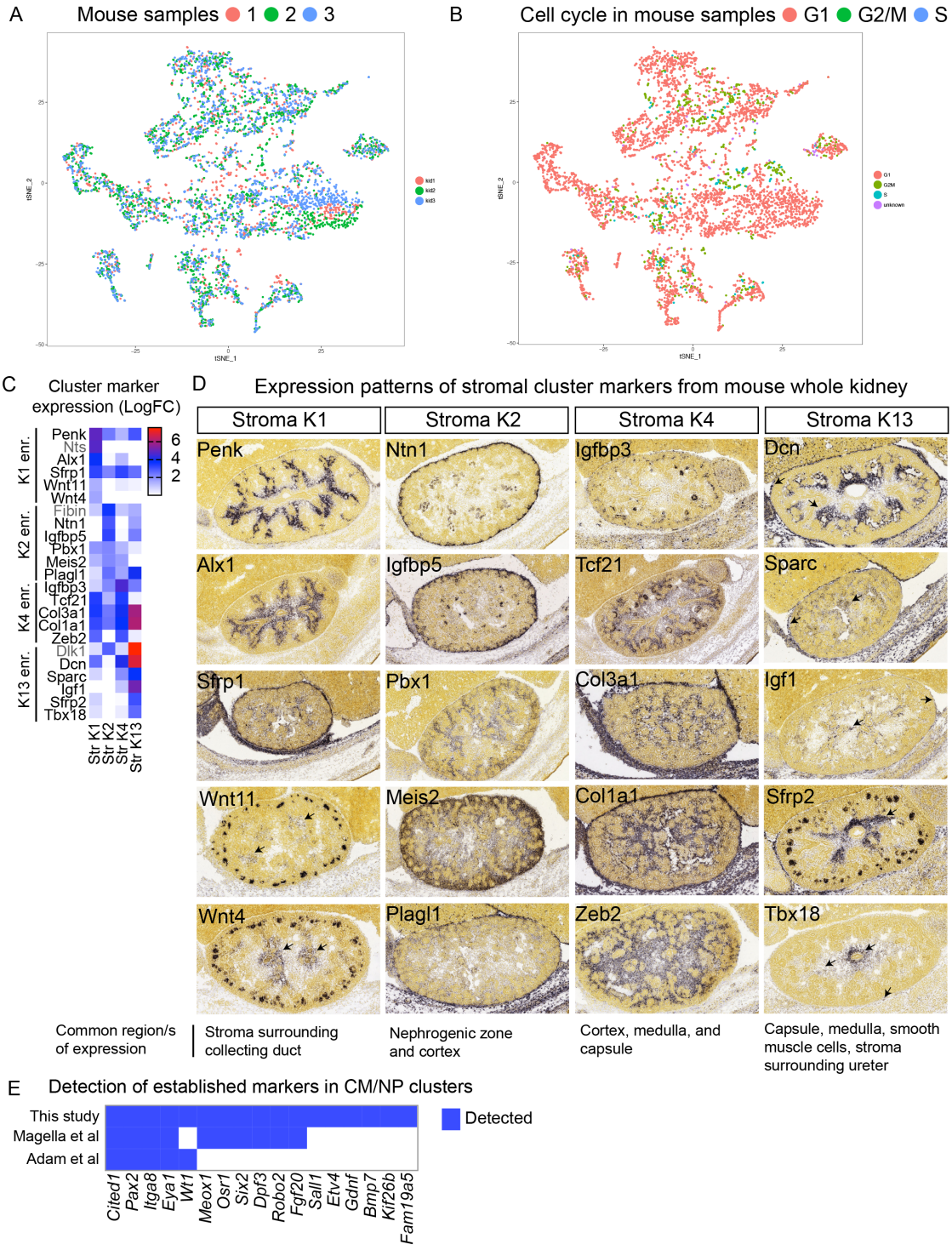


Figure S1. Visualisation of data by sample and cell cycle and supporting information. A. tSNE plot of cells identified by sample of origin shows an even distribution of cell types present within all samples. **B.** tSNE plot of all cells from mouse developing kidney identified by stage of cell cycle (G1, G2/M, S). **C.** Heatmap of stromal cluster markers from whole kidney. Markers in black indicate genes shown in D. **D**. *In situ* hybridisation results from the Allen Developing Mouse Brain Atlas (http://developingmouse.brain-map.org) for stromal marker genes used to aid in cluster mapping. Note that expression domains within each cluster do not completely overlap indicating further heterogeneity. **E.** Detection of established cap mesenchyme (CM)/nephron progenitor (NP) markers in any corresponding clusters from this study, Adam et al (2017) and Magella et al (2017) shows a 40-70% increase in detection of relevant markers in this dataset. Detection in this dataset = LogFC>0.94; Magella = featured in 'cell-type specific gene lists' reported in SuppTable4 for any cap mesenchyme cluster (at any Pearson.rho value); Adam = featured in TableS6 'compartment specific gene lists' for cap mesenchyme. NP expression of *Etv4* (aka *Pea3*) first demonstrated in Lu et al., Nat. Genet. 2009 and Mugford et al., Dev. Biol. 2009.

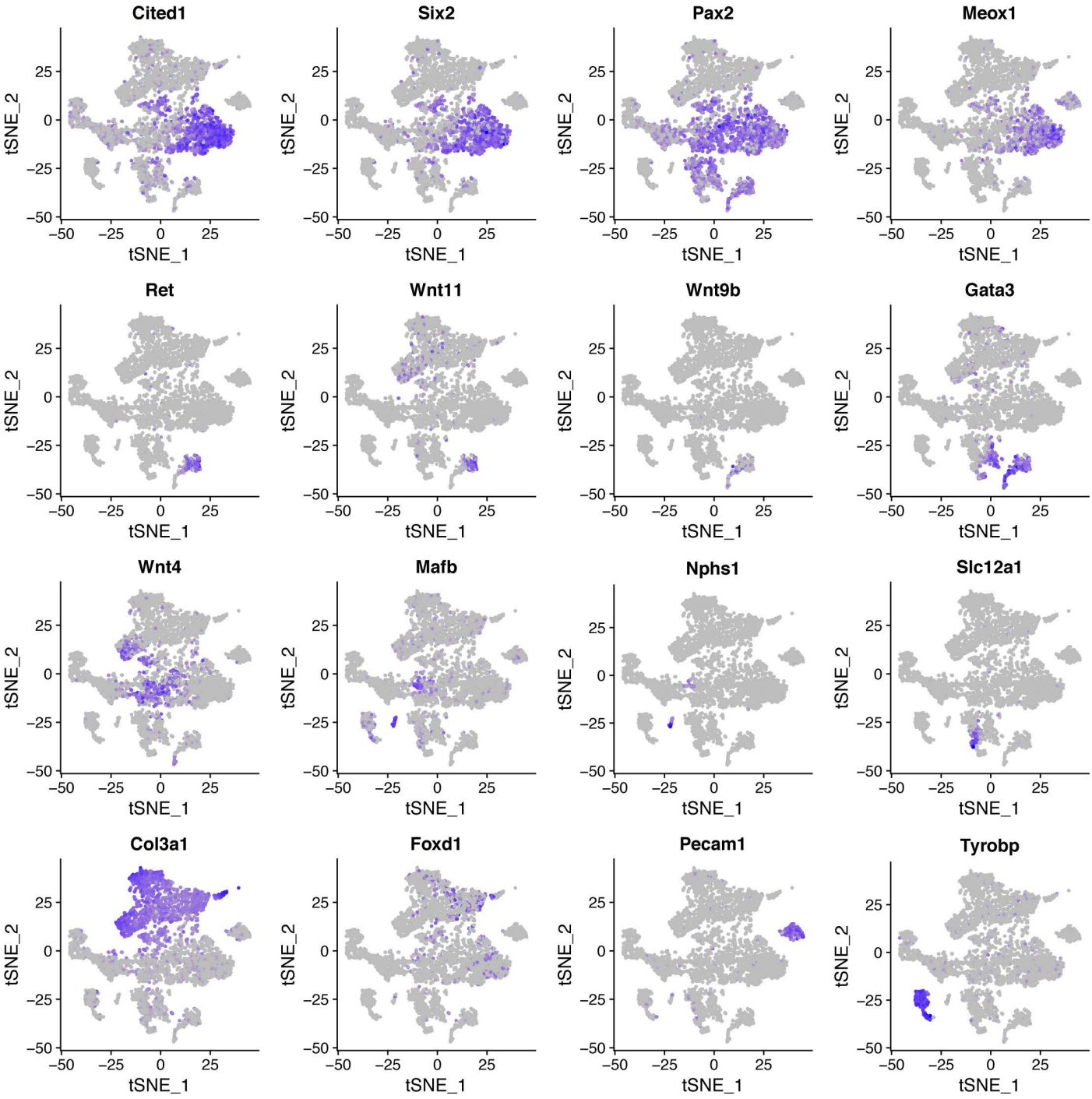


Figure S2. Marker expression across all cells in the developing mouse kidney data. tSNE plot of all cells from E18.5 developing mouse kidney showing the location of cells expressing key marker genes, including markers of nephron progenitor (*Cited1*, *Six2*, *Pax2*, *Meox1*), ureteric epithelium (*Ret*, *Wnt11*, *Wnt9b*, *Gata3*), early nephron (*Wnt4*), podocyte (*Mafb*, *Nphs1*), distal tubule (*Slc12a1*), stroma (*Col3a1*, *Foxd1*), endothelium (*Pecam1*) and immune cells (*Tryobp*).

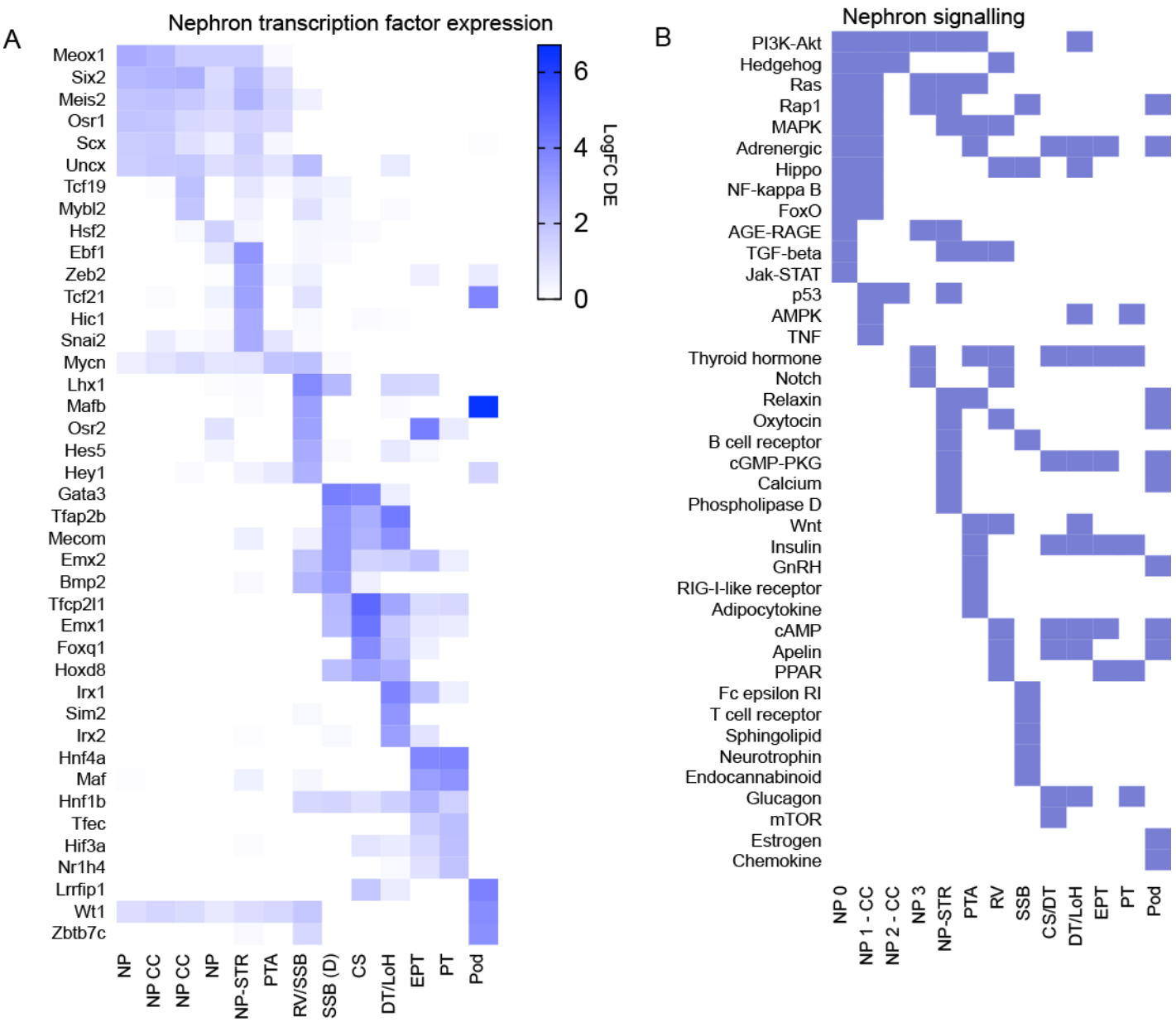
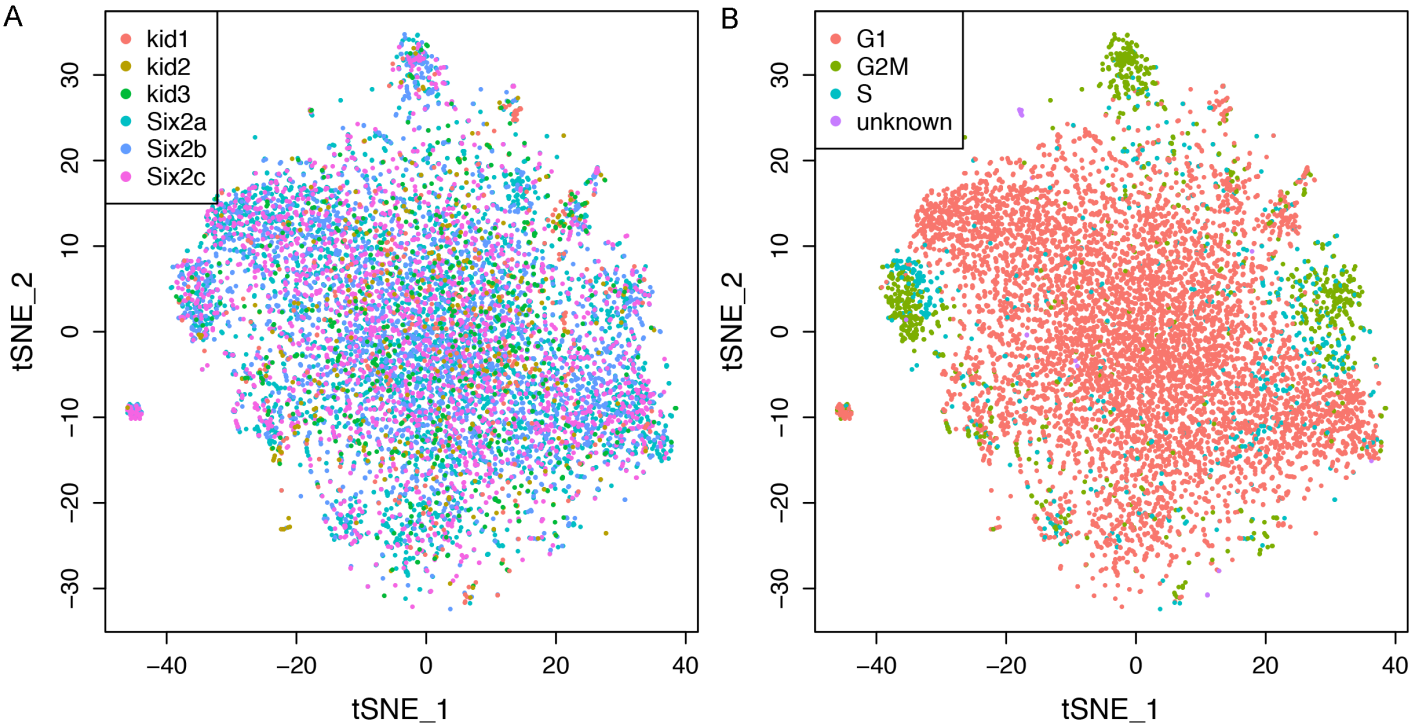
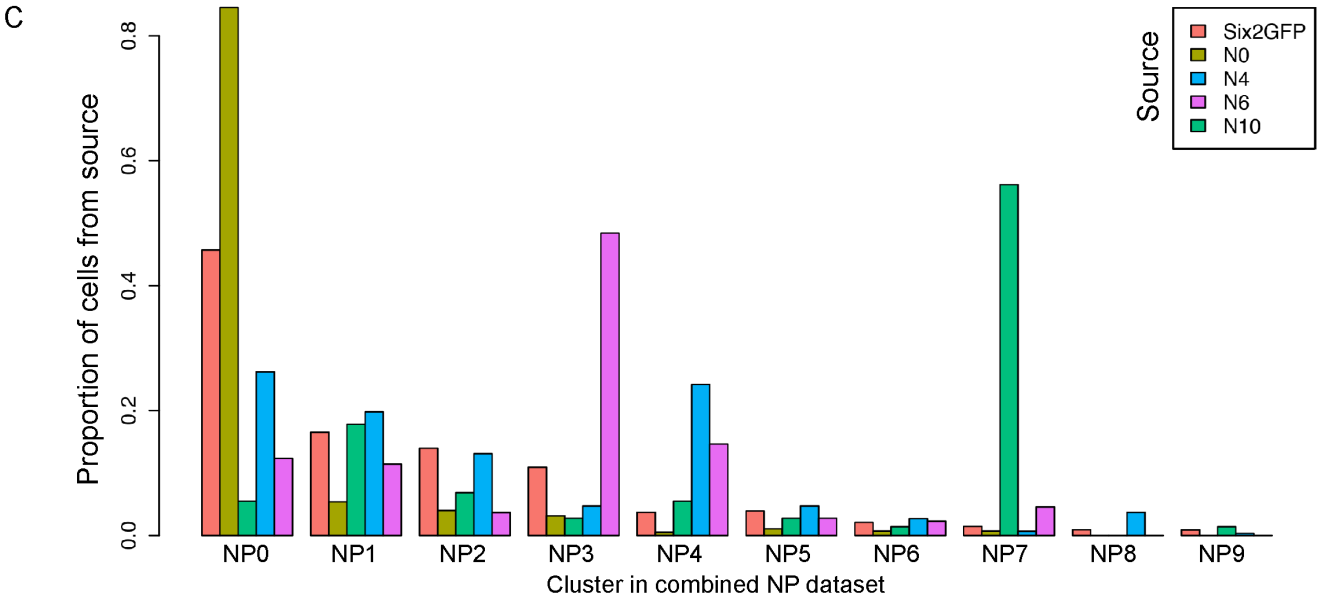


Figure S3. Transcription factors and KEGG analysis of signalling pathways within the nephron lineage. **A.** Top differentially expressed key transcription factors within nephron lineage clusters. **B.** Signalling pathways active within individual nephron lineage clusters identified by GO and KEGG analysis. Information about which ligands, receptors, and effectors are expressed in each cell type can be accessed in Supplementary file 4.





D

Whole kidney cluster ID	source1	source2	N	p.value	lib.size1	lib.size2	prop
14	12	4	11	9.42E-05	0.63807	0.38851	0.01401
7	13	12	133	2.62E-26	1.15620	1.45120	0.05249
0	15	12	179	1.10E-29	0.86902	1.43517	0.15446
1	14	4	222	6.93E-28	2.31078	0.89776	0.12041
2	14	13	289	7.97E-32	2.55341	1.29806	0.10339
4	13	1	366	3.05E-86	1.30849	1.11388	0.08654
12	14	7	375	8.50E-18	1.56723	0.68908	0.03334
5	14	8	491	1.06E-51	1.64427	0.80004	0.07537
6	15	12	506	4.58E-90	0.48300	0.79767	0.06136
13	15	4	629	2.98E-62	0.76002	0.76424	0.01614
3	9	5	681	1.68E-65	0.89895	1.02997	0.09700
9	13	3	760	9.12E-81	0.95771	1.11241	0.04770
15	13	8	852	8.25E-60	1.31576	1.25935	0.00958
11	14	10	922	3.02E-194	2.19966	0.90783	0.03635
10	15	11	1921	5.25E-143	0.93615	1.10153	0.04238
8	15	10	2263	3.67E-162	0.79406	0.84822	0.04948

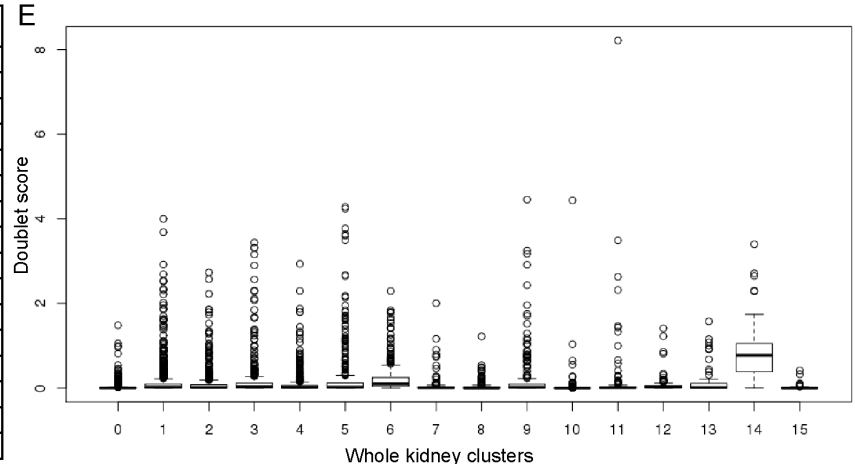


Figure S4. Integration of scRNA-Seq data from nephron progenitors in the nephron lineage clusters N0, N4, N6, N10, and >7800 sorted *Six2*GFP+ cells. **A.** tSNE plot of integrated nephron progenitor data from the e18.5 whole kidney dataset and >7800 sorted *Six2*GFP+ cells. Cells are identified by replicate kidney pairs (kid1-3) or replicated sorted *Six2*GFP populations (Six2a-c). **B.** tSNE plot showing cell cycle state within the integrated nephron progenitor data. **C.** Bar graph showing proportional contributions by source in the integrated

nephron progenitor clusters. Note this does not show the actual number of cells from each source. **D.** Output from doubletCluster algorithm on whole kidney data. Most likely parent clusters are shown in "source1" and "source2" columns. N = number of unique marker genes for each query cluster, p.value = P-value against the doublet hypothesis for query cluster, lib.size1 = ratio of library sizes of parent1 versus query cluster, lib.size2 = ratio of library sizes of parent2 versus query cluster, prop = proportion of cells making up the query cluster compared to the entire dataset. 'Suspicious' clusters have low N, lib.size1 and lib.size2 < 1 and prop < 5%. **E.** Boxplot showing the distributions of doublet scores for the cells in each cluster of the whole kidney dataset. The doubletCells algorithm outputs doublet scores based on simulating pseudo-doublets by randomly selecting two cells in the dataset and adding them together, completely independently of the cluster assignment. High scores indicate higher likelihood of the cell being a doublet. Cluster 14 has markedly higher doublet scores compared to the remaining clusters.

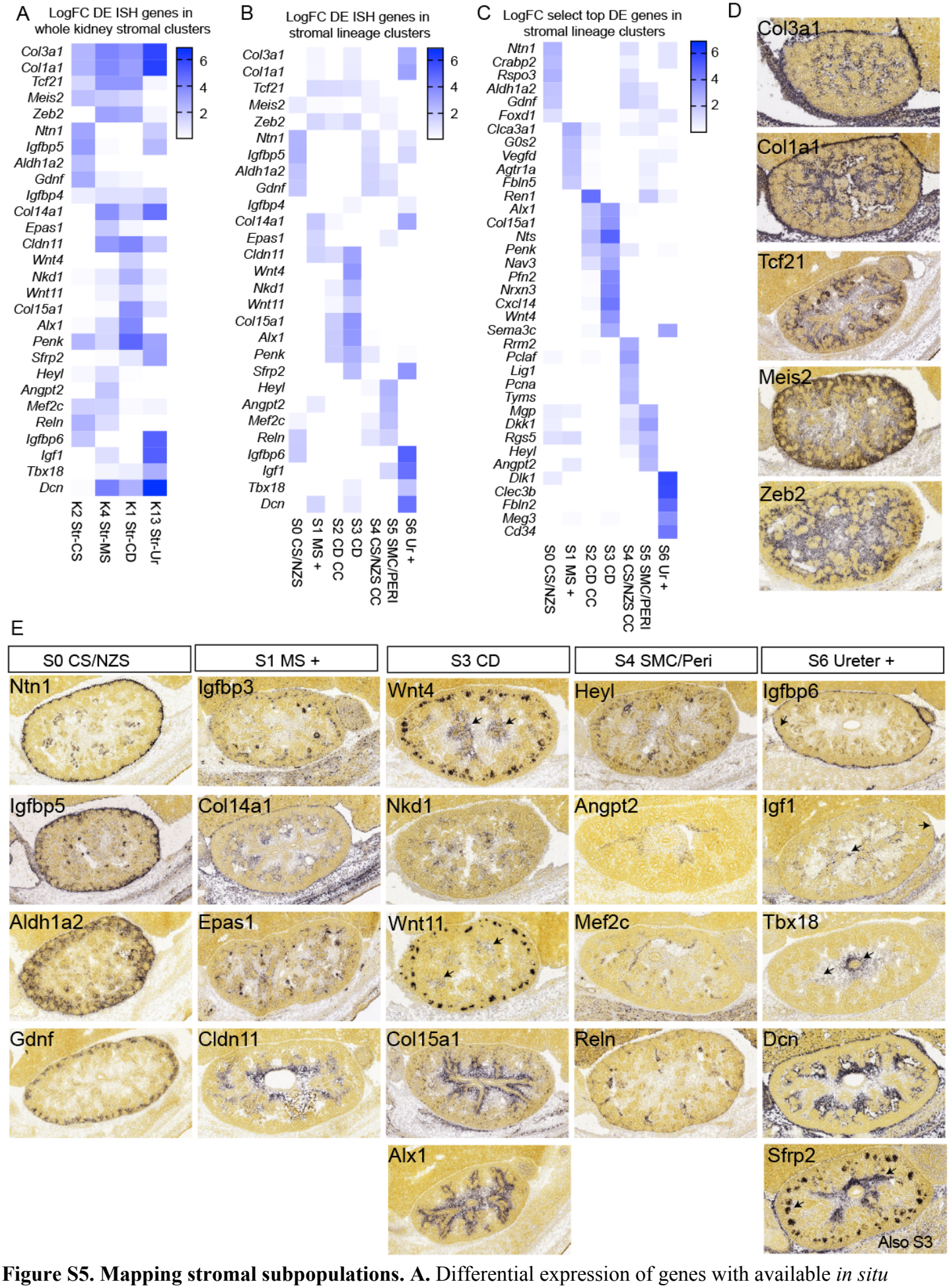


Figure S5. Mapping stromal subpopulations. A. Differential expression of genes with available *in situ* hybridisation (ISH) results* that are also enriched in all stromal clusters or stromal subpopulations in the whole

kidney data. Scale represents log fold change (LogFC) differential expression (DE) within the whole kidney stromal clusters compared to other clusters in the whole kidney. **B.** Differential expression results for the same genes in A within the stromal lineage clustering. Note the low differential expression results for *Col3a1* and *Col1a1* indicate a lack of change in expression rather than an absence of expression. **C.** Expression of select top DE genes within each stromal cluster. **D.** ISH results for markers enriched in all or several stromal populations (refer to B for enriched populations). **E.** *ISH results from the Allen Developing Mouse Brain Atlas (http://developingmouse.brain-map.org) for stromal cluster enriched genes. Several genes are expressed in other cell types within the developing kidney; their stromal expression domain has been taken into account for this analysis. Some genes pictured are expressed in more than one cluster (refer to B).

Table S1. Differentially expressed genes between 16 clusters within the E18.5 developing mouse kidney. Note: For each spreadsheet, use the "lookup" tab to access an interactive sheet. Input an official gene symbol in the left column to retrieve differential expression results for that gene across all clusters.

Click here to Download Table S1

Table S2. Putative ligand-receptor interactions within and between all clusters in the whole kidney dataset.

Click here to Download Table S2

Table S3. Differentially expressed genes between 3 ureteric epithelium subclusters within the E18.5 developing mouse kidney.

Click here to Download Table S3

Table S4. Differentially expressed genes between 8 nephron and 5 nephron progenitor subclusters within the E18.5 developing mouse kidney.

Click here to Download Table S4

Table S5. Integrated nephron progenitor cluster markers and differential expression analysis for clusters NP0, NP3, NP4, and NP7.

Click here to Download Table S5

Table S6. Differentially expressed genes between 6 stromal subclusters within the E18.5 developing mouse kidney.

Click here to Download Table S6

Table S7. Genes that distinguish distal nephron and ureteric epithelium.

Click here to Download Table S7

Table S8: Comparison of this scRNA-seq dataset to other developing mouse kidney scRNA-seq data.

Manuscript	Tissue age / stage	Profiling approach	Cell number	Analysis approach	Conclusion
Brunskill et al, 2014, Development	11.5, 12.5, P4 renal vesicle	Fluidigm C1	33	Genespring 12.6.1	Read through of Hox genes, inappropriate expression of presumed lineage markers within CM, partially degraded non-coding RNAs.
Adam et al, 2017, Development	P1	DropSeq	20,000 (in batches of about 4000 cells for each condition of isolation	Seurat Find All Markers; DEGseq	Use psychrophilic enzymes to avoid c-fos signature, single cell expression profile of the new born mouse kidney.
Magella et al, 2018, Dev Biol	14.5	Drop-Seq, Chromium 10x Genomics and Fluidigm C1	>8000	AltAnalyze	Nephrogenic stroma makes GDNF; stochastic multilineage priming, single cell expression profile of E14.5 kidney.
This study	18.5 whole kidney, 14.5 sorted *Six2*GFP	Chromium 10x	6732 18.5 kidney, 7853 14.5 sorted *Six2*GFP.	Seurat, EdgeR, Monocle 2	Expression profile of E18.5 kidney, improved resolution of known cell type markers and signalling pathway component expression. Identification of congruence and new distinct markers for connecting segment and ureteric epithelium. New insight into mouse nephron progenitor heterogeneity.