Deconstructing the clear cell renal cell carcinoma using single nuclei multiomics

Michael P Mariani[1,2*], Ze Zhang[1], Brock C Christensen[3,4], Lucas A Salas[1,4,5]

[1]Department of Epidemiology, Geisel School of Medicine at Dartmouth, Hanover, NH [2]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Hanover, NH [3]Departments of Molecular and Systems Biology and Community and Family Medicine, Geisel School of Medicine at Dartmouth, Hanover. NH [4]Norris Cotton Cancer Center, Dartmouth Hitchcock Medical Center, Lebanon, NH [5]Center for Molecular Epidemiology at Dartmouth, Lebanon, NH

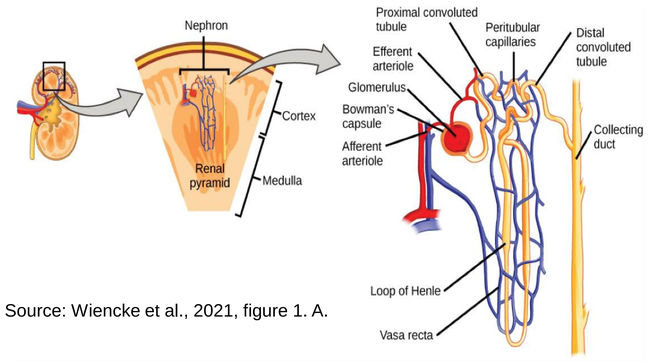
*E-mail: Michael.P.Mariani@dartmouth.edu

Dartmouth GEISEL SCHOOL OF MEDICINE

Background

- Clear cell renal cell carcinoma, ccRCC, microenvironments are highly heterogeneous, containing epithelial cells, vascular endothelial cells, fibroblasts, myofibroblasts and leukocytes[1].
- Single nuclei multiomics sequencing was used to investigate cellular composition of tumors and the relationship of specific cell populations to patient survival groups.
- Samples were from among the first immunotherapy samples ever used in cancer research: from the Renal Tumor Biobank at Dartmouth College.
- Molecular markers have been used to assess the heterogeneity problem using 10X Genomics libraries and Illumina sequencing following by bioinformatics analyses.

Anatomy of the kidney



Source: Wiencke et al., 2021, figure 1. A.

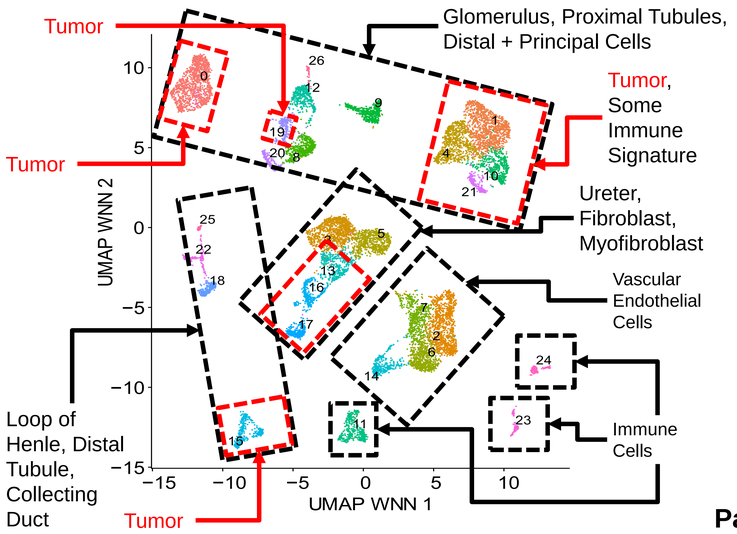
Objectives

1) Investigate the relationship of specific cell populations (clusters) to patient survival.

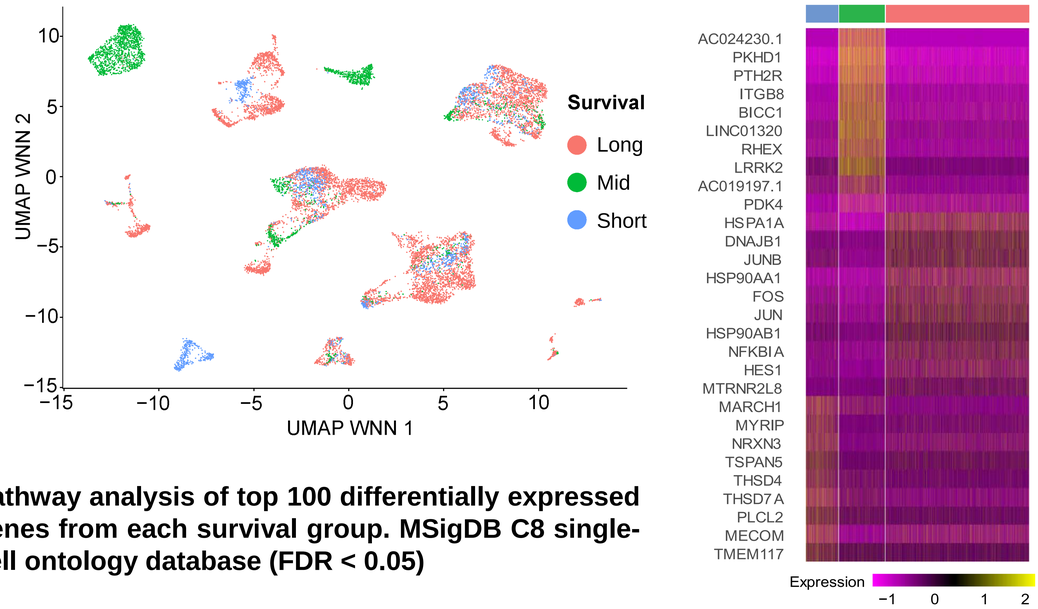
2) Identify key expression (gene) and chromatin (ATAC peaks) markers across the heterogenous tumor microenvironment.

Findings

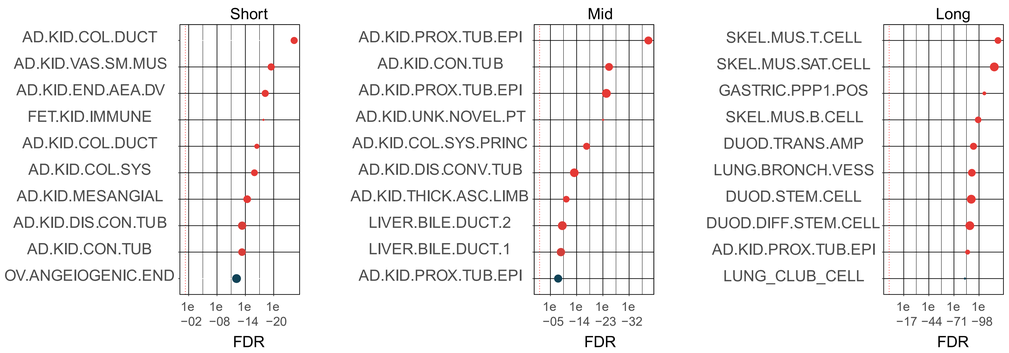
Weighted nearest neighbors (WNN) UMAP dimensionality reduction of single nuclei into distinct clusters based on variation in gene expression and chromatin occupancy



UMAP colored by survival status (bottom) and heatmap of top 10 differentially expressed genes by survival status (right)



Pathway analysis of top 100 differentially expressed genes from each survival group. MSigDB C8 single-cell ontology database (FDR < 0.05)



Data and methods

- 10x Genomic multiomics libraries were prepared from eleven ccRCC patient samples from the Dartmouth College Biobank at NCCC. Among the patient samples, seven were male and four were female with a median age of death of 63.35 years, across all stages and grades.
- A total of 42,794 cells were generated as estimated by 10x Genomics Cell Ranger 3.1.0. Following filtering in Seurat v3.0 (Stuart et al., 2019), 12,571 single nuclei remained for downstream analysis. Marker genes for kidney cells[1,2] and ccRCC[3] already established in literature were used to identify cell types. confirmed cell types approximately grouped into five categories: epithelial, endothelial (vascular), fibroblasts/myofibroblasts, immune cells, and tumor cells.
- Top highly expressed markers across survival groups were identified using Seurat and pathway analysis from top markers was performed using the hypeR R package (Federico and Monti, 2020) and the MSigDB C8 curated single-cell gene ontology set (Subramanian, Tamayo; 2005).

Summary of findings

- **Currently we have 8,121 long survival (>5 yr) nuclei, 2,607 mid survival (1-5 yrs) nuclei, and 1,843 short survival (<1 yr) nuclei. Together they comprise 27 distinct UMAP clusters.**
- **Single-nuclei multiomics sequencing leads to thousands of differentially expressed genes and differential chromatin accessibility regions (ATAC peaks) that are significantly different between survival groups.**
- **Top differentially expressed genes by survival group can be seen to form distinct clusters. These significant genes are part of ontology pathways that are unique to each survival group.**
- **The shorter survival group was enriched with angiogenic pathways while the longer survival group demonstrated more immune cell pathways.**

Survival	# Samples	# Cells pass filter	# Diff. genes (Padj <= 0.05)	# Diff. ATAC peaks (Padj <= 0.05)
Short (<1 yr)	2	1843	864	2080
Mid (1-5 yrs)	2	2607	1893	4761
Long (>5 yrs)	7	8121	1711	4305

Discussion and future directions

- Multiomics single-nuclei sequencings is a new and highly promising tool for investigating the heterogenous tumor microenvironment in ccRCC samples, as well as solid tumors in general.
- We are currently sequencing additional samples to improve confidence of identified markers and will also identify differentially occupied chromatin regions from ATAC data as well. This data along with the differential scRNA-seq data will then be combined with CNV analyses and survival data, comprising a comprehensive systems biology approach to identify ccRCC markers and chromatin states associated with differing ccRCC survival outcomes.

FUNDING: This work was supported by the United States National Institute of General Medical Sciences (Number: P20GM104416 /8299) and the United States Congressionally Directed Medical Research Programs Funding (Number: W81XWH-20-1-0778) to LAS.

References

[1] Young MD, Mitchell TJ, Vieira Braga FA, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science*. 2018;361(6402):594-599. doi:10.1126/science.aat169

[2] Bohnenpoll T, Feraric S, Nattkemper M, et al. Diversification of Cell Lineages in Ureter Development. *J Am Soc Nephrol*. 2017;28(6):1792-1801. doi:10.1681/ASN.2016080849

[3] Yan F, Wang Y, Liu C, et al. Identify clear cell renal cell carcinoma related genes by gene network. *Oncotarget*. 2017;8(66):110358-110366. Published 2017 Nov 30. doi:10.18632/oncotarget.22769