Title: Mapping Immune Landscape in Clear Cell Renal Carcinoma by Single-Cell Genomics

**Authors:** Ajaykumar Vishwakarma?, Nick Bocherding7, Kenneth Nepple10, Aliasger Salem5, 6, Russell W. Jenkins1, 2, 3 \*, Weizhou Zhang 5, 11 \*, Yousef Zakharia5, 10, 12 \*

Also need to add Andrew Belizzi

**Affiliations:**

1 Department of Medicine, Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, MA

2 Laboratory of Systems Pharmacology, Harvard Program in Therapeutic Science, Harvard Medical School, Boston, MA

3 Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA

4 Cancer Biology Graduate Program, Carver College of Medicine, University of Iowa, Iowa City IA

5 Holden Comprehensive Cancer Center, University of Iowa, Iowa City, IA

6 Department of Pharmaceutics and Translational Therapeutics, College of Pharmacy, University of Iowa, Iowa City, IA

7 Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO

10 Department of Urology, University of Iowa Hospitals and Clinics, Iowa City, IA

11 Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, FL

12 Department of Internal Medicine, University of Iowa Hospitals and Clinics, Iowa City, IA

\* Correspondence: [rwjenkins@partners.org](mailto:rwjenkins@partners.org) (R.W.J), [zhangw@ufl.edu](mailto:zhangw@ufl.edu) (W.Z) and [yousef-zakharia@uiowa.edu](mailto:yousef-zakharia@uiowa.edu) (Y.Z)

**Abstract**:

Human clear cell renal cell carcinoma (ccRCC) is one of the most immunologically distinct tumor types due to high levels of tumor-infiltrating immune cells including T cells, yet not every patient responds to immunotherapy. Interestingly, in contrast to other cancers, infiltration with cytotoxic CD8+ T cells is associated with poorer overall survival in ccRCC, suggesting that sub-populations of CD8+ and other immune cells may underlie this observation. To characterize the tumor immune microenvironment of ccRCC, we applied single-cell-RNA sequencing (SCRS) along with T-cell-receptor (TCR) sequencing to map the transcriptomic heterogeneity of 25,688 individual CD45+ lymphoid and myeloid cells in matched tumor and blood from patients with ccRCC. Will need to update based on new findings. This report represents the first such characterization of the ccRCC immune landscape using scRNA-seq. With further characterization and functional validation, these findings may identify novel sub-populations of immune cells amenable to therapeutic intervention.

**Introduction**

Clear cell renal cell carcinoma (ccRCC) is the most common type of renal cell carcinoma arising from epithelial cells of the proximal tubule of the kidney, comprising more than 70% of all renal cancers (1). ccRCC represents an immune sensitive tumor type well known for early advances in systemic immunotherapy using T cell proliferation cytokine IL-2 and interferon (IFN) -ɑ2b therapy (2). Recent novel immunotherapies targeting T cell checkpoints as standard of care has transformed the treatment paradigm of ccRCC (3,4). However, a substantial subset of renal cancer patients still do not respond to these therapies and patients who initially do can eventually progress (5,6). Cytotoxic tumor-infiltrating lymphocytes (TILs), in particular CD8+ T cells are key effectors of the adaptive anti-tumor immune response (7) and abundance of CD8+ T cells in solid cancers is generally associated with better survival in cancer patients (8–10). However, in RCC, immune cell abundance is inversely correlated with survival, specifically TILs (11–13). Biomarker analysis results from recent clinical trials comparing PD-1 blockade versus anti-angiogenic inhibitors and combination therapies in treatment-naïve ccRCC patients also supported the negative prognostic significance of T cell infiltrate in the absence of immunotherapy (14,15). Other abundant immune players in the ccRCC tumor microenvironment include monocytes, dendritic cells and tumor-associated macrophages (TAMs) (16) which are now being harnessed for discovery of novel gene programs but remain far less studied than T cells.

Quantifying and inferring immune cell abundance from transcriptional analysis of primary or metastasized bulk tumor samples is inadequate to provide a clear picture of the immune cell types (17,18). While these studies are suggestive, they lack single cell resolution for characterizing heterogeneous cell subpopulations that ultimately shape anti-tumor response, as has been demonstrated in breast cancer and melanoma (19,20). Single-cell methodologies including flow cytometry, immunohistochemistry, and mass cytometry (13,16,21) have revealed immune cell states in ccRCC but only as discrete phenotypes when in vivo they typically display diverse spectrum of differentiation or activation states. Also, these methods require use of antibody panels targeting known immune cell components, and by design are not capable of identifying novel sub-populations of cells. SCRS has enabled comprehensive characterization of heterogeneous lymphoid and myeloid immune cells in several cancers (22–25), providing an unbiased approach to profiling cells and enabling molecular classification of different subpopulations and identification of novel gene programs. Transcriptome mapping of T lymphocytes coupled with TCR sequencing allows additional measurement of clonal T cell response to cancer at an unprecedented depth (26,27).

Here, we report the single cell RNA profiling of the immune landscape in ccRCC mapping a total 25,688 of immune single cells (5’ gene expression and recombined V(D)J region of the T cell receptor) in matched samples of tumor and peripheral blood isolated from three treatment-naïve ccRCC patients. Will need to update with results. Analysis of myeloid cells revealed a complex mixture of pro- and anti-inflammatory polarized phenotypes across patients. This represents the first such report of the immune landscape of ccRCC using scRNA-seq.

**Methods**

*Subject Details and Tissue Collection*

Fresh blood and primary clear cell renal cell carcinoma (ccRCC) samples were obtained from the University of Iowa Tissue Procurement Core and GUMER repository through the Holden Comprehensive Cancer Center from subjects providing written consent approved by the University of Iowa ethics board committee. The patients ranged from 67 to 74 years old; the tumor samples were of diverse tumor stages and sourced from male subjects. Tumor grades were histologically determined by a pathologist. Three ccRCC tumor specimens paired with individual blood samples were used in the study. Will need IRB FOR publication

*Tumor Dissociation and Isolation of Mononuclear Cells*

Renal tumor samples were dissociated into single cells by a semi-automated combined mechanical/enzymatic process. The tumor tissue was cut into pieces of (2-3mm) in size and transferred to C Tubes (Miltenyi Biotech, Bergisch Gladbach, Germany) containing a mix of Enzymes H, R and A (Tumor Dissociation Kit, human; Miltenyi Biotech). Mechanical dissociation was accomplished by performing three consecutive automated steps on the gentleMACS Dissociator (h\_tumor\_01, h\_tumor\_02 and h\_tumor\_03). To allow for enzymatic digestion, the C tube was rotated continuously for 30 min at 37°C, after the first and second mechanical dissociation step (28). Cells from fresh tumor specimens were incubated with FcR blocking reagent (StemCell Technologies, Vancouver, Canada) for 10 min at 40C and labelled with 1ug/ml of the FITC anti-human CD45 antibody (BioLegend, San Diego, CA) per 107 cells for 20 min at 40C. CD45+ cells were isolated using the EasySepTM FITC Positive Selection Kit (StemCell Technologies). Alternatively, mononuclear cells (MNCs) from whole peripheral blood of paired subjects were isolated using SepMate Tubes (StemCell Technologies) by density gradient centrifugation. Cells were then viably frozen in 5% DMSO in RPMI complemented with 95% FBS. Cryopreserved cells were resuscitated for flow cytometry analyses by rapid thawing and slow dilution.

*Cell Sorting for Single-Cell RNA sequencing*

Viable immune (CD45+ Hoechst-) single cell suspensions generated from three ccRCC tumor samples and blood were FACS sorted on a FACS ARIA sorter (BD Biosciences) for lymphoid and myeloid cells (Ratio 3:1). The cells were sorted into ice cold Dulbecco’s PBS + 0.04% non-acetylated BSA (New England BioLabs, Ipswitch, MA). Sorted cells were then counted and assessed viability MoxiGoII counter (Orflo Technologies, Ketchum, ID) ensuring that cells were re-suspended at 1000 cells/ul with a viability >90%.

Library Preparation, Single-Cell 5’ and TCR Sequencing

Single-cell library preparation was carried out as per the 10X Genomics Chromium Single Cell 5' Library and Gel Bead Kit v2 #1000014 (10x Genomics, Pleasanton, CA). Cell suspensions were loaded onto a Chromium Single-Cell Chip along with the reverse transcription (RT) master mix and single cell 5′ gel beads, aiming for 7,500 cells per channel. Following generation of single-cell gel bead-in-emulsions (GEMs), reverse transcription was performed using a C1000 Touch Thermal Cycler (Bio-Rad Laboratories, Hercules, CA); 13 cycles were used for cDNA amplification. Amplified cDNA was purified using SPRIselect beads (Beckman Coulter, Lane Cove, NSW, Australia) as per the manufacturer’s recommended parameters. Post-cDNA amplification reaction QC and quantification was performed on the Agilent 2100 Bioanalyzer using the DNA High Sensitivity chip. For input into the gene expression library construction, 50ng cDNA and 14 cycles was used. To obtain TCR repertoire profile, VDJ enrichment was carried out as per the Chromium Single Cell V(D)J Enrichment Kit, Human T Cell #1000005 (10x Genomics) using the same input sample. Sequencing libraries were generated with unique sample indices (SI) for each sample and quantified. Libraries were sequenced on an Illumina HiSeq 4000 using a 150-pair-end sequencing kit. Gene expression FASTQ files were aligned to the human genome (GRCh38) using the CellRanger v2.2 pipeline, while clonotype sequencing was aligned to the vdj\_GRCh38\_alts\_ensembl genome build provided by the manufacturer.

*Single-cell Data Analysis*

Initial processing of cells isolated form the heart in myocarditis run 1 (n=2,617), myocarditis run 2 (n=10,618), control run1 (n=1,528), and control run 2 (n=8,201) was performed using the Seurat R package (v3.0.2) (29,30). Samples were normalized using the *SCTtransform* approach (31) with default settings. Preparation for integration used 3,000 anchor features and *PrepSCTIntegration*. The integration of sequencing runs occurred with the SCT-transformed data. The dimensional reduction to form the uniform manifold approximation and project (UMAP) utilized the top 30 calculated dimensions and a resolution of 0.7. The schex R package (v1.1.5) was used to visualize mRNA expression of lineage-specific or highly differential markers by converting the UMAP manifold into hexbin quantifications of the proportion of single-cells with the indicated gene expressed. Default binning was set at 80, unless otherwise indicated in the figure legend.

Cell type identification utilized the SingleR (v1.0.1) R package (32)with correlations of the single-cell expression values with transcriptional profiles from pure cell populations in the ENCODE (33). In addition to correlations, canonical markers for cell lineages (Supplemental Table X) and corresponding TCR sequences were used. Differential gene expression utilized the Wilcoxon rank sum test on count-level mRNA data. For differential gene expression across clusters or subclusters, *FindAllMarkers* function in the Seurat package using the log-fold change threshold > 0.25, minimum group percentage = 10%, and the pseudocount = 0.1. Differential comparisons between condition utilized the *FindMarkers* function in Seurat, without filtering and a pseudocount = 0.1. Multiple hypothesis correction was reported using the Bonferroni method. Cell cycle regression was performed in Seurat using the *CellCycleScoring* function and genes derived from Nestorowa et alia (34). Genes were isolated by calling *cc.genes.updated.2019* in R.Gene set enrichment analysis was performed using the escape R package (v0.99.0). Differential enrichment analysis was performed using the *getSignificance* function in escape that is based on the limma R package linear fit model. TCR analysis utilized our previously described scRepertoire R package (v0.99.3) (35) with clonotype being defined as the combination of the gene components of the VDJ and the nucleotide sequence for both chains. Cell trajectory The code for all analysis is available at <https://github.com/ncborcherding/ccRCC> (this link is private until publication).

*Incorporation of other SCRS data sets*

SCRS and TCR sequencing data processed using Cell Ranger v2.2 for healthy donor peripheral blood immune cells were acquired from the 10x Genomics website on 6/20/2020. Filtered gene matrix and contig annotations were used in the incorporation of the UMAP. SCRS of normal immune populations in the kidney were derived from whole kidney homogenates from 3 patients (36). Gene expression matrices were downloaded from the Gene Expression Omnibus (GSE131685) and processed using the procedure as described above to form the UMAP. Cell types were identified by canonical markers for lineages and immune cells were isolated.

*Statistical Analysis*

Statistical Analyses were performed in R (v3.6.3). Two-sample significance testing utilized Welch’s T test, with significance testing for more than three samples utilizing one-way analysis of variance (ANOVA) with Tukey honest significance determination for correcting multiple comparisons. Two-proportion Z-tests was performed using the total number of cells in each condition as the number of trials and without a prior for proportion.