the three identified clusters (Fig. S2D,E). Cells in sub-cluster 2 are characterized by the specific expression of known parietal epithelial cells such as *CAV2*, *PTRF* (*CAVIN1*), *CLDN1* and *CLDN3* (Kiuchi-Saishin et al., 2002; Krawczyk et al., 2017; Ohse et al., 2008). Cells in sub-cluster 1 express *PODXL*, *NPHS1*, *NPHS2*, *SYNPO* and *VEGF*, all markers of mature human podocytes (Bariety et al., 2006). We have plotted the expression of *NPHS1* and *CLDN1* in this sub-clustering (Fig. S2F), showing that cells in sub-cluster 2 have high expression of *CLDN1* alone or in combination with *NPHS1*, an expression profile compatible with their identity as parietal epithelial cells. Sub-cluster 0 is populated by cells that express specifically *OLFM3*. When we plot the expression of this gene on the trajectory analysis (Fig. 3G), we observe that this gene is not only specific for the podocyte trajectory, but it is also restricted to the more immature section of it. The PAX8 protein has been detected in developing podocytes by immunohistochemistry, but is absent from the mature podocytes that are instead positive for mature marker SYNPO (Fig. 3D,E). *OLFM3* expression pattern overlaps with the immature podocyte marker *PAX8* (Fig. 3F) but is excluded from the mature podocytes that express *SYNPO* (Fig. 3H).

A number of PDZ domain proteins are expressed in the mature podocyte cluster: *HTRA1*, *MAGI2*, *SLC9A3R2*, *AHNAK*, *PDLIM2*, *PARD3B*, *PDLIM5*, *MPP5* and *TJP1*). Sub-cluster 1 (from sub-clustering of initial cluster 4) enrichment in PDZ domain proteins is consistent with a role for these proteins in establishing cell-cell contacts and the slit diaphragm that is characteristic of mature podocytes. Indeed, *MAGI2*, *PDLIM2*, *MPP5*, *PARD3B* and *TJP1* are associated with cytoskeletal and
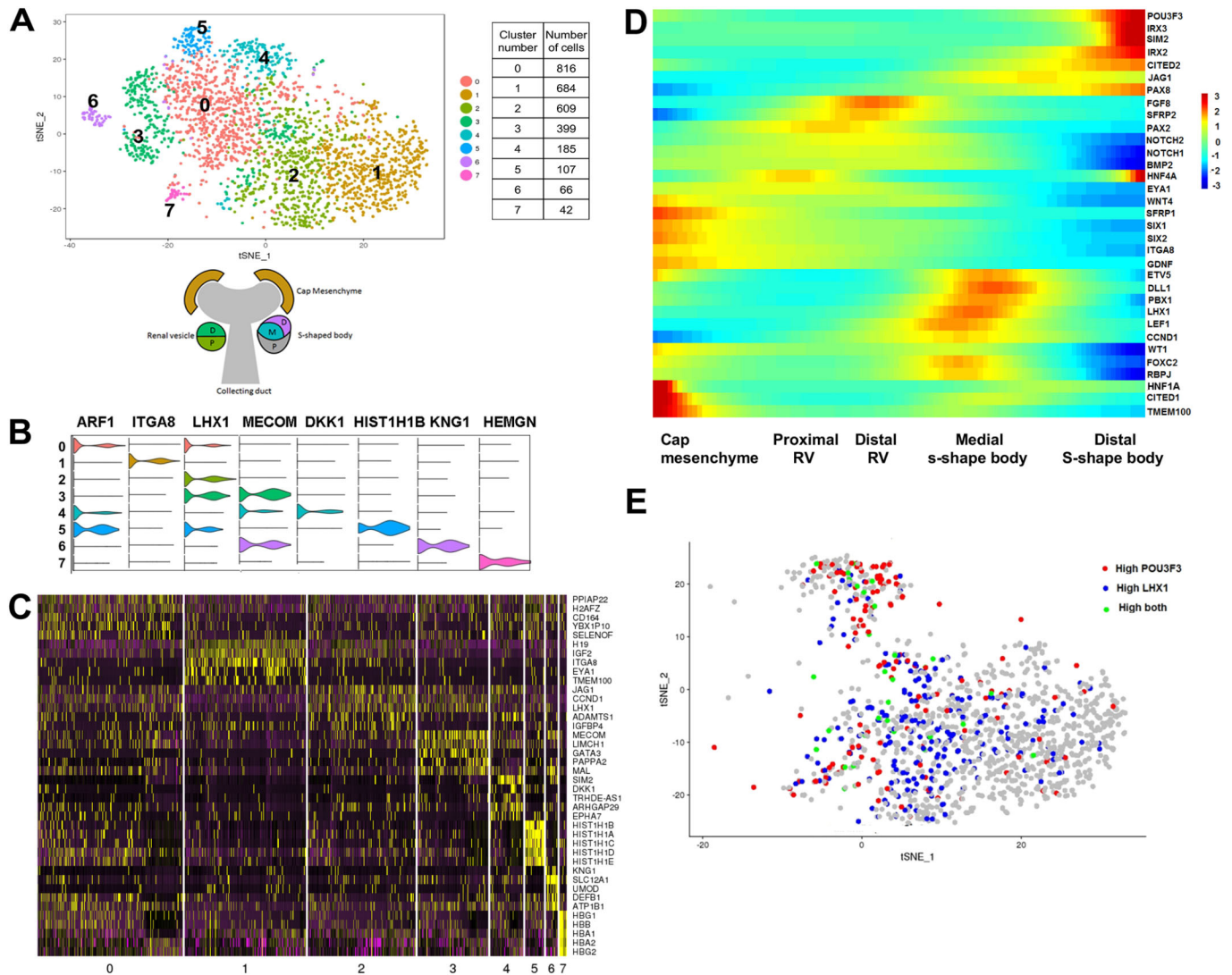
4

Fig. 2. Sub-clustering of nephron progenitor cells. (A) tSNE plot showing the eight clusters from the sub-clustering of 0, 1 and 2 clusters from the initial clustering analyses. The diagram depicts developing nephron structures identified in the tSNE plot. The table provides the number of cells found in each cluster. (B) Violin plots showing the expression pattern of a top differentially expressed (cluster-specific) gene from eight sub-clusters of the original clusters 0, 1 and 2. (C) Heatmap with the expression pattern of the top five cluster-specific genes in the eight sub-clusters of the original clusters 0, 1 and 2. (D) Heatmap showing gene expression dynamics of signaling molecules and cell-specific markers in 1, 2 and 4 sub-clusters (cap-mesenchyme, proximal and distal nephron) of the original clusters 0, 1 and 2. The heatmap was generated using Monocle. Genes (row) are clustered and cells (column) are ordered according to the pseudotime development. (E) Feature plot of *LHX1* and *POU3F3* in 1, 2 and 4 sub-clusters (cap-mesenchyme, proximal and distal nephron) of the original clusters 0, 1 and 2. Cells with high expression of *POU3F3* are in red and the cells with high expression of *LHX1* are in blue. Cells with high expression of both *POU3F3* and *LHX1* are in green. The FeaturePlot function in Seurat R package that shows co-expression of two genes was used to generate this plot. According to this function, for each gene, the cells are divided into two groups (intervals) of equal size based on the range of gene expression using 'cut' R function. The group with higher expression is designated as 'high'.

barrier function in podocytes (Ihara et al., 2014; Itoh et al., 2014; Koehler et al., 2016; Lu et al., 2017a; Sistani et al., 2011).

Sub-clustering of the stromal cluster identifies mesangial cells, as well as cortical and medullary stroma

In order to identify specific cell types within the stromal compartment, we have performed sub-clustering of cluster 3 of our initial analysis. The tSNE plot in Fig. 4A identifies three clusters, with sample contribution and heatmap of specific gene expression shown in Fig. S3A-C. A full list of genes is provided in Table S4. Cells in sub-cluster 2 of this sub-clustering present an expression profile compatible with their identity as mesangial cells. Both *ACTA2* and *ANGPT2* expression is specific for sub-cluster 2, and these genes are

markers of the mesangial cells in humans and mice, respectively (Lu et al., 2017b; Scheinman et al., 1976). We also find expression of *TAGLN* (transgelin), a marker of the migrating mesangial cells that is expressed during glomerular development in the rat embryonic kidney and in human and adult rat kidneys upon injury (Daniel et al., 2012).

Cells in sub-cluster 1 express *SFRP1*, *TNC*, *DCN* and collagens 3A1, 1A2 and 1A1 (*COL3A1*, *COL1A2* and *COL1A1*). This expression pattern confirms that these cells are renal stromal cells. Interestingly, both *SFRP1* and *DCN* are not only highly expressed in sub-cluster 1, but also are very specific to sub-cluster 1. On the other hand, *TNC* and the collagens are strongly expressed in sub-cluster 1 but they are also expressed in the other two sub-clusters.

Table 2. Representative genes from clusters 0-7 from sub-clustering of original clusters 0-2

Gene	Cluster	P	%*	%‡	Location	References
TPT1	0	5.08E-61	0.975	0.947	Ubiq	Fiucci et al., 2003
FTH1	0	2.91E-39	0.841	0.711	n/a	
UBB	0	4.59E-30	0.923	0.876	Ubiq	Goldstein et al., 1975
HBG1	0	8.20E-29	0.397	0.208	RBCs	Marks and Kovach, 1966
H3F3B	0	3.33E-28	0.827	0.747	n/a	
H19	1	5.28E-126	0.886	0.507	MM	GUDMAP
ITGA8	1	4.26E-92	0.406	0.085	CM	Valerius et al., 2002
NR2F2	1	2.88E-64	0.627	0.306	CM, ST	GUDMAP
EYA1	1	2.06E-62	0.287	0.059	CM, PTA	GUDMAP
MEIS2	1	1.63E-60	0.386	0.118	CM, ST	GUDMAP
JAG1	2	1.46E-24	0.432	0.24	PTA, RV, SSB	GUDMAP
CCND1	2	3.62E-22	0.586	0.399	PTA, DRV, SSB	GUDMAP
LHX1	2	2.10E-19	0.435	0.264	PTA, RV, SSB	GUDMAP
ADAMTS1	2	7.74E-17	0.276	0.14	SSB, PT	Thai and Iruela-Arispe, 2002
IGFBP4	2	2.55E-16	0.258	0.126	RV, ST	Rutledge et al., 2017
MECOM	3	4.72E-95	0.599	0.161	CNT, DT, LH	GUDMAP
GATA3	3	1.36E-53	0.296	0.061	CD	GUDMAP
MAL	3	6.45E-45	0.496	0.189	DT, CD	Frank et al., 1998
POU3F3	3	6.97E-43	0.454	0.159	SSB, LH,, DT, CD	GUDMAP
EMX2	3	2.81E-27	0.536	0.278	PTA, RV, SSB	GUDMAP
SIM2	4	3.80E-115	0.427	0.031	SSB	Brunskill et al., 2008
DKK1	4	4.45E-55	0.303	0.037	DRV, SSB	GUDMAP
EPHA7	4	1.73E-42	0.362	0.069	SSB, LH	GUDMAP
TBC1D4	4	2.53E-37	0.368	0.08	LH, DT, CNT, CD	Lier et al., 2012
EPHA4	4	5.00E-30	0.341	0.08	SSB	Peuckert et al., 2016
HIST1H1B	5	1.32E-165	0.729	0.046	Ubiq S-phase	Harris et al., 1991
HIST1H1A	5	9.30E-125	0.822	0.094	Ubiq S-phase	Harris et al., 1991
HIST1H1C	5	8.24E-107	0.748	0.089	Ubiq S-phase	Harris et al., 1991
HIST1H1D	5	2.95E-85	0.963	0.248	Ubiq S-phase	Harris et al., 1991
HIST1H1E	5	6.31E-84	0.785	0.134	Ubiq S-phase	Harris et al., 1991
KNG1	6	9.29E-92	0.53	0.031	rCD	El-Dahr et al., 1998
SLC12A1	6	3.09E-75	0.848	0.128	DT, LH	GUDMAP
UMOD	6	1.17E-57	0.5	0.048	DT, LH	GUDMAP
TMEM213	6	1.02E-34	0.47	0.068	hIC	Lindgren et al., 2017
DEFB1	6	6.83E-34	0.667	0.16	hLH, hDT, hCD	Valore et al., 1998
HEMGN	7	2.01E-149	0.452	0.006	RBCs	Yang et al., 2001
AHSP	7	8.57E-133	0.524	0.011	RBCs	Kihm et al., 2002
ALAS2	7	1.26E-116	0.571	0.017	RBCs	Sadlon et al., 1999
HBG1	7	8.98E-40	0.952	0.251	RBCs	Marks and Kovach, 1966
HBB	7	3.79E-38	0.976	0.27	RBCs	Marks and Kovach, 1966

*Percentage of cells in the cluster of interest.
‡Percentage of cells in all other clusters.
CD, collecting duct; CM, cap mesenchyme; CNT, connecting tubule; DRV, distal renal vesicle; DT, distal tubule; EC, endothelial cell; hCM, human cap mesenchyme; hIC, human intercalated cells; hLH, human loop of Henle; LH, loop of Henle; MM, metanephric mesenchyme; PT, proximal tubule; PTA, pre-tubular aggregate; RBCs, red blood cells; rCD, rat collecting duct; RV, renal vesicle; SSB, S-shaped bodies; ST, stroma; Ubiq, ubiquitous.
Unless stated otherwise, these expression patterns correspond to previously reported expression in mouse tissues.

Indeed, immunofluorescent detection of DCN and TNC in the developing human kidney (Fig. 4B,C) reveals a wide distribution in the renal stroma for TNC but a restriction to the medullary area for DCN. When we plot the expression of *TNC* and *DCN* in the tSNE plot (Fig. 4D), those cells expressing high *DCN* (blue dots) or high *DCN* with high *TNC* (green dots) are predominantly located in sub-cluster 1, indicating that this cluster represents the medullary stroma population. When we plot the expression of *TNC* and *MCAM*, markers of stromal cells that, in mice, differentiate into PECAM-expressing endothelial cells (Halt et al., 2016), we observe that their expression is mutually exclusive (Fig. 4E). These cells also express markers of mesangial cells and their expression pattern is compatible with the defining markers of a reported population of resident mesenchymal stem cells in the human glomeruli (Bruno et al., 2009).

In addition, cells in sub-cluster 0 have few specific genes, including *POSTN* and *CXCL12*. *Postn* (periostin) is expressed in the cortical stroma of the newborn rat kidney (Ito et al., 2002), in the mouse it is expressed in the renal stroma and the ureteric mesenchyme (Sorocos et al., 2011), and is a key player in the progression of renal disease, mediating inflammation and fibrosis (Prakoura and Chatziantoniou, 2017). In a similar fashion, *Cxcl12* is expressed in the cortical stroma of the developing mouse kidney, surrounding the cap mesenchyme and pre-tubular aggregates (Takabatake et al., 2009). This gene expression pattern indicates that cells in cluster 0 correspond to the cortical stroma. In mice, *Foxd1* is expressed in the stromal progenitors located in the cortex of the developing kidney and it is crucial for proper kidney development (Hatini et al., 1996). *FOXD1* expression is detected in a few of the human embryonic kidney cells in both clusters 0 and 1 (Fig. 4F). *PBX1* and *PDGFRA* are expressed at high levels in cluster 1; *Pbx1* is expressed in the developing and mature stroma of the embryonic mouse kidney, in perivascular cells along endothelia (Hurtado et al., 2015), whereas *PDGFRA* is expressed in the developing mouse kidney stroma and in kidney organoids from human iPS cells (Harding et al., 2011; McMahon et al., 2008; Takasato et al., 2015). The reduced number of cells expressing
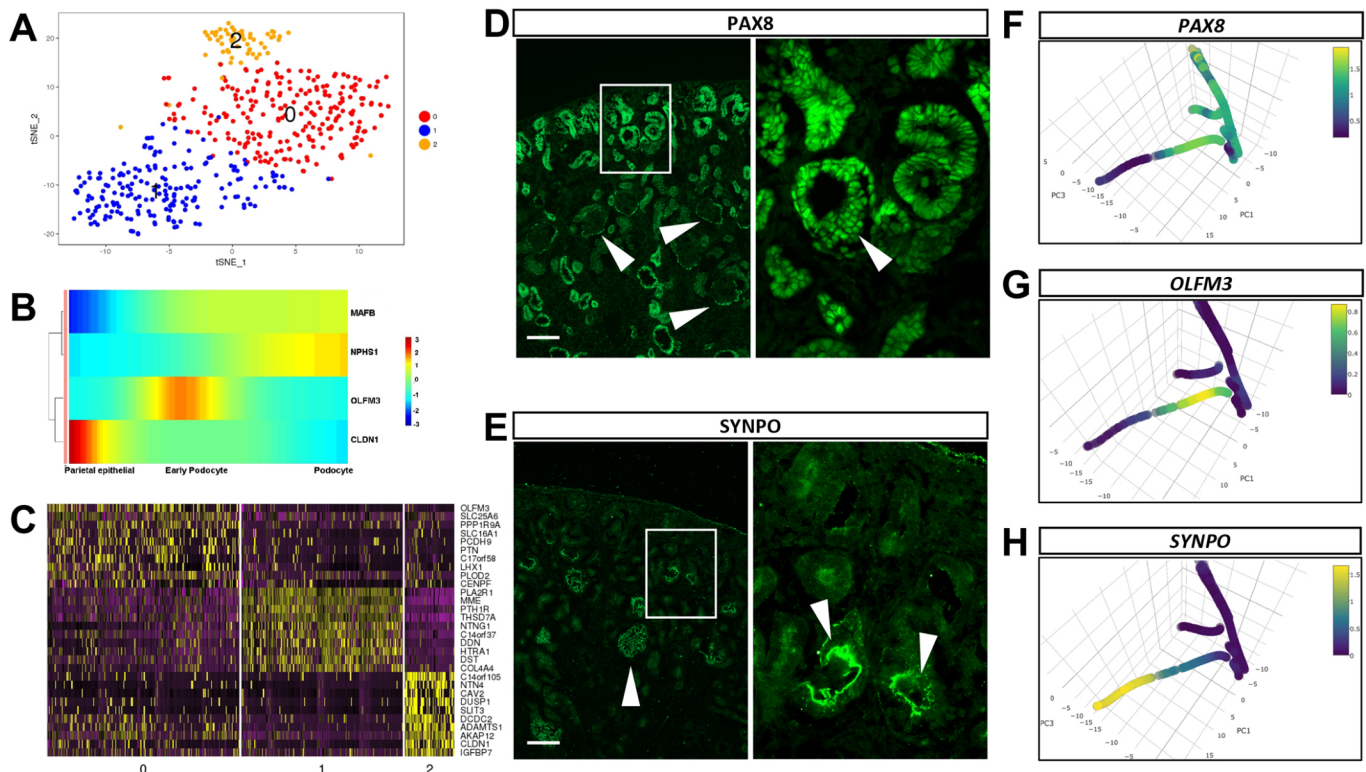
Fig. 3. Podocyte maturation and trajectory. (A) tSNE plot showing the three clusters from the sub-clustering of cluster 4 (podocyte like) from the initial clustering analyses. (B) Heatmap showing the expression levels of *MAFB*, *NPHS1*, *OLFM3* and *CLDN1* in parietal epithelial, early podocyte and mature podocyte cells. (C) Heatmap with the expression pattern of the top 10 cluster-specific genes in the three sub-clusters of the original cluster 4. (D) Immunofluorescent assay detecting PAX8 localization in the human embryonic kidney. Arrowheads indicate mature glomeruli where PAX8 is expressed in the parietal epithelial cells. Arrowhead in the high-magnification inset indicates an immature glomerulus where PAX8 is located in both the parietal epithelial cells and the developing podocytes. Scale bar: 100 μm. Representative image from at least three independent stainings. (E) Immunofluorescent assay detecting SYNPO localization in the human embryonic kidney. Strongest expression is detected in mature glomeruli (arrowheads). Scale bar: 100 μm. Representative image from at least three independent stainings. (F-H) Expression pattern of *PAX8* (F), *OLMF3* (G) and *SYNPO* (H) shown along the trajectory path of podocyte development.

FOXD1 in our analysis is likely due to a methodological cause, rather than reflecting a biological difference, as *FOXD1* has been shown to be expressed in the stromal progenitors in the developing human kidney (Lindstrom et al., 2018a).

The human collecting system maintains a tip/trunk identity
We have performed single-cell sequencing on human embryonic samples using fragments spanning from the nephrogenic zone to the inner medulla. Hence, we expect to identify cells from the collecting duct lineage that span from the progenitor UB tip cells to the mature cells of the collecting duct. Sub-clustering of the original collecting duct cluster 6 identifies three distinct clusters (Fig. S3D-G and Table S5). Cells in sub-cluster 0 express *LHX1*, which in mice is expressed both in the developing nephrons as well as in the UB (Kobayashi et al., 2005). Cells in sub-cluster 1 express *CALB1*, which is expressed in the human collecting duct but mainly excluded from the tip (Fig. S3I). Finally, sub-cluster 2 includes a small number of cells that express *KRT7*, *UPK1A* and *KRT19*, all markers of mature collecting duct cells. When we plot the expression of known markers of the UB and the collecting duct onto the tSNE plot from this sub-clustering, we confirm that nearly all cells express *GATA3*, whereas *SPINK1* marks the mature collecting duct cells and *CALB1* and *RET* expression show little overlap (Fig. S3H), consistent with their expression in the trunk and the tip of the UB, respectively. To further investigate the expression pattern of RET and its downstream targets in human embryonic samples, we have performed immunofluorescence

on human embryonic kidneys with specific antibodies against RET and ETV4 (Fig. 4G). RET is detected in the membrane of the cells at the tips of the branching UB, with some expression extending towards the trunk, and the ETV4 transcription factor is localized in the nucleus of the tip cells as well as in the early renal vesicles, an expression pattern that correlates with that of mouse embryonic kidneys (Lu et al., 2009). In addition, a gene expression heatmap for the collecting duct lineage generated using our trajectory analysis (Fig. 4H) identifies cells with expression of UB tip markers (*RET*, *ETV4*, *ETV5*, *DUSP6*) as well as cells expressing collecting duct markers (*GATA3*, *KRT8*, *KRT19*) and mature cells (*AQP2*, *MUC1*).

Ligand-receptor analysis identifies known and novel signaling networks in the developing human kidney
In order to identify signaling networks between the different compartments of the developing human kidney, we have selected ligands or receptors that were significantly differentially expressed in individual clusters from our initial analysis and search for their ligand/receptor partners in other clusters. This analysis has revealed a list of signaling pathways shown in Fig. S4A.

TNFSF10 [tumor necrosis factor (ligand) superfamily, member 10] is expressed in cluster 5, that corresponds to the proximal tubule cells. The TNFSF10 protein is localized to proximal and distal tubule cells in the adult human kidney (Song et al., 2004; Spierings et al., 2004). Like other tumor necrosis factors, TNFSF10 is involved in induced apoptosis as well as in inflammation, and in humans its increased
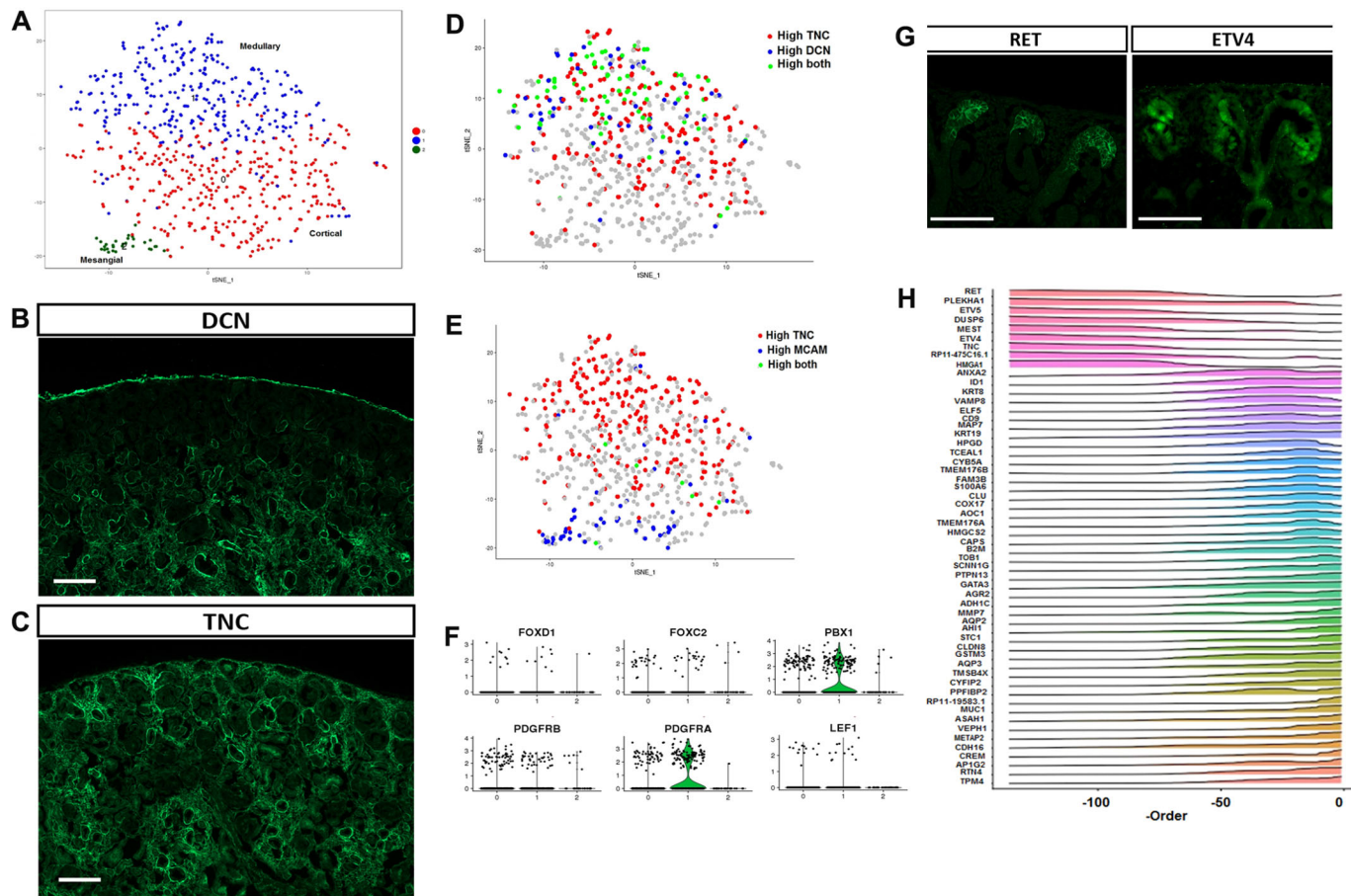
Fig. 4. Stromal and collecting duct cell types. (A) tSNE plot showing the three sub-clusters of cluster 3 (stromal) from the initial clustering. (B) Immunofluorescence assay showing decorin expression exclusively in the medullary stroma of the human developing kidney. Scale bar: 100 μm. Representative image from at least three independent stainings. (C) Tenascin C protein is detected by immunofluorescence in both cortical and medullary areas of the human embryonic kidney. Scale bar: 100 μm. Representative images from at least three independent stainings. (D) Feature plot of TNC and DCN expression in the three sub-clusters of original cluster 3 (stromal). Cells with high expression of TNC are in red and the cells with high expression of DCN are in blue. Cells with high expression of both TNC and DCN are in green. The FeaturePlot function in Seurat R package that shows co-expression of two genes was used to generate this plot. According to this function, for each gene, the cells are divided into two groups (intervals) of equal size based on the range of gene expression. The group of cells with higher expression is labeled as 'High'. (E) Feature plot of TNC and MCAM expression in the three sub-clusters of original cluster 3 (stromal). Cells with high expression of TNC are in red and the cells with high expression of MCAM are in blue. Cells with high expression of both TNC and MCAM are in green. (F) Violin plots depicting the expression levels of stromal genes, including *FOXD1*, *FOXC2*, *PBX1*, *PDGFRB*, *PDGFRA* and *LEF1*, in the three sub-clusters of cluster 3 (stromal) from the initial clustering. High expression of *FOXD1* and *FOXC2* was observed in multiple cells of sub-clusters 0 and 1. None of the cells expressed *LEF1* in sub-cluster 2. (G) Immunofluorescent detection of the receptor tyrosine kinase RET and its downstream target ETV4 in the developing human kidney. Scale bar: 100 μm. Representative images from at least three independent stainings. (H) Gene expression heatmap for the collecting duct differentiation.

expression is linked to the pathogenesis of diabetic nephropathy (Lorz et al., 2008).

Another group of signaling molecules identified in the analysis belong to the TGFβ superfamily and include *BMP7* as well as *TGFB2* and *TGFB3*. We observe significant differential expression of *BMP7* in cluster 4, which we have assigned a podocyte identity. In the mouse, *Bmp7* is expressed in the UB, the nephron mesenchyme, the developing nephron and the podocytes; removing *BMP7* from the podocytes during nephrogenesis results in hypoplastic kidneys and reduced proximal tubules (Kazama et al., 2008), suggesting a role for podocyte-specific BMP7 in promoting nephron differentiation and growth. TGFBR3 is an accessory receptor for the TGFB superfamily that is also significantly differentially expressed in the podocyte cluster. Previous studies have shown the expression of TGFBR3 in the developing mouse kidney and its role in establishing nephron endowment (Walker et al., 2011). Interestingly, TGFBR3 protein is strongly expressed in human adult podocytes (Uhlen et al., 2015) (www.proteinatlas.org/ENSG00000069702-TGFBR3/tissue/

kidney#img). Very little is known, however, about its expression or role in podocyte development or physiology. On the other hand, TGFBR2 is expressed in the endothelial cluster (cluster 9). TGFBR2 is expressed in endothelial cells during the development of several organs, including the heart, the brain and the skin (Jiao et al., 2006; Nguyen et al., 2011; Robson et al., 2010; Yamazaki et al., 2017). In the mouse kidney, deletion of Tgfbr2 in the renal pericytes compromises myofibroblast recruitment (LeBleu et al., 2013).

NOTCH signaling is also identified by our analysis (Fig. S4B). *JAG1* is significantly differentially expressed in the proximal tubule cluster (cluster 5), *NOTCH4* is specifically expressed in the endothelial cells (cluster 9) and NOTCH2 is more widely expressed in the cap mesenchyme, the stroma and the podocytes (clusters 0, 3 and 4, respectively). This expression pattern is compatible with previously described roles of NOTCH signaling in the developing mouse kidney (Cheng et al., 2007; Cheng et al., 2003) and with its role in the development of glomerular disease (Niranjan et al., 2008). Altogether, this analysis identifies signaling pathways that have been

described in animal models, as well as new pathways relevant to human kidney development and pathophysiology.

DISCUSSION

The analysis of single-cell RNA sequencing of human embryonic kidneys provides a powerful tool with which to dissect the gene expression profile during renal organogenesis. The fact that during most of development the kidney contains progenitors and differentiated cells, as well as cells at intermediate developmental stages, precludes the use of conventional high-throughput gene expression techniques. On the other hand, single-cell sequencing is uniquely poised to provide such information. We have used five independent samples and, despite the difference in age between them, they all belong to midpoints in renal development; the histology of the samples confirms the presence of progenitors as well as differentiated cell types (Fig. 1A, Fig. S1H). In addition, all five human developing kidney specimens contribute to each of the clusters (Fig. S1C,D), confirming that those cell types are present in all samples and that our clustering reflects biological differences between these cells rather than sample-to-sample differences.

We have performed an initial clustering followed by sub-clustering of several of the identified populations. This approach has allowed us to investigate the specific gene expression pattern of the developing nephron as shown in Fig. 3D. Although gene expression does not imply a lineage relationship, our data suggest the existence of several intermediate stages in the early developing nephron that are defined by gene expression. Some of these genes present unexpected expression patterns when compared with the mouse developing kidney. For example, both *IRX2* and *IRX3* are co-expressed with *POU3F3* in the distal segment of the s-shaped body; however, in the mouse, they are expressed in the medial segment of the s-shaped body and are absent from the more distal segment (Reggiani et al., 2007). Another example is *HNF4A*, which is also strongly expressed in a subpopulation of *POU3F3* distal cells but, in mice, is restricted to the medial proximal tubule progenitors (Heliot et al., 2013). POU3F3 has been recently identified as a distal marker in both mice and human developing nephrons (Lindstrom et al., 2018b); hence, it is a reliable marker of that segment. Interestingly, the most proximal lineages in the early nephron, i.e. podocytes and parietal epithelial cells, are absent from this sub-clustering. The development trajectory analysis (Fig. 1D) suggests that these are the first cells to express markers that separate them from the nephron progenitor lineage. Hence, the early expression of these genes caused these cells to cluster separately in cluster 4 of our initial analysis and were therefore not included in our sub-clustering analysis of clusters 0-2.

The trajectory analysis has also revealed a significant proximity between lineages that are very distinct in either their origin or their role in development, at least in mouse models. The proximity between the stromal trajectory and the nephron progenitor trajectory reflects a significant degree of gene expression overlap, despite the fact that the epithelial nephron is entirely derived from the nephron progenitors whereas the stroma provides survival cues as well as some non-epithelial cell types. This overlap is likely a reflection of their shared origin: the metanephric mesenchyme. On the other hand, the proximity between the collecting duct and the distal/connecting tubule trajectory reflects the expected proximity in gene expression, as well as the similar functional properties between these two closely located segments, despite originating from distinct renal compartments: the UB and the nephron progenitors, respectively.

The identity of each cluster is assigned based on known gene expression patterns either in human or in mouse developing kidneys; however, new markers for these clusters are also revealed by the analysis. Sub-cluster 1 derived from original clusters 0-2 corresponds to the nephrogenic mesenchyme and is characterized by the expression of *ITGA8*, *EYA1* and *TMEM100*, among other genes (Müller et al., 1997; Rumballe et al., 2011; Xu et al., 1999 and Fig. S1). In addition, we detected strong expression of *COL2A1*, which is a well-established chondrocyte marker (Ng et al., 1993) and *MEG3* (maternally expressed 3) in the nephrogenic mesenchyme. *Col2a1* is expressed in the developing mouse kidney (Wada et al., 1997) but its cell specificity and functional role remain to be elucidated. In humans, mutations in *COL2A1* cause Stickler syndrome, a genetic disorder affecting the connective tissue (Knowlton et al., 1989); however, no renal defects are known to be associated with this syndrome. *MEG3* has not been reported to be expressed in the nephron progenitors and mouse expression in GUDMAP appears to be stromal. Interestingly, *MEG3* is one of several imprinted genes found in sub-cluster 1. A total of five genes specific to the nephrogenic mesenchyme cluster (*H19*, *IGF2*, *MEG3*, *NNAT* and *PTPN14*) are known imprinted genes (Kikyo et al., 1997; Luedi et al., 2007; Miyoshi et al., 2000; Zemel et al., 1992). In addition, *Xist* is imprinted in mice but not in humans (Norris et al., 1994) and *WT1* is imprinted as its alternative form, *AWT1*, in Wilms tumor (Dallosso et al., 2004; Little et al., 1992). Of these, *MEG3*, *NNAT* and *PTPN14* have not been reported to be expressed in the developing kidney and their possible role in nephron progenitors has not been elucidated. Interestingly, Ptpn14 is a negative regulator of Yap (Liu et al., 2013) and Yap is essential for nephron induction and differentiation (Reginensi et al., 2013), and its subcellular localization is dysregulated in Wilms tumor (Murphy et al., 2014).

In the collecting duct (cluster 6 of the initial clustering), the genes most significantly defining the cluster were *GATA3*, *ELF5*, *AGR2*, *ADH1C* and *BCAT1*. GATA3 is a well-known marker of the collecting duct system in mice and human developing kidneys (Labastie et al., 1995; Oosterwegel et al., 1992), and Elf5 has been recently reported in mice as a specific transcription factor in the principal cell lineage (Grassmeyer et al., 2017). On the other hand, the expression of *AGR2*, *ADH1C* and *BCAT1* in the human developing kidney has not been previously reported. CagAgr2 (the fish homolog of AGR2) is expressed in the renal collecting system from gibel carp (Xia et al., 2009) and in the intrahepatic, hilar and extrahepatic biliary tree in embryonic and adult human samples (Lepreux et al., 2011), and Adh1 is expressed in the urothelium of the developing mouse kidney (Mitchell et al., 2006). *BCAT1* (branched chain aminotransferase 1) has not been studied in the kidney nor its expression reported; it has been shown to regulate early liver bud growth in the developing mice and in human embryonic stem cell-derived liver organoids (Koike et al., 2017).

Specific to the immature podocytes (sub-cluster 0 of the sub-clustering of initial cluster 4), we find the expression of olfactomedin 3 (*OLFM3*) (Fig. 3B,C,G). The function of OLFM3 has not been elucidated but it has been suggested to be involved in cell adhesion (Hillier and Vacquier, 2003). Also enriched in the immature podocytes, we find *SLC16A1*, *PCDH9* and *C17orf58* (*1810010H24Rik*). None of these genes has been characterized in the human developing kidney context. Genetic abnormalities in *SLC16A1* have been identified to cause congenital hyperinsulinism (Otonkoski et al., 2007); PCDH9 is a non-clustered protocadherin, mutations in humans are associated with autism spectrum disorders and its expression in glioblastoma suggest a role as tumor suppressor (Kim et al., 2011). On the other hand, mature podocytes (sub-cluster 1 of the sub-clustering of initial cluster 4) show specific expression of netrin G1 (*NTNG1*), a member of the

Netrin family. Netrins are extracellular, laminin-related proteins that provide guidance for migrating cells, and NTNG1 is a membrane-tethered glycophosphatidylinositol (GPI)-linked Netrin (Lai Wing Sun et al., 2011). Expression of *NTNG1* has been reported in the adult human kidney by northern blot and semi-quantitative PCR (Meerabux et al., 2005) but its specific podocyte expression in the developing kidney has not been reported. The analysis also unveiled a novel list of PDZ domain proteins expressed in the mature podocytes, suggesting a role for these proteins in the establishment or maintenance of the slit diaphragm in mature human podocytes. This clear distinction between immature and mature podocytes could reflect a transition between pre-functional and functional filtering structures, and the identification of novel markers for each of those populations provides additional tools in the study of podocyte function during development, in homeostasis and during renal disease.

These are just a few examples of the novel cluster-specific genes identified by our single-cell transcriptomics analysis that open new avenues to advance our understanding of human kidney development. Additional studies will be required to characterize the role of these expressed genes, and to determine their functional relevance, possibly through the use of cultured human tissue or via human renal organogenesis.

MATERIALS AND METHODS
Single-cell dissociation of human fetal kidney tissue using a cold active protease (subtilisin)
All research using human fetal tissue was approved by the University of Michigan institutional review board. Normal human fetal kidneys at 87, 105, 110, 115 and 132 days of gestation were obtained from the University of Washington Laboratory of Developmental Biology. Gestation age is estimated based on the date of the last period. All tissues were shipped overnight in Belzer's solution at 4°C and were processed immediately upon arrival to the laboratory. Single-cell dissociation was performed using a cold active protease, as described recently (Adam et al., 2017). The embryonic kidneys were decapsulated and cut in half. All tissue samples collected for digestion spanned from the cortical nephrogenic zone to the inner medulla, and were dissected in ice-cold PBS and finely minced in a petri dish on ice using razor blades. About 20 mg of tissue were added to 1 ml of ice-cold active protease solution [PBS, 10 mg of *Bacillus Licheniformis* protease (Sigma, #P5380), 5 mM CaCl$_2$, 20 U DNAse I (Roche, #4716728001)]. The tissue was incubated in a 2 ml reaction tube for 15-20 min on a slow-moving shaker (nutator) in a coldroom at 4°C with repeated trituration steps for 20 s every 5 min. Single-cell dissociation was confirmed with a microscope. The dissociation was stopped with 1 ml ice-cold PBS supplemented with 10% fetal bovine serum (FBS). Afterwards, the cells were immediately pelleted at 300 *g* for 5 min at 4°C. Subsequently, the supernatant was discarded and cells were suspended in 2 ml PBS/10% FBS and pelleted again at 300 *g* for 5 min at 4°C. Then cells were suspended in PBS/0.01% BSA and pelleted again (300 *g* for 5 min at 4°C), suspended in 1 ml PBS/0.01%BSA and passed through a 30 µM filter mesh (Miltenyi MACS smart strainer). Viability was then investigated with the Trypan-blue exclusion test and cell concentration was determined using a hemocytometer and adjusted to 200,000 cells/ml for the Drop-seq procedure.

Drop-seq
Uniformly dispersed 1 nl-sized droplets were generated using self-built polydimethylsiloxane (PDMS) microfluidic co-flow devices on the basis of the AutoCAD design provided by the McCarroll group. The devices were treated with a water repellant solution (Aquapel) to create a hydrophobic channel surface. Drop-Seq runs followed closely the procedure published by Macosko et al. (online Dropseq protocol v. 3.1, mccarrolllab.com/dropseq/). Barcoded beads (ChemGenes), suspended in lysis buffer, were co-flown with a single-cell suspension and a droplet generation mineral oil (QX200, Bio-Rad Laboratories). Resulting droplets were collected in a 50 ml tube and immediately disrupted after adding 30 ml high-salt saline-sodium citrate

buffer (6×SSC) and 1 ml perfluorooctanol. Subsequently, captured mRNAs were reverse transcribed for 2 h using 2000 U of the Maxima H Minus Reverse Transcriptase (ThermoFisher) followed by an exonuclease treatment for 45 min to remove unextended primers. After two washing steps with 6× SSC buffer, about 70,000 remaining beads (60% of input beads) were aliquoted (5000 beads per 50 µl reaction) and PCR amplified (five cycles at 65°C and 12 cycles at 67°C annealing temperature). Aliquots of each PCR reaction were pooled and double-purified using 0.5× volume of Agencourt AMPure XP beads (# A63881, Beckman Coulter) and finally eluted in 10 µl EB buffer. Quality and quantity of the amplified cDNAs were analyzed on a BioAnalyzer High Sensitivity DNA Chip (Agilent Technologies). About 600 pg cDNA was fragmented and amplified (17 cycles) to generate a next-generation sequencing library by using the Nextera XT DNA sample preparation kit (Illumina). The libraries were purified, quantified (Agilent High sensitivity DNA chip) and then sequenced (paired end 26×115 bases) on the Illumina HiSeq2500 platform. A custom primer (5′-GCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC-3′) was used for the first sequence read to identify all different cell barcodes und unique molecular identifier (UMI) sequences.

Data analysis of scRNASeq data
We processed the sequencing data using the Drop-seq pipeline (mccarrolllab. com/wp-content/uploads/2016/03/Drop-seqAlignmentCookbookv1.2Jan2016. pdf) and generated an expression data matrix. Each element of the matrix is the number of Unique Molecular Identifier (UMIs) associated with a gene in a barcoded cell. Further analysis of the expression matrix was performed using R package Seurat (Butler et al., 2018). We filtered out cells with fewer than 500 genes per cell and with more than 25% mitochondrial read content. The downstream analysis steps include log normalization, identification of highly variable genes across the single cells, scaling, PCA dimensionality reduction, unsupervised clustering and the discovery of differentially expressed cell-type-specific markers. In the scaling step, we regressed out technical variables, including mitochondrial read content, the number of UMI per cell and batch effect using Seurat ScaleData function. We used the default non-parametric Wilcoxon rank sum test for the differentially expression analysis. The clusters were viewed in a two-dimensional frame using tSNE clustering (Shekhar et al., 2016). The initial unsupervised clustering was performed at a resolution level of 0.6 and it resulted in 11 clusters. Next, we did sub-clustering on individual cell clusters. In the sub-clustering analysis, we regressed out one more parameter, cell cycle gene expression effect. The list of the cell cycle genes was downloaded from satijalab.org/seurat/cell_cycle_vignette.html. We used 0.6 as resolution for all of the unsupervised sub-clustering analyses. As the markers of cell clusters 0, 1 and 2 indicated these cell types to be less differentiated than the other cell clusters, we carried out sub-clustering of these three clusters together. The R package Monocle (Trapnell and Cacchiarelli, 2014) was used to find the expression dynamics of the signaling molecules across sub-clusters of the initial clusters. Cells with more than 1000 genes and the set of variable genes (genes that are more than one standard deviation apart from the average dispersion within a bin) from the Seurat sub-clustering analyses were used to order the cells using Monocle. Gene Ontology enrichment analysis was performed on the cluster-specific differentially expressed genes from the initial clustering analysis using iPathwayGuide (apps.advaitabio.com/oauth-provider/).

Ligand-receptor interaction
We used receptors and ligand pairs from the public resource Database of Interacting Partners (dip.doe-mbi.ucla.edu/dip/DLRP.cgi). Any ligand or receptor was classified as expressed if it was significantly differentially in a cell cluster (adjusted *P* value<0.001). Owing to extremely high or low cell numbers, to be more accurate, the expression of the corresponding partner had to be greater than the mean expression across all clusters.

Correlation analysis
The postnatal day 1 mouse kidney scRNASeq dataset (GSE94333) (Adam et al., 2017) was downloaded from NCBI GEO (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE94333) and analyzed in the same way as the human fetal kidney using the Seurat R package. All clusters published by the

authors have been identified. In addition, we found an immune cell cluster with 76 distinct transcriptomes. Correlation analysis was performed on the average gene expression in all 11 clusters from the initial unsupervised clustering of the human fetal kidney data with the clusters of postnatal day 1 mouse kidney data.

Trajectory estimation from single-cell RNA-seq data

Single-cell RNA-seq data was normalized and log-transformed by R package Seurat. The transformed data was filtered by the following criteria: (1) include only genes with detected expression in at least three cells; and (2) include only cells with at least 1000 genes detected and with less than 5% of total reads being mitochondria reads.

To remove technical variations, we regressed out the number of genes expressed, the percentage of mitochondria reads and batch variables via a linear model. To control for confounding effects of cell cycle to trajectory estimation, we normalized expression levels of non-cell cycle genes to a baseline estimated using expression of only cell cycle genes. Specifically, we used a curated cell cycle gene list of 1946 genes from Barron and Li (2016); for each non-cell cycle gene, we fitted a ridge regression model to predict its log-transformed expression level from cell-cycle genes, using the cv.glmnet function in R package glmnet. The regularization parameter was automatically selected based on 10-fold crossvalidation. With the fitted models, we computed the residuals of each non-cell cycle gene and used the residuals as input for trajectory analysis.

Trajectory analysis was performed with a method described by J.Z. and O.T. (unpublished) that uses a nonparametric estimation method based on probability distribution. To identify genes with significant expression change along the trajectory, for each trajectory segment of interest, we tested the significance of each gene by fitting a generalized additive model (GAM) with thin plate regression spline to predict expression level from the cell order in the trajectory segment using the gam function of R package mgcv. The significance level of the gene is the significance level of cell order in the GAM model.

Immunofluorescence

Kidneys were fixed overnight in 4% PFA and processed for frozen sectioning and incubation as previously reported (Cebrián, 2012). Sections (5 μm) were incubated at 4°C overnight with primary antibodies against the following epitopes: decorin (R&D systems, AF143), ETV4 (Proteintech, 10684-1-AP), PAX8 (Proteintech, 10336-1-AP), RET (R&D, AF1485), synaptopodin (Proteintech, 21064-1-AP) and tenascin C (R&D, MAB2138). Alexa 488-conjugated secondary antibodies were purchased from Jackson ImmunoResearch (705-546-147, 711-546-152 and 712-546-153) and incubated at 37°C for 3 h. All primary and secondary antibodies were used at 1:100 dilution. Imaging was performed on an Olympus IX71 inverted microscope or a Nikon A1 confocal microscope.

References

Adam, M., Potter, A. S. and Potter, S. S. (2017). Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. *Development* **144**, 3625-3632.

Assémat, E., Châtelet, F., Chandellier, J., Commo, F., Cases, O., Verroust, P. and Kozyraki, R. (2005). Overlapping expression patterns of the multiligand endocytic receptors cubilin and megalin in the CNS, sensory organs and developing epithelia of the rodent embryo. *Gene Expr. Patterns* **6**, 69-78.

Astier, A. L., Xu, R., Svoboda, M., Hinds, E., Munoz, O., de Beaumont, R., Crean, C. D., Gabig, T. and Freedman, A. S. (2003). Temporal gene expression profile of human precursor B leukemia cells induced by adhesion receptor: identification of pathways regulating B-cell survival. *Blood* **101**, 1118-1127.

Aufderheide, E., Chiquet-Ehrismann, R. and Ekblom, P. (1987). Epithelial-mesenchymal interactions in the developing kidney lead to expression of tenascin in the mesenchyme. *J. Cell Biol.* **105**, 599-608.

Bariety, J., Mandet, C., Hill, G. S. and Bruneval, P. (2006). Parietal podocytes in normal human glomeruli. *J. Am. Soc. Nephrol.* **17**, 2770-2780.

Barron, M. and Li, J. (2016). Identifying and removing the cell-cycle effect from single-cell RNA-Sequencing data. *Sci. Rep.* **6**, 33892.

Bruno, S., Bussolati, B., Grange, C., Collino, F., di Cantogno, L. V., Herrera, M. B., Biancone, L., Tetta, C., Segoloni, G. and Camussi, G. (2009). Isolation and characterization of resident mesenchymal stem cells in human glomeruli. *Stem Cells Dev.* **18**, 867-880.

Brunskill, E. W., Aronow, B. J., Georgas, K., Rumballe, B., Valerius, M. T., Aronow, J., Kaimal, V., Jegga, A. G., Yu, J., Grimmond, S. et al. (2008). Atlas of gene expression in the developing kidney at microanatomic resolution. *Dev. Cell* **15**, 781-791.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411.

Cacalano, G., Fariñas, I., Wang, L.-C., Hagler, K., Forgie, A., Moore, M., Armanini, M., Phillips, H., Ryan, A. M., Reichardt, L. F. et al. (1998). GFRalpha1 is an essential receptor component for GDNF in the developing nervous system and kidney. *Neuron* **21**, 53-62.

Carroll, T. J., Park, J.-S., Hayashi, S., Majumdar, A. and McMahon, A. P. (2005). Wnt9b plays a central role in the regulation of mesenchymal to epithelial transitions underlying organogenesis of the mammalian urogenital system. *Dev. Cell* **9**, 283-292.

Cebrián, C. (2012). Fluorescent immunolabeling of embryonic kidney samples. *Methods Mol. Biol.* **886**, 251-259.

Cebrian, C., Asai, N., D'Agati, V. and Costantini, F. (2014). The number of fetal nephron progenitor cells limits ureteric branching and adult nephron endowment. *Cell Rep.* **7**, 127-137.

Cheng, H.-T., Miner, J. H., Lin, M., Tansey, M. G., Roth, K. and Kopan, R. (2003). Gamma-secretase activity is dispensable for mesenchyme-to-epithelium transition but required for podocyte and proximal tubule formation in developing mouse kidney. *Development* **130**, 5031-5042.

Cheng, H.-T., Kim, M., Valerius, M. T., Surendran, K., Schuster-Gossler, K., Gossler, A., McMahon, A. P. and Kopan, R. (2007). Notch2, but not Notch1, is required for proximal fate acquisition in the mammalian nephron. *Development* **134**, 801-811.

Chi, X., Michos, O., Shakya, R., Riccio, P., Enomoto, H., Licht, J. D., Asai, N., Takahashi, M., Ohgami, N., Kato, M. et al. (2009). Ret-dependent cell rearrangements in the Wolffian duct epithelium initiate ureteric bud morphogenesis. *Dev. Cell* **17**, 199-209.

Crosnier, C., Attié-Bitach, T., Encha-Razavi, F., Audollent, S., Soudy, F., Hadchouel, M., Meunier-Rotival, M. and Vekemans, M. (2000). JAGGED1 gene expression during human embryogenesis elucidates the wide phenotypic spectrum of Alagille syndrome. *Hepatology* **32**, 574-581.

Dallosso, A. R., Hancock, A. L., Brown, K. W., Williams, A. C., Jackson, S. and Malik, K. (2004). Genomic imprinting at the WT1 gene involves a novel coding transcript (AWT1) that shows deregulation in Wilms' tumours. *Hum. Mol. Genet.* **13**, 405-415.

Daniel, C., Lüdke, A., Wagner, A., Todorov, V. T., Hohenstein, B. and Hugo, C. (2012). Transgelin is a marker of repopulating mesangial cells after injury and promotes their proliferation and migration. *Lab. Invest.* **92**, 812-826.

Das, A., Tanigawa, S., Karner, C. M., Xin, M., Lum, L., Chen, C., Olson, E. N., Perantoni, A. O. and Carroll, T. J. (2013). Stromal-epithelial crosstalk regulates kidney progenitor cell differentiation. *Nat. Cell Biol.* **15**, 1035-1044.

El-Dahr, S. S., Dipp, S., Yosipiv, I. V. and Carbini, L. A. (1998). Activation of kininogen expression during distal nephron differentiation. *Am. J. Physiol.* **275**, F173-F182.

Fiucci, G., Lespagnol, A., Stumptner-Cuvelette, P., Beaucourt, S., Duflaut, D., Susini, L., Amson, R. and Telerman, A. (2003). Genomic organization and expression of mouse Tpt1 gene. *Genomics* **81**, 570-578.

Frank, M., van der Haar, M. E., Schaeren-Wiemers, N. and Schwab, M. E. (1998). rMAL is a glycosphingolipid-associated protein of myelin and apical membranes of epithelial cells in kidney and stomach. *J. Neurosci.* **18**, 4901-4913.

Fujii, T., Pichel, J. G., Taira, M., Toyama, R., Dawid, I. B. and Westphal, H. (1994). Expression patterns of the murine LIM class homeobox gene lim1 in the developing brain and excretory system. *Dev. Dyn.* **199**, 73-83.

Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D. H., Niall, H. D. and Boyse, E. A. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc. Natl. Acad. Sci. USA* **72**, 11-15.

Grassmeyer, J., Mukherjee, M., deRiso, J., Hettinger, C., Bailey, M., Sinha, S., Visvader, J. E., Zhao, H., Fogarty, E. and Surendran, K. (2017). Elf5 is a principal cell lineage specific transcription factor in the kidney that contributes to Aqp2 and Avpr2 gene expression. *Dev. Biol.* **424**, 77-89.

Grieshammer, U., Cebrian, C., Ilagan, R., Meyers, E., Herzlinger, D. and Martin, G. R. (2005). FGF8 is required for cell survival at distinct stages of nephrogenesis and for regulation of gene expression in nascent nephrons. *Development* **132**, 3847-3857.

Grobstein, C. (1953). Inductive epitheliomesenchymal interaction in cultured organ rudiments of the mouse. *Science* **118**, 52-55.

Halt, K. J., Pärssinen, H. E., Junttila, S. M., Saarela, U., Sims-Lucas, S., Koivunen, P., Myllyharju, J., Quaggin, S., Skovorodkin, I. N. and Vainio, S. J. (2016). CD146(+) cells are essential for kidney vasculature development. *Kidney Int.* **90**, 311-324.

Harada, H., Kimura, A., Dong, R.-P., Xu, X. P., Bhatia, K. and Sasazuki, T. (1992). Sequencing and population analysis of four novel HLA-DPA1 alleles. *Hum. Immunol.* **35**, 173-178.

Harding, S. D., Armit, C., Armstrong, J., Brennan, J., Cheng, Y., Haggarty, B., Houghton, D., Lloyd-MacGilp, S., Pi, X., Roochun, Y. et al. (2011). The GUDMAP database–an online resource for genitourinary research. *Development* **138**, 2845-2853.

Harris, M. E., Böhni, R., Schneiderman, M. H., Ramamurthy, L., Schümperli, D. and Marzluff, W. F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. *Mol. Cell. Biol.* **11**, 2416-2424.

Hatini, V., Huh, S. O., Herzlinger, D., Soares, V. C. and Lai, E. (1996). Essential role of stromal mesenchyme in kidney morphogenesis revealed by targeted disruption of Winged Helix transcription factor BF-2. *Genes Dev.* **10**, 1467-1478.

Heliot, C., Desgrange, A., Buisson, I., Prunskaite-Hyyrylainen, R., Shan, J., Vainio, S., Umbhauer, M. and Cereghini, S. (2013). HNF1B controls proximal-intermediate nephron segment identity in vertebrates by regulating Notch signalling components and Irx1/2. *Development* **140**, 873-885.

Hillier, B. J. and Vacquier, V. D. (2003). Amassin, an olfactomedin protein, mediates the massive intercellular adhesion of sea urchin coelomocytes. *J. Cell Biol.* **160**, 597-604.

Holzinger, I., de Baey, A., Messer, G., Kick, G., Zwierzina, H. and Weiss, E. H. (1995). Cloning and genomic characterization of LST1: a new gene in the human TNF region. *Immunogenetics* **42**, 315-322.

Huang, L., Xiao, A., Choi, S. Y., Kan, Q., Zhou, W., Chacon-Heszele, M. F., Ryu, Y. K., McKenna, S., Zuo, X., Kuruvilla, R. et al. (2014). Wnt5a is necessary for normal kidney development in zebrafish and mice. *Nephron Exp. Nephrol.* **128**, 80-88.

Hudson, C. R., Bellew, T., Briggs, J. A., Casey, S. B. and Briggs, R. C. (1988). Production and use of rat monoclonal antibodies to the human myeloid cell nuclear differentiation antigen. *Hybridoma* **7**, 541-553.

Hurtado, R., Zewdu, R., Mtui, J., Liang, C., Aho, R., Kurylo, C., Selleri, L. and Herzlinger, D. (2015). Pbx1-dependent control of VMC differentiation kinetics underlies gross renal vascular patterning. *Development* **142**, 2653-2664.

Ihara, K.-I., Asanuma, K., Fukuda, T., Ohwada, S., Yoshida, M. and Nishimori, K. (2014). MAGI-2 is critical for the formation and maintenance of the glomerular filtration barrier in mouse kidney. *Am. J. Pathol.* **184**, 2699-2708.

Ito, T., Suzuki, A., Imai, E., Horimoto, N., Ohnishi, T., Daikuhara, Y. and Hori, M. (2002). Tornado extraction: a method to enrich and purify RNA from the nephrogenic zone of the neonatal rat kidney. *Kidney Int.* **62**, 763-769.

Itoh, M., Nakadate, K., Horibata, Y., Matsusaka, T., Xu, J., Hunziker, W. and Sugimoto, H. (2014). The structural and functional organization of the podocyte filtration slits is regulated of Tjp1/ZO-1. *PLoS ONE* **9**, e106621.

Jiao, K., Langworthy, M., Batts, L., Brown, C. B., Moses, H. L. and Baldwin, H. S. (2006). Tgfbeta signaling is required for atrioventricular cushion mesenchyme remodeling during in vivo cardiac development. *Development* **133**, 4585-4593.

Kalatzis, V., Sahly, I., El-Amraoui, A. and Petit, C. (1998). Eya1 expression in the developing ear and kidney: towards the understanding of the pathogenesis of Branchio-Oto-Renal (BOR) syndrome. *Dev. Dyn.* **213**, 486-499.

Kazama, I., Mahoney, Z., Miner, J. H., Graf, D., Economides, A. N. and Kreidberg, J. A. (2008). Podocyte-derived BMP7 is critical for nephron development. *J. Am. Soc. Nephrol.* **19**, 2181-2191.

Kihm, A. J., Kong, Y., Hong, W., Russell, J. E., Rouda, S., Adachi, K., Simon, M. C., Blobel, G. A. and Weiss, M. J. (2002). An abundant erythroid protein that stabilizes free alpha-haemoglobin. *Nature* **417**, 758-763.

Kikyo, N., Williamson, C. M., John, R. M., Barton, S. C., Beechey, C. V., Ball, S. T., Cattanach, B. M., Surani, M. A. and Peters, J. (1997). Genetic and functional analysis of neuronatin in mice with maternal or paternal duplication of distal Chr 2. *Dev. Biol.* **190**, 66-77.

Kim, S.-Y., Yasuda, S., Tanaka, H., Yamagata, K. and Kim, H. (2011). Non-clustered protocadherin. *Cell Adh. Migr.* **5**, 97-105.

Kiuchi-Saishin, Y., Gotoh, S., Furuse, M., Takasuga, A., Tano, Y. and Tsukita, S. (2002). Differential expression patterns of claudins, tight junction membrane proteins, in mouse nephron segments. *J. Am. Soc. Nephrol.* **13**, 875-886.

Knowlton, R. G., Weaver, E. J., Struyk, A. F., Knobloch, W. H., King, R. A., Norris, K., Shamban, A., Uitto, J., Jimenez, S. A. and Prockop, D. J. (1989). Genetic linkage analysis of hereditary arthro-ophthalmopathy (Stickler syndrome) and the type II procollagen gene. *Am. J. Hum. Genet.* **45**, 681-688.

Kobayashi, A., Kwan, K. M., Carroll, T. J., McMahon, A. P., Mendelsohn, C. L. and Behringer, R. R. (2005). Distinct and sequential tissue-specific activities of the LIM-class homeobox gene Lim1 for tubular morphogenesis during kidney development. *Development* **132**, 2809-2823.

Kobayashi, A., Valerius, M. T., Mugford, J. W., Carroll, T. J., Self, M., Oliver, G. and McMahon, A. P. (2008). Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development. *Cell Stem Cell* **3**, 169-181.

Koehler, S., Tellkamp, F., Niessen, C. M., Bloch, W., Kerjaschki, D., Schermer, B., Benzing, T. and Brinkkoetter, P. T. (2016). Par3A is dispensable for the function of the glomerular filtration barrier of the kidney. *Am. J. Physiol. Renal. Physiol.* **311**, F112-F119.

Koike, H., Zhang, R.-R., Ueno, Y., Sekine, K., Zheng, Y.-W., Takebe, T. and Taniguchi, H. (2017). Nutritional modulation of mouse and human liver bud growth through a branched-chain amino acid metabolism. *Development* **144**, 1018-1024.

Kolset, S. O. and Gallagher, J. T. (1990). Proteoglycans in haemopoietic cells. *Biochim. Biophys. Acta* **1032**, 191-211.

Krawczyk, K. M., Hansson, J., Nilsson, H., Krawczyk, K. K., Swärd, K. and Johansson, M. E. (2017). Injury induced expression of caveolar proteins in human kidney tubules - role of megakaryoblastic leukemia 1. *BMC Nephrol.* **18**, 320.

Labastie, M.-C., Catala, M., Gregoire, J.-M. and Peault, B. (1995). The GATA-3 gene is expressed during human kidney embryogenesis. *Kidney Int.* **47**, 1597-1603.

Lai Wing Sun, K., Correia, J. P. and Kennedy, T. E. (2011). Netrins: versatile extracellular cues with diverse functions. *Development* **138**, 2153-2169.

LeBleu, V. S., Taduri, G., O'Connell, J., Teng, Y., Cooke, V. G., Woda, C., Sugimoto, H. and Kalluri, R. (2013). Origin and function of myofibroblasts in kidney fibrosis. *Nat. Med.* **19**, 1047-1053.

Leimeister, C., Bach, A. and Gessler, M. (1998). Developmental expression patterns of mouse sFRP genes encoding members of the secreted frizzled related protein family. *Mech. Dev.* **75**, 29-42.

Lepreux, S., Bioulac-Sage, P. and Chevet, E. (2011). Differential expression of the anterior gradient protein-2 is a conserved feature during morphogenesis and carcinogenesis of the biliary tree. *Liver Int.* **31**, 322-328.

Lier, N., Gresko, N., Di Chiara, M., Loffing-Cueni, D. and Loffing, J. (2012). Immunofluorescent localization of the Rab-GAP protein TBC1D4 (AS160) in mouse kidney. *Histochem. Cell Biol.* **138**, 101-112.

Lindgren, D., Eriksson, P., Krawczyk, K., Nilsson, H., Hansson, J., Veerla, S., Sjölund, J., Höglund, M., Johansson, M. E. and Axelson, H. (2017). Cell-type-specific gene programs of the normal human nephron define kidney cancer subtypes. *Cell Rep.* **20**, 1476-1489.

Lindstrom, N. O., Guo, J., Kim, A. D., Tran, T., Guo, Q., De Sena Brandine, G., Ransick, A., Parvez, R. K., Thornton, M. E., Basking, L. et al. (2018a). Conserved and divergent features of mesenchymal progenitor cell types within the cortical nephrogenic niche of the human and mouse kidney. *J. Am. Soc. Nephrol.* **29**, 806-824.

Lindstrom, N. O., Tran, T., Guo, J., Rutledge, E., Parvez, R. K., Thornton, M. E., Grubbs, B., McMahon, J. A. and McMahon, A. P. (2018b). Conserved and divergent molecular and anatomic features of human and mouse nephron patterning. *J. Am. Soc. Nephrol.* **29**, 825-840.

Little, M. H., Dunn, R., Byrne, J. A., Seawright, A., Smith, P. J., Pritchard-Jones, K., van Heyningen, V. and Hastie, N. D. (1992). Equivalent expression of

paternally and maternally inherited WT1 alleles in normal fetal tissue and Wilms' tumours. *Oncogene* **7**, 635-641.

Liu, X., Yang, N., Figel, S. A., Wilson, K. E., Morrison, C. D., Gelman, I. H. and Zhang, J. (2013). PTPN14 interacts with and negatively regulates the oncogenic function of YAP. *Oncogene* **32**, 1266-1273.

Lorz, C., Benito-Martin, A., Boucherot, A., Ucero, A. C., Rastaldi, M. P., Henger, A., Armelloni, S., Santamaria, B., Berthier, C. C., Kretzler, M. et al. (2008). The death ligand TRAIL in diabetic nephropathy. *J. Am. Soc. Nephrol.* **19**, 904-914.

Lu, B. C., Cebrian, C., Chi, X., Kuure, S., Kuo, R., Bates, C. M., Arber, S., Hassell, J., MacNeil, L., Hoshi, M. et al. (2009). Etv4 and Etv5 are required downstream of GDNF and Ret for kidney branching morphogenesis. *Nat. Genet.* **41**, 1295-1302.

Lu, Y., Ye, Y., Bao, W., Yang, Q., Wang, J., Liu, Z. and Shi, S. (2017a). Genome-wide identification of genes essential for podocyte cytoskeletons based on single-cell RNA sequencing. *Kidney Int.* **92**, 1119-1129.

Lu, Y., Ye, Y., Yang, Q. and Shi, S. (2017b). Single-cell RNA-sequence analysis of mouse glomerular mesangial cells uncovers mesangial cell essential genes. *Kidney Int.* **92**, 504-513.

Luedi, P. P., Dietrich, F. S., Weidman, J. R., Bosko, J. M., Jirtle, R. L. and Hartemink, A. J. (2007). Computational and experimental identification of novel human imprinted genes. *Genome Res.* **17**, 1723-1730.

Marks, P. A. and Kovach, J. S. (1966). Development of mammalian erythroid cells. *Curr. Top. Dev. Biol.* **1**, 213-252.

McCright, B., Gao, X., Shen, L., Lozier, J., Lan, Y., Maguire, M., Herzlinger, D., Weinmaster, G., Jiang, R. and Gridley, T. (2001). Defects in development of the kidney, heart and eye vasculature in mice homozygous for a hypomorphic Notch2 mutation. *Development* **128**, 491-502.

McMahon, A. P., Aronow, B. J., Davidson, D. R., Davies, J. A., Gaido, K. W., Grimmond, S., Lessard, J. L., Little, M. H., Potter, S. S., Wilder, E. L. et al. GUDMAP Project. (2008). GUDMAP: the genitourinary developmental molecular anatomy project. *J. Am. Soc. Nephrol.* **19**, 667-671.

Meerabux, J. M. A., Ohba, H., Fukasawa, M., Suto, Y., Aoki-Suzuki, M., Nakashiba, T., Nishimura, S., Itohara, S. and Yoshikawa, T. (2005). Human netrin-G1 isoforms show evidence of differential expression. *Genomics* **86**, 112-116.

Metsuyanim, S., Levy, R., Davidovits, M. and Dekel, B. (2009). Molecular evaluation of circulating endothelial progenitor cells in children undergoing hemodialysis and after kidney transplantation. *Pediatr. Res.* **65**, 221-225.

Mitchell, E. K. L., Taylor, D. F., Woods, K., Davis, M. J., Nelson, A. L., Teasdale, R. D., Grimmond, S. M., Little, M. H., Bertram, J. F. and Caruana, G. (2006). Differential gene expression in the developing mouse ureter. *Gene Expr. Patterns* **6**, 519-538.

Miyoshi, N., Wagatsuma, H., Wakana, S., Shiroishi, T., Nomura, M., Aisaka, K., Kohda, T., Surani, M. A., Kaneko-Ishino, T. and Ishino, F. (2000). Identification of an imprinted gene, Meg3/Gtl2 and its human homologue MEG3, first mapped on mouse distal chromosome 12 and human chromosome 14q. *Genes Cells* **5**, 211-220.

Mugford, J. W., Sipilä, P., McMahon, J. A. and McMahon, A. P. (2008). Osr1 expression demarcates a multi-potent population of intermediate mesoderm that undergoes progressive restriction to an Osr1-dependent nephron progenitor compartment within the mammalian kidney. *Dev. Biol.* **324**, 88-98.

Müller, U., Wang, D., Denda, S., Meneses, J. J., Pedersen, R. A. and Reichardt, L. F. (1997). Integrin alpha8beta1 is critically important for epithelial-mesenchymal interactions during kidney morphogenesis. *Cell* **88**, 603-613.

Murphy, A. J., Pierce, J., de Caestecker, C., Libes, J., Neblett, D., de Caestecker, M., Perantoni, A. O., Tanigawa, S., Anderson, J. R., Dome, J. S. et al. (2014). Aberrant activation, nuclear localization, and phosphorylation of Yes-associated protein-1 in the embryonic kidney and Wilms tumor. *Pediatr. Blood Cancer* **61**, 198-205.

Ng, L. J., Tam, P. P. L. and Cheah, K. S. E. (1993). Preferential expression of alternatively spliced mRNAs encoding type II procollagen with a cysteine-rich amino-propeptide in differentiating cartilage and nonchondrogenic tissues during early mouse development. *Dev. Biol.* **159**, 403-417.

Nguyen, H.-L., Lee, Y. J., Shin, J., Lee, E., Park, S. O., McCarty, J. H. and Oh, S. P. (2011). TGF-beta signaling in endothelial cells, but not neuroepithelial cells, is essential for cerebral vascular development. *Lab. Invest.* **91**, 1554-1563.

Niranjan, T., Bielesz, B., Gruenwald, A., Ponda, M. P., Kopp, J. B., Thomas, D. B. and Susztak, K. (2008). The Notch pathway in podocytes plays a role in the development of glomerular disease. *Nat. Med.* **14**, 290-298.

Norris, D. P., Patel, D., Kay, G. F., Penny, G. D., Brockdorff, N., Sheardown, S. A. and Rastan, S. (1994). Evidence that random and imprinted Xist expression is controlled by preemptive methylation. *Cell* **77**, 41-51.

O'Brien, L. L., Guo, Q., Lee, Y. J., Tran, T., Benazet, J.-D., Whitney, P. H., Valouev, A. and McMahon, A. P. (2016). Differential regulation of mouse and human nephron progenitors by the Six family of transcriptional regulators. *Development* **143**, 595-608.

Ohse, T., Pippin, J. W., Vaughan, M. R., Brinkkoetter, P. T., Krofft, R. D. and Shankland, S. J. (2008). Establishment of conditionally immortalized mouse glomerular parietal epithelial cells in culture. *J. Am. Soc. Nephrol.* **19**, 1879-1890.

Oosterwegel, M., Timmerman, J., Leiden, J. and Clevers, H. (1992). Expression of GATA-3 during lymphocyte differentiation and mouse embryogenesis. *Dev. Immunol* **3**, 1-11.

Otonkoski, T., Jiao, H., Kaminen-Ahola, N., Tapia-Paez, I., Ullah, M. S., Parton, L. E., Schuit, F., Quintens, R., Sipilä, I., Mayatepek, E. et al. (2007). Physical exercise-induced hypoglycemia caused by failed silencing of monocarboxylate transporter 1 in pancreatic beta cells. *Am. J. Hum. Genet.* **81**, 467-474.

Pachnis, V., Mankoo, B. and Costantini, F. (1993). Expression of the c-ret proto-oncogene during mouse embryogenesis. *Development* **119**, 1005-1017.

Park, J.-S., Valerius, M. T. and McMahon, A. P. (2007). Wnt/beta-catenin signaling regulates nephron induction during mouse kidney development. *Development* **134**, 2533-2539.

Peuckert, C., Aresh, B., Holenya, P., Adams, D., Sreedharan, S., Porthin, A., Andersson, L., Pettersson, H., Wölfl, S., Klein, R. et al. (2016). Multimodal Eph/Ephrin signaling controls several phases of urogenital development. *Kidney Int.* **90**, 373-388.

Pichel, J. G., Shen, L., Sheng, H. Z., Granholm, A.-C., Drago, J., Grinberg, A., Lee, E. J., Huang, S. P., Saarma, M., Hoffer, B. J. et al. (1996). Defects in enteric innervation and kidney development in mice lacking GDNF. *Nature* **382**, 73-76.

Pietilä, I., Prunskaite-Hyyryläinen, R., Kaisto, S., Tika, E., van Eerde, A. M., Salo, A. M., Garma, L., Miinalainen, I., Feitz, W. F., Bongers, E. M. H. F. et al. (2016). Wnt5a deficiency leads to anomalies in ureteric tree development, tubular epithelial cell organization and basement membrane integrity pointing to a role in kidney collecting duct patterning. *PLoS ONE* **11**, e0147171.

Pode-Shakked, N., Pleniceanu, O., Gershon, R., Shukrun, R., Kanter, I., Bucris, E., Pode-Shakked, B., Tam, G., Tam, H., Caspi, R. et al. (2016). Dissecting stages of human kidney development and tumorigenesis with surface markers affords simple prospective purification of nephron stem cells. *Sci. Rep.* **6**, 23562.

Prakoura, N. and Chatziantoniou, C. (2017). Periostin in kidney diseases. *Cell. Mol. Life Sci.* **74**, 4315-4320.

Reggiani, L., Raciti, D., Airik, R., Kispert, A. and Brandli, A. W. (2007). The prepattern transcription factor Irx3 directs nephron segment identity. *Genes Dev.* **21**, 2358-2370.

Reginensi, A., Scott, R. P., Gregorieff, A., Bagherie-Lachidan, M., Chung, C., Lim, D.-S., Pawson, T., Wrana, J. and McNeill, H. (2013). Yap- and Cdc42-dependent nephrogenesis and morphogenesis during mouse kidney development. *PLoS Genet.* **9**, e1003380.

Riccio, P., Cebrian, C., Zong, H., Hippenmeyer, S. and Costantini, F. (2016). Ret and Etv4 promote directed movements of progenitor cells during renal branching morphogenesis. *PLoS Biol.* **14**, e1002382.

Robson, A., Allinson, K. R., Anderson, R. H., Henderson, D. J. and Arthur, H. M. (2010). The TGFbeta type II receptor plays a critical role in the endothelial cells during cardiac development. *Dev. Dyn.* **239**, 2435-2442.

Rumballe, B. A., Georgas, K. M., Combes, A. N., Ju, A. L., Gilbert, T. and Little, M. H. (2011). Nephron formation adopts a novel spatial topology at cessation of nephrogenesis. *Dev. Biol.* **360**, 110-122.

Rutledge, E. A., Benazet, J.-D. and McMahon, A. P. (2017). Cellular heterogeneity in the ureteric progenitor niche and distinct profiles of branching morphogenesis in organ development. *Development* **144**, 3177-3188.

Sadlon, T. J., Dell'Oso, T., Surinya, K. H. and May, B. K. (1999). Regulation of erythroid 5-aminolevulinate synthase expression during erythropoiesis. *Int. J. Biochem. Cell Biol.* **31**, 1153-1167.

Scheinman, J. I., Fish, A. J., Brown, D. M. and Michael, A. J. (1976). Human glomerular smooth muscle (mesangial) cells in culture. *Lab. Invest.* **34**, 150-158.

Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. and Pachnis, V. (1994). Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. *Nature* **367**, 380-383.

Shakya, R., Watanabe, T. and Costantini, F. (2005). The role of GDNF/Ret signaling in ureteric bud cell fate and branching morphogenesis. *Dev. Cell* **8**, 65-74.

Shekhar, K., Lapan, S. W., Whitney, I. E., Tran, N. M., Macosko, E. Z., Kowalczyk, M., Adiconis, X., Levin, J. Z., Nemesh, J., Goldman, M. et al. (2016). Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308-1323.e30.

Sistani, L., Dunér, F., Udumala, S., Hultenby, K., Uhlen, M., Betsholtz, C., Tryggvason, K., Wernerson, A. and Patrakka, J. (2011). Pdlim2 is a novel actin-regulating protein of podocyte foot processes. *Kidney Int.* **80**, 1045-1054.

Song, C. J., Liu, X. S., Zhu, Y., Chen, L. H., Jia, W., Li, Y. N., Cao, Y. X., Xie, X., Zhuang, R., Zhu, C. S. et al. (2004). Expression of TRAIL, DR4, and DR5 in kidney and serum from patients receiving renal transplantation. *Transplant. Proc.* **36**, 1340-1343.

Sorocos, K., Kostoulias, X., Cullen-McEwen, L., Hart, A. H., Bertram, J. F. and Caruana, G. (2011). Expression patterns and roles of periostin during kidney and ureter development. *J. Urol.* **186**, 1537-1544.

Spierings, D. C., de Vries, E. G., Vellenga, E., van den Heuvel, F. A., Koornstra, J. J., Wesseling, J., Hollema, H. and de Jong, S. (2004). Tissue

distribution of the death ligand TRAIL and its receptors. *J. Histochem. Cytochem.* **52**, 821-831.

Surendran, K., Boyle, S., Barak, H., Kim, M., Stomberski, C., McCright, B. and Kopan, R. (2010). The contribution of Notch1 to nephron segmentation in the developing kidney is revealed in a sensitized Notch2 background and can be augmented by reducing Mint dosage. *Dev. Biol.* **337**, 386-395.

Takabatake, Y., Sugiyama, T., Kohara, H., Matsusaka, T., Kurihara, H., Koni, P. A., Nagasawa, Y., Hamano, T., Matsui, I., Kawada, N. et al. (2009). The CXCL12 (SDF-1)/CXCR4 axis is essential for the development of renal vasculature. *J. Am. Soc. Nephrol.* **20**, 1714-1723.

Takasato, M., Er, P. X., Chiu, H. S., Maier, B., Baillie, G. J., Ferguson, C., Parton, R. G., Wolvetang, E. J., Roost, M. S., Chuva de Sousa Lopes, S. M. et al. (2015). Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature* **526**, 564-568.

Thai, S. N.-M. and Iruela-Arispe, M. L. (2002). Expression of ADAMTS1 during murine development. *Mech. Dev.* **115**, 181-185.

Trapnell, C. and Cacchiarelli, D. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381-386.

Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A. et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419.

Valerius, M. T. and McMahon, A. P. (2008). Transcriptional profiling of Wnt4 mutant mouse kidneys identifies genes expressed during nephron formation. *Gene Expr. Patterns* **8**, 297-306.

Valerius, M. T., Patterson, L. T., Feng, Y. and Potter, S. S. (2002). Hoxa 11 is upstream of Integrin alpha8 expression in the developing kidney. *Proc. Natl. Acad. Sci. USA* **99**, 8090-8095.

Valore, E. V., Park, C. H., Quayle, A. J., Wiles, K. R., McCray, P. B., Jr. and Ganz, T. (1998). Human beta-defensin-1: an antimicrobial peptide of urogenital tissues. *J. Clin. Invest.* **101**, 1633-1642.

Wada, J., Kumar, A., Ota, K., Wallner, E. I., Batlle, D. C. and Kanwar, Y. S. (1997). Representational difference analysis of cDNA of genes expressed in embryonic kidney. *Kidney Int.* **51**, 1629-1638.

Walker, K. A., Sims-Lucas, S., Caruana, G., Cullen-McEwen, L., Li, J., Sarraj, M. A., Bertram, J. F. and Stenvers, K. L. (2011). Betaglycan is required for the establishment of nephron endowment in the mouse. *PLoS ONE* **6**, e18723.

Wallgard, E., Larsson, E., He, L., Hellström, M., Armulik, A., Nisancioglu, M. H., Genove, G., Lindahl, P. and Betsholtz, C. (2008). Identification of a core set of 58 gene transcripts with broad and specific expression in the microvasculature. *Arterioscler. Thromb. Vasc. Biol.* **28**, 1469-1476.

Xia, J.-H., Jiang, J., Shi, Y.-H. and Gui, J.-F. (2009). Predominant expression and cellular distribution of fish Agr2 in renal collecting system. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **152**, 397-404.

Xu, P.-X., Adams, J., Peters, H., Brown, M. C., Heaney, S. and Maas, R. (1999). Eya1-deficient mice lack ears and kidneys and show abnormal apoptosis of organ primordia. *Nat. Genet.* **23**, 113-117.

Yamazaki, T., Nalbandian, A., Uchida, Y., Li, W., Arnold, T. D., Kubota, Y., Yamamoto, S., Ema, M. and Mukouyama, Y.-S. (2017). Tissue myeloid progenitors differentiate into pericytes through TGF-beta signaling in developing skin vasculature. *Cell Rep.* **18**, 2991-3004.

Yang, L. V., Nicholson, R. H., Kaplan, J., Galy, A. and Li, L. (2001). Hemogen is a novel nuclear factor specifically expressed in mouse hematopoietic development and its human homologue EDAG maps to chromosome 9q22, a region containing breakpoints of hematological neoplasms. *Mech. Dev.* **104**, 105-111.

Yun, K. and Perantoni, A. O. (2017). Hydronephrosis in the Wnt5a-ablated kidney is caused by an abnormal ureter-bladder connection. *Differentiation* **94**, 1-7.

Zemel, S., Bartolomei, M. S. and Tilghman, S. M. (1992). Physical linkage of two mammalian imprinted genes, H19 and insulin-like growth factor 2. *Nat. Genet.* **2**, 61-65.

Zimmermann, K., Opitz, N., Dedio, J., Renne, C., Muller-Esterl, W. and Oess, S. (2002). NOSTRIN: a protein modulating nitric oxide release and subcellular distribution of endothelial nitric oxide synthase. *Proc. Natl. Acad. Sci. USA* **99**, 17167-17172.

A.

Number of Genes/cell	Number of UMI per cell	Number of Cells	Percentage of Mitochondrial Reads
Min. : 501	Min. : 504.0	FK_105days : 1813	Min. : 0.001202
1st Qu. : 654	1st Qu. : 886.2	FK_110days : 1799	1st Qu. : 0.026760
Median : 842	Median : 1206.0	FK_115days : 977	Median : 0.036854
Mean : 1024	Mean : 1617.1	FK_132days : 801	Mean : 0.041930
3rd Qu. : 1182	3rd Qu. : 1818.0	FK_87days : 1024	3rd Qu. : 0.050069
Max. : 4978	Max. : 14622.0	NA	Max. : 0.247021

B.
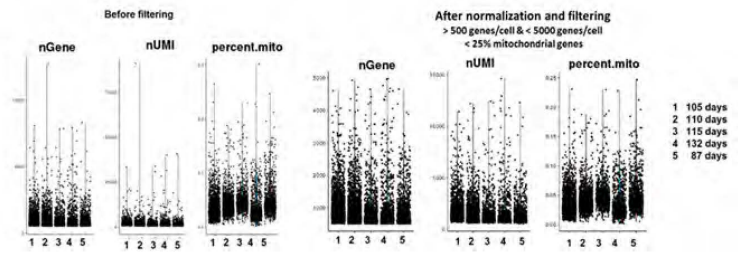
C.
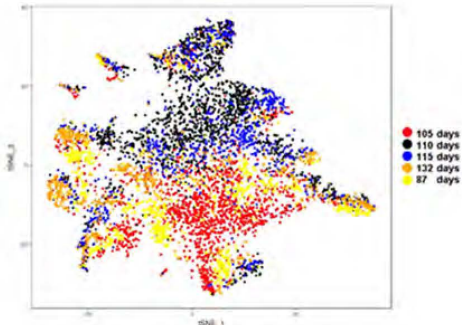
D.
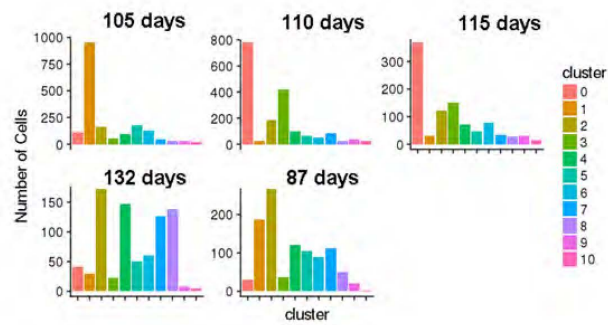
E.
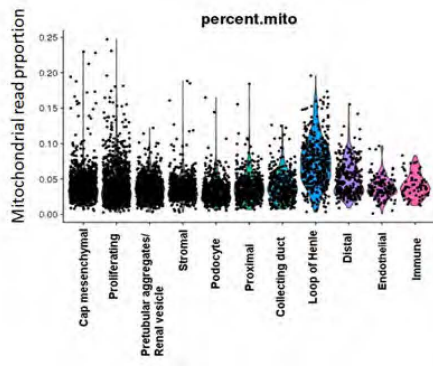
F.
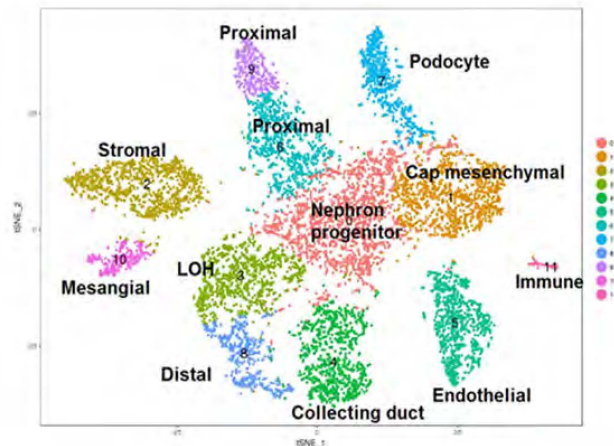
G.
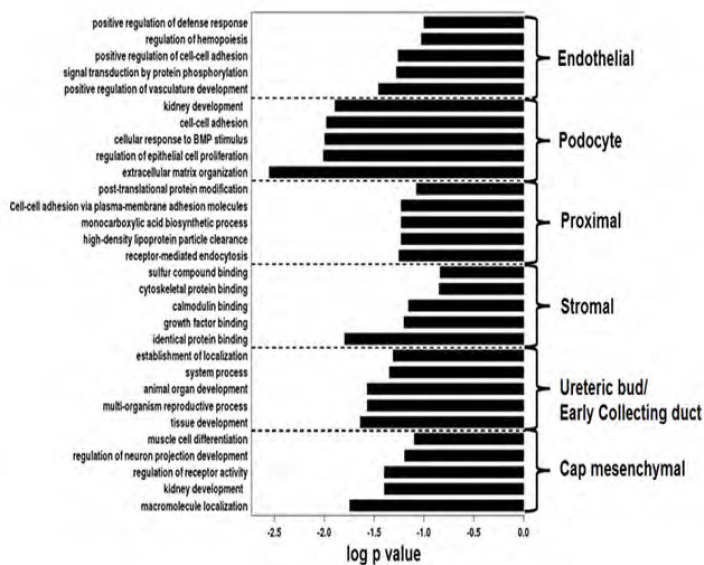
H.
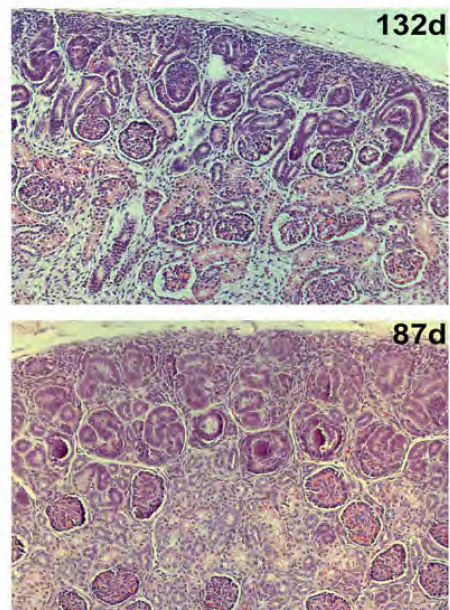
Figure S1: A) Table showing the summary of the quality control (QC) parameters including number of genes/cell (nGene), number of unique molecular identifier (nUMI) and mitochondrial read content for the 5 fetal kidney scRNASeq datasets **B)** Violin plots of the QC parameters in the 5 datasets before and after filtering out cells with < 500 genes and > 25% mitochondrial read content. **C)** tSNE plot depicting the distribution of cells from the 5 fetal kidney datasets in the 11 clusters from the initial unsupervised clustering. **D)** Bar plots depicting the contribution of the cells from the 5 fetal kidney datasets in the 11 clusters from the initial clustering **E)** Violin plots of mitochondrial read content per cell in the 11 clusters from the initial clustering. . **F)** tSNE plot showing the clusters from the unsupervised clustering of the postnatal day 1 mouse kidney samples (GSE94333). **G)** Top 5 Gene Ontology biological processes for the cluster-specific significantly differentially expressed (p < 0.05) genes of 11 clusters from the initial clustering. **H)** H&E staining of thin paraffin sections of human embryonic kidneys of 87 and 132 days post period.
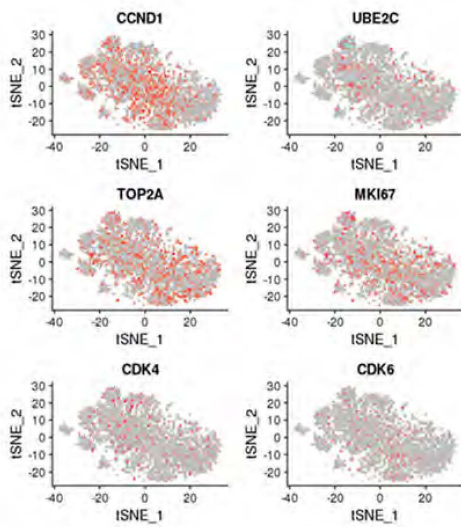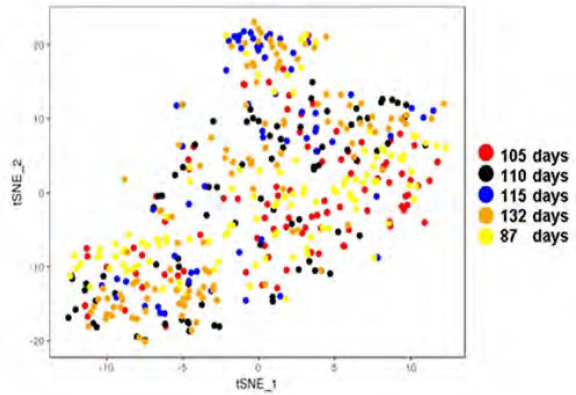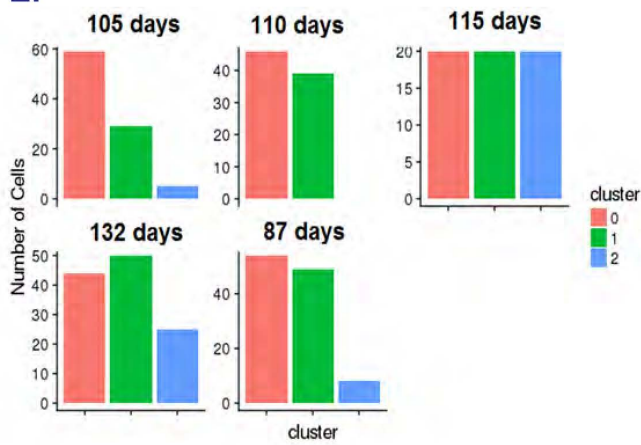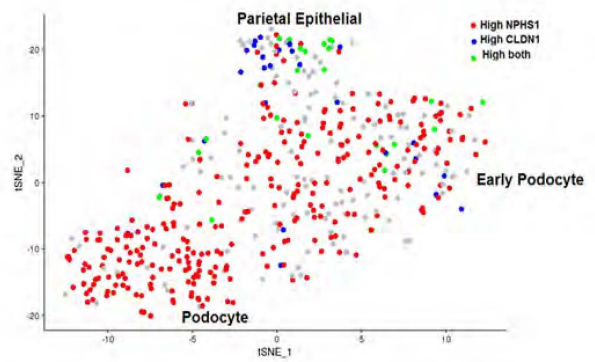
A.
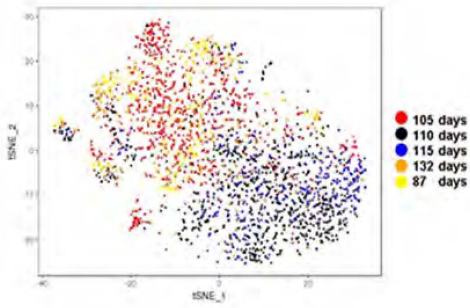
B.
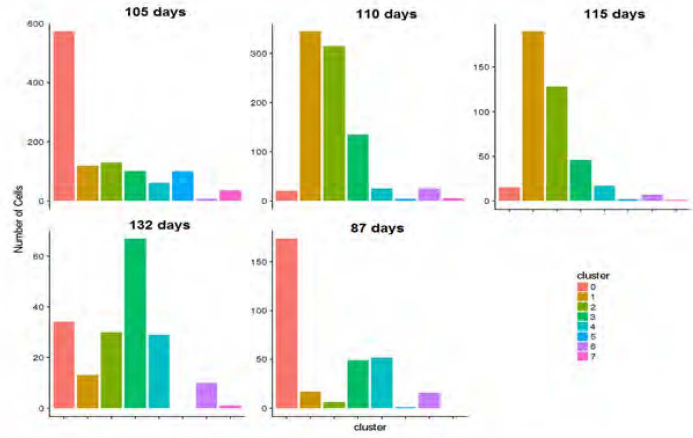
C.

D.

E.

F

Figure S2: A) tSNE plot depicting the distribution of cells from the 5 fetal kidney datasets in the 8 sub-clusters of initial clusters 0, 1 and 2   **B)** Bar plots depicting the contribution of the cells from the 5 fetal kidney datasets in the 8 sub-clusters from the initial clusters 0,1 and 2. **C)** Feature plot of proliferation-related genes including CCND1, UBE2C, TOP2A, MKI67, CDK4 and CDK6 in the sub-clustering from the initial unsupervised clustering of clusters 0,1 and 2. **D)** tSNE plot depicting the distribution of cells from the 5 fetal kidney datasets in the 3 sub-clusters of cluster 4 from the initial unsupervised clustering. **E)** Bar plots depicting the contribution of the cells from the 5 fetal kidney datasets in the 3 sub-clusters from the initial cluster 4. **F)** Feature plot of NPHS1, CLDN1 expression in the three sub-clusters of original cluster 4. Cells with high expression of NPHS1 are in red and the cells with high expression of CLDN1 are in blue. Cells with high expression of both NPHS1 and CLDN1 are in green.
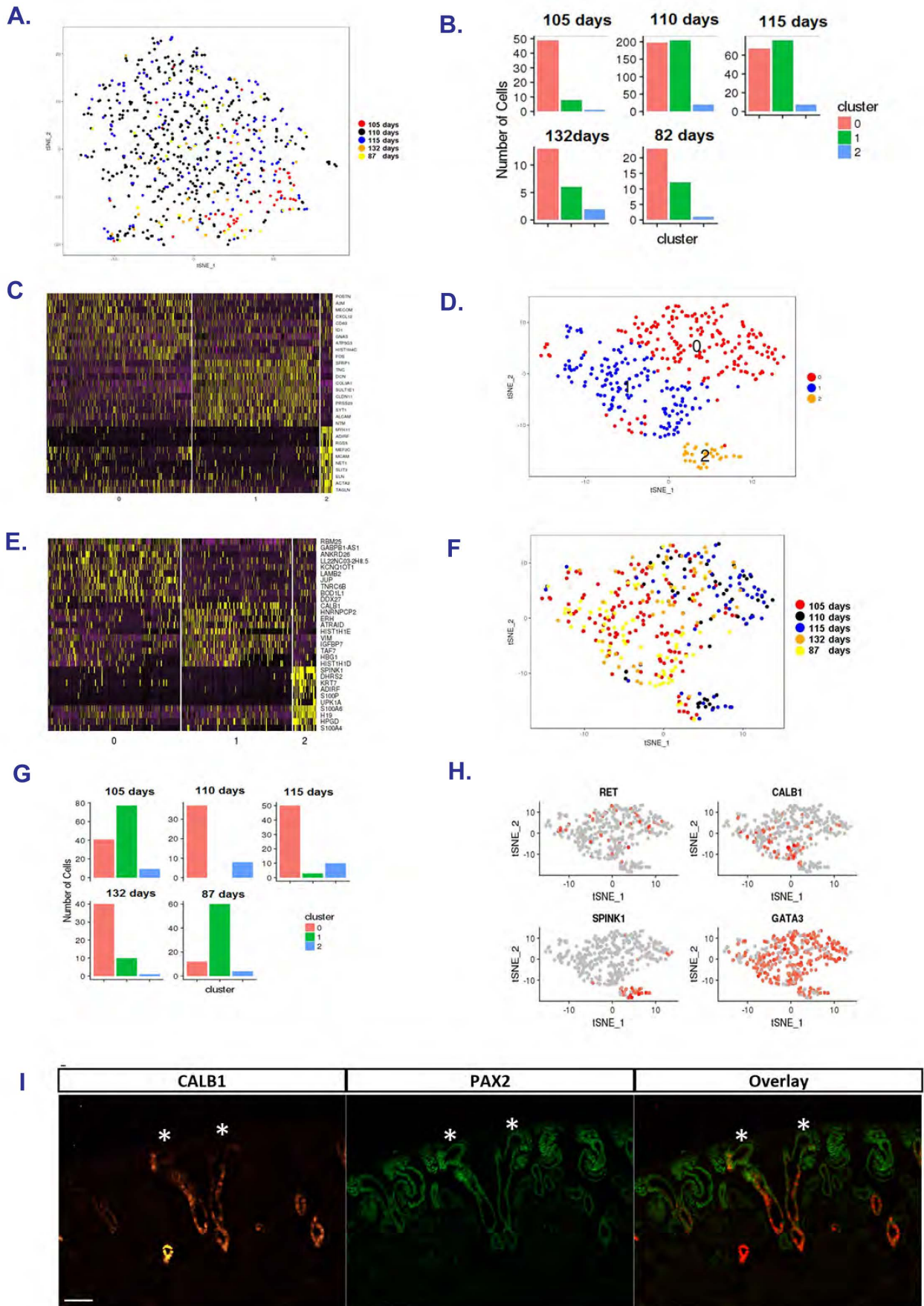
Figure S3: A) tSNE plot depicting the distribution of cells from the 5 fetal kidney datasets in the 3 sub-clusters of cluster 3 (Stromal) from the initial unsupervised clustering. **B)** Bar plots depicting the contribution of the cells from the 5 fetal kidney datasets in the 3 sub-clusters from the initial cluster 3 (Stromal) **C)** Heatmap with the expression of the top 10 cluster-specific genes in the 3 sub-clusters of the original cluster 3 (Stromal). **D)** tSNE plot showing the three sub-clusters of cluster 6 (UB) from the initial clustering. **E)** Heatmap with the expression of the top 10 cluster-specific genes in the 3 sub-clusters of the original cluster 6 (UB). **F)** tSNE plot depicting the distribution of cells from the 5 fetal kidney datasets in the 3 sub-clusters of cluster 6 (UB) from the initial unsupervised clustering. **G)** Bar plots depicting the contribution of the cells from the 5 fetal kidney datasets in the 3 sub-clusters from the initial cluster 6 (UB) **H)** Expression of known markers (RET, CALB1, SPINK1, GATA3) in cluster 6 (UB) sub-clusters. **(I)** Immunofluorescent staining of human embryonic sections with anti-Calbindin1 and anti-PAX2 antibodies. Calbindin staining is excluded from the tips of the ureteric buds (asterisks). Scale bar: 100um. Representative image of 3 independent stainings.

A

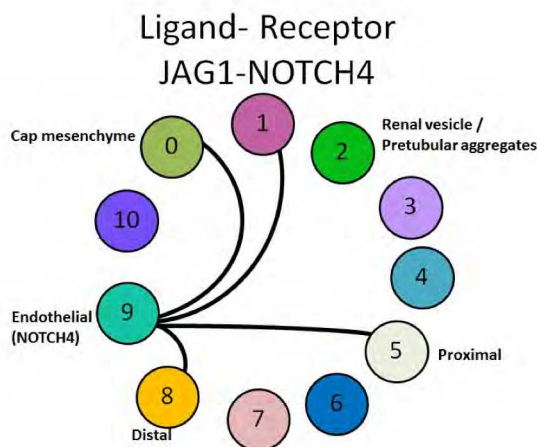| Significantly differentially expressed gene (ligand/receptor)	cluster	Corresponding interactor (ligand/receptor)	cluster (> mean expression across all clusters)
BMP7	4	ACVR1	0, 3
EFNB2	4, 6, 9	EPHB3	0, 2, 7, 8
EPHA7	5	EPHB3	0, 2, 7, 8
ERBB2	8	NRG1	4, 6, 7
ERBB4	2, 7, 8	BTC	4, 5, 6
FLT1	9	VEGFC	4, 8, 9
IGF2	0, 3	IGF2R	4, 7, 8, 9
IL6ST	9	CTF1	2, 4, 7, 8
JAG1	5	NOTCH4	9
JAG1	5	NOTCH2	0, 3, 4
KDR	9	VEGFC	4, 8, 9
LIFR	9	CTF1	2, 4, 7, 8
MDK	0, 3	PTPRB	9
NOTCH4	9	JAG1	0, 1, 2, 5, 8
NTRK2	0	BDNF	0, 3, 4, 5, 6
PTN	3	PTPRB	9
TGFBR2	9	TGFB3	0, 3, 6
TGFBR2	9	TGFB2	0, 3
TGFBR3	4	TGFB3	0, 3, 6, 9
TGFBR3	4	TGFB2	0, 3
TNFSF10	5	TNFRSF10B	2, 4, 5, 7, 8

B

Figure S4: A) Table showing the list of either ligand or receptor that was significantly differentially expressed in a cell cluster compared to all other clusters (adjusted p value < 0.001). The corresponding interacting partner expression is greater than the mean expression across all other clusters. **B)** The figure illustrates the possible interaction of ligand JAG1-NOTCH4 interaction. NOTCH4 expression was specifically seen in endothelial cell cluster.

Supplementary Tables

Table S1: The table provides list of significantly expressed genes for each of the 11 clusters compared to all other clusters. pct.1 is the proportion of cells expressing the gene in the cluster of interest and pct.2 is the proportion of cells expressing the gene in all other clusters. avg_logFC is the log fold-change of the average expression between the cluster of interest and all other clusters. Positive values indicate that the gene is more highly expressed in the first group.

Click here to Download Table S1

Table S2: The cells from the progenitor cell clusters (clusters 0, 1 and 2) in the first clustering were further sub-clustered. The table provides list of significantly expressed genes for each of the 8 sub-clusters compared to all other sub-clusters. pct.1 is the proportion of cells expressing the gene in the cluster of interest and pct.2 is the proportion of cells expressing the gene in all other clusters. avg_logFC is the log fold-change of the average expression between the cluster of interest and all other clusters. Positive values indicate that the gene is more highly expressed in the first group.

Click here to Download Table S2

Table S3: The cells from the podocyte-like cluster (cluster 4) in the first clustering were further sub-clustered. The table provides list of significantly expressed genes for each of the 3 sub-clusters compared to all other sub-clusters. pct.1 is the proportion of cells expressing the gene in the cluster of interest and pct.2 is the proportion of cells expressing the gene in all other clusters. avg_logFC is the log fold-change of the average expression between the cluster of interest and all other clusters. Positive values indicate that the gene is more highly expressed in the first group.

Click here to Download Table S3

Table S4: The cells from the stromal cluster (cluster 3) in the first clustering were further sub-clustered. The table provides list of significantly expressed genes for each of the 3 sub-clusters compared to all other sub-clusters. pct.1 is the proportion of cells expressing the gene in the cluster of interest and pct.2 is the proportion of cells expressing the gene in all other clusters. avg_logFC is the log fold-change of the average expression between the cluster of interest and all other clusters. Positive values indicate that the gene is more highly expressed in the first group.

Click here to Download Table S4

Table S5: The cells from the collecting duct-like cluster (cluster 6) in the first clustering were further sub-clustered. The table provides list of significantly expressed genes for each of the 3 sub-clusters compared to all other sub-clusters. pct.1 is the proportion of cells expressing the gene in the cluster of interest and pct.2 is the proportion of cells expressing the gene in all other clusters. avg_logFC is the log fold-change of the average expression between the cluster of interest and all other clusters. Positive values indicate that the gene is more highly expressed in the first group.

Click here to Download Table S5
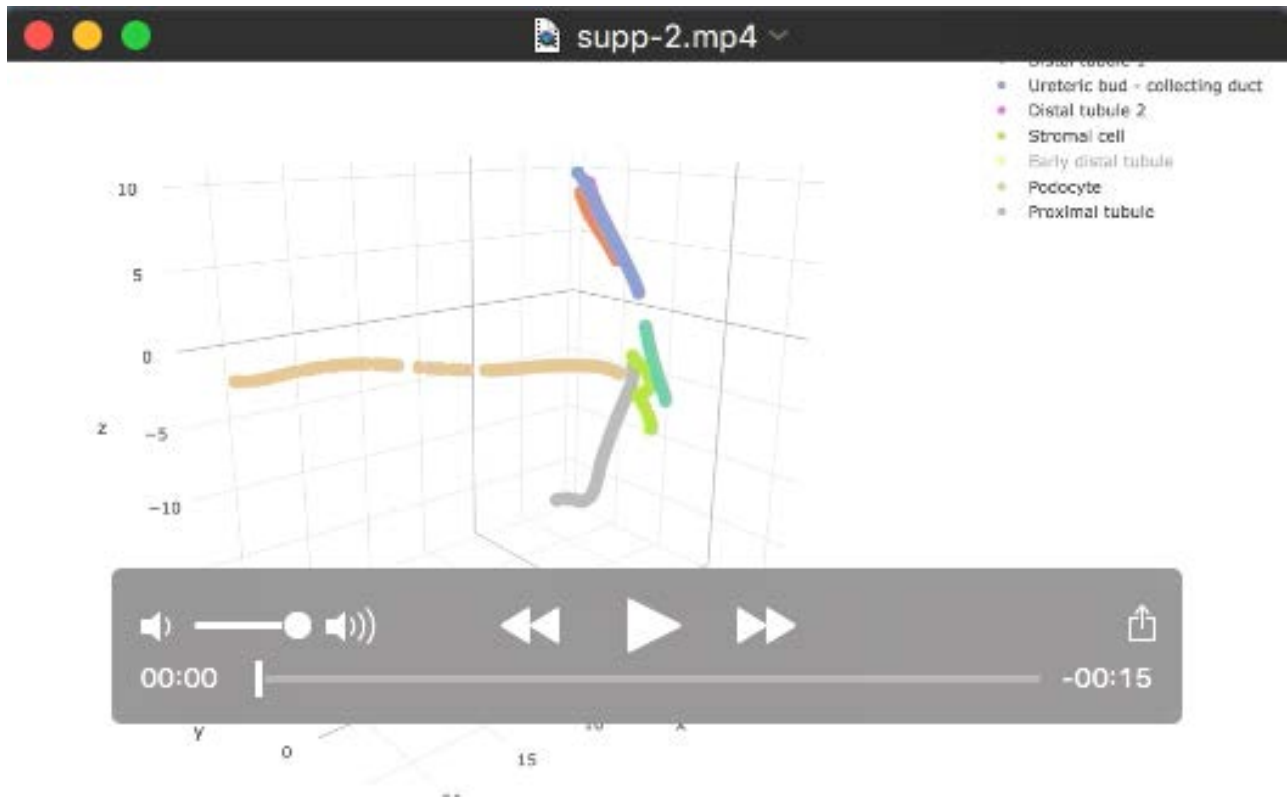
Movie 1: 3-Dimensional view of the pseudo-temporal trajectory analysis of the fetal kidney single cell expression data. Single cell RNASeq data from 5 human fetal kidney datasets (87, 105, 110, 115, and 132 days old) were used in this analysis. Each developmental trajectory is color-coded and each dot represents a cell. Cells with > 1000 genes /cell and < 25% mitochondrial read content were used this cell-ordering analysis. 3 independent developmental trajectories were identified: the stroma, the collecting duct and the nephron lineage. The proximity between the stromal trajectory and the nephron progenitor trajectory reflects a significant degree of gene expression overlap. Similarly, the proximity between the collecting duct and the distal/connecting tubules trajectory reflects the expected proximity in gene expression, as well as function between these two closely located segments, despite originating from distinct renal compartments, the UB and the nephron progenitors respectively.