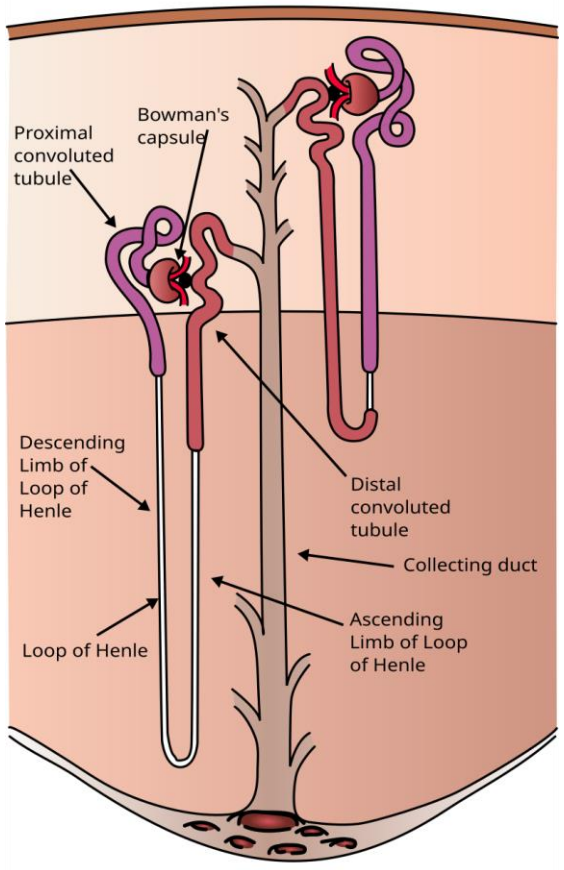
About RCC (renal cell carcinoma, kidney cancer)



Type of Renal cell carcinoma	Gene involved	Cell of origin	Microscopic features
Clear cell RCC	VHL gene (3p25.3)	Proximal tubular epithelial cells	Tumor cells are polygonal or rounded having abundant clear or granular cytoplasm containing glycogen and lipids
Papillary RCC	MET a proto-oncogene on chromosome 7.	Distal convoluted tubular epithelial cells	Tumor cells are cuboidal to low columnar cells lining the papillae having fibro vascular are with foamy macrophages
Chromophobe RCC	Multiple chromosome losses and hypoploidy	Intercalated cells of collecting ducts	Tumor cells have pale eosinophilic cytoplasm with perinuclear halo
Xp11translocation carcinoma	Translocations of TFE3 gene located at Xp11.2	Tubular epithelial cells	
Collecting duct carcinoma	Chromosomal losses & deletions without distinct pattern	Collecting duct cells in medulla	Tumor are arranged as irregular channels lined by highly atypical epithelium with hobnail pattern

Glomerulus → **Proximal** tubule → Loop of Henle → **Distal** tubule → Collecting duct

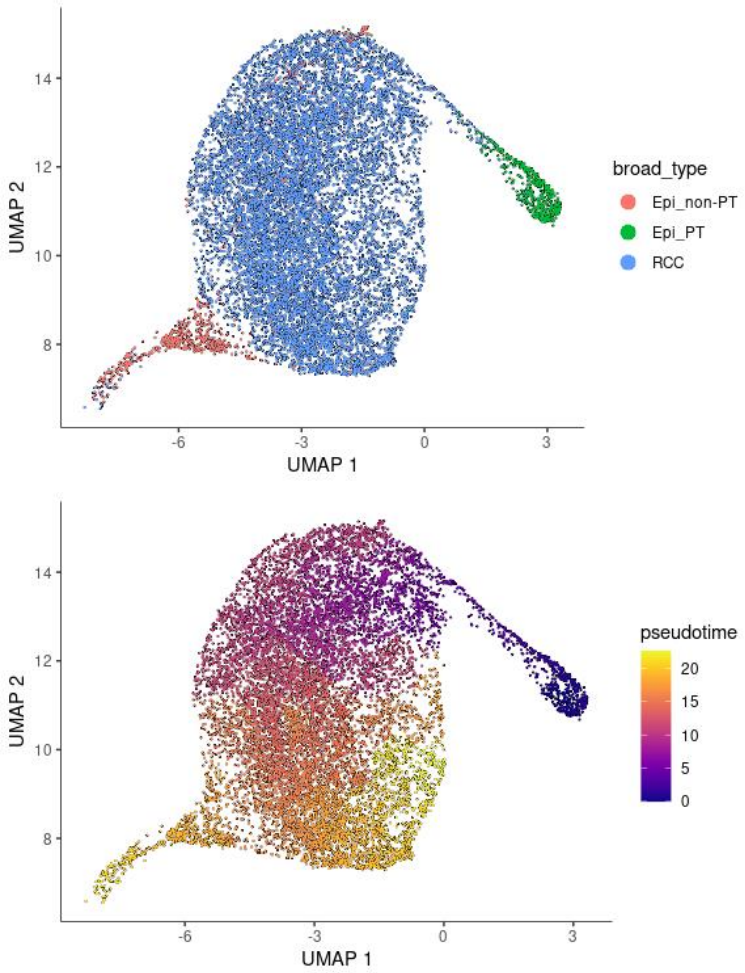
Proximal tubule: PT

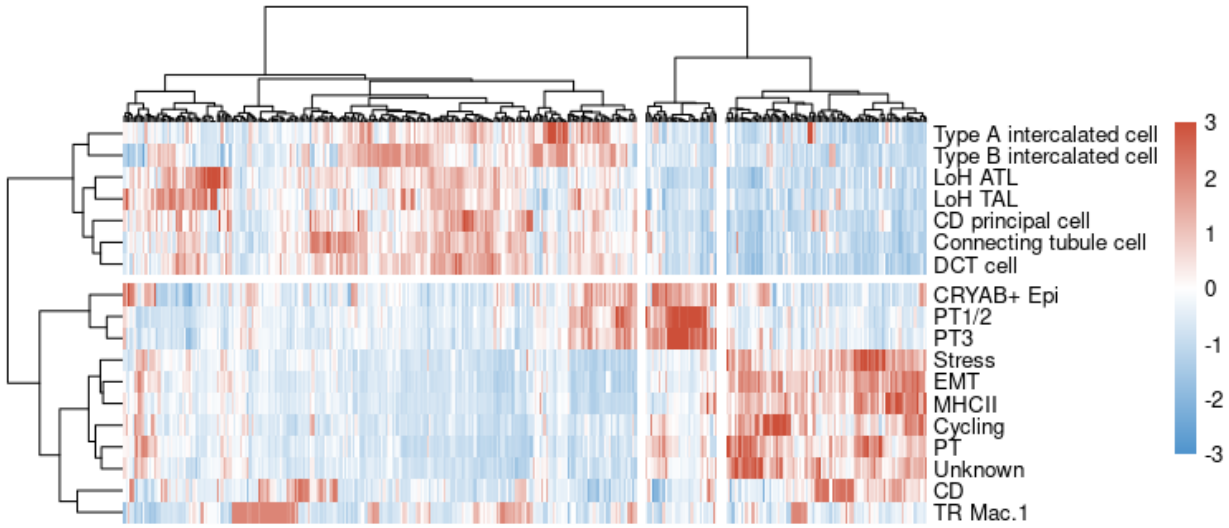
Distal tubule: DCT cell (distal convoluted tubule)
Collecting duct: IC cell (intercalated cell)

RCC is common type (about 70 % of RCC) and
originate from proximal tubular epithelial cells

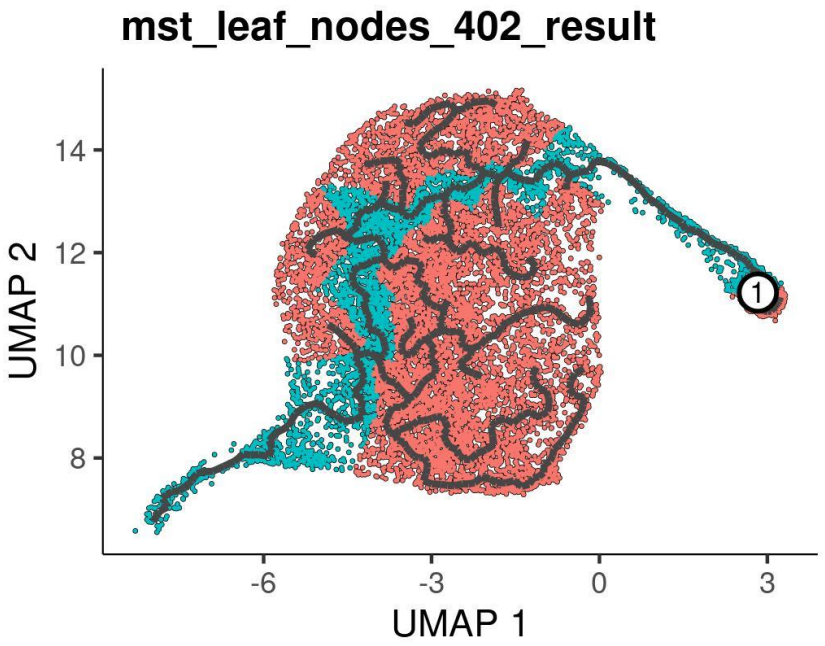
Tan S.K. et al., 2016, Nature review urology
Data retrieved from https://www.histopathology.guru/academics/undergraduate-academics/the-kidney/renal-cell-carcinoma/

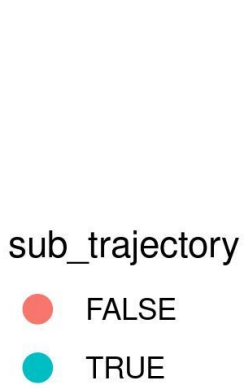
Preprocessing

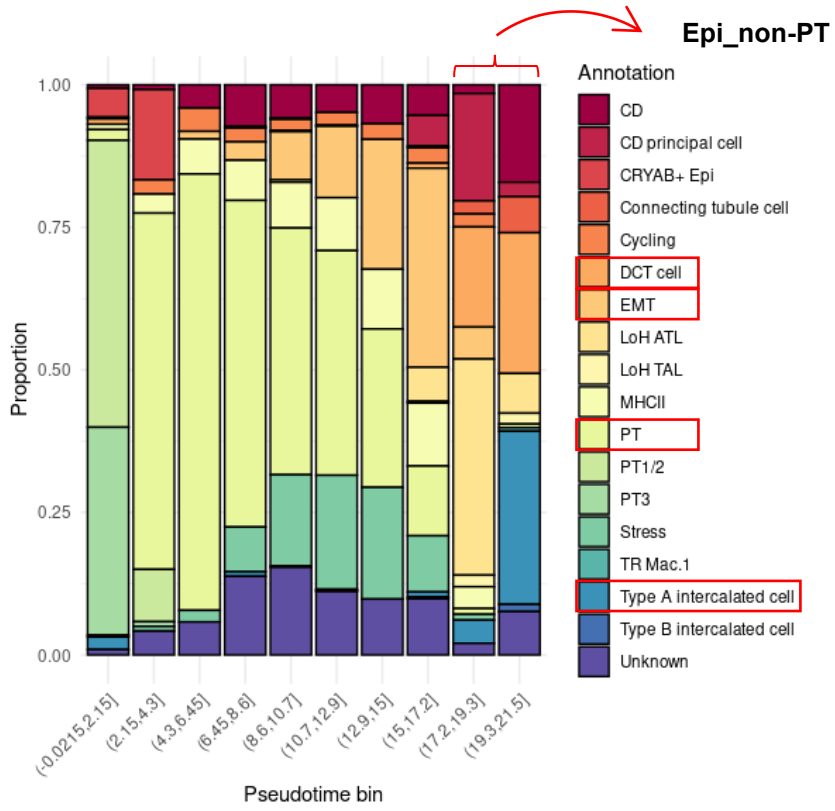




Preprocessing

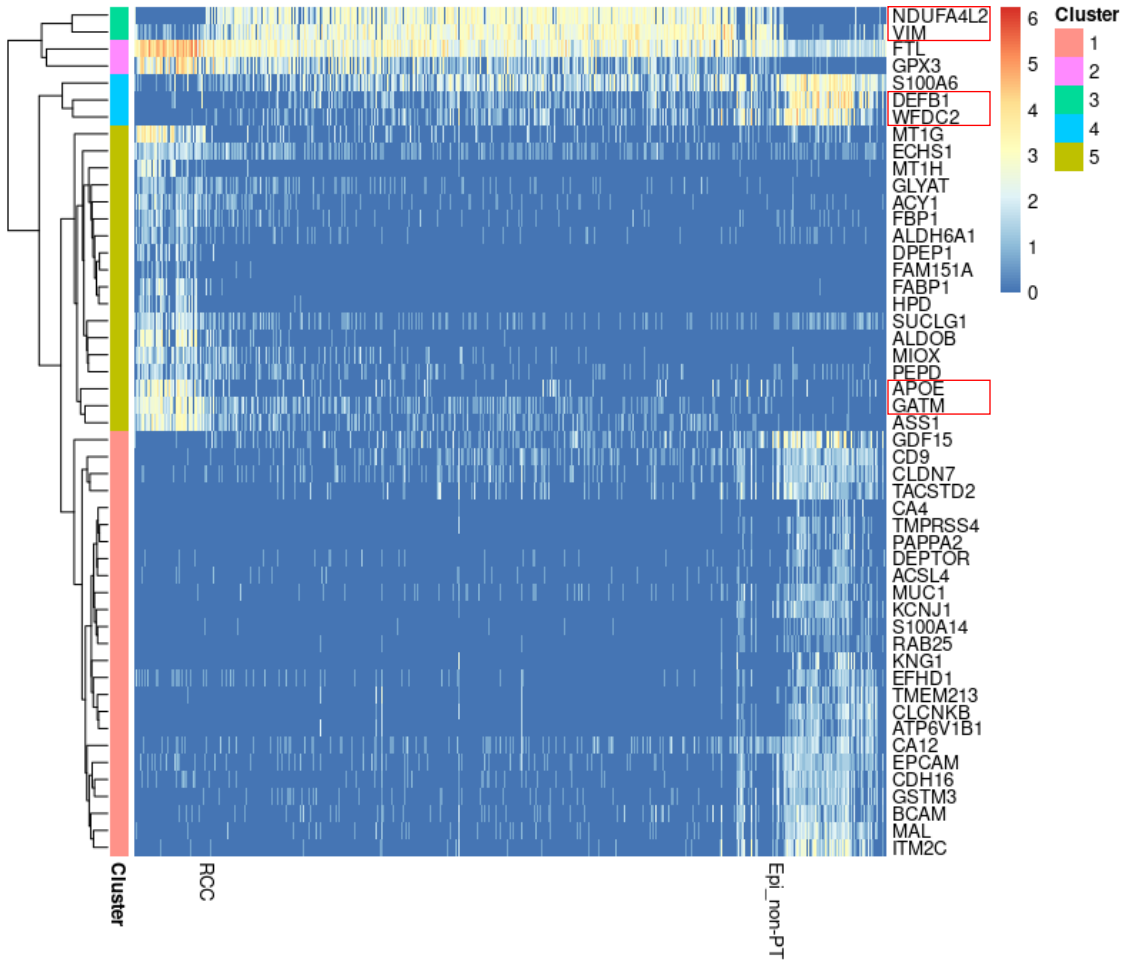


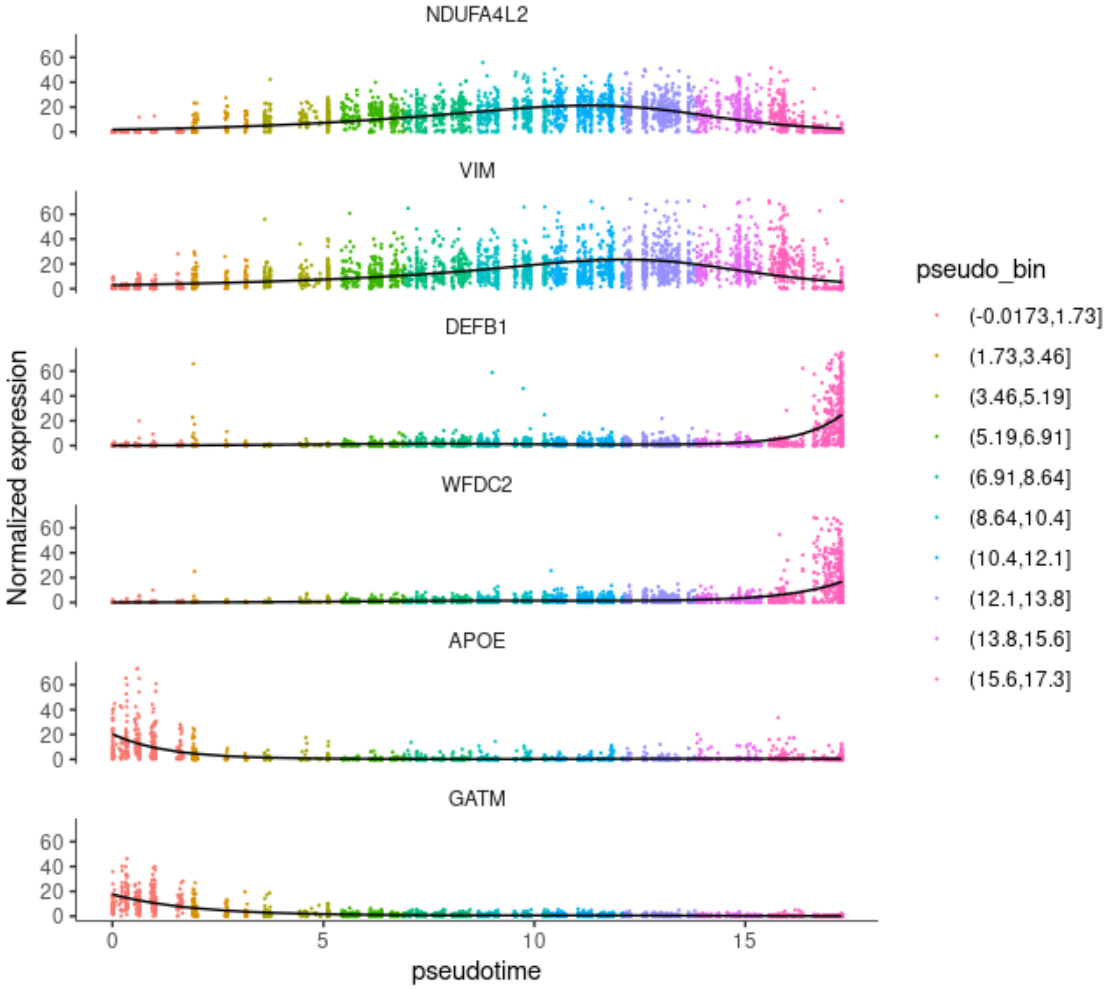




- ✓ During pseudo-time progression, a shift from PT identity to EMT-like cell states was evident
- ✓ Proportion of DCT cell and Type A intercalated cell increased at Epi_non-PT cells dramatically

Preprocessing





NDUFA4L2 & VIM → RCC marker
DEFB1 & WFDC2 → Epi_non-PT marker
APOE & GATM → Epi_PT marker

Preprocessing

A minimal amount of filtering is needed before running a SCENIC analysis. On a cell level, we examine the number of expressed genes and remove cells that fell into the distribution extremes. In the peripheral blood mononuclear cell (PBMC) study case (Table 1), we discard cells with <200, and more than ~5,000, expressed genes; however, these thresholds must be determined empirically. We further filter out cells that have a large fraction of mitochondrial gene transcripts; these cells are thought to be of lower quality as this is indicative of cell membrane breach[9]. We again use the empirical distribution to select an upper threshold on mitochondrial genes expressed, and this is highly dependent on cell type but typically 5–15%. Finally, on the gene level, genes with low overall expression are removed; with our default settings, we remove genes expressed in fewer than three cells in the data set.

Raw UMI count data **without** ln(x+1) transformation

↓

Delete ribosomal gene and exclude cell with percent.mt over 10

↓

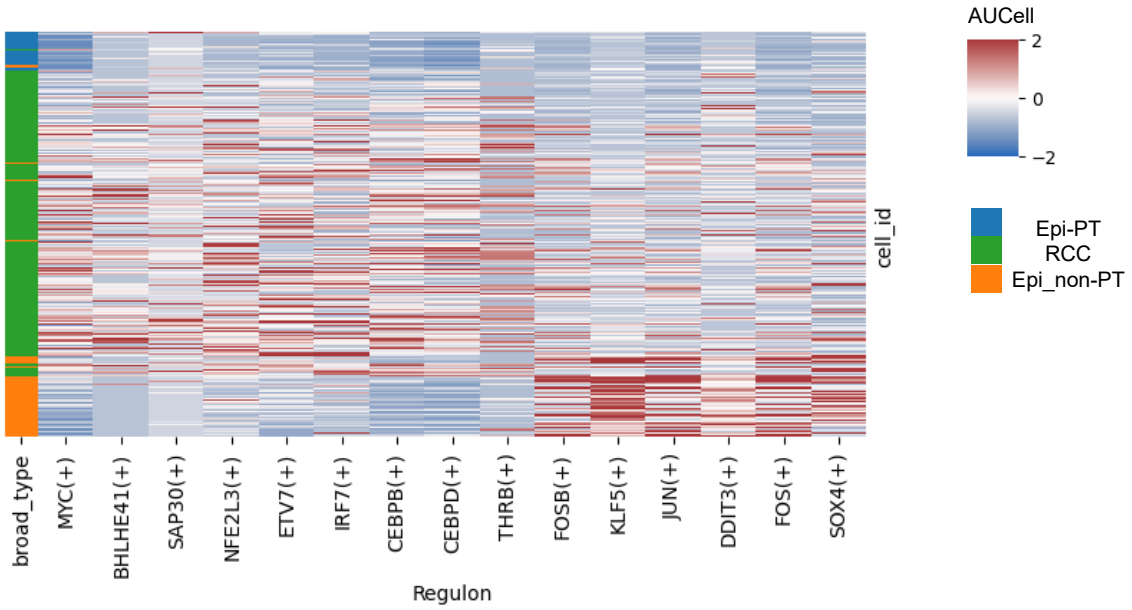
Auxiliary data info :

[FeatherRankingDatabase(name="hg38__refseq-r80__10kb_up_and_down_tss.mc9nr.genes_vs_motifs.rankings")]

Bram Van de Sande, *et al*., 2020, Nature Protocols

Epi_PT = 330 cells, Epi_non-PT = 530 cells, and RCC = 2524 cells after monocle3 (Y_402)
5460 genes (Moran's I index > 0 and q-value <5e-2)

GRN inference



	TF	target	importance
317	SMARCA4	TUBA4B	15.1029649
16	NFIA	RCBTB1	13.2023111
164	CEBPD	AC090498.1	12.97238503
222	TPI1	LDHA	11.29414422
271	ZNF276	NRP1	10.94957279
101	BDP1	ZNF430	10.81409148
222	TPI1	NDUFA4L2	10.76638321
406	FOXI1	SLC26A7	10.35831408
406	FOXI1	ATP6AP2	10.33240095
92	H2AFZ	STMN1	10.28568309
44	SUCLG1	PBLD	10.0510242
345	ATF5	NUPR1	10.04877133
378	U2AF1	KCNMA1	9.89096574
406	FOXI1	KIF21A	9.786802135
92	H2AFZ	PTTG1	9.710261459

Tree-based approaches are able to capture combinatorial transcriptional control and complex interaction patterns