

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

Authors: Nick Bocherding^{1,2}, Ajaykumar Vishwakarma^{3,4}, Andrew P. Voigt², Andrew Bellizzi⁵, Jacob Kaplan⁵, Kenneth Nepple⁶, Aliasger Salem³, Russell W. Jenkins^{4, 7, 8*}, Yousef Zakharia^{6, 9*}, Weizhou Zhang^{10*},

Affiliations:

¹ Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO

² Medical Science Training Program, University of Iowa, Iowa City, IA

³ Department of Pharmaceutics and Translational Therapeutics, College of Pharmacy, University of Iowa, Iowa City, IA

⁴ Department of Medicine, Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, MA

⁵ Department of Pathology, University of Iowa Hospitals and Clinics, Iowa City, IA

⁶ Department of Urology, University of Iowa Hospitals and Clinics, Iowa City, IA

⁷ Laboratory of Systems Pharmacology, Harvard Program in Therapeutic Science, Harvard Medical School, Boston, MA

⁸ Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA

⁹ Department of Internal Medicine, University of Iowa Hospitals and Clinics, Iowa City, IA

¹⁰ Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, FL

* Correspondence:

rwjenkins@partners.org (R.W.J)

yousef-zakharia@uiowa.edu (Y.Z)

zhangw@ufl.edu (W.Z)

Abstract:

Purpose: Human clear cell renal cell carcinoma (ccRCC) is one of the most immunologically distinct tumor types due to low levels of mutational load, high levels of tumor-infiltrating immune cells and response to immunotherapies. In contrast to other cancers where higher CD8⁺ T cells are generally associated with better prognosis, increased CD8⁺ T cell infiltration is associated with poorer overall survival in ccRCC, indicating that sub-populations of CD8⁺ T cells and/or the interaction with other cell types may underlie this observation. Here we will use single cell genomics to explore the immune cell composition and explain the contradictory connection between CD8⁺ T cell infiltration and ccRCC prognosis.

Experimental Design: 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. We also included 11,367 additional immune cells from normal kidney and peripheral blood to facilitate the identification and assessment of ccRCC-specific differences.

Results: There is an overall increase in the CD8⁺ T cell and macrophage populations in tumor-infiltrated immune cells. We further demonstrate the divergent cell transcriptional states for tumor-infiltrating CD8⁺ T cells across ccRCC patients that correspond to clonal expansion. Tumor-associated macrophages (TAMs) are separated into three distinct clusters, with distinct difference in the expression of angiogenic and secretory genes.

Conclusions: Using the SCRS gene expression, we found preferential prediction of clinical outcomes and pathological diseases by subcluster assignment. With further characterization and functional validation, our findings may reveal certain sub-populations of immune cells amenable to therapeutic intervention.

Translational Relevance:

ccRCC is the most common form of renal cancer and nearly 50% of ccRCC patients develop metastasis. Although anti-CTLA-4 and anti-PD-1 antibodies have become first-line therapeutics for ccRCC, only a fraction of patients will respond. Using single-cell sequencing approaches, we found increased CD8⁺ T cells and macrophages in ccRCC compared to peripheral blood and normal renal parenchyma. We thoroughly characterized the immune cell populations, identifying proliferative CD8⁺ T cells and CD207⁺ macrophage signatures able to discriminate overall survival and pathological parameters in ccRCC using machine-learning approaches. These findings may offer insight into identifying immune cell subsets that can serve as prognostic or therapeutic indicators and serve as a data resource for both scientists and clinicians.

Introduction

ccRCC is the most common type of renal cell carcinoma, comprising more than 70% of all renal cancers (1). ccRCC represents an immune sensitive tumor type and is known for early advances in systemic immunotherapy using T cell proliferation cytokine IL-2 and interferon (IFN)- α 2b therapy (2). Recent novel immunotherapies targeting immune checkpoints as standard of care have transformed the treatment paradigm of ccRCC (3,4). However, a substantial subset of renal cancer patients do not respond to these therapies and patients who initially do 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 (8–11). However, in RCC, immune cell abundance is inversely correlated with survival, specifically TILs including CD8⁺ T cells (12–15). Biomarker analysis results from recent clinical trials also supported the negative prognostic significance of T cell infiltrates in the absence of immunotherapy within treatment-naïve ccRCC patients (16,17). Other abundant immune players in the ccRCC tumor microenvironment include monocytes, dendritic cells, and tumor-associated macrophages (TAMs) (18) that are now just starting to be studied.

Quantifying and inferring immune cell abundance from transcriptional analysis of bulk tumor samples is inadequate to provide a clear picture of the immune cell types (19,20). 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 (21,22). Single-cell methodologies including flow cytometry, immunohistochemistry, and mass cytometry (14,18,23) have revealed immune cell states in ccRCC 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 (24–27), providing an unbiased approach to profiling cells and enabling molecular classification of different subpopulations and identification of novel gene programs. Transcriptomic mapping of T lymphocytes coupled with TCR sequencing allows additional measurement of clonal T cell response to cancer at an unprecedented depth (28,29).

Here, we report the single-cell mRNA profiling of the immune landscape in ccRCC mapping 25,688 of immune single cells (5'-sequencing and recombined V(D)J region of the T cell receptor) in matched tumor samples and peripheral bloods from three treatment-naïve ccRCC patients. We further integrated an additional, 11,367 immune cells isolated from peripheral blood and renal parenchyma (30) providing controls to evaluate tumor-specific transcriptional and clonal changes in immune populations at the single-cell level. Analysis of tumor-infiltrating T cells demonstrated distinct expression changes compared to peripheral blood and normal renal parenchyma. Clonal structure of T cells differed – with marked expansion seen in CD8⁺ T cells but not CD4⁺ T cells – and was associated with transcriptional patterning revealed by cell trajectory analysis. In myeloid cells, we observed an overall increase in macrophage populations with mixed polarization across patients. Predictive models, derived from the CD8⁺ T cells and TAMs, identified worse overall survival associated with proliferative CD8⁺ T cells and CD207⁺ TAMs. This represents the first such report of the immune landscape of ccRCC using SCRS for both transcriptional and clonal assessment.

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 were males with an age range of 67 to 74 years old. Tumor grades were histologically determined by a pathologist. Primary tumor stage for Patient 1 and Patient 2 were reported as pT1b without extension, while Patient 3 was reported as pT3a with vascular invasion into the renal vein. Three ccRCC tumor specimens paired with individual blood samples were used in the study. The current study was approved by the University of Iowa Institutional Review Board (IRB) under the IRB number 201304826 and conducted under the Declaration of Helsinki Principles. De-identified renal patients were recruited by Dr. Yousef Zakharia in the Department of Internal Medicine at the University of Iowa.

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 (31). Cells from fresh tumor specimens were incubated with FcR blocking reagent (StemCell Technologies, Vancouver, Canada) for 10 min at 4°C and labelled with 1ug/ml of the FITC anti-human CD45 antibody (BioLegend, San Diego, CA) per 10^7 cells for 20 min at 4°C. CD45⁺ cells were isolated using the EasySep™ 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, Ipswich, MA). Sorted cells were then counted and assessed viability MoxiGoll counter (Orflo Technologies, Ketchum, ID) ensuring that cells were resuspended at 1,000 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 Cell Ranger v2.2 pipeline, while clonotype sequencing was aligned to the vdj_GRCh38_alts_ensembl genome build provided by the manufacturer.

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 uniform manifold approximation and project (UMAP). Total number of cells from healthy peripheral blood control were 7,726. SCRS of normal immune populations in the kidney were derived from previously published data (30). Gene expression matrices were downloaded from the EGAS00001002325 and filtered for normal renal parenchyma cells using the provided cell manifest for the samples RCC1, RCC2, and RCC3. These samples were processed using the procedure as described below to form a UMAP. Immune cells were identified using canonical markers for lineage and were then isolated. Isolated immune cells for normal renal parenchyma were: RCC1 (n=1,011), RCC2 (n=888), and RCC3 (n=1,757).

SCRS Integration

Initial processing of cells isolated from ccRCC patients; Patient 1 (n=10,694), Patient 2 (n=5,174) and Patient 3 (n=9,805) were processed and integrated with the above samples using the Seurat R package (v3.0.2) (32,33). Samples were normalized using the *SCTtransform* approach (34) 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 UMAP utilized the top 30 calculated dimensions and a resolution of 0.7. Data characteristics by sequencing run can be found in Supplemental Table 1.

Cell type subclustering used the SCTtransform approach as described above, but by integrating the data across samples instead of individual sequencing runs. The adjusted dimensional inputs for the subclustering analysis can be found in Supplemental Table 2.

SCRS Data Analysis and Visualizations

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 bins across all cells was 80 and 40 for subcluster analyses, unless otherwise indicated in the figure legend. 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 conditions 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 (35). Genes were isolated by calling *cc.genes.updated.2019* in R.

Cell type identification utilized the SingleR (v1.0.1) R package (36) with correlations of the single-cell expression values with transcriptional profiles from pure cell populations in the ENCODE (37). In addition to correlations, canonical markers for cell lineages (Supplemental Table 3) and corresponding TCR sequences were used. Gene set enrichment analysis was performed using the *escape* R package (v0.99.0). Gene sets were derived from the Hallmark library of the Molecular Signature Database and from previous publications (22,25). Enrichment for anti-PD-1 therapy response was derived from Sade-Feldmen et alia to develop gene signatures for the CD8_B (nonresponsive) and CD8_G (responsive) single-cell populations(22). TCR analysis

utilized our previously described scRepertoire R package (v0.99.3) (38) with clonotype being defined as the combination of the gene components of the VDJ and the nucleotide sequence for both chains and assigned on the integrated Seurat object. Cell trajectory analysis used the slingshot (v1.6.0) R package (39) with default settings for the *slingshot* function and using the embedding from the subclustering for each cell type. Inferred start and end clusters were applied in the CD8⁺ T cell trajectory based on gene expression markers. Ranked importance of genes were calculated using the top 300 variable genes and rsample (v0.0.9) and tidymodels (v0.1.0) R packages were used to generate random forest models based on a training data set of 75% of the cells. The *rand_forest* function in the parsnip (v0.1.1) R package was used, with mtry set to 200, trees to 1400, and minimum number of data points in a node equal to 15 across all cell types. The code for all analyses are available at <https://github.com/ncborcherding/ccRCC>.

Mass Cytometry Analysis

Flow cytometry standard files were downloaded for 78 samples utilizing a previously-defined T-cell and tumor-associated macrophage panel (18). Subsequent loading and analyses of the data was based on the accompanying published methods (18). These files were loaded into R using the flowCore (v2.0.1) R package. Protein signal was arcsinh transformed using a cofactor of 5, filtered for previously identified T or myeloid cells. Further data visualization utilized ggplot2 (v3.3.1) and pheatmap (v1.0.12).

Machine Learning Modeling

The renal clear cell carcinoma (KIRC) log2 gene expression data from the downloaded from the University of California Santa Cruz Xena Browser and filtered for only primary tumor samples. To the expression data, updated clinical information was assigned by tumor barcode (40). Gene signatures from subcluster analysis were generated by filtering differential gene expression results for genes with > 0.5 log-fold change and 15% difference in cell expression. Training and

testing sample cohorts were divided using the *sample* function with *set.seed* set to 10, splitting the data into 1:1 ratio. Feature selection was performed using recursive feature selection using the cross-validation method to optimize feature selection for 5, 10, 15, and 20 features in the caret (v6.0-86) R package. Models were trained across several models, including support vector machines, bagged trees, and k-nearest neighbors, with selection of final models were blinded and based on the performance of trained results. Selected models were then used to predict survival in the testing cohort and testing parameters were then calculated. Survival analyses utilized the survival (3.1-12) and survMiner (v0.4.7) R packages.

Statistical Analysis

Statistical Analyses were performed in R (v4.0.1). 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.

Results

Single-cell expression profiling of immune cells in ccRCC

In order to define the immune microenvironment of human ccRCC, we obtained tumors and peripheral bloods – from three treatment-naïve ccRCC patients – that were subject to flow-sorted lymphoid and myeloid cells to SCRS. The general workflow for isolation and sequencing is available in Supplemental Figure 1. To these samples, we added immune cells from peripheral blood and normal renal parenchyma (30) to obtain an integrated UMAP projection of 22 clusters across 37,055 primary immune cells (Figure 1A). Across the three tissues, peripheral blood (n=21,160), tumor (n=12,239) and normal kidney (n=3,556), we found a number of clusters – notably clusters 0, 1, 3 and 7 – sharing similar gene expression (Figure

1B). Despite the integration of sequencing runs to reduce tissue-type divergence, each tissue type had enrichment for distinct clusters: peripheral blood formed the majority of Cluster 2, tumor tissues were enriched within Clusters 14, 17, 18, and 19, and normal kidney was enriched within Cluster 11. Based on gene expression, we assigned cell lineages to each cluster using a three method approach: 1) expressions of canonical markers for T cells (*CD3E*, *CD8A*, *CD4*, and *IL7R* or *CD127*), B cells (*CD19* and *MS4A1*), myeloid cells (*CD14* and *FCGR3A* or *CD16*) and natural killer (NK) cells (*KLRD1* and *NKG7*) (Figure 1C), 2) correlations with gene signatures derived from purified cell populations deposited by ENCODE (37) (Figure 1D), and 3) assignments of T cell clonotypes based on the TCR sequencing. Based on these approaches we annotated clusters as monocytes (Clusters 0, 5, 11, 12, and 16), CD4⁺ T Cells (Clusters 4, 6, 10, 13, 15, and 20), CD8⁺ T cells (Clusters 1, 8, 9 and 17), NK cells (Clusters 3 and 7), B cells (Cluster 2), macrophages (Cluster 14) and dendritic cells (DC, Clusters 18 and 19) (Figure 1E). Based on these annotations, we also examined the relative proportion of cell types comprising the sequencing runs by tissue type (Figure 1E). We observed a decrease of CD4⁺ T cells and B cells within normal kidneys or tumors relative to peripheral blood (Figure 1E). Conversely and as expected, we also found an increase of CD8⁺ T cells and macrophages in tumors relative to peripheral bloods (Figure 1E).

Preferential overlap between peripheral blood and tumor CD8⁺ T lymphocytes

With the extensive literature demonstrating the role of TCR expansion in anti-tumor immunity and immunotherapy (11), we first wanted to investigate the dynamics of CD4⁺ and CD8⁺ T cell clonal species in ccRCC. Using our previously described scRepertoire software (38), we assigned productive TCR sequences for TCRA and TCRB and defined clonotypes by the combination of both the genes and nucleotide sequences. For the identified T cells in ccRCC patients, recovering of at least one TCR chain, ranged from 74.8% to 87.6% after filtering and

clonotype reconstruction. The complete table of clonotype information for the ccRCC samples is available in Supplemental Table 4. T cell clonotypes had a clear distribution along the UMAP, with principal enrichment within Clusters 1, 4, 6, 8, 9, 13, 15, 17, and 20 (Figure 2A). The frequency of clonotypes was assigned across patient samples, allowing for the quantification of clonotype numbers in the context of individual patients. We observed an increase in clonotype frequency principally in the CD8⁺ T cell clusters (Figure 2A). Separating the T cell classes, we noted a stark difference in clonotype space occupied by the top 10 clones in the CD8⁺ T cells compared to CD4⁺ T cells across ccRCC patients (Figure 2B). This trend was consistent between the tumor-infiltrating and peripheral blood CD8⁺ T cells (Figure 2B). We next asked if this consistency in CD8⁺ T cell expansion was a result of shared expanded clonotypes between tumor and peripheral blood. We found a relative patient-specific increase in shared clonotypes in CD8⁺ T cells compared to CD4⁺ T cells (Figure 2C). We also noted that there was minimal overlap between patient clonotypes for both CD8⁺ and CD4⁺ T cells (Figure 2C). The patient-specific overlap of CD8⁺ clonotypes showed relatively larger pools in peripheral blood clonotypes contributing to the tumors (Figure 2D). Interestingly, Patient 3 – with the more advanced tumor stage (pT3a compared to T1 of Patient 1 and 2) – showed expansion in tumor-specific clonotypes that was not seen in the blood (Figure 2D). In Patient-3 with more advanced ccRCC, two clonotypes accounted for a total of 619 CD8⁺ T cells and were distributed across UMAP clusters (Figure 2E), which supports the notion that T cell clonotype is neither a determinant for UMAP clustering nor for functional indication.

CD8⁺ T cells in ccRCC tumors exhibit a transcriptional continuum with distinct populations

Sub-clustering of CD8⁺ T cells revealed 8 distinct clusters (Figure 3A) with relative tissue-specific distribution (Figure 3B). To understand the distribution of these new CD8⁺ clusters along the UMAP, we first examined the relative percent of single-cells represented in each

cluster by tissue type. Tissue-infiltrating CD8⁺ T cells (both tumor and normal kidney) comprised the majority of CD8_0, CD8_1, CD8_3, CD8_5, CD8_6 and CD8_7. Only Clusters CD8_2 and CD8_4 had increased relative levels of peripheral blood cells (Figure 3B). Going from right to left across the x-axis of the UMAP, there is a change in tissue-specific contribution starting from peripheral blood (right) to kidney (middle) to increasing levels of ccRCC tumor CD8⁺ T cells (left), which may represent the process of tissue infiltration itself. Within SCRS literature, there are concerns for variations in cell cycle leading to increased heterogeneity or obscure subpopulations (42), however proliferation of CD8⁺ T cells is an important surrogate marker of anti-tumor immune response (7). We next examined the variation in proliferative gene signatures, finding a similar distribution to the tissue-type with increasing cells in S or G2M phases from right to left, peaking with Cluster CD8_6 (Figure 3C).

In order to better characterize the CD8⁺ clusters, we next used canonical and differential T cells markers to examine gene expression differences along the UMAP (Figure 3D) with several patterns. The first pattern was the discovery of a naïve *CCR7*⁺ *SELL*⁺ *TCF7*⁺ being seen in CD8_4 (Figure 3D). Looking for effector CD8⁺ T cells, we next observed two populations of *IFNG*⁺ *PRF1*⁺ T cells, principally in CD8_1 and CD8_0 (Figure 3D). The latter also expressed immune checkpoints, such as *CTLA4*, *HAVCR2*, *PDCD1*, and *TIGIT* (Figure 3D). These immune checkpoints were expressed at more moderate levels in both CD8_5 and CD8_6; however, CD8_6 exclusively expressed a number of proliferation markers, such as *CDK1*, *MKI67*, *STMN1*, and *TOP2A* (Figure 3D). In order to examine gene expression patterns above single or selected genes, we used slingshot (39) to build minimum spanning trees between clusters, generating curves based on the most varied genes (Figure 3E). We identified 5 distinct curves (labeled B₁ to B₅) with the origin in the *CCR7*⁺ *SELL*⁺ *TCF7*⁺ CD8_4. With the exception of B₁ extending into CD8_2, the remaining curves graphed along a similar trajectory with a

common node of CD8_1 and branching at distinct clusters with increased levels of tumor-infiltrating CD8⁺ T cells (Figure 3E). These curves also varied by CD8⁺ T clonotypes based on TCR sequencing, with the root having no clonal expansion and the B₃, B₄ and B₅ curves terminating into regions with higher levels of clonal expansion compared to B₁ or B₂ (Figure 3E). This clonotype relationship was also observed in specific clonotype sequences with overlapping clonotypes seen in subclusters CD8_0, CD8_6, CD8_5 and CD8_3 (Figure 3F). In contrast, CD8_7 had minimal overlapping clonotypes with other subclusters (Figure 3F). This relationship was seen also independent of the individual patient sequenced (Supplemental Figure 2). In order to assess possible functional differences based on these branching, we performed gene set enrichment analysis (Figure 3G) (43). As expected based on the immune checkpoint inhibitors expression (Figure 3D), Clusters CD8_0 and CD8_5 showed increased terminal differentiation and exhaustion (Figure 3G). Cytolytic gene enrichment was seen in CD8_1, the PRF1⁺ IFNG⁺ population lacking immune checkpoints (Figure 3G). The highly proliferative CD8_6 population was enriched for metabolic activity, such as the tricarboxylic acid cycle and glycolysis, and DNA repair (Figure 3G). The B₂ curve termination cluster, CD8_7, has preferential enrichment of cytokine signaling, such as IL-2/STAT5, TGFβ, and type 1 interferon (Figure 3G). With immune checkpoint inhibitor responsiveness associated with distinct CD8⁺ T cell populations (22), we next examined enrichment of signatures associated with response or nonresponse to anti-PD-1 therapies (Figure 3H). Using the ordinal construction of the trajectories, we created a pseudotime variable for cells, allowing us to see the difference in the enrichment along the curves. This approach allowed us to see an overall enrichment in responsiveness to anti-PD-1 at the terminal points of curve B₂ and midpoints of B₃, B₄ and B₅, corresponding to cells in CD8_1 (Figure 3H, blue lines). Likewise, we observed an overall increase in gene expression associated with no response or progression was observed at the terminal points of curves B₃, B₄, and B₅ (Figure 3H, red lines).

Single-cell CD4⁺ T cell characterization in ccRCC identifies disparate intratumoral populations.

CD4⁺ T cells can influence cancer pathogenesis in various ways, either directly through cytolytic mechanisms or indirectly by modulating the tumor immune microenvironment. Subclustering of CD4⁺ T cells revealed 9 distinct clusters (Figure 4A), with a similar pattern – as seen in CD8 T cells – of tissue distribution with predominantly peripheral blood CD4⁺ T cells on the right leading to tissue-infiltrating CD4⁺ T cells on the left (Figure 4B). The CD4_8 was composed solely of peripheral blood cells from the healthy donor and was eliminated from the remaining analysis. Like the CD8⁺ T cells, we next examined the canonical and differential T cell markers along the UMAP (Figure 4C). The first pattern that emerged was a naïve *CCR7⁺ SELL⁺ TCF7⁺* being seen in CD4_1 and CD4_3 (Figure 4C). Within the tumor-infiltrating CD4_4 cluster, we observed increased expression of the Th1 driver *TBX21* (T-bet), activation marker *LAG3* and *NR4A2* and cytokine expression (Figure 4C). Both CD4_5 and CD4_7 had expression of regulatory T (Tregs) cell markers (Figure 4C), with higher levels of *FOXP3*, *IL2RA* (CD25), *CTLA4* and *TNFRSF18* (GITR) in the tumor-predominant CD4_5 (Figure 4C).

Constructing the cell trajectory curves based on the CD4⁺ subclustering, we observed two root points of the *CCR7⁺ SELL⁺ TCF7⁺* Clusters CD4_1 and CD4_3 leading to a common CD4_4 termination (Figure 4D). Unlike the other CD4⁺ T cells, the curve generated for Tregs was divergent, starting at CD4_5 through CD4_7 and into CD4_4 (Figure 4D). This likely represents a distinct expression pattern for Tregs (shared by CD4_5 and CD4_7) compared to other tumor-infiltrating CD4⁺ T cells. In addition, compared to the CD8⁺ subclustering, modest clonal expansion was seen in CD4_4 and CD4_5 and was not a clear pattern for cell trajectory (Figure 4E). With the common termination point for the curves at Cluster CD4_4, we next wanted to examine if there are common markers for CD4⁺ T cell infiltration in ccRCC by comparing tumor-

infiltrating to peripheral-blood CD4⁺ T cells. Within the tumor-infiltrating CD4⁺ T cell, 203 genes adjusted p-value < 0.05, log-fold change \geq 0.5 and Δ cell percent > 10% (Supplemental Table 5). Upregulated within the tumor-infiltrating CD4⁺ T cells were heat shock proteins (*HSPA1A* and *HSPA1B*), Jun and FOS constituents (*FOS*, *JUN*, *JUNB*), MHC-II molecules (HLA-DRB), and secreted molecules (*CCL5*, *GZMA*, *GZMK*) (Figure 4F). Several of the upregulated genes are shared across all the tumor-predominant CD4 Clusters (Figure 4F); however, each cluster also had unique expression markers. Both CD4_2 and CD4_4 had increased levels of *IFNG* (Figure 4F), but CD4_2 was enriched for heat shock proteins, while CD4_4 had cytotoxic component and there was expression of *CD8A*, which likely represents modest contamination of CD8⁺ T cells (Figure 4F). The tumor-infiltrating Tregs, CD4_5, had high levels of *CTLA4*, *GITR* (*TNFRSF18*) and *TIGIT*. In addition CD4_5 had the highly-specific expression *CCR8* and *LAYN*, corresponding to previous reports (44,45). The CD4_6 cluster had increased expression of the IL-6 cytokine, *OSM6*, and *AREG* and *SOCS3*, downstream of interleukin signaling (Figure 4F). The differential expression closely matched the pathway analysis, with CD4_4 enriched for cytolytic and type I interferon signaling (Figure 4G). The CD4_5 and CD4_7 Treg cluster had preferential enrichment for metabolic pathways, with high levels of terminal differentiation in tumor-infiltrated CD4_5 (Figure 4G). The *OSM*-high CD4_6 was enriched for IL-6/JAK/STAT3 signaling and inflammatory response genes (Figure 4G).

Prominent infiltrating macrophages in ccRCC have transcriptional divergence.

With the previous observation of an overall increase in macrophages and decreased monocytes in the integrated UMAP (Figure 1E), we next focused on differential analyses of the myeloid populations (Figure 5A). Across monocytes, macrophages and dendritic cells, subclustering found 20 distinct clusters (Figure 5A). Tissue-specific distribution was observed, with the majority tumor-infiltrating myeloid cells in subclusters 0, 3, 7, 8 and 15 (Figure 5B). In contrast,

both normal kidney parenchyma and peripheral blood were comprised of a majority of monocytic subclusters (Figure 5B). Populations were assigned using canonical markers and in addition to the previously described singleR approach with macrophage subclusters (0, 3, 7, 13 and 14) identified using markers such as *CD1C*, *CLEC10A* (CD301), *FCER1A*, and *MSR1* (Figure 5C). In total, we observed a decreased proportion in CD14⁺ monocytes in tissue-infiltrating myeloid cells compared to peripheral blood and an increase in macrophages (Figure 5D). Normal renal parenchyma had a variable increase in CD14⁺ CD16⁺ monocytes, which was not significant (Figure 5D). As previously seen, we found a small number of DC (subclusters 15, 16, and 19) with distinct expression profiles associated with conventional DC1 (cDC1), plasmacytoid DC (pDC), and monocyte-derived DCs (moDC), respectively (Supplemental Figure 3).

Next we isolated the five macrophage subclusters, relabeling them tumor-associated macrophage 1 (TAM_1), TAM_2, TAM_3, resident macrophage (rM) and peripheral macrophage (pM) based on the relative percent of cells derived from the respective tissue (Figure 5E). Although similar in distribution along the UMAP, which preserves global structure of expression, these five clusters had distinct expression patterns (Figure 1E, 5E). For example, the CD88^{high} (*C5AR1*) CD54⁺ (*ICAM1*) TAM_1 expressed increased levels of chemokines and cytokines, like *CCL3*, *CCL4*, *CXCL2*, and *IL10*; CD64^{high} (*FCGR1A*) CD16^{high} (*FCGR3A*) TAM_2 subcluster expressed the apolipoprotein gene *APOE*, lysosomal lipase (*LIPA*), and ferroportin (*SLC40A1*); and CD1C^{high} CD86⁺ TAM_3 had high levels of *IL1R2* and Langerhin (*CD207*), a marker of the skin-resident Langerhan cells (Figure 5F). The pM subcluster had the highest level of the cell adhesion molecules *CLEC10A*, *SELL* (CD62L), and *ITGB7*, which can dimerize with *ITGA4* (CD49d) or *ITGAE* (CD103). Like the CD8⁺ T cells, we built cell trajectories based on varied genes and found two distinct curves converging into TAM_3 and pM (Figure 5G). In

order to assess potential functional differences for the macrophages, we performed gene set enrichment analysis (Figure 5H). As previously observed in single-cell data (25), no subclusters were distinctly M1 or M2 polarized. For example, TAM_1 had enrichment for gene sets commonly associated with the M2 macrophage compartment, such as angiogenesis and the production of lipid mediators, while also having the highest levels of TNF α signaling enrichment, a common M1 macrophage characteristic. Across the three TAM subclusters, modest enrichment in both type 1 and type 2 interferon signaling was observed (Figure 5H). The non-TAM subclusters, rM and pM, had relatively lower levels of enrichment with the exception of proinflammatory signaling and glycolysis, respectively (Figure 5H). We also found an increase in antigen processing and presentation of lipid antigens via MHC-I in TAM_2 and TAM_3, while TAM_1 had higher enrichment for polysaccharide antigens (Supplemental Figure 4).

Differential prognostic significance in CD8⁺ T cell and TAM subclusters.

As our data demonstrates transcriptional and suggests functional differences in CD8⁺ T cells and TAMs in ccRCC, we investigated whether gene signatures can be developed from our SCRS data with prognostic values (Figure 6A). Using the Cancer Genome Atlas dataset for ccRCC (19), we separated the cohort in half, giving a training and testing set. We isolated significantly upregulated genes from each subcluster of CD8⁺ T cells and macrophages selecting the top models for each cell type based on training results. Interestingly, we saw a consistently high performance for overall survival discrimination in CD8_6 and TAM_3-based signatures across all models we trained and different sizes of gene signatures. Applying the models to the testing cohort of 267 primary tumors, we found that both signatures had similar performance with 15-genes derived from CD8_6 and TAM_3, categorized roughly 25% of ccRCC into poor prognostic groups and equated to hazard ratios of 3.44 and 2.59, respectively

(Figure 6B). We also observed the distribution of the categories were associated with increasing histological grades (Figure 6C).

Moving beyond mRNA expression, we wanted to see if we could identify the proliferative CD8_6 subcluster in tumor samples. Using mass cytometry data for T cells isolated from 4 healthy tissue samples, 68 ccRCC primary tumors, and 4 ccRCC metastasis, we identified a PD-1⁺ Ki-67^{Hi} subset in 14.6% of a CD45⁺ CD3⁺ CD8⁺ T cells (Figure 6D). The majority of CD45⁺ CD3⁺ CD8⁺ T cells were either PD-1⁺ Ki-67⁻ (42.4%) or PD-1⁻ Ki-67⁻ (39%) (Figure 6D). In addition to PD-1, this proliferative subset of CD8⁺ T cells in ccRCC had increased levels of CTLA-4, ICOS, 4-1BB (CD137), TIM-3, HLA-DR, and CD38 compared to the other CD45⁺ CD3⁺ CD8⁺ T cells (Figure 6E). Calculating the proportion of PD-1⁺ Ki-67^{Hi} cells to total CD8⁺ T cells by sample, we categorized samples into thirds. We observed a similar distribution to the CD8_6 gene signature assignments of the highest tertile for PD-1⁺ Ki-67^{Hi} cells in more advanced histological grades (Figure 6F).

Discussion

With the improved understanding on how immunotherapies work, the phenotypic and functional profile of immune cells in the tumor microenvironment is now well known to influence prognosis and disease outcome. Comprehensive knowledge of gene expression signatures to fully understand the roles of specific immune populations in cancer is of high biological relevance; not only in identifying dysregulated immune determinants of cancer progression but also as a useful tool for the selection of patients, evaluating the likelihood of benefit from immunotherapy and further identifying clinically significant subpopulations.

Despite immunotherapy being a mainstay of treatment for advanced and treatment-naïve ccRCC (3,46,47), ccRCC tumors have numerous counterintuitive immune findings compared with other immunotherapy-responsive tumors (48). For example, unlike other tumors that respond to immune checkpoint blockade, ccRCC has a relatively low tumor mutational load, which is thought to drive T-cell infiltration (19,49), and is not associated with response to anti-PD-1 therapy (15). Despite this low mutational burden, ccRCC has the highest T cell infiltration score among tumor types within the TCGA (50). Similarly, therapeutic responses to anti-PD-1 therapy have been correlated with HLA heterozygosity in lung cancers and melanoma (51), which does not seem to be the case for ccRCC (15). Taken together, these disparate findings suggest a more complex interrelationship of the immune compartment for ccRCC tumors. While a few recent studies have explored human ccRCC at a single-cell level (30,52–54), the technique has not yet been applied to tumor-infiltrating immune cells to characterize the global transcriptional immune and T cell receptor landscape along with underlying mechanisms contributing to this unique tumor environment.

T cells are recognized as key effectors of the adaptive anti-tumor immune response. Several studies have demonstrated association of these cells with an unfavorable response to therapy and poor patient survival in ccRCC (6,13). In a comprehensive study, T cells represented the dominant lymphocytic population in most ccRCC cases and B cells were rarely detected (23), consistent with our findings of increased CD4⁺ and CD8⁺ T cells (Figure 1). We find that the blood CD8⁺ T cells are non-heterogeneous and poorly reflect tumor-infiltrating CD8⁺ T cell transcriptional profiles (Figure 3A,D). Organizing the structure of the CD8⁺ T cell manifold, we found four distinct branches that may represent transcriptional states upon tumor-infiltration, two associated with a PD-1⁺ TIM-3⁺ exhausted subcluster, a proliferative subcluster, and a fourth with the higher levels of cytokine signaling (Figure 3E,G). The latter cluster, CD8_7 was

also unique with minimal overlap in clonotypes compared to the other tumor-infiltrating predominant subcluster. Recent single-cell analyses in melanoma showed CD8⁺ T cells with lower activation and exhausted expression patterns were associated with improved anti-PD-1 responses (22). These responsive T cells had minimal shared clonotypes, similar to CD8_7 (22). Other studies have found the ccRCC tumors polyclonal CD8⁺ T cells with an “immune-regulated” phenotype and lower cytotoxicity compared to tumors with oligoclonal CD8⁺ T cells (55). Recent SCRS studies of pre- versus post-treatment of anti-PD-1 in basal cell carcinoma found have increased number and clonal expansion of CD39⁺ CD8⁺ T cells after immunotherapy (56). However, CD39⁺ CD8⁺ T cells in ccRCC has been shown to associated increased pathological stage and with poor overall survival (57). Based on gene expression, our CD8_0 and CD8_6 subclusters most closely fit this population of cells and these clusters had 57% and 46.5% of cells from the advanced-stage Patient 3, respectively. In developing the CD8 signature, we found the model discriminated overall survival, but also was associated with increasing histological grade (Figure 6B,C). Interestingly, patients with higher numbers of CD39⁺ CD8⁺ T cells had improved responses to the multi-tyrosine kinase inhibitor, sunitinib, suggesting that evaluation of exhausted phenotype for CD8⁺ T cells may help in clinical decision making or therapy selection (57). This is particularly interesting as we found shared CD8⁺, but not CD4⁺, T cell clonotypes in the corresponding peripheral blood of ccRCC patients (Figure 2). Although we find a stable overlap coefficient of around 13% for CD8⁺ clonotypes, more work is needed to assess the dynamics of infiltration versus exfiltration on the CD8⁺ T lymphocytes into the tumor bed.

The exhausted CD8⁺ T cell phenotype has been associated with advanced histological features and increased risk of disease progression (50,55,57), increased dysfunctional DC (6), and increased macrophage populations (18). However, controversy surrounds the role of myeloid

populations in ccRCC tumor prognosis and progression. This may, in part, be a result of transcriptional and phenotypic plasticity of tumor-infiltrating myeloid cells (18,25). Our analysis demonstrated distinct CD16⁺ myeloid population derived within tumor compared to peripheral blood or normal renal parenchyma and an overall increase in tumor-associated macrophages (Figure 5A, D). M2 markers, like CD163 and CD204, have been associated with poor clinical outcomes in ccRCC (18,58) and were the highest in the TAM_1 and TAM_2 subclusters (Figure 5G). This is despite no clear identification of canonical M1 or M2 macrophages subclusters (Figure 5H). Model training for gene signatures for TAMs found better overall discrimination using genes derived from TAM_3 (Figure 6B), a subcluster that was unique for having lower levels of gene enrichment for M2 macrophages, angiogenesis, and lipid mediator production (Figure 5H). The increased immunogenicity of ccRCC has been tied to upregulation of the antigen presenting machinery expression through MHC-I (50). Across the myeloid subclustering, there was gene enrichment for MHC class I processing and presentation machinery in the DC subsets, while the macrophages had increased MHC class II enrichment (Supplemental Figure 4). Although several distinct DC populations were detected and there was a trend for increase in cDC1 subset (subcluster 15) with known function of tumor antigen cross-presentation, further analysis was limited by the total number of DCs isolated.

Our strategy of single cell analysis performed on immune cells taking into consideration the frequencies of lymphoid and myeloid cells during flow sorting provides a powerful way to identify the relationship between proportion of cell types and corresponding immune cell states. With the population structure and gene programs defined in our study for multiple immune populations, we showed that though the patients showed variable proportions of lymphoid and myeloid in each cell state, the number of states remain limited. Our study has limitations, including the small total number of samples (n=7), the rarity of certain cell populations, and the need for further functional

characterizations of these immune populations. However, taken together, we provide a transcriptional and clonotypic map of ccRCC immune cells that in hope to gain insight into biomarkers and therapeutic targets in ccRCC.

Acknowledgments

We thank Michael Knudson, Rita Sigmund, Joe Galbraith, Janice Cook-Granroth, Bethany Kilburg and Celeste Charchalac from University of Iowa Carver College of Medicine, Tissue Procurement Core (TPC) and Genito-Urologic Tissue Repository (GUMER) for receiving biological samples and clinical data. We thank Justin Fishbaugh, Heath Vignes and Michael Shey from the University of Iowa Flow Cytometry Facility. We thank Kevin Knudtson, Mary Boes, Garry Hauser and Mari Eyestone from the Iowa Institute of Human Genetics (IIHG) Genomics Division for planning and assisting use of Next Gen Sequencing (NGS) platforms, Diana Kolb from the IIHG Bioinformatics Division and the University of Iowa High Performance Computing (HPC) facility.

Funding

Funding for this project was provided by Rock 'N' Ride Foundation (PI: Y.Z.) and from the National Institute of Health under the R01 CA200673 (PI: W.Z.), R01 CA203834 (PI: W.Z.), K08 CA226391 (PI: R.W.J) and F30 CA206255 (PI: N.B.). W.Z. was also supported by DOD/CDMRP grant BC180227 (PI :W.Z.), and an endowment from the Dr. and Mrs. James Robert Spencer Family Cancer Research Fund (PI: W.Z.) The flow cytometry and sequencing facilities are funded in part, by the National Cancer Institute of the National Institutes of Health under Award Number P30CA086862. The FACSria Fusion high-speed cell sorter was supported with funds from the National Center for Research Resources of the National Institutes of Health under Award Number 1 S10 OD016199-01A1. The content is solely the responsibility

of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

Conception and design: AV, YZ, WZ

Development of methodology: AV NB WZ

Acquisition of data: KN, YZ, AV

Analysis and interpretation of data: NB AV AS RWJ WZ YZ

Writing, review, and/or revision of the manuscript: AV NB AS RWJ WZ YZ

Supervision: YZ, WZ, RWJ

Declaration of interests

Dr. Russell W. Jenkins has a financial interest in XSphera Biosciences Inc., a company focused on using ex vivo profiling technology to deliver functional, precision immune-oncology solutions for patients, providers, and drug development companies. Dr. Jenkins' interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies.

Data and materials availability:

Quantified gene expression counts and V(D)J T cell receptor sequences for single-cell RNA sequencing are available at the Gene Expression Omnibus (GEO) at [GSE121638](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE121638). Code for the analysis and visualizations are available at <https://github.com/ncborcherding/ccRCC>

References

1. Saad AM, Gad MM, Al-Husseini MJ, Ruhban IA, Sonbol MB, Ho TH. Trends in Renal-Cell Carcinoma Incidence and Mortality in the United States in the Last 2 Decades: A SEER-

- Based Study. Clin Genitourin Cancer. 2019;17:46–75.
2. Koneru R, Hotte SJ. Role of cytokine therapy for renal cell carcinoma in the era of targeted agents. Curr Oncol. Multimed Inc.; 2009;16:S40.
 3. Motzer RJ, Penkov K, Haanen J, Rini B, Albiges L, Campbell MT, et al. Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. N Engl J Med. 2019;380:1103–15.
 4. Dudani S, Graham J, Wells C, Pal SK, Dizman N, Donskov F, et al. First-line (1L) immuno-oncology (IO) combination therapies in metastatic renal cell carcinoma (mRCC): Preliminary results from the International Metastatic Renal Cell Carcinoma Database Consortium (IMDC). J Clin Oncol. American Society of Clinical Oncology; 2019;37:584–584.
 5. Sharma P, Allison JP. Immune checkpoint targeting in cancer therapy: Toward combination strategies with curative potential. Cell. 2015;161:205–14.
 6. Giraldo NA, Becht E, Pagès F, Skliris G, Verkarre V, Vano Y, et al. Orchestration and prognostic significance of immune checkpoints in the microenvironment of primary and metastatic renal cell cancer. Clin Cancer Res. 2015;21:3031–40.
 7. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJM, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature. 2014;515:568–71.
 8. Galon J, Fox BA, Bifulco CB, Masucci G, Rau T, Botti G, et al. Immunoscore and Immunoprofiling in cancer: An update from the melanoma and immunotherapy bridge 2015. J Transl Med. 2016;4.
 9. Ziai J, Gilbert HN, Foreman O, Eastham-Anderson J, Chu F, Huseni M, et al. CD8+ T cell infiltration in breast and colon cancer: A histologic and statistical analysis. PLoS One. 2018;13:e0190158.
 10. Shimizu S, Hiratsuka H, Koike K, Tsuchihashi K, Sonoda T, Ogi K, et al. Tumor-infiltrating

CD8+ T-cell density is an independent prognostic marker for oral squamous cell carcinoma. *Cancer Med.* 2019;8:80–93.

11. Borchering N, Kolb R, Gullicksrud J, Vikas P, Zhu Y, Zhang W. Keeping Tumors in Check: A Mechanistic Review of Clinical Response and Resistance to Immune Checkpoint Blockade in Cancer. *J Mol Biol.* 2018;430:2014–29.
12. Patel HD, Puligandla M, Shuch BM, Leibovich BC, Kapoor A, Master VA, et al. The future of perioperative therapy in advanced renal cell carcinoma: How can we PROSPER? *Futur Oncol.* 2019;15:1683–95.
13. Nakano O, Naito Y, Nagura H, Ohtani H, Nakano O, Sato M, et al. Proliferative activity of intratumoral CD8+ T-lymphocytes as a prognostic factor in human renal cell carcinoma: Clinicopathologic demonstration of antitumor immunity. *Cancer Res.* 2001;61:5132–6.
14. Baine MK, Turcu G, Zito CR, Adeniran AJ, Camp RL, Chen L, et al. Characterization of tumor infiltrating lymphocytes in paired primary and metastatic renal cell carcinoma specimens. *Oncotarget.* 2015;6:24990.
15. Braun DA, Hou Y, Bakouny Z, Ficial M, Sant' Angelo M, Forman J, et al. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat Med.* 2020;26:909–18.
16. Choueiri TK, Albiges L, Haanen JBAG, Larkin JMG, Uemura M, Pal SK, et al. Biomarker analyses from JAVELIN Renal 101: Avelumab + axitinib (A+Ax) versus sunitinib (S) in advanced renal cell carcinoma (aRCC). *J Clin Oncol.* 2019;37:101.
17. Choueiri TK, Larkin J, Oya M, Thistlethwaite F, Martignoni M, Nathan P, et al. Preliminary results for avelumab plus axitinib as first-line therapy in patients with advanced clear-cell renal-cell carcinoma (JAVELIN Renal 100): an open-label, dose-finding and dose-expansion, phase 1b trial. *Lancet Oncol.* 2018;19:451–60.
18. Chevrier S, Levine JH, Zanutelli VRT, Silina K, Schulz D, Bacac M, et al. An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell.* 2017;169:736–49.

19. Creighton CJ, Morgan M, Gunaratne PH, Wheeler DA, Gibbs RA, Robertson G, et al. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*. 2013;499:43–9.
20. Van Den Heuvel CNAM, Van Ewijk A, Zeelen C, De Bitter T, Huynen M, Mulders P, et al. Molecular profiling of druggable targets in clear cell renal cell carcinoma through targeted RNA sequencing. *Front Oncol*. 2019;9:117.
21. Savas P, Virassamy B, Ye C, Salim A, Mintoff CP, Caramia F, et al. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat Med*. 2018;24:986–93.
22. Sade-Feldman M, Yizhak K, Bjorgaard SL, Ray JP, de Boer CG, Jenkins RW, et al. Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell*. 2018;175:998–1013.
23. Geissler K, Fornara P, Lautenschläger C, Holzhausen HJ, Seliger B, Riemann D. Immune signature of tumor infiltrating immune cells in renal cancer. *Oncoimmunology*. 2015;4:e985082.
24. Chung W, Eum HH, Lee HO, Lee KM, Lee HB, Kim KT, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun*. 2017;8:1–12.
25. Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell*. 2018;174:1293–308.
26. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med*. 2018;24:978–85.
27. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* (80-). 2016;352:189–96.

28. Beausang JF, Wheeler AJ, Chan NH, Hanft VR, Dirbas FM, Jeffrey SS, et al. T cell receptor sequencing of early-stage breast cancer tumors identifies altered clonal structure of the T cell repertoire. *Proc Natl Acad Sci U S A*. 2017;114:e10409-10417.
29. Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell*. 2017;169:1342–56.
30. Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* (80-). 2018;361:594–9.
31. Baldan V, Griffiths R, Hawkins RE, Gilham DE. Efficient and reproducible generation of tumour-infiltrating lymphocytes for renal cell carcinoma. *Br J Cancer*. 2015;112:1510–8.
32. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. *Cell*. 2019;177:1888–902.
33. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411–20.
34. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol*. 2019;20:1–15.
35. Nestorowa S, Hamey FK, Pijuan Sala B, Diamanti E, Shepherd M, Laurenti E, et al. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood*. 2016;128:e20-31.
36. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol*. 2019;20:163–72.
37. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57–74.
38. Borchertding N, Bormann NL. scRepertoire: An R-based toolkit for single-cell immune

receptor analysis. *F1000Research*. 2020;9:47.

39. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*. 2018;19:477.
40. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell*. 2018;173:400–16.
41. Apetoh L, Smyth MJ, Drake CG, Abastado JP, Apte RN, Ayyoub M, et al. Consensus nomenclature for CD8+ T cell phenotypes in cancer. *Oncoimmunology*. 2015;4:e998538.
42. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol*. 2019;15.
43. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. 2009;
44. Plitas G, Konopacki C, Wu K, Bos PD, Morrow M, Putintseva E V., et al. Regulatory T Cells Exhibit Distinct Features in Human Breast Cancer. *Immunity*. 2016;45:1122–34.
45. De Simone M, Arrigoni A, Rossetti G, Gruarin P, Ranzani V, Politano C, et al. Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells. *Immunity*. 2016;45:1135–47.
46. Motzer RJ, Tannir NM, McDermott DF, Arén Frontera O, Melichar B, Choueiri TK, et al. Nivolumab plus Ipilimumab versus Sunitinib in advanced renal-cell carcinoma. *N Engl J Med*. 2018;378:1277–90.
47. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med*. 2015;373:1803–13.
48. Vuong L, Kotecha RR, Voss MH, Hakimi AA. Tumor microenvironment dynamics in clear-cell renal cell carcinoma. *Cancer Discov*. 2019;9:1349–57.
49. Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ, et al. Genomic correlates of

- response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* (80-). 2018;359:801–6.
50. Şenbabaoğlu Y, Gejman RS, Winer AG, Liu M, Van Allen EM, de Velasco G, et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* 2016;71:1–25.
 51. Chowell D, Morris LGT, Grigg CM, Weber JK, Samstein RM, Makarov V, et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* (80-). 2018;359:582–7.
 52. Kim KT, Lee HW, Lee HO, Song HJ, Jeong DE, Shin S, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome Biol.* 2016;17:1–17.
 53. Zhou W, Yang F, Xu Z, Luo M, Wang P, Guo Y, et al. Comprehensive Analysis of Copy Number Variations in Kidney Cancer by Single-Cell Exome Sequencing. *Front Genet. Frontiers*; 2020;10:1379.
 54. Wu TD, Madireddi S, de Almeida PE, Banchereau R, Chen YJJ, Chitre AS, et al. Peripheral T cell expansion predicts tumour infiltration and clinical response. *Nature.* 2020;579:274–8.
 55. Giraldo NA, Becht E, Vano Y, Petitprez F, Lacroix L, Validire P, et al. Tumor-infiltrating and peripheral blood T-cell immunophenotypes predict early relapse in localized clear cell renal cell carcinoma. *Clin Cancer Res.* 2017;23:4416–28.
 56. Yost KE, Satpathy AT, Wells DK, Qi Y, Wang C, Kageyama R, et al. Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat Med.* 2019;25:1251–9.
 57. Qi Y, Xia Y, Lin Z, Qu Y, Qi Y, Chen Y, et al. Tumor-infiltrating CD39+ CD8+ T cells determine poor prognosis and immune evasion in clear cell renal cell carcinoma patients. *Cancer Immunol, Immunother. Springer*; 2020;1565–76.

58. Komohara Y, Hasita H, Ohnishi K, Fujiwara Y, Suzu S, Eto M, et al. Macrophage infiltration and its prognostic relevance in clear cell renal cell carcinoma. *Cancer Sci.* 2011;102:1424–31.

Figures

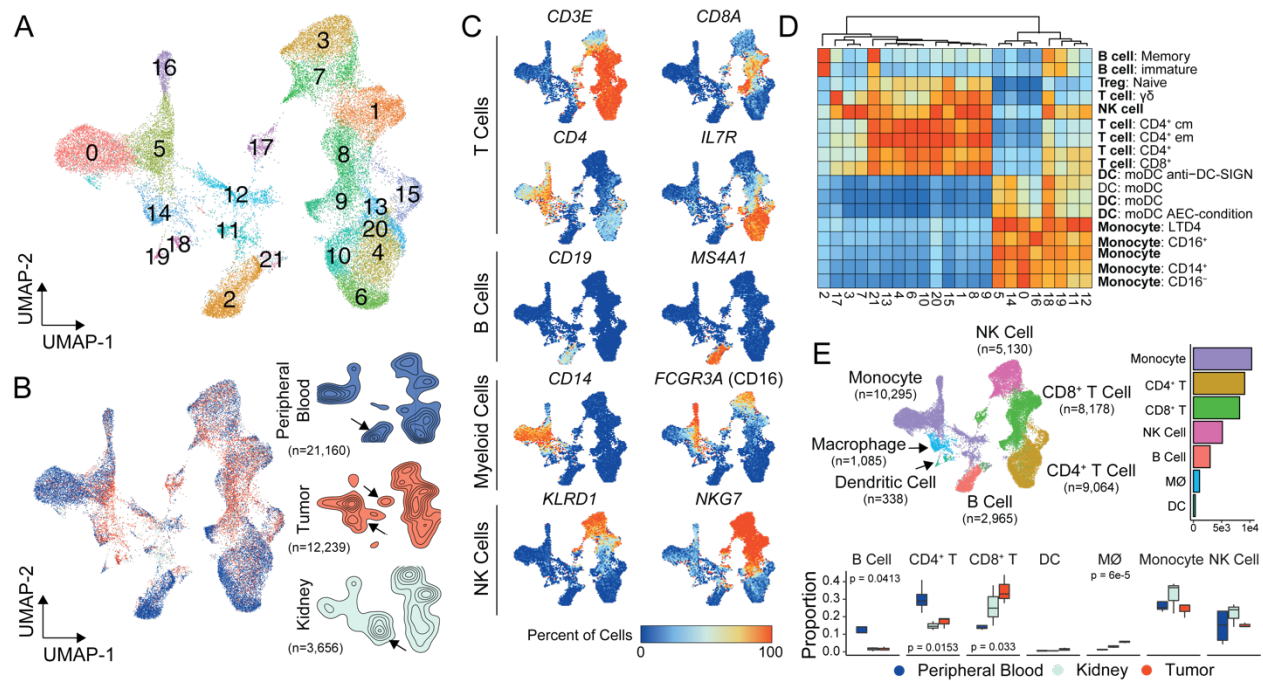


Figure 1: Single-cell RNA sequencing results from immune cells in ccRCC. **A.** UMAP of 37,055 primary immune cells of peripheral blood, normal renal parenchyma and tumor-infiltrating ccRCC patients. **B.** Distribution of cells by tissue type, peripheral blood (blue), tumor (red), and kidney (light blue). Arrows indicated potential enriched or unique immune cells populations for tissue type. **C.** Percent of cells expressing canonical immune cell markers across the UMAP. **D.** Normalized correlation values for predicted immune cell phenotypes based on the SingleR R package for each cluster; dendrogram based on Euclidean distance. **E.** UMAP demonstrating inferred immune cell types in ccRCC, clusters are colored by cell type and proportion of single-cell per sequencing run by tissue type. P values based on one-way ANOVA; lack of p-values equates to value > 0.05.

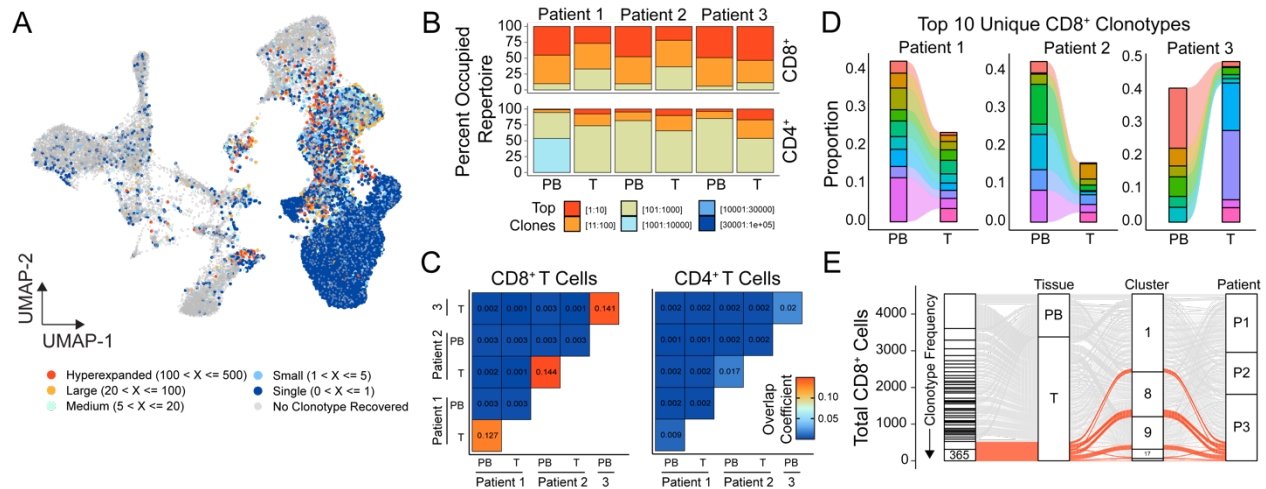


Figure 2: Clonal dynamics vary by T cell types and patients. **A.** UMAP of 37,055 primary immune cells overlaid with the frequency of clonotypes assigned by sample identification. **B.** Occupied repertoire space for the indicated clonotype groups for CD8⁺ and CD4⁺ T cells by sample and tissue type in ccRCC patients. **C.** Clonal overlap quantification by overlap coefficient for CD8⁺ and CD4⁺ T cells by sample and tissue type in ccRCC patients. **D.** The top 10 clonotypes for each patient as a relative proportion of clonotypes for corresponding peripheral or tumor populations. Each color represents a unique clonotype by patient. **E.** Distribution of clonotypes by tissue, UMAP cluster and ccRCC patient with highlighted (red) the top 2 clonotypes.

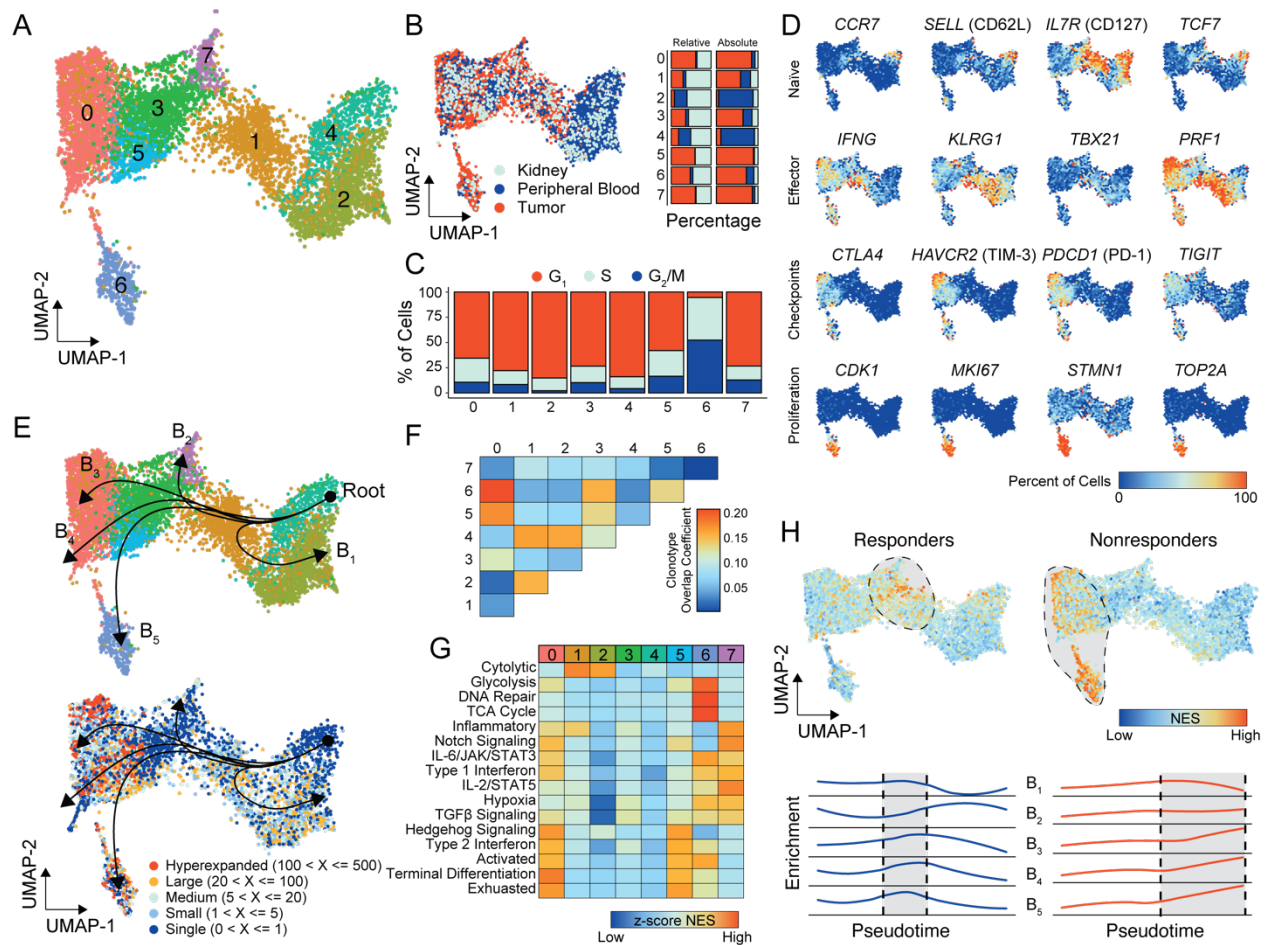


Figure 3: CD8⁺ T cells in ccRCC tumors exhibit a transcriptional continuum with distinct populations. **A.** UMAP subclustering of CD8⁺ T cells (original clusters 1, 8, 9, and 17). **B.** UMAP distribution of single cells by tissue type with relative and absolute percent of cells by tissue in each cluster. **C.** Cell cycle regression assignments for CD8⁺ T cells by cluster assignment. **D.** Percent of cells expressing selected markers for T cell biology. **E.** CD8⁺ UMAP of clusters (upper panel) and clonotype frequency (lower panel) overlaid with slingshot-based (39) cell trajectory starting at CD8_4 and proceeding into 5 distinct curves: branch 1 (B₁), B₂, B₃, B₄, and B₅. **F.** Clonotype overlap coefficients between subclusters. **G.** Z-transformed normalized enrichment scores from ssGSEA for selected gene sets by subcluster. **H.** Normalized enrichment scores for therapeutic response or lack of response to anti-PD-1 therapy across the CD8⁺ T cells (upper panel) and by pseudotime of each branch (lower panel).

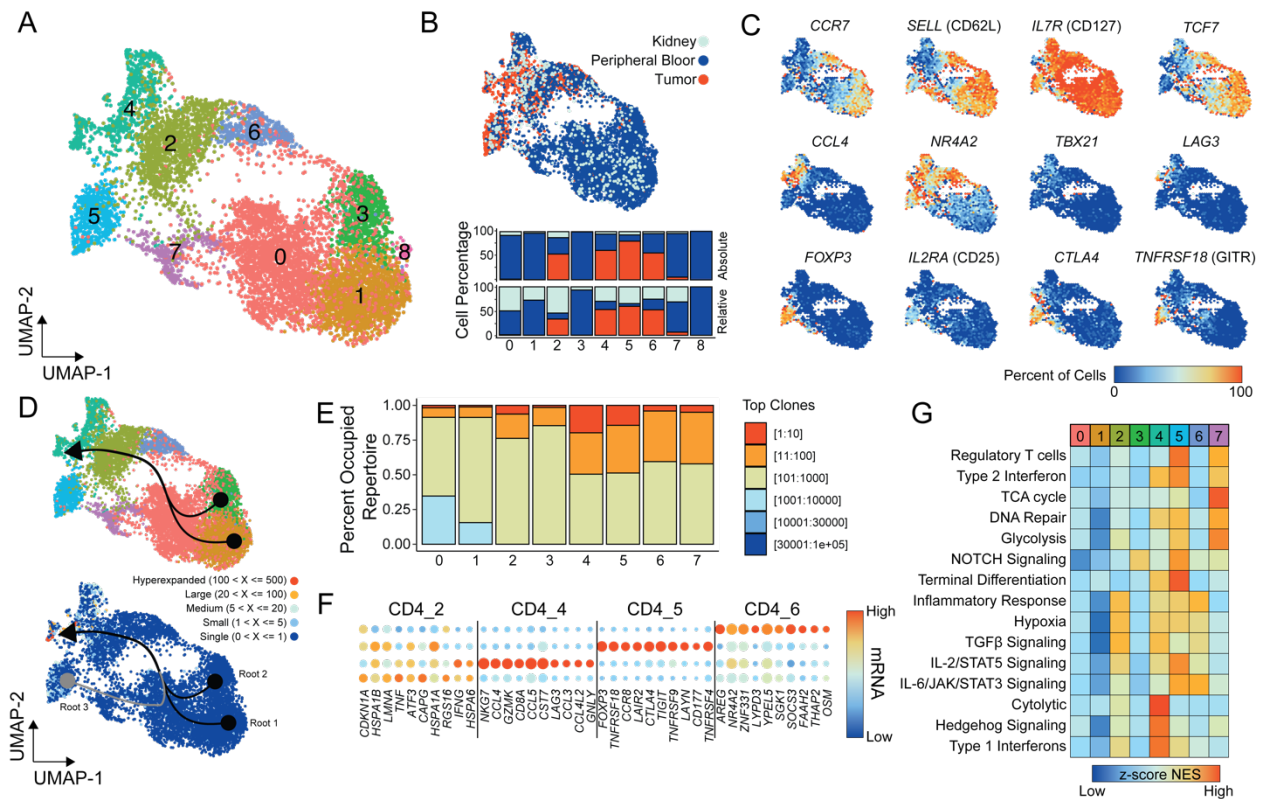


Figure 4: Single-cell CD4⁺ T cell characterization within ccRCC. **A.** UMAP subclustering of CD4⁺ T cells (original clusters 4, 6, 10, 13, 15, and 20). **B.** UMAP distribution of single cells by tissue type with relative and absolute percent of cells by tissue in each cluster. **C.** Percent of cells expressing selected markers for T cell biology. **D.** CD4⁺ UMAP of subclusters (upper panel) and clonotype frequency (lower panel) overlaid with slingshot-based (39) cell trajectory starting from CD4_1 (root 1) and CD4_3 (root 2) with relative pseudotime for all curves calculated using slingshot. **E.** Occupied repertoire space for CD4⁺ subclusters. **F.** Top 10 markers for TI-predominant CD4⁺ subclusters. Size of points are relative to percent of cells in the subcluster expressing the indicated mRNA species. **G.** Z-transformed normalized enrichment scores from ssGSEA for selected gene sets by subcluster.

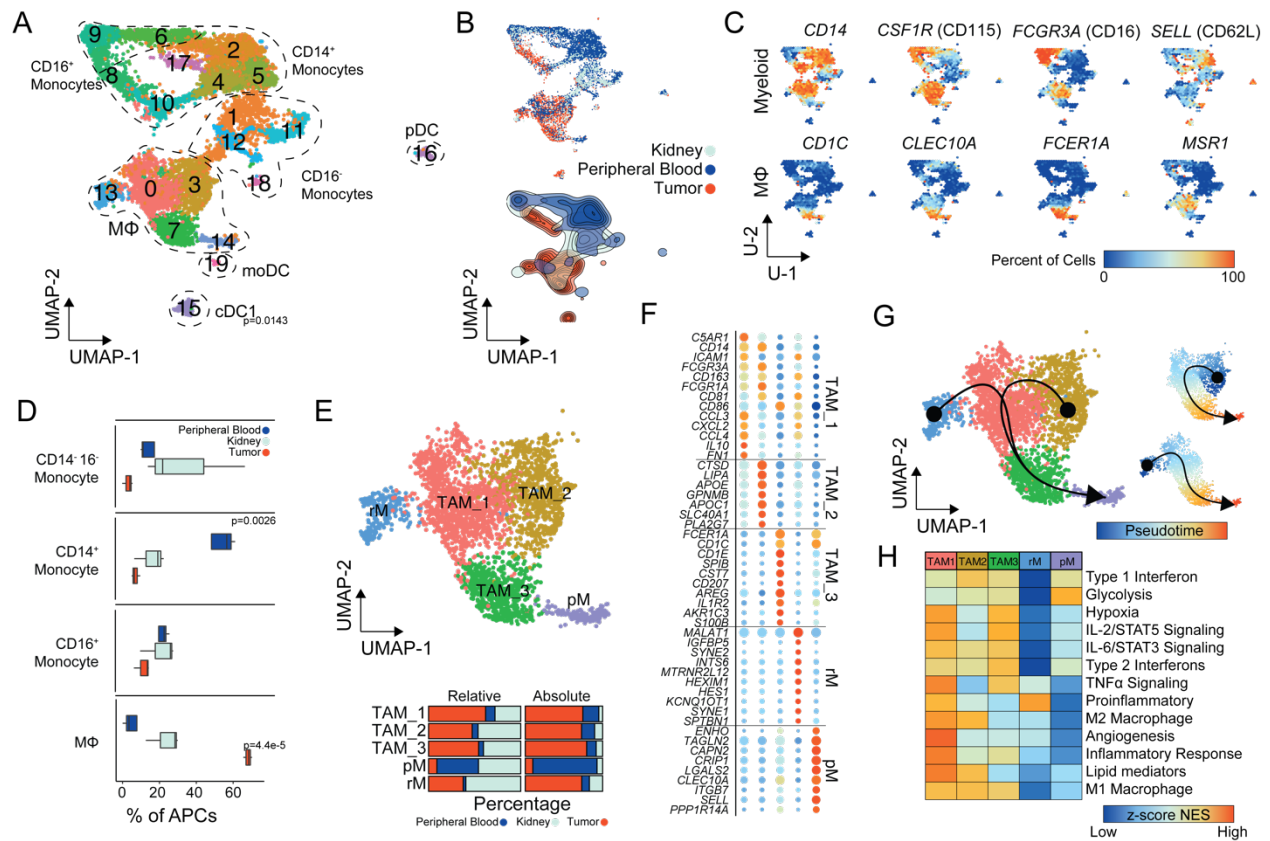


Figure 5: Single-cell myeloid characterization in ccRCC. **A.** UMAP subclustering of CD4⁺ T cells (original clusters 4, 6, 10, 13, 15, and 20). **B.** UMAP distribution of single cells by tissue type with relative percent of cells by tissue in each cluster. **C.** Percent of cells expressing selected markers for myeloid and macrophage markers. **D.** Proportion of assigned cell types compared to total antigen presenting cells by tissue type. **E.** Macrophage subclusters: TAM_1 (n=1,262), TAM_2 (n=840), TAM_3 (n=594), pM (n=275), and rM (n=194) with relative and absolute percent of cells by tissue in each cluster. **F.** Top differential expression markers for macrophage subclusters. **G.** Macrophage UMAP overlaid with slingshot-based (39) cell trajectories starting at rM and TAM_2 and proceeding into pM. Smaller UMAP shows pseudotime created by the cell trajectories. **H.** Z-transformed normalized enrichment scores from ssGSEA for selected gene sets by subcluster.

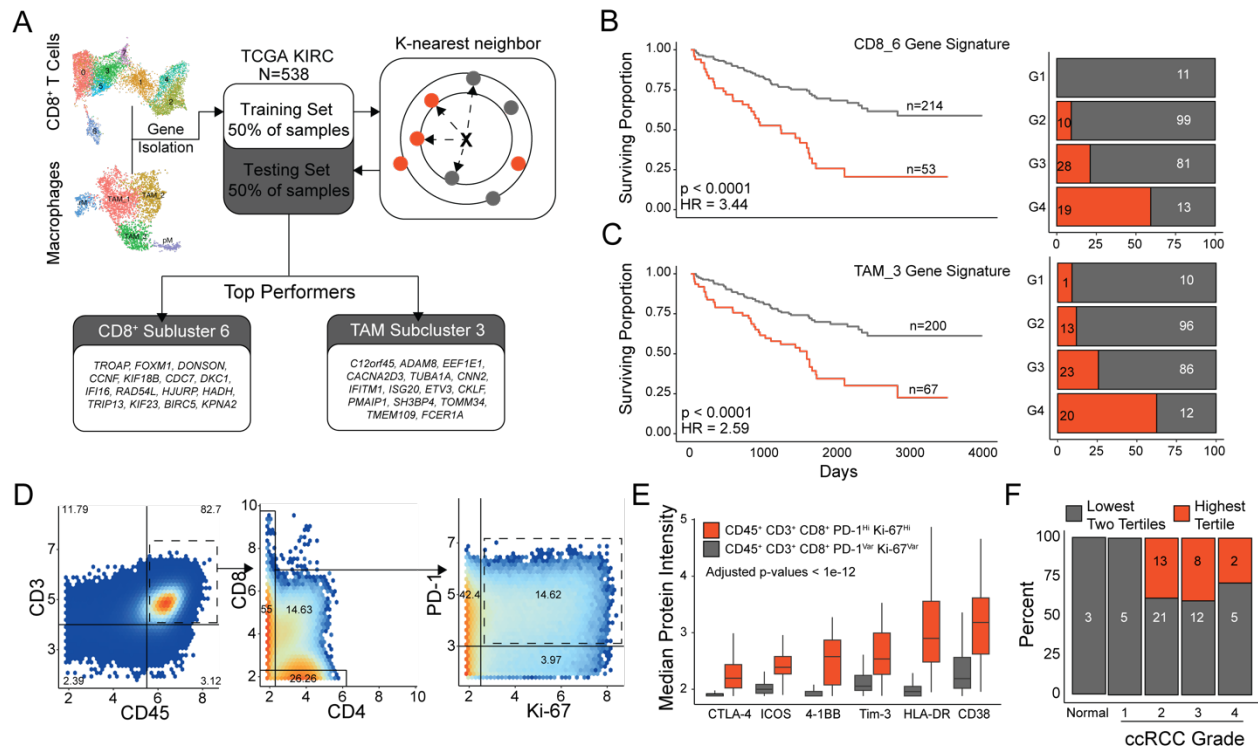


Figure 6: Prognostic values of gene signatures derived from CD8⁺ T cell and TAM subclusters. **A.** Schematic diagram of the machine-learning approach for signature development, selection and testing based on the k-nearest neighbors algorithm using the TCGA renal clear cell carcinoma data set. **B.** Kaplan-Meier curves for overall survival in testing subset for the CD8_6 subcluster signature with corresponding distribution of histological grades by model assignment. **C.** Kaplan-Meier curves for overall survival in testing subset for the TAM_3 subcluster signature with corresponding distribution of histological grades by model assignment. **D.** Density plots for 937,713 T cells isolated from 4 healthy samples, 68 primary ccRCC, 4 metastasis and quantified using mass cytometry (18). **E.** Markers with increased median expression in CD45⁺ CD3⁺ CD8⁺ PD-1⁺ Ki-67⁺ cells compared to other CD8⁺ cells. Adjusted p-values < 1e-12 for all indicated markers. **F.** ccRCC primary tumor and healthy samples subdivided into tertiles by the proportion of CD45⁺ CD3⁺ CD8⁺ PD-1⁺ Ki-67⁺ relative to the entire CD45⁺ CD3⁺ CD8⁺ pool by histological grade.