Discovering Master Regulators of Single-Cell Transcriptional States in the Tumor Immune Microenvironment to Reveal Immuno-Therapeutic Targets and Synergistic Treatments

Aleksandar Obradovic

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

Abstract

Discovering Master Regulators of Single-Cell Transcriptional States in the Tumor Immune Microenvironment to Reveal Immuno-Therapeutic Targets and Synergistic Treatments

Aleksandar Obradovic

The development of checkpoint immunotherapy has been a paradigm shift in the treatment of cancer, leading to dramatic improvement in treatment outcomes across a broad range of tumor types. Nevertheless, our current understanding of the tumor immune microenvironment and mediators of resistance to therapy are limited. The recent development of high-throughput single-cell RNA-Sequencing (scRNA-Seq) technology has opened up an unprecedented window into the transcriptional states of distinct tumor-infiltrating immune and stromal cells. However, even this technology has its biological limitations, with very high levels of data dropout induced by low total mRNA molecules and capture efficiency.

This thesis explores the application of a transcriptional regulatory protein activity inference approach to single-cell data in order to resolve gene dropout and more deeply characterize upstream drivers of cell state within the micro-environment of several distinct tumor types. To this end, algorithms for inference of protein activity, drug sensitivity, and cell-cell interaction have been adapted to scRNA-Seq data, along with an approach for querying enrichment of single-cell-derived population marker gene sets patient-by-patient in larger bulk-RNA-Seq cohorts. By applying these tools systematically, we have identified distinct cellular sub-populations associated with clinical outcome in different tumor types, including a novel population of C1Q+/TREM2+/APOE+ macrophages associated with post-surgical tumor recurrence in clear cell renal carcinoma, a sub-population of fibroblasts associated with

improved response to immunotherapy in head and neck squamous cell carcinoma, tumor cell subpopulations with distinct inferred drug sensitivities in cholangiocarcinoma and prostate cancer, as well as tumor-specific regulatory T-cells (Tregs), active as a mechanism of immunotherapy resistance across a range of tumor types. In ongoing clinical trials from both primary and metastatic prostate cancer as well as clear cell renal carcinoma, we are able to assess which of these populations are enriched in non-responders to checkpoint immunotherapy. The proteomic master regulators of each of these single-cell types have direct utility as potential biomarkers for treatment response, but they may also be therapeutically modulated as novel targets for combination immunotherapy, potentially improving treatment response rates and treatment outcomes in future clinical trials.

Finally, this thesis also presents a discovery-to-validation platform to accelerate micro-environment-directed drug repurposing in the context of immunotherapy resistance and rapid CRISPRko validation of novel therapeutic targets. This platform has been developed specifically to validate newly identified master regulators of tumor-specific immunosuppressive regulatory T-cells (Tregs), resulting in discovery of low-dose gemcitabine as a tumor-specific Treg-modulating drug synergistic with anti-PD1 checkpoint immunotherapy and TRPS1 as a proteomic master regulator with clinically significant effect on tumor Treg-infiltrating and tumor growth rate. However, the platform itself may be readily extended in future work to prioritize agents against immunosuppressive macrophage and fibroblast populations for clinical development and trials. As we have discovered, different cancers have different populations of cells driving therapy response and resistance. Taken together, the analytical and validation tools presented in this thesis represent an opportunity to tailor future immuno-therapies at the single-cell level to particular tumor types and to individual patients.

Table of Contents

List of Charts, Graphs, Illustrations

Acknowledgments

I would first and foremost like to thank my mentors, Dr. Andrea Califano and Dr. Charles Drake, for fostering my growth and development as a scientist with their trust, encouragement, and excellent advice. Through their example I have developed my skills and experiences in systems biology, immunology, and clinical research, and I have been honored every day to work with them. I would also like to thank my committee members, Dr. Yufeng Shen, Dr. Benjamin Izar, and Dr. Steven Reiner, for their advice and insight, as well as my close collaborators and friends Dr. Matthew Dallos, Dr. Jessica Hawley, Dr. Mark Stein, Dr. Casey Ager, Dr. Mikko Turunen, Dr. Katherine Spina, and Dr. Nivedita Chowdhury. I would further like to thank everyone in the Drake and Califano Labs who has contributed to my personal and scientific growth and made my research possible, particularly Lukas Vlahos, Aaron Griffin, and Lorenzo Tomassoni, from whom I have learned much over the past several years. Finally, I thank all my colleagues and collaborators at Vanderbilt University, Memorial Sloane Kettering, and Johns Hopkins University, who have each shown me great trust and respect.

Dedication

This thesis work is dedicated to my wife, Kayla, who has been my constant support through graduate school and through life, and to Luka, my little buddy who has been motivating me since before he was born. I am truly thankful and lucky to have such an amazing family. This work is dedicated also to my parents, whose good examples first exposed me to scientific research, and who taught me to work hard for the things that I aspire to achieve.

Introduction

Checkpoint Inhibitors and the Role of the Tumor Microenvironment

Solid tumors consist not only of tumor cells, but also of many diverse stromal and immune cell types infiltrating the tumor micro-environment. Traditional approaches to cancer therapy have commonly focused on killing tumor cells directly, but recent immune checkpoint inhibitor immunotherapies have aimed instead to activate anti-tumor immune cells in the tissue. The development of checkpoint immunotherapies has been transformative in clinical oncology over the past several years, motivating efforts to better profile specific immune cell types in tumors under various treatment conditions, in the hope that this will reveal novel therapeutic targets and combination therapies.

Within any solid tumor, there is a complex milieu of stromal cells, including fibroblasts and vascular endothelium, as well as immune cells ranging from macrophages and dendritic cells which present cancer antigens, to B-cells, cytotoxic T-cells, regulatory T-cells, and various immunosuppressive populations. Anti-tumor response relies on infiltration of immune cells into the tumor compartment, tumor-cell recognition, and immune-mediated cell death, as well as local production of cytokines and clonal expansion of cytotoxic T-cells [1]. Mechanisms of immune tolerance in tumors are common, most notably the activity of immune checkpoints that inhibit cytotoxic T-cell activity.

In particular, the interaction between programmed cell death 1 (PD-1) on T-cells and its corresponding ligand PD-L1 on tumor cells induces functional T-cell exhaustion and has been successfully targeted in the clinic by the anti-PD1 antibodies pembrolizumab and nivolumab to reactivate exhausted T-cells and re-engage their tumor-killing capacity. These drugs are now

widely used across a range of tumor types. For example, objective responses are seen in 40-45% of patients with melanoma given pembrolizumab or nivolumab in the first line setting and in 20% of patients with non-small cell lung cancer [2]. Although these and other checkpoint immunotherapies can elicit dramatic therapeutic responses, the characteristics of the tumor-immune microenvironment responsible for response or treatment resistance are incompletely described.

Clinical Predictors of Response to Immunotherapy

To date, the most-utilized biomarker of response to checkpoint immunotherapy has been expression of PD-L1 in tumor tissue. However, both PD-L1 negative and PD-L1 positive tumors respond to immunotherapy, with different trials showing different levels of biomarker correlation with Response Rate (RR) and Progression-Free Survival (PFS) [3, 4]. Negative predictive value of PD-L1 staining as a marker for treatment response is as low as 58% for nivolumab, and 45% for nivolumab plus ipilimumab in patients with melanoma [2]. Other investigational predictors of response to immunotherapy have included Tumor Mutational/Neoantigen Burden, Bulk Immune Gene Expression Signatures, and Multiplex Immunohistochemistry staining [2]. These capture different potential mechanisms of therapy resistance, but each has significant drawbacks.

A study of mutational load based on tumor whole-exome sequencing in advanced melanoma patients treated with anti-CTLA4 immunotherapy identified that a load of >100 non-synonymous somatic mutations was associated with longer overall survival (log-rank p-value=0.04 in discovery cohort, 0.10 in validation cohort) [2]. In non-small-cell lung cancer higher mutational burdens were found to be associated with clinical benefit, but intra-tumoral heterogeneity of neoantigen load further affects response rate [2].

Immunohistochemistry is limited to investigating small sets of pre-specified markers and can be very laborious to expand to novel markers. A study of melanoma tumor slides stained for CD3, CD8, FOXP3, CD163, and PD-L1 showed that presence of CD8+ T-cells alone was insufficient to predict growth of tumor-infiltrating lymphocytes ex vivo, although the ratio of CD8+ T-cells to immunosuppressive CD3+FOXP3+ Tregs was associated with tumor-infiltrating lymphocyte growth [2].

Finally, gene expression has been investigated as a predictor of response to immunotherapy, identifying panels of immune-related genes that capture multiple aspects of the immune response. One 28-gene signature developed in melanoma significantly correlated with overall response to immunotherapy and progression-free survival, optimized to a positive predictive value of 59% and negative predictive value of 90% [2]. However, such approaches suffer from an inability to distinguish which cell types present in the tumor micro-environment are responsible for expression of any given gene and are unable to distinguish the distinct contributions of different cell types to immunotherapy response across patients, or to identify transcriptional regulators specific to those cell types.

Mechanisms of Resistance to Immunotherapy

Cell-intrinsic mechanisms of immunotherapy resistance are those active in the tumor cells themselves, rendering them less susceptible to immune-mediated cell death or less visible to the immune system. These include downregulation of antigen presentation, engagement of alternative immune checkpoints, secretion of immunosuppressive cytokines, or decrease in tumor mutational burden under selective pressure on heterogeneous tumors [5] [6] [7]. Cell-extrinsic mechanisms largely involve exclusion of cytotoxic cells and recruitment of immunosuppressive cell types,

including regulatory T-cells (Tregs) as well as myeloid-derived suppressor cells (MDSCs), each of which interact with each other and with the remainder of the tumor microenvironment and which have been associated with poor prognosis across various tumor types [6] [8] [9] [10] [11]. In reality, these two categories of resistance mechanisms likely share common regulatory pathways and may be modulated by drugs targeting key tumor cell regulators of immune evasion, immunosuppressive cell regulators of tumor-infiltration, or interactions between the two.

Immunomodulatory Effects of Conventional and Targeted Cancer Therapies

A large number of commonly used conventional and targeted cancer therapies have been shown to exert immunomodulatory effects on the tumor microenvironment, but the mechanisms of these effects and the clinical relevance of each effect in relation to immunotherapy efficacy have not been well-studied [12] [1] [13]. This presents an area of great opportunity for exploration of drug synergies that may overcome mechanisms of resistance to immunotherapy. Chemotherapies can promote tumor immunity by inducing immunogenic cell death, or by disrupting cell-intrinsic mechanisms by which tumors evade immune recognition. It may also render tumor cells more sensitive to T-cell-mediated lysis through fas, perforin, and granzyme dependent mechanisms [14]. Chemotherapy has additionally been shown to exert immunomodulatory effects directly on immune cells. For example, low-dose cyclophosphamide depletes circulating naïve and activated Tregs [14], and paclitaxel has been shown in mouse models to reprogram tumor macrophages to a pro-inflammatory phenotype in a TLR4-dependent manner [15]. Targeted therapies also have potentially significant immunomodulatory effects [12]. In mouse models, intra-tumoral Treg depletion with an anti-CTLA4 antibody synergizes with androgen deprivation where it would otherwise fail to show survival benefit as monotherapy [4, 4]. This leads to the hope that even in

cancers otherwise unresponsive to immunotherapy, synergistic approaches that combine immunotherapy with drugs targeting resistance mechanisms may provide survival benefit.

Single-Cell RNA-Sequencing as a Tool for Dissecting the Tumor Micro-Environment

The development of high-throughput droplet-based single-cell RNA sequencing (scRNA-Seq) technology [16] presents a significant opportunity in study of the tumor-immune microenvironment. scRNA-Seq has an advantage in resolution over traditional bulk RNA sequencing and can distinguish the transcriptional states of individual cell types contributing to the overall tumor phenotype. This is particularly important for immune populations present at relatively low frequency, such as Tregs, and for highly heterogeneous tumors.

The potential of scRNA-Seq has been shown in recent studies of melanoma, where scRNA-Seq in one trial of checkpoint immunotherapy identified a distinct TCF7+CD8+ T-cell state associated with outcome [17] and scRNA-Seq in another trial identified a tumor-cell-intrinsic transcriptional program predictive of response to immunotherapy [18]. In a pre-clinical trial of CDK4/CDK6 resistant breast cancer, scRNA-Seq identified a novel population of myeloid cells associated with poor response to immunotherapy and enriched in genes targeted by cabozantinib [19]. Application to other tumor types and different treatment conditions presents a broad area of novel opportunity for scRNA-Seq to elucidate mechanisms of immunotherapy resistance and identify context-specific drug synergies.

However, scRNA-Seq is not without problems. Due to low input RNA, it is common for individual cells to capture expression of fewer than 1000 unique genes, yielding gene expression matrices that are >90% sparse. This may adequately capture variation in highly expressed genes but miss variation in key transcriptional regulators and signaling molecules that have high

biological relevance but low gene expression. The Seurat pipeline commonly used for scRNA-Seq analysis attempts to resolve this issue by normalizing the raw expression matrix with regularized binomial regression and implements anchor-integration for batch-correcting multiple samples [20], but remains noisy at the level of individual genes.

Transcriptomics and the Master Regulator Paradigm of Tumor Biology

Upstream of gene expression at the RNA level, the actual drivers of cell phenotype are interactions between proteins and their regulatory/transcriptional targets. The Califano lab has developed and widely applied a method to infer transcriptional regulatory protein activity from gene expression data using two algorithms- the Algorithm for Reconstruction of Accurate Cellular Networks (ARACNe), and Virtual Inference of Protein activity by Enriched Regulon analysis (VIPER) [21] [22] [23]. First, ARACNe uses mutual information between genes at the expression level to build and prune a gene regulatory network, such that each transcription factor, co-transcriptional regulator, signaling molecule, and surface marker has an inferred "regulon" of downstream targets with directionality and strength of regulation encoded [23]. VIPER can then be used to infer protein activity from the enrichment of each regulon in a gene expression signature (Figure 1).

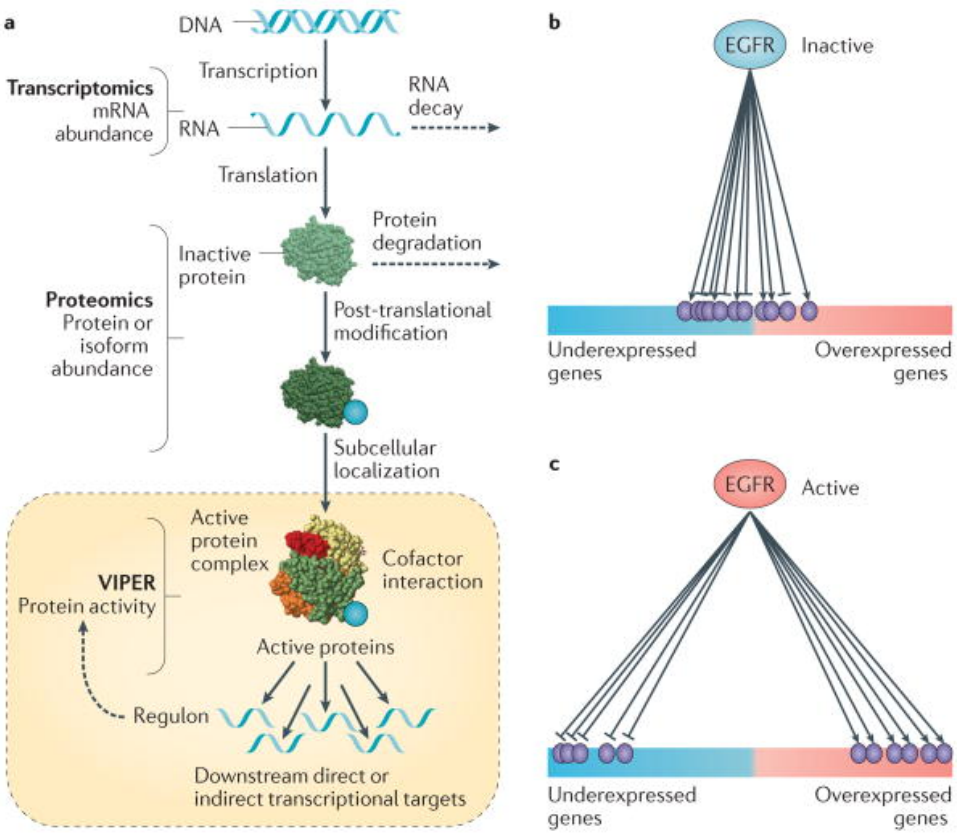


Figure 1: Conceptual Logic of the VIPER Algorithm.

Where a) indicates the sequence of events between gene expression and biological protein activity, b) represents a context in which EGFR (a master regulatory protein) is not active, and its downstream targets are consequently not over- or under-expressed, and c) represents a context in which EGFR is active, which can be inferred from the upregulation of its positively regulated targets and the downregulation of its negatively regulated targets.

The ability to infer protein activity allows for identification of differentially active protein master regulators (MRs) that regulate clinically relevant phenotypes. Clustering of tumor samples on inferred protein activity is significantly more robust to noise than clustering on raw gene expression [21] and also allows for drug discovery by identification of known drug targets (OncoTarget) and by identifying drugs that perturb the transcriptional state in a way that inverts MR activity for a specific tumor phenotype (OncoTreat) [24]. Furthermore, an extension of the

ARACNe/VIPER pipeline has been developed for Conditional Inference of Network Dynamics (CINDy), which infers surface receptors that modulate the activity of MRs based on mutual information between MRs and their downstream targets, conditioned on specified surface receptors [25].

These protein activity inference tools have thus far been applied primarily on bulk RNASeq data, but they provide significant potential benefit in scRNA-Seq data. Because ARACNe regulons infer many downstream targets for each protein, VIPER is less sensitive to single-cell dropout of specific genes, and has been shown to robustly infer protein activity from single-cell gene expression data using a metaVIPER approach that integrates protein activity inferred from ARACNe networks in multiple tissue contexts [26]. By clustering single-cells from patients' tumor microenvironment at the gene expression level, distinct patient-specific and cell-type-specific ARACNe networks can be constructed, and leveraged to infer protein activity in tumor cells and immune cells for improved clustering resolution, as well as to run the OncoTarget/OncoTreat algorithms for identification of druggable targets in tumor cells or immunosuppressive cells involved in treatment resistance, and to identify tumor-immune interactions involving receptors with high VIPER-inferred activity and matched ligands with high expression in an interacting cell type.

These methods require optimization of parameters for effective application to single-cell datasets in a way that is robust to varying levels of dropout and scalable to very high numbers of cells produced by modern droplet-based scRNA-Seq technologies. Approaches for optimization include repeated sub-sampling as well as pooling transcriptional nearest-neighbor cells for network inference, and optimal analytic pipelines may need to be tailored to characteristics of particular tumor types. Ultimately, leveraging the Califano Lab suite of network-based protein activity

inference methods on clinical datasets of patients profiled by scRNA-Seq with and without immunotherapy treatment allows for improved identification of cell-intrinsic and cell-extrinsic mechanisms of therapy resistance at single-cell resolution, and may lead to identification of clinically relevant drug synergies.

Chapter 1: Validation and Benchmarking of an Analysis Pipeline for Systematic Protein-Activity Inference from Single-Cell RNA Sequencing Data

The following is adapted from:

Obradovic, A.*, Vlahos, L.*, Laise, P., Worley, J., Tan, X., Wang, A.-L., & Califano, A. (2021). PISCES: A pipeline for the systematic, PROTEIN Activity-based analysis of single Cell Rna sequencing data. https://doi.org/10.1101/2021.05.20.445002
*These authors contributed equally

1.1 Summary

While single-cell RNA sequencing provides a new window on physiologic and pathologic tissue biology and heterogeneity, it suffers from low signal-to-noise ratio and a high dropout rate at the individual gene level, thus challenging quantitative analyses. To address this problem, we introduce PISCES (**P**rotein-activity **I**nference for **S**ingle **C**e**l**l **S**tudies), an integrated analytical framework for the protein activity-based analysis of single cell subpopulations. PISCES leverages the assembly of lineage-specific gene regulatory networks, to accurately measure activity of each protein based on the expression its transcriptional targets (regulon), using the ARACNe and metaVIPER algorithms, respectively. It implements novel analytical and visualization functions, including activity-based cluster analysis, identification of cell state repertoires, and elucidation of master regulators of cell state and cell state transitions, with full interoperability with Seurat's single-cell data format. Accuracy and reproducibility assessment, via technical and biological validation assays and by assessing concordance with antibody and CITE-Seq-based measurements, show dramatic improvement in the ability to identify rare subpopulations and to assess activity of key lineage markers, compared to gene expression analysis.

1.2 Introduction

High-throughput, droplet-based single-cell RNA Sequencing (scRNASeq) has recently emerged as a valuable tool to elucidate the diverse repertoire of cellular subpopulations comprising a broad range of mammalian tissues. Applications of this technology range from study of tissue development [27] and tumor micro-environment [28], to the elucidation of tissue heterogeneity [29] and even of tissue-level response to infectious diseases, such as COVID-19 [30] [31]. More specifically, scRNASeq data allows identification of representative gene expression signatures for thousands of individual cells dissociated from a tissue sample [16] [32], thus providing fine-grain characterization of the transcriptional state of individual cell types contributing to the emergence of complex phenotypes, which would be impossible from bulk profiles. This can help elucidate the role of rare populations, for instance, whose gene expression signature would be diluted below detection limits in bulk samples [20]. Moreover, in contrast to flow cytometry or CyTOF, scRNASeq generates genome-wide single cell profiles, without requiring *a priori* selection of a limited number of antibody-based markers. The value of scRNASeq in tumor biology has been broadly demonstrated in recent studies of melanoma [17] [18], pancreatic cancer [33], breast cancer [34], and renal cell carcinoma [35].

The key drawback of scRNAseq technologies is that the total number of mRNA molecules per cell, combined with low capture efficiency, fundamentally limits the number of distinct mRNA molecules that can be detected in each single cell (UMI reads). As a result, scRNASeq profiles are extremely sparse, with as many as 90% of all genes producing no reads in any given cell and the majority of detected genes producing one or two reads. This phenomenon, commonly known

as gene dropout, greatly hinders downstream analysis, making quantitative assessment of differential gene expression extremely challenging. For instance, while broadly different cell types can be classified, a majority of biologically relevant genes, including the established lineage markers of specific cellular subpopulations, are undetected. As a result, cellular subpopulations presenting more subtle differences, such as different fibroblast or macrophage subpopulations, may be impossible to differentiate [33] [35]. Even with cutting edge analysis tools such as the Seurat analysis pipeline [36], which can often identify individual subpopulations, scRNAseq gene expression data remains limited in its ability to elucidate fine-grain biological mechanisms due to its sparseness. Additionally, interrogation of individual genes of interest across cells is significantly impaired, particularly for transcription factors and signaling molecules, which do not need to be abundantly transcribed in order to fundamentally drive cell phenotype through their downstream effects on transcriptional state.

To address these limitations, we have shown that network-based analysis of protein activity, using the VIPER and metaVIPER algorithms [26] [35], can provide accurate, quantitative assessment for >6,000 proteins, including transcription factors, co-factors, chromatin remodeling enzymes, and signaling proteins. Moreover, we have shown that protein activity-based analysis can help identify rare subpopulations that are responsible for the presentation of key macroscopic phenotypes, ranging from immune evasion [37] to relapse following surgery [35]. It can also help identify master regulator proteins representing mechanistic, causal determinants of cell state and cell state transitions, such as to de-differentiation to a pluripotent stem cell state [38] or transdifferentiation between distinct tumor cell states [39]. However, these analyses can be extremely complex because they require assembly of lineage specific regulatory networks and

master regulator analyses that are challenging for biologists who are not trained in network biology.

To allow broad access to these methodologies to biologists with relatively limited network-based analyses expertise, we introduce a comprehensive pipeline for **P**rotein Activity **I**nference for **S**ingle **Cel**l **S**tudies **(PISCES)**, which is made available to the research community via a general-use R package. The pipeline automates the optimal generation of lineage specific regulatory networks, via ARACNe (Algorithm for Reconstruction of Accurate Cellular Networks) [40] [23], measurement of protein activity via VIPER (Virtual Inference of Protein Activity by Enriched Regulon Analysis) [21], as well as the identification of molecularly distinct subpopulations via a variety of clustering methodologies, and the identification of Master Regulators of cell state and cell state transitions (**Figure 2**).

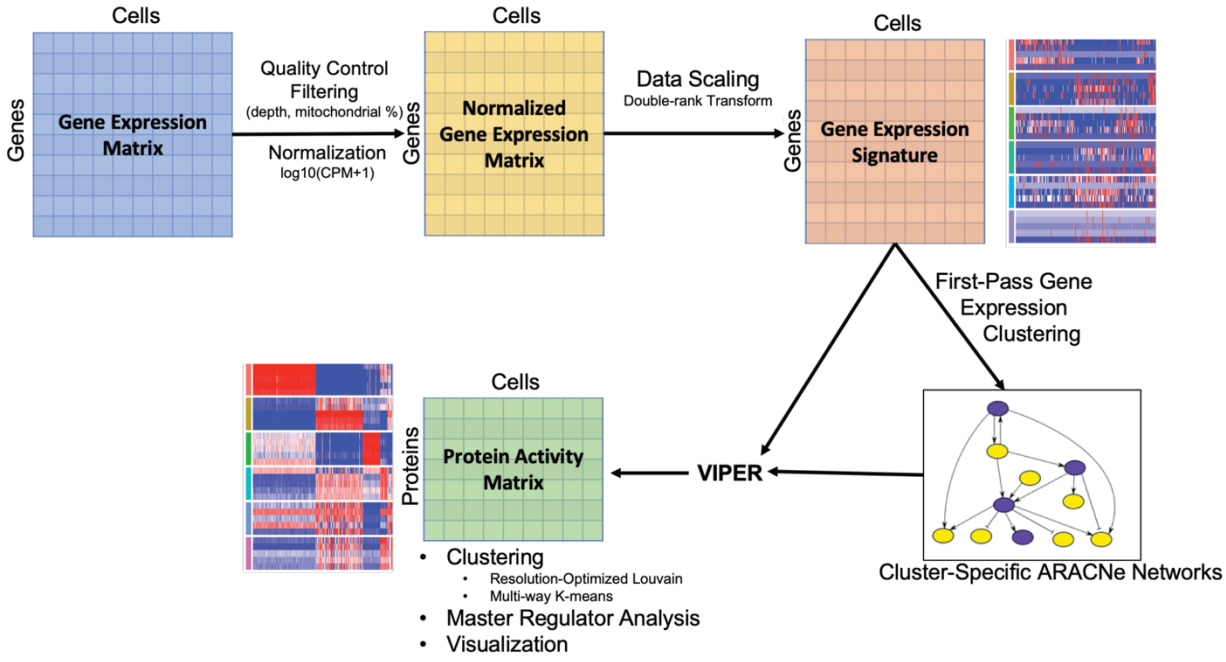


Figure 2: Single-Cell RNA-Seq Protein Activity Analysis Workflow.

Flowchart of overall analysis pipeline, showcasing sequential data transformations from original raw RNA-Seq gene expression counts matrix (blue) followed by Quality Control Filtering and Normalization (yellow) and data scaling (red), followed by cluster-specific ARACNe and final VIPER transformation to generate a single-cell VIPER-inferred protein activity matrix (green).

ARACNe is an information theoretic algorithm for the inference of the direct transcriptional targets of transcriptional regulator proteins, as well as the least indirect targets of signal transduction proteins. This allows reconstructing the tissue specific repertoire of transcriptional targets (regulon) of ~6,500 regulatory and signaling proteins, including surface markers (SMs). VIPER computes the activity of each protein based on the differential expression of the genes in its regulon, as assessed by weighted gene set enrichment analysis. Since regulons are generally large, containing up to several hundred genes, we prune them to include the same number of the most likely targets (between 50 and 100), to avoid biasing the statistical significance of the gene set enrichment analysis, as discussed in [21]. As a result, even when the specific gene encoding for a protein of interest is undetected, VIPER can still quantitatively assess its activity (**Figure 3**).

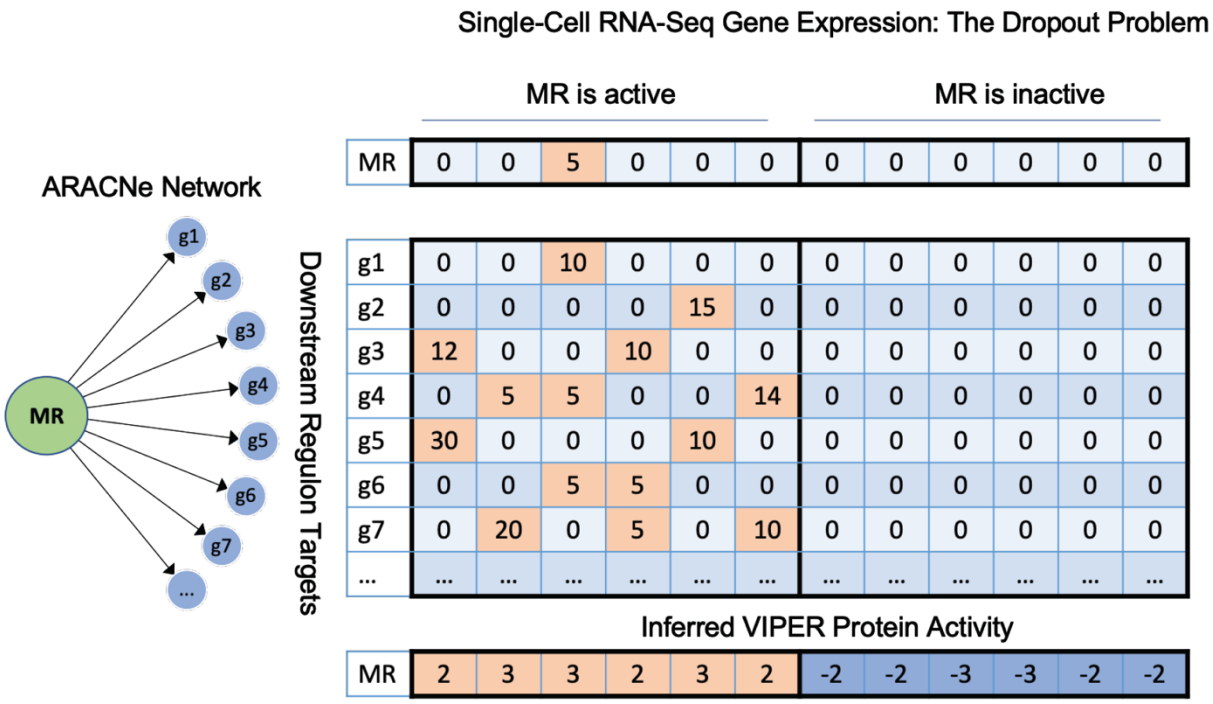


Figure 3: Single-Cell RNA-Seq Protein Activity Analysis Workflow.

Graphical of the gene expression dropout mitigation effect. A theoretical ARACNe-inferred regulon of a proteomic master regulator of cell state (MR) and its downstream transcriptional targets (g1,g2,g3,g4,….) is shown, along with a matrix showing sparseness of expression for MR and each of its targets both in cells with high real activity of MR and cells with low activity. From MR expression alone, only a single sample with high MR-activity would be correctly identified. However, by integrating the expression values from each target gene, high protein activity of MR can be correctly inferred despite the high dropout rate of any single gene target.

Previous work in the Califano lab has shown the accuracy and reproducibility of these algorithms when used to analyze bulk data. Indeed, ARACNe and VIPER have been used extensively to identify master regulators (MRs) that were experimentally validated as mechanistic determinants of diverse biological states, many of which have been extensively validated, see [41] [42] [43] [24], just to cite a few, and resulting in two CLIA-approved clinical tests to predict tumor drug sensitivity, including OncoTreat [24] and OncoTarget [44]. Most critically, when comparing

30M read RNASeq profiles to down-sampled profiles with 10K to 50K reads (similar to typical scRNASeq profiles), VIPER-measured protein activity profiles retain high Spearman correlation ($\rho \geq 0.8$), while correlation of the raw gene expression profiles is extremely poor ($\rho \leq 0.3$) [21].

To adapt these tools to the analysis of scRNASeq profiles, PISCES implements three major modifications. First, an initial gene expression-based cluster analysis is used to identify molecularly distinct cellular subpopulations representing distinct sub-lineages. Fine grain cluster analysis is not necessary as we have shown that regulatory networks for closely lineage-related cells are virtually indistinguishable [45]. ARACNe is then used to generate distinct regulatory networks for each cluster containing at least N = 500 cells. Second, to increase regulon coverage, cells within each sub-lineage-related cluster are combined into "meta-cells" using a K-nearest-neighbors graph analysis. This creates pseudo-bulk samples that can then be analyzed by ARACNe, producing networks with more accurate edges, larger regulons, and greater coverage of regulatory proteins. Finally, rather than using VIPER for protein activity measurement, we use its derivative metaVIPER [26], which is designed to optimally integrate protein activity inferences from multiple networks. This allows for the use of multiple single-cell and, when available, lineage-matched bulk-tissue-derived networks. Downstream of the ARACNe and metaVIPER analyses, PISCES provides access to a variety of novel protein-activity based clustering and data visualization algorithms, in addition to implementing interoperability with the popular Seurat single-cell data format.

In order to establish the efficacy of these tools and optimal parameters for future benchmarking and improvement, we have performed both technical and biological validation experiments, first

by evaluating reproducibility of protein activity assessment from progressively down sampled data, and then by assessing concordance of gene expression and protein activity to antibody-based measurements using multiplexed FACS (Cytek) and CITE-Seq [46].

Taken together, the results of these benchmarks show that the PISCES analytical pipeline dramatically outperforms gene expression-based analyses and even outperforms experimental assessment via selected antibodies, while allowing essentially proteome-wide activity quantitation. As such, these data suggest that PISCES provides a valuable and highly flexible tool for the analysis of scRNA-Seq datasets, which greatly improves the granularity of cell subpopulation detection, allowing detection of rare yet biologically relevant subpopulations that would be missed by gene expression analysis, due to gene dropout issues, and supports accurate assessment of Master Regulators of single-cell states.

1.3 Results

Analytical Pipeline Overview: The PISCES pipeline takes a single-cell Unique Molecular Identifier (UMI) count matrix as input, with genes organized by row and cells by column. Initial Quality Control filtering is adjustable, with user-defined parameters. By default, it will remove cells with fewer than 1,000 UMIs or more than 25% mitochondrial gene UMIs. The gene expression matrix is then normalized and scaled to generate a matrix of gene expression signatures. By default, this is accomplished by converting counts to $Log_{10}(CPM + 1)$, where CPM indicates counts per million. However, it can also be implemented via the Seurat SCTransform algorithm [20] or any other third-party methods of choice.

Following normalization, a first-pass clustering is performed on scaled gene expression using one of several clustering approaches implemented in the pipeline, including partition around

medioids (PAM) [47]or Louvain clustering with resolution-optimization [35]. For each gene expression cluster with ≥ 500 cells, by default, metaCells are computed by first selecting 250 unique cells at random and then transforming their scRNASeq profile into a metaCell by adding the UMI counts from the *k* nearest neighbors (*k* = 5 by default). Independent ARACNe networks are generated from each cluster using the $Log_{10}(CPM + 1)$ values of each metaCell.

In parallel, the normalized gene expression profile is transformed into a gene expression signature (GES). This can be done in a number of ways, either with an internal normalization against mean and standard deviation of all cells to query differences within the dataset or with an external reference to answer experiment-specific questions (i.e. the differences between cancerous and healthy cells). By default, PISCES will perform a standard internal normalization to generate the gene expression signature, which is then transformed into a matrix of protein activity using MetaVIPER. MetaVIPER takes as input the GES and the previously generated cluster-specific networks and identifies the best network matches to each sample by maximum regulon consensus. Enrichment scores from each matched network are then integrated using a weighted average to produce a final enrichment value that can then be used for downstream visualization and analysis. The entire pipeline is visualized in Figure 2.

Since every scRNAseq experiment is unique—depending on the specific cell types, the quality of the data, or the overarching question driving the research—PISCES allows users to fine tune the pipeline to match their specific requirements. For instance, since Seurat represents a widely used platform for scRNAseq analysis at the gene expression level, the Seurat batch-correction and SCTransform data scaling approach are incorporated as optional pre-processing steps to generate

gene expression signatures before they are analyzed by PISCES. These may, however, be substituted by any user defined normalization and data scaling routine, such that effect of alternative normalization or pre-processing methods may be tested using PISCES's default technical and biological benchmarks. Output from the PISCES pipeline is converted to a Seurat object for convenient export into a variety of external visualization or processing tools and analyzed by other commonly used tools. In particular, cell type annotation is implemented in PISCES at the single-cell level using SingleR [48], which infers cell types represented in the dataset by correlation of gene expression to expression of sorted bulk-RNASeq reference datasets and stores these labels as metadata for downstream analysis.

Technical Validation Shows Improved Recovery of Data Structure from Low-Depth Cells: To benchmark PISCES reproducibility relative to gene expression and to establish an optimal UMI depth for user-driven adjustment of metacell parameters, we executed the entire pipeline using progressively down-sampled profiles from relatively high-depth scRNAseq data. For this purpose, we used the SNU-16 cell line, a relatively homogenous stomach adenocarcinoma model that is transcriptionally complex and produces high UMI counts per cell (i.e., 40,000-50,000), on the high end of the typical yield for cell lines and significantly above the yield produced by clinical samples. Average UMI count in our dataset was 41,915 across 6157 single cells. To create synthetic data with lower depth, we down-sampled this data by first drawing each cell's total UMI-count from a multinomial distribution with mean target depth manually specified and a uniform probability weight over all cells, then drawing the gene-specific counts from a second multinomial with probabilities given by the proportions of genes in the original, full depth profile for each cell. This procedure was applied with target depths between one and ten thousand UMIs

at a step-size of 1,000 and between 10,000 and 40,000 UMIs at a step-size of 5,000. We then generated meta-cells using a consistent sub-set of 500 cells for each down-sampled matrix with depth of 10,000 UMIs or fewer. These data were used to generate 27 ARACNe networks in total; one for the full data, 16 from each of the down-sampled gene expression profiles, and 10 from each of the meta-cell matrices.

To generate gene expression signatures, we normalized each down-sampled matrix against the Cancer Cell Line Encyclopedia (CCLE) from The Broad. Because this data is from bulk-sequencing, we first had to apply the previously described down-sampling scheme in order to generate depth-matched reference samples for each single-cell matrix. Gene expression profiles were then normalized gene-by-gene by subtracting the mean expression from CCLE, then dividing by the standard deviation of the expression in CCLE. Finally, we generated VIPER matrices for all pairwise combinations of GES and regulatory networks, culminating in 459 VIPER matrices. A flowchart illustrating this experimental design is shown in Figure 4.

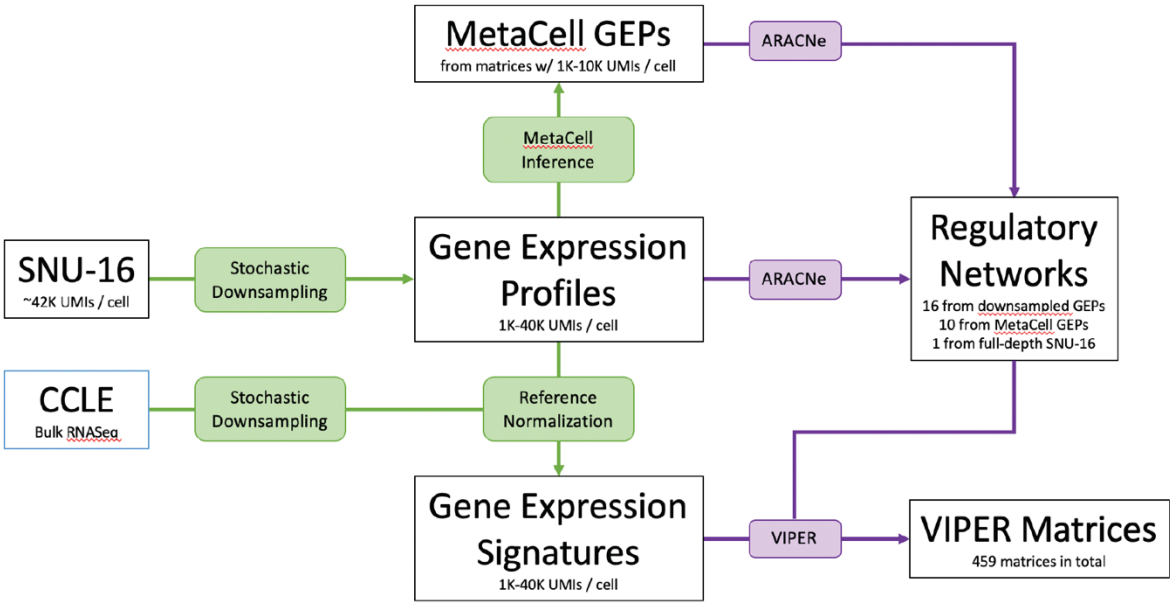


Figure 4: Flowchart of Technical Validation Down-Sampling Approach

To assess the reproducibility of gene expression and protein activity signatures at different depths, we computed the cell-by-cell Pearson correlation between each down-sampled matrix and the full depth data. In each cell, we subset the comparison to those genes or proteins with significantly different expression or activity (p-value < 0.05 with Bonferroni correction) in the full-depth data, then computed the correlation coefficients cell-by-cell between full-depth and down-sampled data using this subset. This reduction was performed in order to avoid inflation of correlation values based on non-significant data. In protein activity signatures generated fully from down-sampled data (down-sampled GEP as input to ARACNe, down-sampled GES as input to VIPER), we observe a statistically significant improvement in correlation to full-depth data relative to gene expression signature at all depths above 5,000 UMIs (Figure 5A; p-value < 0.05 by Wilcoxon signed rank test). Strikingly, when an ARACNe network generated from full-

depth GEP is applied to down-sampled GES as input to VIPER, correlation to original full—depth VIPER signature is strongly conserved even at extremely low UMI counts, remaining above 0.75 on average at UMI depth of 1000, where average correlation of gene expression signature to full-depth data is below 0.1. This emphasizes the importance of constructing a high-quality ARACNe network in the VIPER inference pipeline, such that applying high-quality networks inferred for a given cell type from one dataset to a matched cell type in lower-quality data is likely to provide a significant boost to the power of protein activity inference even from very-low-depth data. Additionally, we find a significant improvement in correlation values when constructing metaCell-based ARACNe networks from lower-depth data (Figure 5B), such that metaCell networks applied to run VIPER on GES matrices with mean UMI count of 3000 approach the inference quality seen when running ARACNe and VIPER on gene expression matrices with a mean UMI count of 20,000. However, at the very low mean depth of 1000 UMI/cell this breaks down, and metaCell ARACNe network inference no longer offers any statistically significant improvement over inference on low-depth data. Therefore, we strongly recommend applying the metaCell ARACNe network inference option in PISCES for any datasets with data quality between 1000 and 5000 mean UMIs/cell, which is common in clinical datasets.

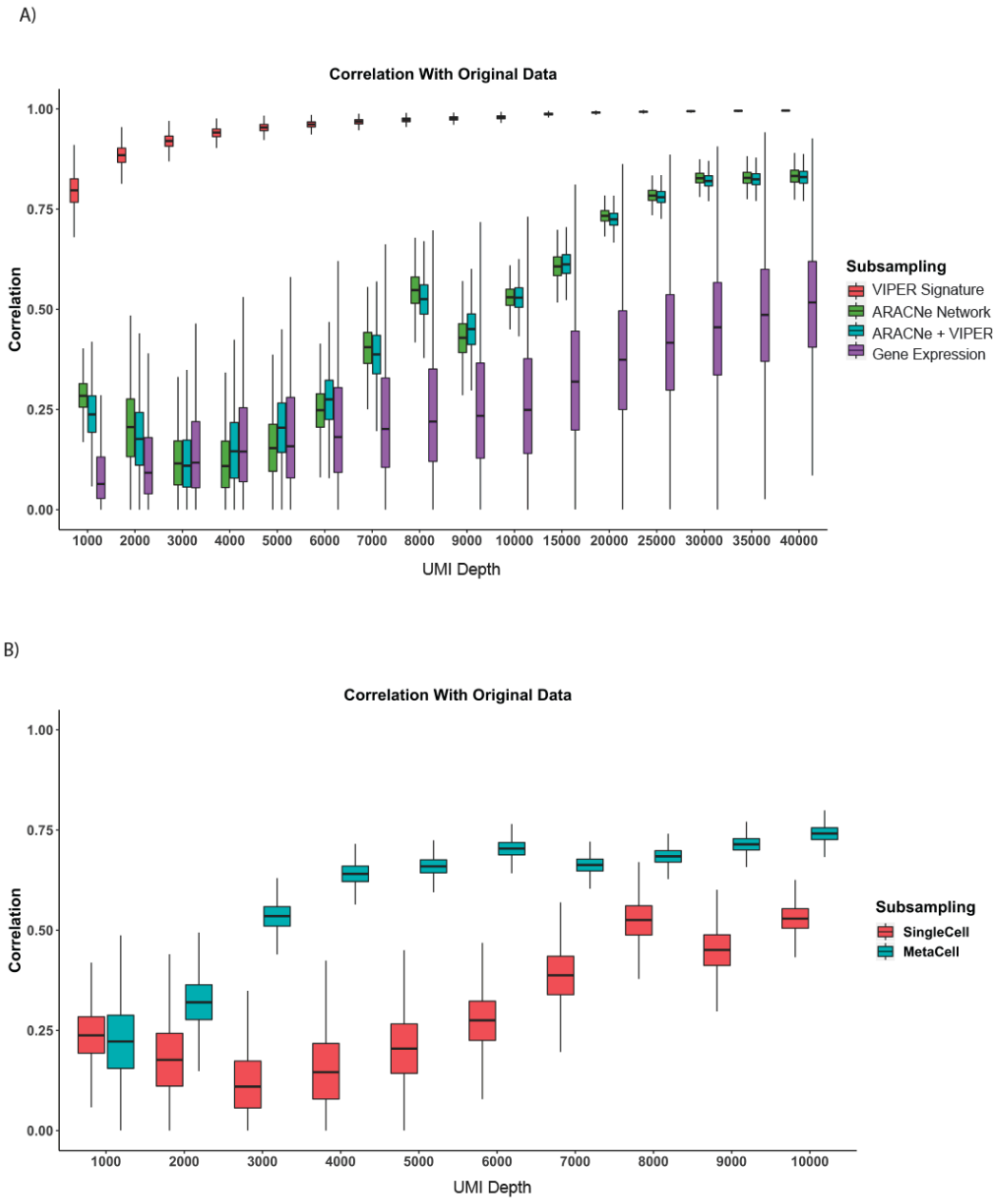


Figure 5: Technical Benchmarking Shows Increased Recovery of Original Data Structure from Down-sampled Matrices by VIPER vs Gene Expression

A) Boxplot showing distribution across single cells of Pearson correlation between sub-sampled and original full-depth cells. Along the x-axis is the UMI/cell downsampling quotient. In purple, correlation between downsampled and original gene expression is shown to rapidly degrade, to a median consistently below 0.5, and below 0.25 even by the relatively high depth of 10,000 UMI/cell. In red, correlation is shown between VIPER inference on down-sampled gene

expression signature with full-depth ARACNe network vs VIPER inference on full-depth gene expression signature using full-depth ARACNe network, such that correlation remains high even at extremely low sample depth, with a median above 0.75 even at 1000 UMI/cell. In green, correlation is shown between VIPER inference on full-depth gene expression signature using ARACNe networks derived from full-depth vs down-sampled data, and in blue correlation is shown between full-depth VIPER inference using full-depth ARACNe networks and VIPER inference on down-sampled gene expression signature using down-sampled ARACNe network. In all cases protein activity improves on gene expression, and down-sampling of both VIPER and ARACNe simultaneously still improves correlation relative to gene expression down to a depth of 5000 UMI/cell, with Bonferroni-corrected p-values by paired Wilcox test < 0.05. **B)** For UMI depths ranging from 1000 to 10000, correlation between full-depth VIPER matrix using full-depth ARACNe network and VIPER matrices computed on on down-sampled gene expression signatures with either full-depth or metaCell ARACNe. metaCell ARACNe significantly improves on correlation with full-depth data for all depths >1000 UMI/cell, by paired Wilcox test p-values < 0.05. Mean correlation at low-depth with metaCell ARACNe network approaches 0.75, seen only at UMI depths >20000 without applying the metaCell ARACNe inference approach.

Overall, these data show that the correlation between full-depth and down-sampled gene expression signatures is poor even at relatively high depth, and decays rapidly to a median value of less than 0.25 even at depths of 10,000 UMIs/cell (purple bars, Figure 5A). Protein activity, by comparison, is much more robust, significantly outperforming gene expression at all depths above 5,000 UMIs/cell. Interestingly, down-sampling only the gene expression signature input to VIPER while retaining a full-depth ARACNe network had little effect (red bars, Figure 5A) on protein activities robustness, while down-sampling either the data using to generate ARACNe networks or both ARACNe data and gene expression signature (green and blue bars respectively) had a much more significant effect on correlation to original full-depth VIPER matrix, which was partially rescued by metaCell ARACNe. The full heatmap showing mean correlation across

cells comparing all VIPER matrices against full depth data is available in the Figure 6. These findings indicate that the quality of the ARACNe networks is the driving force behind protein activity signatures' ability to retain signal at low UMI depths and supports the idea of using metacells to rescue signal within the ARACNe network or use context-appropriate bulk networks where available.

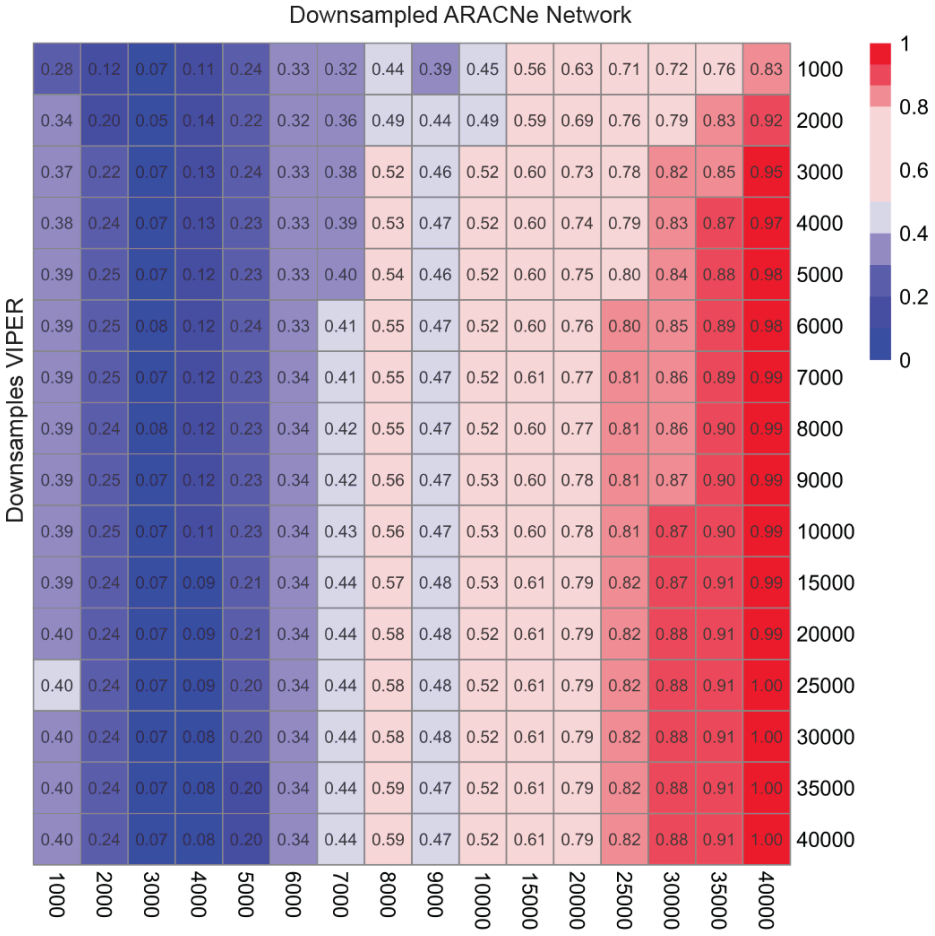


Figure 6: Pairwise Down-sampling Correlation Matrix

Heatmap of mean correlation values compared to original full-depth VIPER matrix with full-depth ARACNe network for each combination of down-sampled ARACNe and VIPER gene expression signature depth. Each row corresponds to depth of gene expression signature input to VIPER, and each column corresponds to depth of gene expression input to ARACNe. Correlation is subset to proteins differentially up-regulated or down-regulated ($p<0.05$) within

original full-depth VIPER matrix, on a cell-by-cell basis, and mean correlation across all cells is plotted for each box on the heatmap corresponding to a particular down-sampling approach.

Biological Validation Shows Improved Concordance with Antibody Profiling: To assess whether protein activities measured by PISCES effectively track with direct assessment of protein abundance in single cells, thus providing improved mechanistic understanding of single cell processes, we compared PISCES-measured protein activity to CITE-Seq single-cell measurements of protein abundance in a publicly available dataset of cord blood-derived mononuclear cells (CBMCs) [46].

Single cell clustering based on CITE-Seq measurements, using a pre-selected antibody panel, yields six major cell type clusters, including CD4 T-cells, CD8 T-cells, Monocytes, NK Cells, B-cells, and Hematopoietic Stem Cells (HSCs) (**Figure 7C**). In sharp contrast, gene expression-based clustering by Seurat identified only four distinct cell clusters, with NK cells and HSCs subsumed into the other major cell types. Protein activity-based clustering by PISCES not only recapitulated all six clusters identified by antibody measurement but also identified many additional proteins representing established lineage markers of these sub-populations, which were completely missed by gene expression analysis. Indeed, the most differentially active proteins in each cluster present a highly cluster-specific activity pattern not visible by gene expression alone.

Furthermore, when gene expression-based clustering was limited only to the genes encoding for the proteins in the CITE-Seq panel, the single-cell RNA-Seq dropout problem was so severe that cluster structure was completely lost (**Figure 7D**). This suggests that critical proteins, whose role

in the biology of these populations is extremely well established, are completely missed in terms of their gene expression. In sharp contrast, PISCES analysis fully recapitulated the experimentally assessed cluster structure when the analysis was limited to the proteins represented on the CITE-Seq panel (**Figure 7E**).

Critically, the coefficients of variation (i.e., $COV = \sigma/\mu$), as computed for gene-expression, antibody-measured protein abundance, and VIPER-measured protein activity, shows that VIPER-measured activity dramatically outperforms gene expression (p=0.0004 by paired t-test across the entire panel) and even antibody measurements for most proteins (p = 0.0083 across the entire panel), indicating a significant improvement in reproducibility and signal-to-noise ratio (**Figure 7A**). Finally, we assessed correlation between either gene expression or VIPER-measured protein activity against protein abundance as assessed by CITE-Seq. Across the board VIPER significantly outperformed gene expression (**Figure 7B**), with strong visual cluster-separation even on single genes (**Figure 7F**), and pairwise plots of VIPER activity vs paired CITE-Seq antibody staining resembling flow cytometry plots (**Figure 8**).

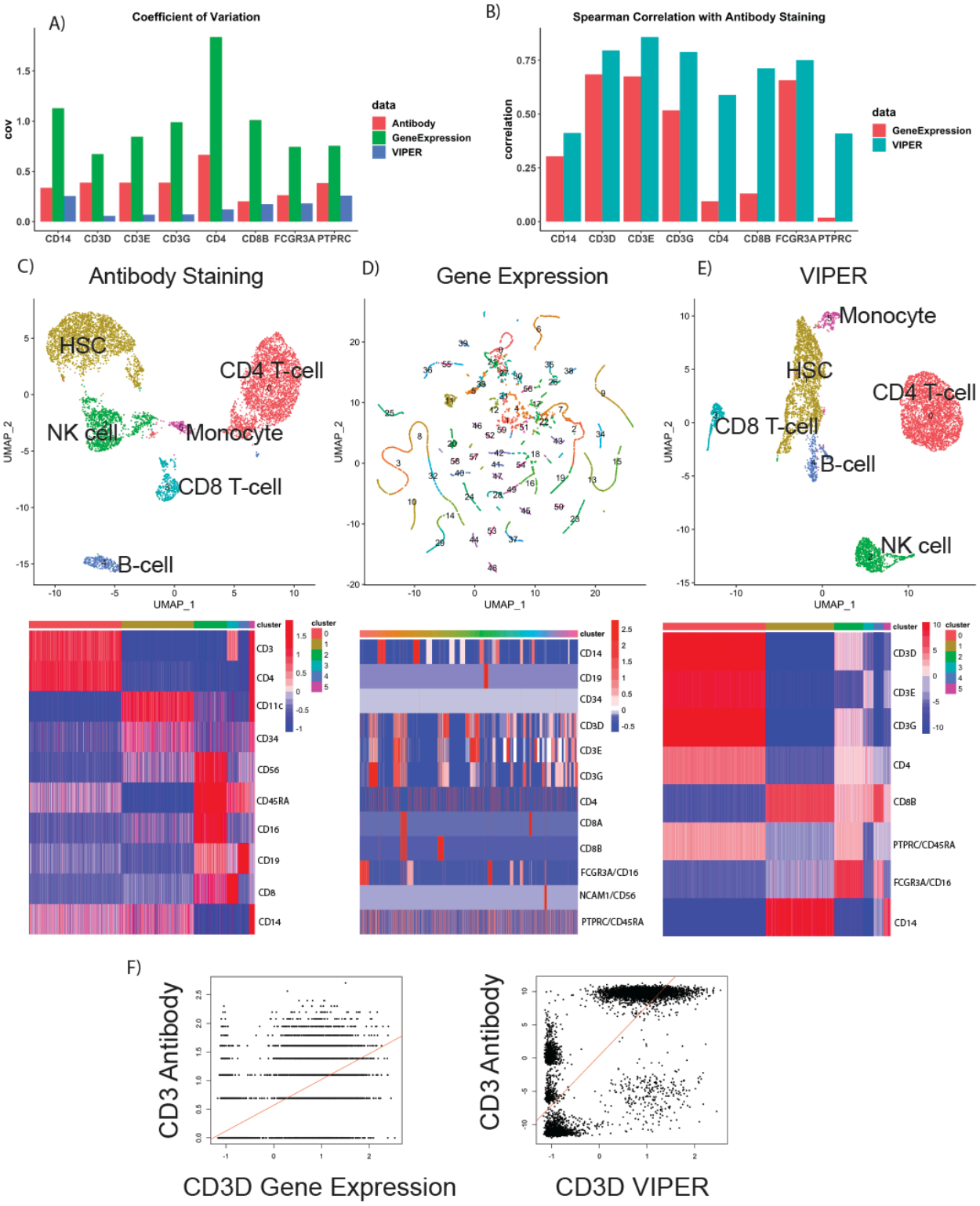


Figure 7: Biological Benchmarking Shows Dramatically Increased Concordance with CITE-Seq Antibody Profiling by VIPER vs Gene Expression

A) Coefficient of Variation (computed as σ/μ) for each gene profiled by the CITE-Seq antibody panel, shown for antibody staining (red), Gene Expression (green), and VIPER-inferred protein activity (blue), with higher Coefficient indicating lower signal-to-noise ratio. **B)** Spearman Correlation between Gene Expression vs Antibody (red) and VIPER vs Antibody (blue) computed across cells for each gene profiled by the CITE-Seq antibody panel. **C)** UMAP projection and clustering of CITE-Seq antibody staining panel, labelled with cell types inferred from SingleR and validated by staining for known markers. Row-scaled heatmap is shown below with antibody staining intensity grouped by cluster. **D)** UMAP projection and clustering of Gene Expression for the subset of genes concurrently profiled by CITE-Seq antibody staining panel. Row-scaled heatmap is shown below, with excessive noise for meaningful clustering due to single-cell RNA-Seq dropout effect. **E)** UMAP projection and clustering of VIPER protein activity, labelled with cell types as in 3C. Row-scaled heatmap is shown below with VIPER activity grouped by cluster, for the subset of genes concurrently profiled by CITE-Seq antibody staining panel with activity inferred by VIPER. **F)** Representative Correlation plots of Gene Expression vs Antibody and VIPER vs Antibody, showing greater concordance of CD3D VIPER activity with Antibody intensity, relative to CD3D Gene Expression.

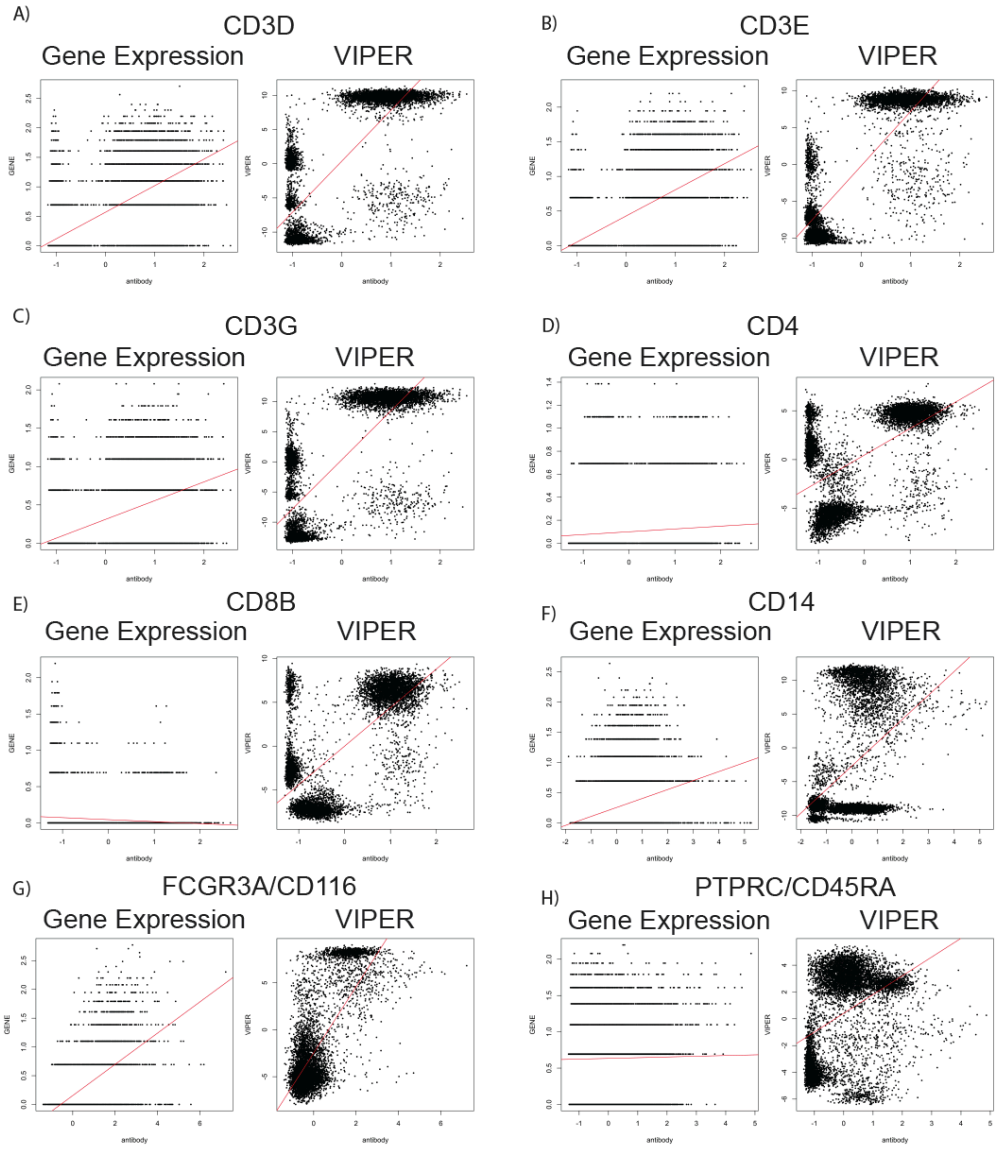


Figure 8: Pairwise CITE-Seq Antibody vs VIPER Correlation Plots

A) Correlation Plots of CD3D Gene Expression vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right). **B)** Correlation Plots of CD3E Gene Expression vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right). **C)** Correlation Plots of CD3G Gene Expression vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right). **D)** Correlation Plots of CD4 Gene Expression vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right). **E)** Correlation Plots of CD8B Gene Expression vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right). **F)** Correlation Plots of CD14 Gene Expression vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right). **G)** Correlation Plots of FCGR3A (CD16) Gene Expression

vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right). **H)** Correlation Plots of PTPRC (CD45) Gene Expression vs Antibody Intensity (left) and VIPER vs Antibody Intensity (right).

Furthermore, we would like to point out that protein abundance, as assessed by antibodies, is a poor proxy for protein activity. This is because, even after a protein is expressed, its activity is manifested only when it is effectively post-translationally modified, it is translocated into the appropriate sub-cellular compartment, and it has formed complexes with critical cognate binding partners. By measuring activity via expression of highly multiplexed gene reporter assay, VIPER can effectively report on the activity of proteins, which has been so far elusive, especially in single cells. In a separate analysis of CD45+ cells that were isolated from renal clear cell carcinoma, then split and profiled at the single cell level using both scRNA-Seq and a CyTEK high-throughput flow cytometry panel of 19 lymphoid and 19 myeloid antibodies [35], the de-noising effect of PISCES was even more obvious. Not only did these results completely recapitulate the results obtained for the CITE-Seq comparison, but, given the larger number of experimentally assessed proteins, they provide further evidence of the dramatic improvement offered by PISCES analysis over both gene expression and antibody-measured protein abundance. This is reflected in three key findings. First, experimentally assessed protein abundance (e.g., using the 19 lymphoid markers) was unable to identify the clusters that could be identified by VIPER-based measurement of the same 19 proteins, including splitting of the myeloid cluster into monocytes and macrophages, the CD8 T cell cluster into CD8 T cells and NK cells, and the CD4 T cell cluster into CD4 T cells and Regulatory T cells (Figure 9).

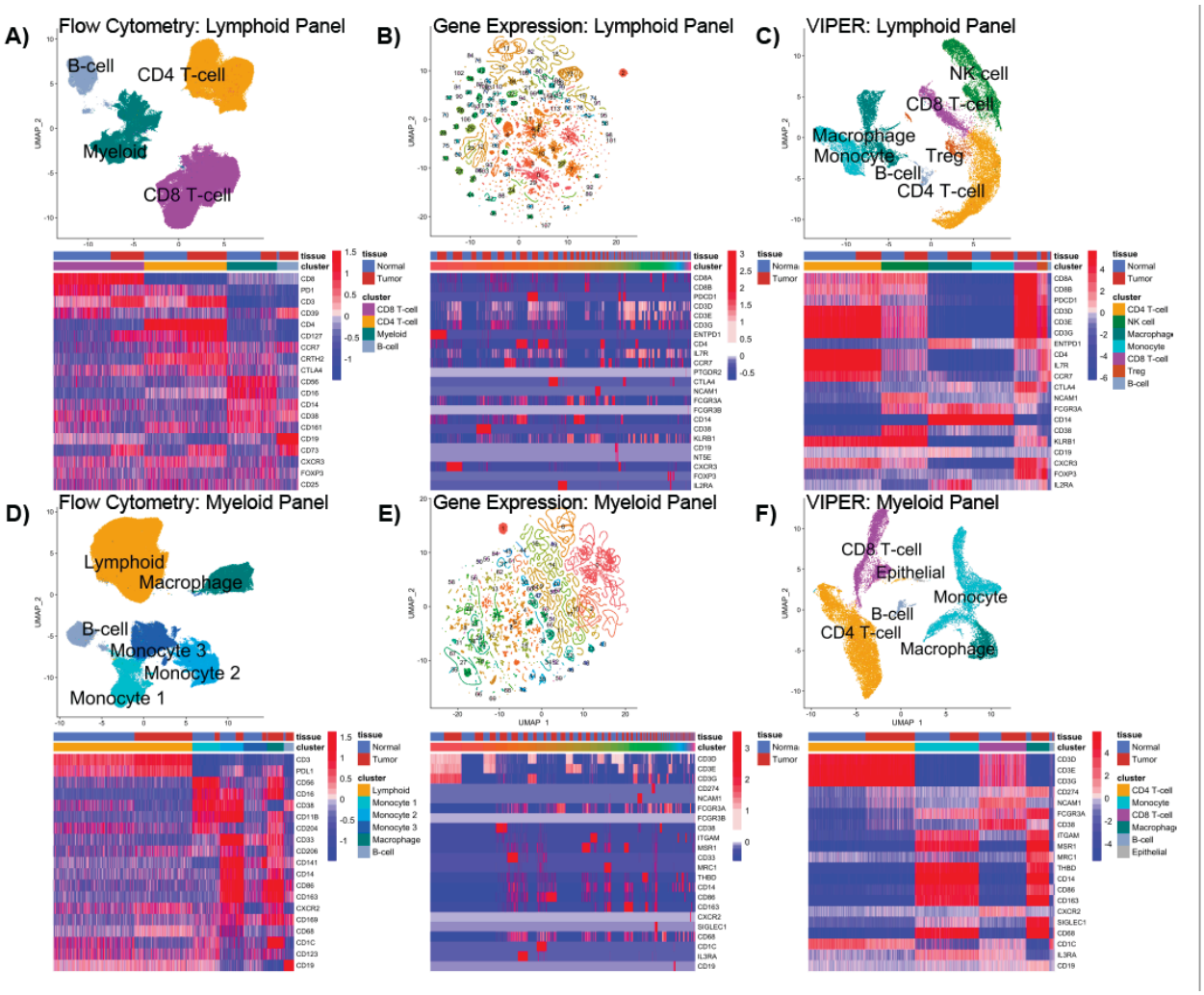


Figure 9: Comparison of VIPER Inferences and Gene Expression to Flow Cytometry in Renal Clear Cell Carcinoma Dataset

A) UMAP projection, clustering, and heatmap by flow cytometry proteins profiled in CyTEK Lymphoid Panel. **B)** UMAP and clustering by scRNASeq gene expression subset to the proteins profiled in A, showing noise-induced decrease in clustering resolution. **C**) UMAP and clustering by VIPER-inferred protein activity using PISCES, subset to the proteins profiled in A. **D**) UMAP and clustering by flow cytometry proteins profiled in CyTEK myeloid panel. **E)** UMAP and clustering by scRNA-Seq gene expression, subset to the proteins profiled in D. **F)** UMAP and clustering by VIPER-inferred protein activity using PISCES, subset to the proteins profiled in D. partially reproduced with permissions from Obradovic et al., 2021.

Second, proteins not expressed on the surface of the cell, such as FOXP3, a canonical marker of regulatory T cells, could not be reliably detected by antibody measurements but were clearly detected in the correct sub-population by VIPER. Indeed, taken together, only 4 of 38 proteins assessed by VIPER and antibody measurement were not effectively and correctly detected by VIPER in the specific cellular sub-populations for which they represent an established lineage marker (NT5E/CD73, FCGR3B/CD16b, PTGDR2/CD294, CD33). In contrast, 9 of 38 proteins could not be consistently detected by antibody measurement or were not restricted to the associated sub-populations due to noisy background staining (CD14, CD127, FOXP3, CD38, CD25, CXCR3, CD161, CTLA4, CD39). Indeed, clustering on the full set of proteins identified by PISCES on this dataset [35] led to identification of rare cellular subpopulations that play a critical role in post-surgical tumor recurrence, and for which PISCES-inferred markers were validated by immunohistochemistry.

This indicates amplification of biologically meaningful rather than artifactual signal from single cells by PISCES, and its ability to enable interrogation of individual genes of interest without data dropout. In fact, while CITE-Seq is limited by time-consuming antibody titration and panel optimization, ultimately profiling relatively few proteins in most experiments, PISCES typically captures several orders of magnitude more unique proteins, enabling interrogation of intracellular proteins which would otherwise be difficult to stain for without losing cellular RNA, as well as select surface markers of interest. Nevertheless, the cell-matched profiling of both gene expression and protein abundance by CITE-Seq enables direct comparison of PISCES inferences to measured protein abundance for a subset of proteins within the same cells, which may be used

as a benchmark of the high concordance between PISCES and measured protein abundance, and the degree to which PISCES improves signal-to-noise with respect to antibody-based measurements.

1.4 Discussion

The PISCES package for analysis of single-cell RNA-Sequencing data represents a comprehensive and highly generalizable pipeline for inference of protein activity to maximize utility of single-cell datasets. We have demonstrated its ability to mitigate the single-cell RNA-Seq data dropout problem and recapitulate high-depth data structure even from low UMI counts. We have also demonstrated its ability to recapitulate biological structure from CITE-Seq antibody-based protein profiling with much better gene-by-gene signal than gene expression. These technical and biological validations also serve as benchmarks for further refinement of the pipeline by which any changes can be comprehensively assessed.

For biological validation benchmarking, protein selection was based on pre-defined protein panels from CITE-Seq experiments. As a result, this represents a completely unbiased set of proteins that was not selected to skew performance in VIPER's favor. While we limited the comparison only to the CITE-Seq panel of proteins, PISCES produced activity profiles for 6,500 proteins. Thus, if these results are further confirmed in follow-up studies, PISCES would provide the equivalent of a single cell FACS with 6,500 antibodies, remedying the need to select and validate antibodies for specific cellular populations. Indeed, VIPER was originally developed for the analysis of proteins that directly control gene expression on the chromatin (i.e., TFs and co-TFs). As a result, accuracy and reproducibility of VIPER-based measurement of surface markers

is likely to be significantly outperformed for TFs and co-TFs, which represent the most critical class of lineage markers.

In addition to the technical benchmarking of correlation between down-sampled and full-depth data, the extent of improvement by PISCES in coefficient of variation, number of genes recovered, and gene-by-gene correlation to matched antibody profiling represent a critical biological benchmark for alternative workflows by PISCES users as new pre-processing methods are incorporated and existing algorithms are refined. The pipeline has been consciously designed to be highly modular, with customizable workflows and parameter optimization enabled by separate pre-processing, meta-cell, and clustering steps and interoperability with the popular Seurat workflow. We recommend targeting a median UMI depth / cell of no less than 5000, with the crucial step being inference of ARACNe network from high-depth data, applying the metaCell algorithm to improve sample depth for ARACNe network inference. Wherever a high-depth-derived ARACNe net is available, inference fidelity is high even on extremely low-depth datasets, so the increased availability of single-cell RNA-Seq datasets across a broad range of tissue contexts will continually allow construction of an expanding library of ARACNe networks which can be broadly applied to new data.

PISCES is chiefly limited by the fraction of 6,500+ total proteins recoverable at low UMI depth, although the number of proteins recovered nearly always compares favorably to CITE-Seq, which requires time-consuming antibody titration and is limited to predefined cell surface proteins, whereas PISCES captures proteins with the strongest signal-to-noise from the data and can infer both cell surface and intracellular protein activity. Applying metaCell ARACNe

network inference addresses this to some degree, such that nearly 100% of all proteins recoverable at full depth in SNU-16 cell line sequencing data were recovered at a UMI depth of 10,000, where only half of the proteins inferred at full-depth were recoverable without metaCell, and over half of proteins remained recoverable with metaCell even at critically low UMI depth of 1,000 (Figure 10). Future iterations of the pipeline will continue to improve on the fraction of recoverable proteins by integrating and testing novel pre-processing procedures and optimization of the ARACNe and VIPER inference steps. The development version of the PISCES R package will be continually available at https://github.com/califano-lab/PISCES.

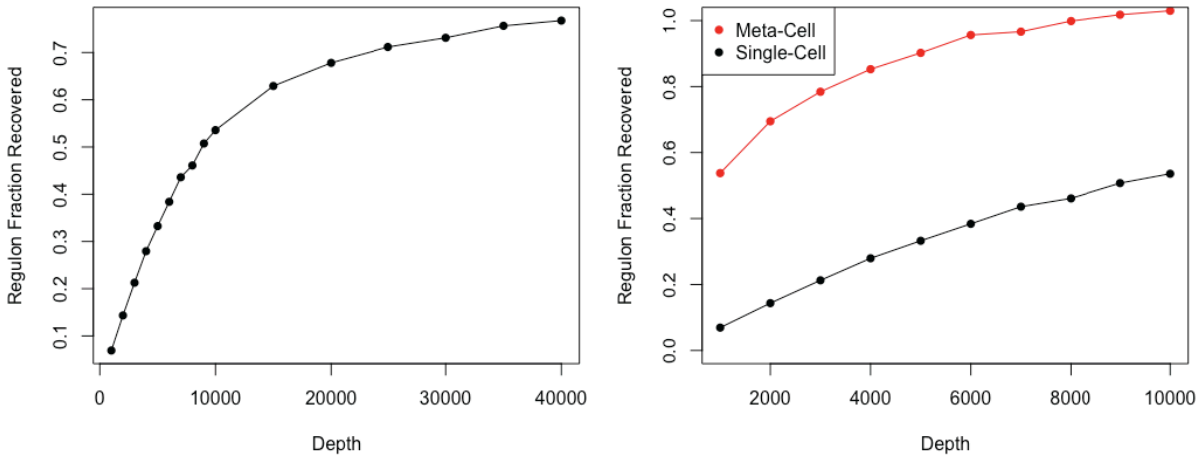


Figure 10: ARACNe Network Size at Variable Depth with and without Meta-Cell

On the left, fraction of Total ARACNe network regulons (y-axis) recovered at each down-sampling depth (x-axis) relative to full-depth data, such that fraction decreases log-linearly with down-sampling depth. On the right, fraction of Total ARACNe network regulons relative to full-depth data (y-axis) recovered at each down-sampling depth from 1000 to 10000 UMI/cell, with metaCell approach (red) or without metaCell approach (black).

1.5 Methods

Quality Control, Normalization, and Scaling: As a pre-processing step, low quality cells and genes lacking enough data to be useful are removed from the analysis. Cell quality is determined by two primary factors – read depth and mitochondrial gene percentage. Samples with too many or too few reads are likely sequencing errors (doublets or empty droplets), while a high mitochondrial gene percentage is indicative of cell stress or damage. This latter group of cells will typically have a biased transcriptome not representative of the actual cell state. For most data sets, PISCES will simply remove genes with no reads at all. For larger data sets, genes that appear in less than 1% of the total cells will be removed in order to optimize computational complexity. Cells with fewer than 1000 total UMIs or mitochondrial transcript fraction greater than 25% are also removed in quality-control filtering. Filtered data are then normalized to log10(counts per million + 1). A gene expression signature is then generated from the normalized data using either double rank transformation or Seurat SCTransform scaling function.

Seurat Pre-Processing Workflow: Gene Expression UMI count matrices for each sample are processed in R using the Seurat SCTransform command to perform a regularized negative binomial regression based on the 3000 most variable genes. For datasets combining samples across multiple patients, normalized datasets may be integrated using the FindIntegrationAnchors and IntegrateData functions in Seurat. The resulting data are projected into their first 50 principal components, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Differential Gene Expression between clusters is computed by the MAST hurdle model for single-cell gene expression modeling, as implemented in the Seurat

FindAllMarkers command, with log fold change threshold of 0.5 and minimum fractional expression threshold of 0.25, indicating that the resulting gene markers for each cluster are restricted to those with log fold change greater than 0 and non-zero expression in at least 25% of the cells in the cluster.

Initial Clustering and MetaCells: In order to generate accurate, robust networks, ARACNe requires data from a population that shares the majority of its transcriptional architecture. In the context of single cells, this requires separating the data into coarse cell type clusters before network generation. These clusters can be generated in a number of ways; any of the popular gene expression methods for clustering will work, as will a simple clustering based on the first 30 principle components in gene expression space. We have implemented clustering on gene expression signature by Partition Around Medioids (PAM), Multi-Way K-Means, and Louvain with Resolution Optimization. Once the data have been clustered, meta-cells can be generated for input to ARACNe. By pooling cells that are close together in either gene expression or VIPER space within a cluster, the number of interactions inferred using ARACNe can be increased. PISCES uses a simple K-nearest-neighbors approach to pool cells, then sums reads across neighbors and re-normalizing. This data then serves as the input to ARACNe.

ARACNe Network Generation: A full guide for utilizing ARACNe is available on the Califano Lab Github at https://github.com/califano-lab/PISCES. For each gene expression cluster, 250 metaCells are sampled to compute a regulatory network. All networks are reverse engineered by the ARACNe algorithm, run with 100 bootstrap iterations using 1785 transcription factors (genes annotated in gene ontology molecular function database as GO:0003700, “transcription factor

activity", or as GO:0003677, "DNA binding" and GO:0030528, "transcription regulator activity", or as GO:0003677 and GO:0045449, "regulation of transcription"), 668 transcriptional cofactors (a manually curated list, not overlapping with the transcription factor list, built upon genes annotated as GO:0003712, "transcription cofactor activity", or GO:0030528 or GO:0045449), 3455 signaling pathway related genes (annotated in GO biological process database as GO:0007165, "signal transduction" and in GO cellular component database as GO:0005622, "intracellular" or GO:0005886, "plasma membrane"), and 3620 surface markers (annotated as GO:0005886 or as GO:0009986, "cell surface"). Each regulator set is run separately, as different types of proteins will have different mutual information thresholds. Once a set of regulons has been inferred for each group of regulators, the results are combined into a single network. ARACNe is only run on these gene sets so as to limit protein activity inference to proteins with biologically meaningful downstream regulatory targets, and we do not apply ARACNe to infer regulatory networks for proteins with no known signaling or transcriptional activity, for which protein activity may be difficult to biologically interpret. Parameters are set to zero DPI (Data Processing Inequality) tolerance and MI (Mutual Information) p-value threshold of 10^{-8}, computed by permuting the original dataset as a null model. Each gene list used to run ARACNe is available on github.

VIPER Analysis and Re-clustering: Once cluster-specific networks have been generated, they will serve as the input to a final VIPER run. More accurate networks will naturally lead to more accurate inferences of protein activity, which in turn allows for more robust downstream analyses. Bulk networks can also be incorporated to fill in any gaps present in the single-cell networks, as ARACNe will typically be unable to infer regulons for some proteins even with the

implementation of MetaCells. These protein activities inferred from bulk should be considered less accurate, but they can be used to follow-up on previously known proteins of interest, for instance. Once a final VIPER matrix has been inferred, the data can be re-clustered. VIPER-space will typically allow for the parsing of smaller, more transcriptionally distinct populations. These classifications can then be used for a master regulator analysis that identifies the driving regulators of the differential cell state. This can be done in several ways, with a Bootstrapped Mann Whitney-U test being the most robust. Cluster-specific Stouffer integration or a data-wide ANOVA or Kruskal-Wallis test are also viable alternatives and implemented within PISCES.

Weighted VIPER: Previously, MetaVIPER was developed as an initial adaptation of VIPER to single-cell data. By using multiple networks, MetaVIPER sought to accurately recapitulate protein activity in populations for which no context-specific network was available. To briefly explain this method, protein activity would be inferred from a given gene expression signature using multiple networks, which would then be integrated on a protein-by-protein basis using the square of the NES. Since a non-relevant network would generate a protein activity NES close to zero under the null model, networks that generate more extreme NES's can be interpreted to more accurately match the given biological context and were thus weighted more heavily. This approach has been improved on further in PISCES. Rather than relying on the square of the NES to integrate networks in a protein-by-protein manner, Weighted VIPER utilizes all the proteins in a given sample to determine network accuracy. For each sample, the NES's generated by the set of networks for each protein are ranked, and the ranks are totaled across proteins. Networks are then weighted based on their frequency as the most-accurate network. As an example, if network A generates the most extreme NES for 50% of the proteins

in a sample and network B generates the most extreme NES for 25% of the proteins, network A will be weighted twice as heavily in the integration. This technique utilizes all proteins as a multiplexed reporter of network accuracy, allowing for more accurate matching of samples and the most-context specific network available.

Single Cell Visualization Functionality: Visualizing data with thousands of dimensions is a fundamental challenge of transcriptomics. PISCES has a number of pre-built plotting functions to aid in the visualization of results. Scatter plots are based in UMAP coordinates, with the starting features filtered by the most significant proteins within each sample. Functions within PISCES allow for the visualization of clustering schemes, protein activity, or gene expression in UMAP space, along with density plotting. Additionally, we provide heatmap functionality for more tractable succinct visualization of a set of genes or proteins grouped by cluster, such as a set of known markers or a list of candidate master regulators.

Resolution-Optimized Louvain Clustering Algorithm: The default clustering method implemented in Seurat is Partitioning Around Medioids (PAM). However, for large datasets aggregating hundreds of thousands of single-cells, PAM is computationally slow, requiring more computational power than is available to the average user and computation of pairwise distance matrices exceeding the vector size limit in R. In such cases, it is preferable to run a network-based Louvain clustering, as implemented in Seurat, which optimizes network modularity score. However, practical implementations of Louvain clustering include a user-adjustable resolution parameter which allows over-clustering and under-clustering without an objective cluster quality metric. To solve this problem, we have implemented a hybrid clustering approach in PISCES

which performs cluster assignment in two steps. First, Seurat Louvain clustering is performed with resolution values ranging from 0.01 to 1.0 at intervals of 0.01, then cluster quality is evaluated at each resolution value to select an optimum in this range. For each resolution value, clustered cells are subsampled to 1000, and silhouette score is computed for these 1000 cells and their corresponding cluster labels, with correlation distance metric. This procedure is repeated for 100 random samples to compute a mean and standard deviation of average silhouette score at each resolution value. The highest resolution value that maximizes mean silhouette score is selected as the optimal resolution at which to cluster the data.

Multi-Way K-Means Clustering Algorithm: In addition to PAM and Louvain with Resolution Optimization, PISCES further implements a Multi-Way K-Means Clustering approach. Transitioning populations, such as in a differentiation pathway, are extremely common, and such relationships will not be accurately characterized by a discrete clustering scheme. To handle this set of problems, we adapted the Multiway K-Means algorithm for use in biological settings, where samples can be thought of as linear combinations of related phenotypes rather than simply belonging to totally distinct populations. Originally developed for clustering speciating microbiome populations, Multiway K-Means technique has two primary advantages. First, it more accurately captures cluster center (in biological terms, a representative phenotype) for each population endpoint. Second, it places cells along a trajectory between cluster centers, providing a more accurate representation of cell state and allowing for additional inferences into the drivers of transitional populations.

Semi-Supervised Cell Type Calling: For each single cell gene expression sample, cell-by-cell identification of cell types is performed using the SingleR package and the preloaded Blueprint-ENCODE reference, which includes normalized expression values for 259 bulk RNASeq samples generated by Blueprint and ENCODE from 43 distinct cell types representing pure populations of stroma and immune cells [49] [50]. The SingleR algorithm computes correlation between each individual cell and each of the 259 reference samples, and then assigns both a label of the cell type with highest average correlation to the individual cell and a p-value computed by wilcox test of correlation to that cell type compared to all other cell types. Cell-by-cell SingleR labels with $p<0.05$ are added as metadata and may be projected onto PISCES-generated UMAP space. Unsupervised clusters may then be labelled as a particular cell type based on the most-represented SingleR cell type label within that cluster.

Data Collection and Down-sampling for Technical Validation: SNU-16, a stomach adenocarcinoma cell line, was dissociated into a single-cell suspension and scRNAseq was performed using 10X Genomics Chromium platform (3'v3). Libraries were sequenced on an illumina Novaseq 6000 according to 10X Genomics' protocol. In mid-log growth, SNU-16 is a transcriptionally complex cell line that will typically have 40,000-50,000 UMIs/cell with 134,000 reads sequencing. These data were then down-sampled to depths of 10-40,000 at 5,000 UMI intervals and 1-10,000 at 1,000 UMI intervals. Sample depths were first drawn from a uniformly distributed multinomial with $n = N*x$ and $p1,...,pn = 1 / N$, where N was the number of cells and x is the target mean depth. Once sample depths were drawn, UMI counts were drawn from a sample-specific multinomial with $n = di$ and p1...pg = 1 / G, where di is the sample depth and G is the number of genes detected in the original UMI matrix.

Biological Validation Analysis: A highly used public CITE-Seq dataset of cord blood mononuclear cells was downloaded from Gene Expression Omnibus (GEO), and subset to human cells only. RNA counts were processed by the standard PISCES workflow, and antibody dependent tags (ADTs) were concurrently analyzed. ADT matrix was normalized by Seurat Centered Log Ratio "CLR" function, and clustered by PISCES resolution-optimized Louvain algorithm. Two-dimensional data representation was computed by RunUMAP, and antibody staining of all markers was visualized in a heatmap, with cells grouped by ADT cluster. For single-cell sequencing data, both gene expression signature and PISCES-inferred VIPER matrix were subset to genes encoding proteins represented in the ADT panel, and data were re-clustered on those gene subsets. For genes shared across all three modalities, coefficient of variation was computed as standard deviation divided by mean across all cells, and Spearman correlation was computed between gene expression or VIPER and corresponding protein-targeting antibody.

Data Availability: The PISCES pipeline is implemented as an R package with all dependencies listed and a usage tutorial available at https://github.com/califano-lab/PISCES. All data, ARACNe networks, and VIPER matrices referenced in this manuscript are also available at https://github.com/califano-lab/PISCES-validation.

Chapter 2: Profiling the Renal Cell Carcinoma Microenvironment by Protein Activity Inference Reveals Novel Recurrence-Associated Macrophages

The following is adapted from:
Obradovic, A.*, Chowdhury, N.*, Haake, S. M.*, Ager, C., Wang, V., Vlahos, L., Guo, X. V., Aggen, D. H., Rathmell, W. K., Jonasch, E., Johnson, J. E., Roth, M., Beckermann, K. E., Rini, B. I., McKiernan, J., Califano, A., & Drake, C. G. (2021). Single-cell protein activity analysis identifies recurrence-associated renal tumor macrophages. *Cell*, *184*(11). https://doi.org/10.1016/j.cell.2021.04.038
*These authors contributed equally

2.1 Summary

Clear Cell Renal Carcinoma (ccRCC) is a heterogeneous disease with a variable post-surgical course. To assemble a comprehensive ccRCC tumor microenvironment (TME) atlas, we performed single-cell RNA sequencing (scRNASeq) of hematopoietic and non-hematopoietic subpopulations from tumor and tumor-adjacent tissue of treatment-naïve ccRCC resections. We leveraged the VIPER algorithm to quantitate single-cell protein activity and validated this approach by comparison to flow cytometry. The analysis identified key TME sub-populations, as well as their master regulators and candidate cell-cell interactions, revealing clinically relevant populations, undetectable by gene expression analysis. Specifically, we uncovered a tumor-specific macrophage subpopulation characterized by upregulation of TREM2/APOE/C1Q - validated by spatially resolved, quantitative multispectral immunofluorescence. In a large clinical validation cohort, these markers were significantly enriched in tumors from patients who recurred following surgery. The study thus identifies TREM2/APOE/C1Q-positive macrophage infiltration as a potential prognostic biomarker for ccRCC recurrence, as well as a candidate therapeutic target.

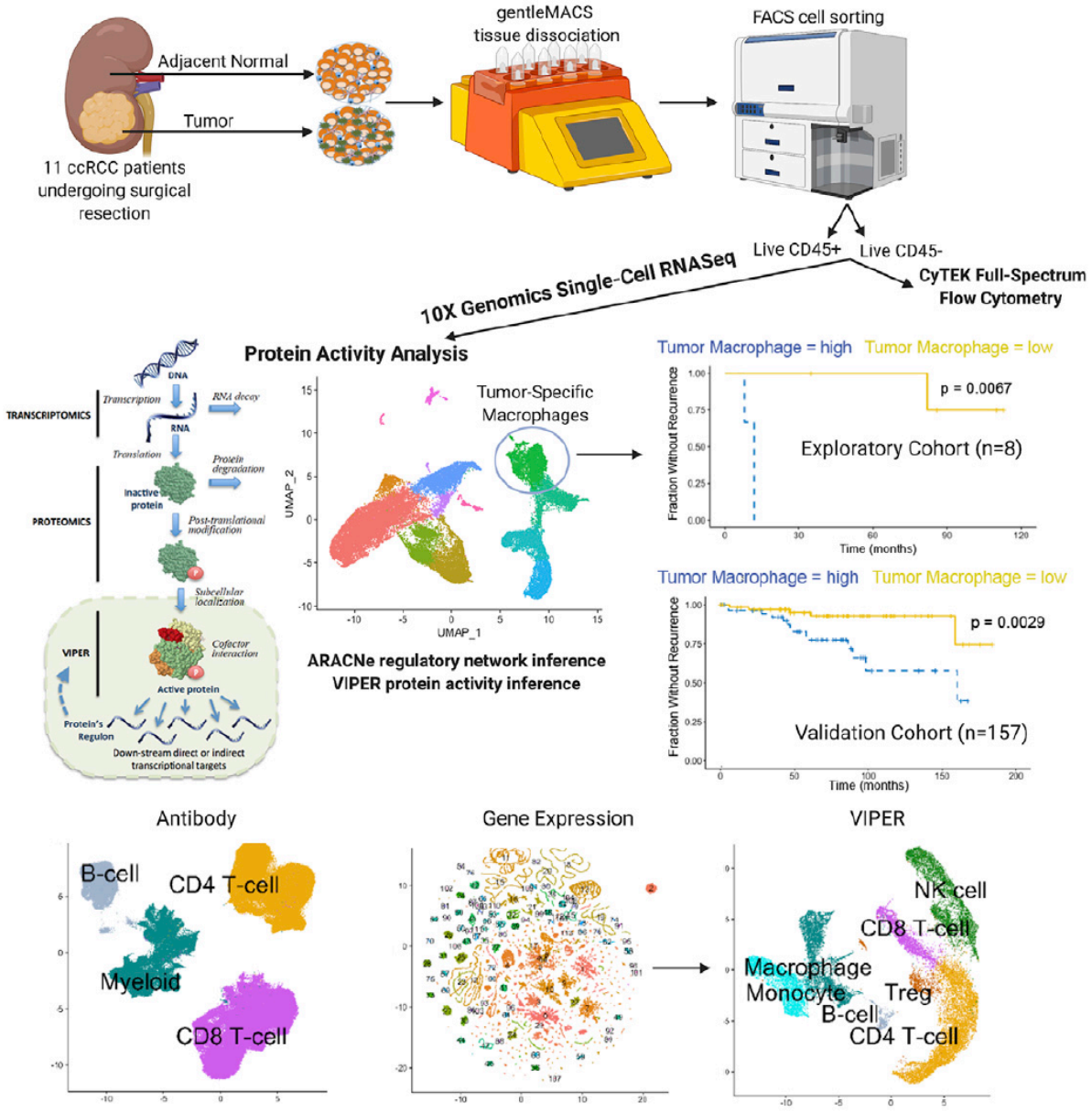


Figure 11: Experimental Design and Major Findings (Graphical Abstract)

2.2 Introduction

Clear cell renal carcinoma (ccRCC) is the most common histological subtype of renal carcinoma. Although primary disease is treated surgically, approximately 40% of resected ccRCC patients will relapse and develop metastases [51]. With a 5-year survival of 10% [52] metastatic ccRCC

is a lethal disease, underscoring the need to understand the cellular and molecular mechanisms in primary lesions that are prognostic for recurrence, both as biomarkers and as potential targets for intervention. Although ccRCC is an immunogenic tumor, the tumor-immune cell dynamics that regulate effective anti-tumor responses remain incompletely characterized. Consistent with other immunogenic tumors, overall immune infiltration and tumor mutation burden are partially predictive of response to therapy; yet, the value of these biomarkers in clinical decision making remains elusive. Indeed, the complete picture of anti-tumor immune response drivers is complex [53] [54] [55]. Predictors of post-surgical disease recurrence are also limited, with previous gene expression studies suggesting CD44 as a marker of recurrence [56].

To date, the most comprehensive studies of the primary ccRCC Tumor Micro-Environment (TME) used cytometry by time of flight (CyTOF) to interrogate markers of innate and adaptive immunity [57]. These studies showed that expression of T cell exhaustion markers and CD38+ myeloid cell infiltration was associated with worse overall outcome. High-throughput droplet-based single-cell RNA Sequencing (scRNASeq) has recently emerged as a valuable tool to catalog the diverse cellular subpopulations that comprise the TME, with the ability to identify representative gene expression signatures from thousands of individual cells in a single sample [16] [32]. In contrast to bulk RNA-Sequencing, scRNASeq can characterize the transcriptional state of individual cell types, highlighting the role of rare populations whose gene expression signature would be diluted below the limits of detection in bulk samples [20]. In contrast to flow cytometry or CyTOF, scRNASeq generates a genome-wide profile of each individual cell's transcriptome, without requiring selection of predefined markers. The value of scRNASeq has been demonstrated in recent studies of melanoma [17] [18] and breast cancer [34]. However, no

systematic, single-cell studies have been performed to study tumor tissue vs. adjacent normal in ccRCC.

A key technical limitation of scRNASeq is that gene expression profiles are extremely sparse, with ~80% - 90% of genes undetected in every cell, a phenomenon known as *gene dropout*. While such data are effective in characterizing more molecularly distinct cellular subpopulations, they are not well suited to study specific genes and may also fail to detect more subtle differences, for instance due to activation of a few critical lineage markers [33]. Although dimensionality reduction tools, such as the Seurat pipeline [36], are successful in identifying individual subpopulations, the sparse and noisy nature of the data often prevents elucidation of finer-grain biological mechanisms.

To address this issue, we have developed the metaVIPER algorithm [26] , which leverages highly-multiplexed, tissue-specific gene-reporter assays to accurately measure the activity of up to ~6,500 regulatory proteins on a single-cell basis, including transcription factors (TFs), co-factors (co-TFs), signaling proteins (SPs), and surface markers (SMs), based on the expression of their downstream regulatory targets (*regulon*). MetaVIPER extends the VIPER algorithm [21] to single-cells, independent of lineage. For simplicity, here we will use the term VIPER to refer to its single-cell implementation. Single-cell, tissue-specific regulons are inferred using ARACNe, an information theoretic algorithm that has been experimentally validated in multiple tissue contexts, with a >70% accuracy in target identification [40].

To comprehensively characterize the interaction of immune- and non-immune cells in the ccRCC TME, we generated scRNASeq data from FACS-purified hematopoietic and non-hematopoietic cells dissociated from tumor and adjacent non-tumor tissue of 11 treatment-naïve primary ccRCC patients. To analyze these data, we developed a VIPER-based scRNASeq analysis pipeline to assess single-cell protein activity from single-cell ARACNe networks followed by an optimized single-cell clustering approach. These studies revealed a population of tumor-specific C1Q+TREM2+APOE+ macrophages associated with early post-surgical disease recurrence, as well as a potential target for therapeutic intervention. To validate VIPER predictions, we generated spectral flow cytometry and scRNASeq from matched patient samples, as well as quantitative, multi-spectral immunofluorescence (qmIF) data for a set of proteins significantly activated in a macrophage subpopulation prognostic for post-surgical disease recurrence. Taken together, these data provide a comprehensive atlas of primary ccRCC TME subpopulations—including the Master Regulator (MR) proteins that control their transcriptional state, lineage markers, and predicted cell-cell interactions.

2.3 Results

Protein Activity Analysis of CD45+ TME Cells Reveals Tumor-Specific Immune Subpopulations: To study hematopoietic and non-hematopoietic populations in the primary ccRCC TME at single cell resolution, we isolated live cells from 11 treatment-naïve resected tumors, along with adjacent normal tissue. Expression-based clustering of scRNASeq profiles revealed populations broadly consistent across patients (Figure 12).

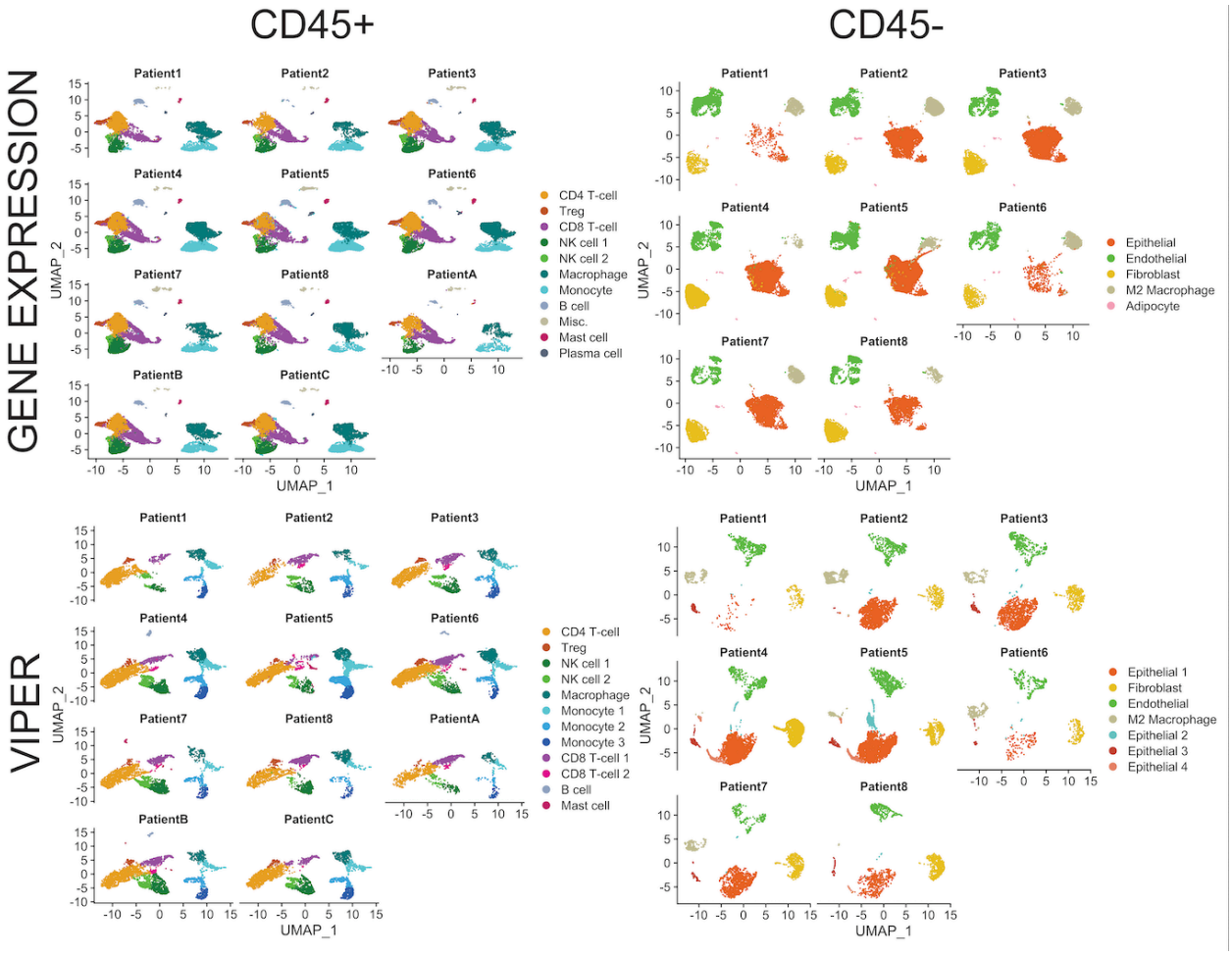


Figure 12: Consistency of Cell Type Clustering Across Patients

UMAP plots of global gene expression clustering and global VIPER clustering for both CD45+ and CD45- cells, split by individual patient identity, such that the overall cell types identified are consistent across patients, with minimal batch effect.

We initially focused on the hematopoietic compartment (CD45+), which was visualized following UMAP dimensionality reduction and clustered using the Seurat Louvain algorithm (Stuart et al., 2019). To optimize often arbitrary clustering while retaining scalability to hundreds of thousands of cells, we performed Louvain clustering across a range of 100 resolution values and selected optimal clustering resolution by optimizing a bootstrapped mean silhouette score (see methods). This clustering approach resolved CD4 and CD8 lymphocytes, regulatory T cells

(Tregs), NK cells (two populations), macrophages, monocytes, and small populations of B cells, mast cells, and plasma cells (**Figure 13A**), which were represented in all patient samples. For visualization purposes, we show a heatmap for the top five transcripts most uniquely upregulated in each cluster (**Figure 13C**). These data confirmed SingleR-inferred cellular identify of each cluster — including expression of IL7R in CD4 T cells, CD3 and granzyme in CD8 T cells, and S100A8/S100A9 in monocytes.

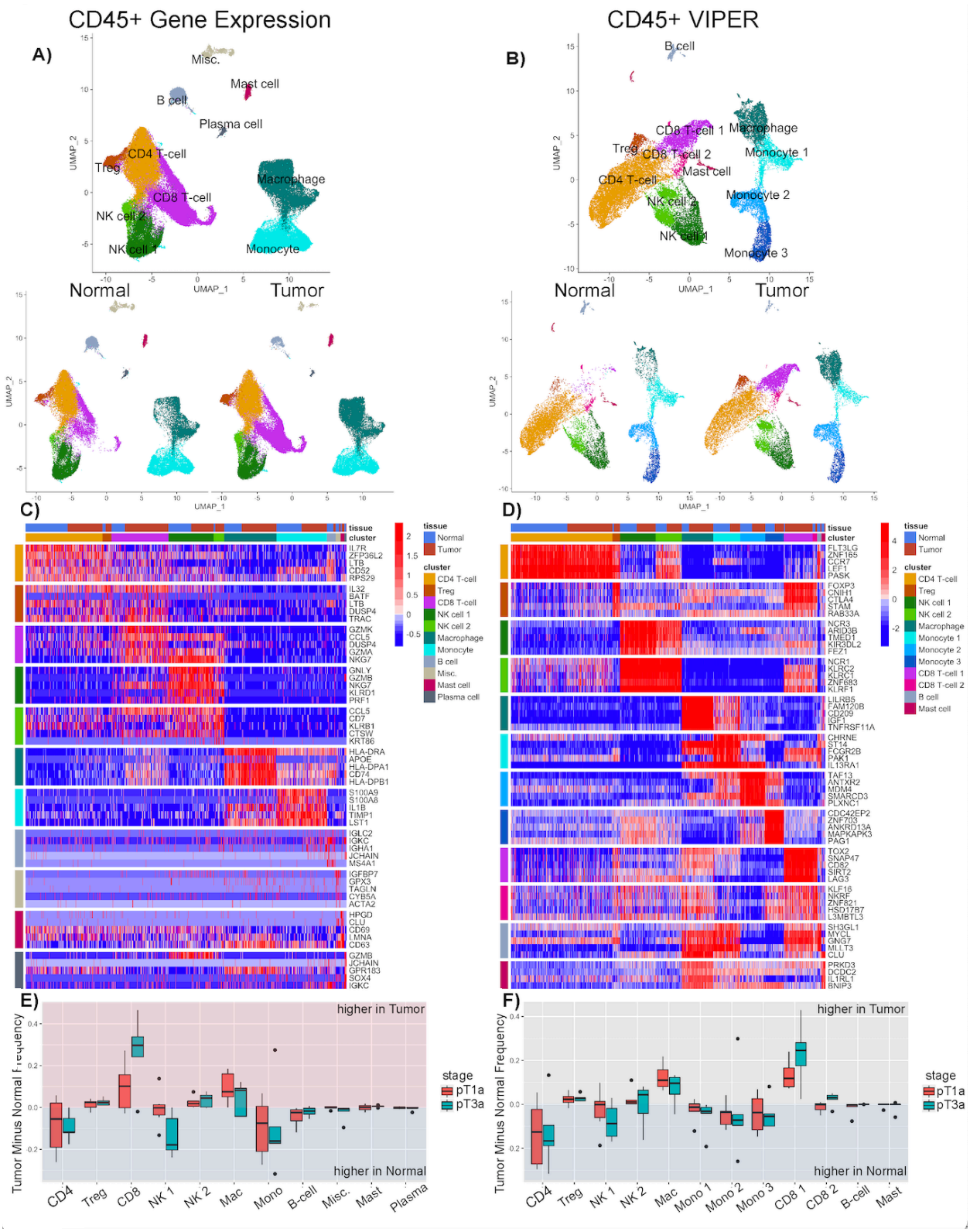


Figure 13: Deep Profiling of CD45+ Microenvironment by Gene Expression and Protein Activity Reveals Tumor-Specific Immune Populations

A) UMAP Plots for single-cell gene expression pooled across CD45+ samples, clusters visualized and labelled by cell type. Bottom plot is split by Tumor vs Adjacent Normal label B) UMAP Plots for VIPER-Inferred protein activity pooled across CD45+ samples. Bottom plot is split by Tumor vs Adjacent Normal label C) Heatmap of top5 upregulated genes for each cluster by expression; each row represents a gene and each column represents a cell. Legend shows cluster identity with cell type inferred by SingleR and Tumor (red) or Adjacent Normal (blue) tissue source. D) Heatmap of top5 differentially upregulated proteins for each cluster by VIPER-inferred activity. Legend as in C. E) Bar plots of patient-by-patient cluster frequency in Tumor minus frequency in Adjacent Normal for each Gene Expression cluster, grouped by stage; values < 0 (blue) indicate higher frequency in Adjacent Normal and values > 0 (red) indicate higher frequency in Tumor. F) Bar plots of patient-by-patient cluster frequency in Tumor minus frequency in Adjacent Normal for each VIPER cluster, grouped by stage, as in E.

However, expression-based clustering missed multiple established markers of these populations. For example, Tregs did not show differential expression of the canonical FOXP3 transcription factor; rather the most overexpressed gene was IL-32. While other genes in this cluster, e.g., BATF, TIGIT, and TNFRSF18— are expressed in Tregs [58], none is considered a canonical marker. Further, expression-based clustering failed to recapitulate the heterogeneity of these subpopulations; for example, it distinguished only two populations of myeloid cells (**Figure 13**). Finally, considering intra-cluster statistics, differentially expressed genes had poor reproducibility. For instance, average within-cluster standard deviation of classical markers such as IL7R, KLRD1, and CD8B was quite high, $\sigma_{IL7R} = 3.19$, $\sigma_{KLRD1} = 3.79$, and $\sigma_{CD8B} = 3.01$ respectively, exceeding the mean expression values of the gene: $\mu_{IL7R} = 2.52$, $\mu_{KLRD1} = 3.17$, and $\mu_{CD8B} = 1.41$.

We next proceeded to assess whether protein activity-based clustering, using VIPER, would yield additional robustness and biological insight. Since we previously showed that regulatory networks of lineage-related cells have >95% overlap [59], we generated (*n = 69*) ARACNe-inferred regulatory models—one for each gene expression cluster in each patient—as we expected that finer differences would likely be present within primary lineages. For each single cell, we then used the cluster-specific regulatory networks to perform VIPER-based protein activity inference. While the resulting clusters were generally consistent with those derived by gene expression, protein activity-based clusters showed multiple critical differences (**Figure 13B, Figure 13D**). For example, the most differentially active protein in the Treg cluster was FOXP3, consistent with well-established Treg biology [58] (**Figure 13D**), and Cytotoxic T Lymphocyte Antigen–4 (CTLA-4), which is up-regulated on the surface of tumor infiltrating Tregs [60], was inferred as differentially active by VIPER. VIPER identified a distinct CD8 T cell population with markers consistent with exhaustion, including differential activation of LAG-3, TOX2, and PD1, which had been missed by expression-based analysis (**Figure 13D, Figure 14A**). Additionally, myeloid cells were further stratified by VIPER into macrophages and three distinct monocyte subpopulations.

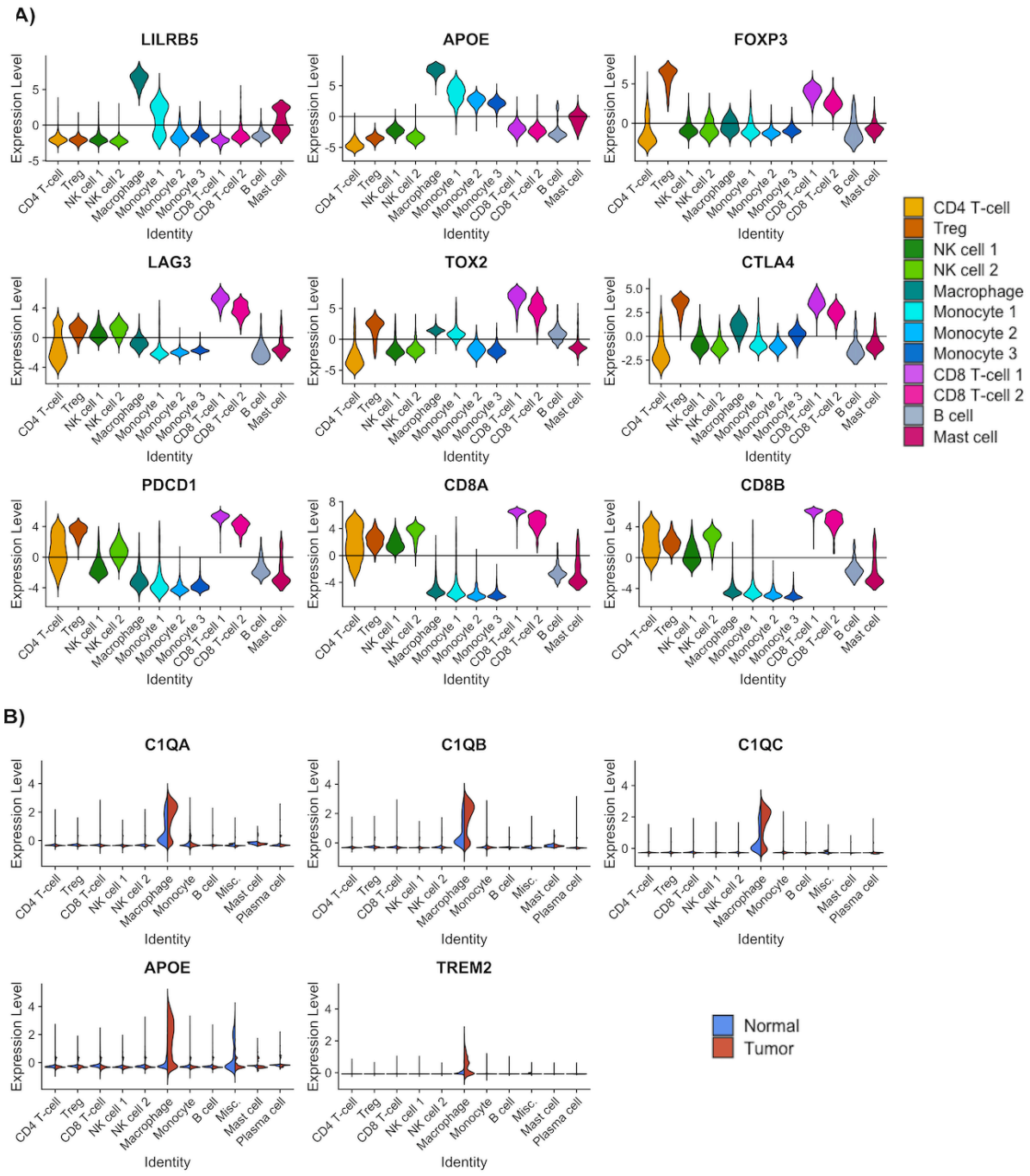


Figure 14: Known and Novel Tumor-Infiltrating Immune Population Markers Discovered from Single-Cell Transcriptomic and Inferred Proteomic Data

A) Violin plots of VIPER-inferred Proteins upregulated in CD45+ cell subsets corresponding to Tregs (FOXP3, CTLA4), Exhausted CD8s (TOX2, LAG3, PD1, CTLA4), and Tumor-specific Macrophages (LILRB5, APOE). B) Violin plots of top transcriptional markers (C1Q, APOE,

TREM2) specifically up-regulated in Tumor-Infiltrating Macrophages as compared to other cell populations as well as non-tumor macrophages.

VIPER analysis identified several populations that were differentially represented as a function of tumor stage and localization in tumor vs. adjacent non-tumor; these were undetectable by gene expression analysis (**Figures 13E,F**). Specifically, Treg, CD8 T cell, and Macrophage normalized counts were higher in the tumor compared to normal adjacent tissue ($p = 0.012$, $p = 0.006$, $p = 0.013$, respectively). In contrast, Monocyte, B cell, and CD4 T cell counts were higher in adjacent normal ($p = 0.097$, $p = 0.017$, $p = 0.018$, respectively). Two NK cell clusters were identified, one with higher counts in the adjacent normal (NK cell 1) and a second with higher counts in tumor (NK cell 2), ($p = 0.09$, $p = 0.008$, respectively). Consistent with prior data [61], activity-based, but not expression-based clustering identified higher counts of exhausted CD8 T cells in tumor vs. adjacent normal ($p = 0.0005$), and also in stage pT3a vs. pT1a tumors ($p = 0.015$) (**Figure 13F**). Further, the tumor-specific macrophage population identified by inferred protein activity was more significantly enriched in tumor as compared to adjacent normal than the coarse macrophage population identified by gene expression, i.e., $p = 0.0006$ vs. $p = 0.013$.

Reproducibility of individual markers was also significantly improved by VIPER (**Figure 13D**), compared to gene expression (**Figure 13C**). For example, the standard deviation of the classical markers (IL7R, KLRD1, and CD8B) was much lower, $\sigma_{IL7R} = 0.75$, $\sigma_{KLRD1} = 0.48$, and $\sigma_{CD8B} = 0.49$ respectively, whereas their mean activity value was substantially larger $\mu_{IL7R} = 5.28$, $\mu_{KLRD1} = 6.54$, and $\mu_{CD8B} = 5.88$, as further confirmed by significant improvement in silhouette scores, SC = 0.7 by VIPER as compared to a SC = 0.35 by gene expression based clustering

(**Figure 15**). Thus, based on the coefficient of variation (i.e., σ/μ), reproducibility was increased between 8.9-fold (IL7R) and 25.6-fold (CD8B) by VIPER-based clustering.

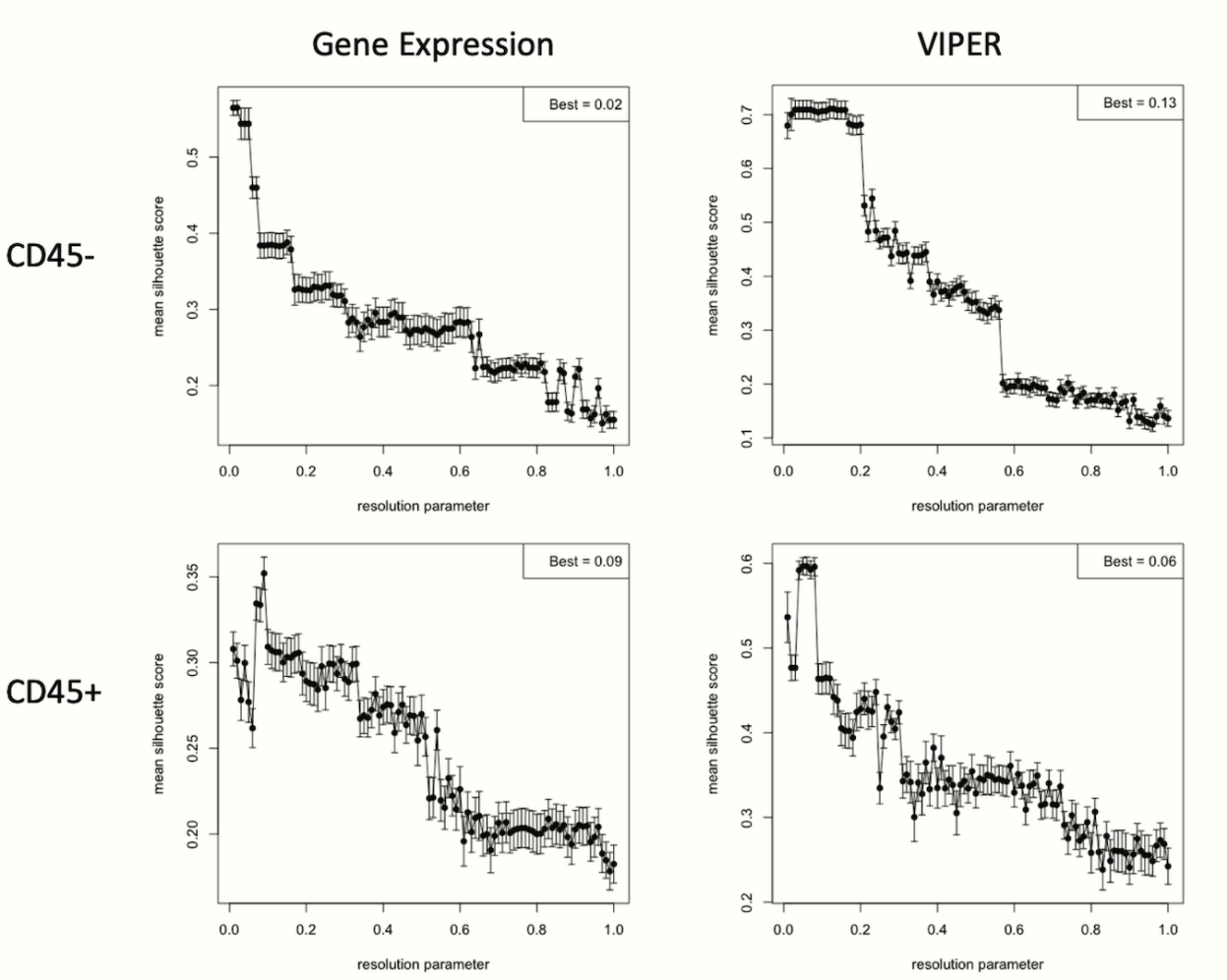


Figure 15: Resolution-Optimized Louvain Clustering Silhouette Scores

Mean and Standard Deviation of Silhouette Score by Resolution-Optimized Louvain algorithm for each resolution value ranging along the x-axis from 0 to 1.0 at intervals of 0.01, showing "best" resolution in the top-right as the resolution that maximizes mean silhouette score. Includes CD45+ Gene Expression clustering, CD45- Gene Expression clustering, CD45+ VIPER clustering, and CD45- VIPER clustering.

Relative Specificity of Tumor Infiltrating Immune Cell Populations: To visualize protein activity differences in key cell type markers over-represented in tumor vs. adjacent normal, including Macrophages, Tregs, and exhausted CD8 T cells, we generated violin plots (**Figure 14A**). As shown, LILRB5 was identified as the most differentially active among VIPER-inferred markers of tumor-specific macrophages. These macrophages, as well as the three monocyte clusters, showed high APOE activity. Among T-cells, FOXP3 was identified as the most activated protein in TIL Treg; relative activation was also noted in tumor associated CD8 T cells. CTLA-4 followed a similar pattern, consistent with previous data from bulk TIL Treg studies [60]. LAG-3 and PD-1 (PDCD1) showed similar inferred protein activity distributions, with higher activity in CD8 T cells and, in particular, in the CD8 TIL cluster (CD8 T cell 1) vs. the (CD8 T cell 2) cluster that was also represented in adjacent normal. PD-1 also showed significant expression in TIL Treg. Interestingly, we found significant activation of TOX2—a protein related to TOX1 which plays a critical role in epigenetic reprogramming of exhausted CD8 T cells [62] [63]—in the CD8 TIL cluster, consistent with an exhausted phenotype. We complemented these differential activity analyses by examining genes that were differentially expressed in the tumor-specific macrophage population, which were only be identified by VIPER-based protein activity clustering (**Figure 13C**). Once a finer-grain cluster structure was revealed, with several genes found to be over-expressed in specific clusters, including APOE, C1QA-C and TREM2 - demonstrating the ability to integrate both differential protein activity and differential gene expression in the analysis (**Figure 14B**). Notably, differential expression of these genes would have gone undetected if the cluster structure produced by gene expression-

based clustering had been used. As discussed below, we subsequently validated the tumor-specific macrophage marker proteins identified by these analyses by immunofluorescence and correlated those data with clinical outcome.

Master Regulators of Sub-Population Transcriptional State: In addition to recapitulating differential activity of established subpopulation markers, VIPER analysis identified novel proteins that were differentially active in both common and rare subpopulations. For TFs and co-TFs, these Master Regulator (MR) proteins represent novel mechanistic drivers of the transcriptional state of these cells, while signaling and surface marker proteins may represent novel lineage markers for FACS-based isolation, as further confirmed by the protein-based validation assays discussed below. Of note, we previously showed that a majority (>70%) of VIPER-inferred proteins control the transcriptional state of a cell by direct, physical regulation of their transcriptional targets and can be used to efficiently reprogram cell state by ectopic expression/co-expression [42] [64] [65] [66]. As a result, these proteins may also represent attractive drug targets for modulation of specific sub-populations. Candidate MR proteins of each VIPER-inferred cluster are reported in supplement at Obradovic et al; the top MRs are also shown in Figure 13D. In particular, MRs of tumor-enriched cell populations (i.e., Tregs, CD8 T 1 cells, and Macrophages) included both established drug targets for Tregs or exhausted CD8 T-cells, such as CTLA-4 and PD-1, as well as less well-characterized markers of Tregs (CNIH1, STAM, RAB33A, etc.), exhausted CD8 T cells (TOX2, SNAP47, CD82, SIRT2, LAG-3, etc.), and Tumor-Infiltrating Macrophages (LILRB5, FAM120B, CD209, IGF1, TNFRSF11A, etc.). As a result, these data provide a valuable resource of proteomic regulators for the full complement of cell phenotypes in the ccRCC TME.

***VIPER-inferred Protein Activity Recapitulates Flow Cytometry, Overcoming scRNASeq-related Gene Dropout*:** To more fully characterize the proteomic profile of the ccRCC TME and to benchmark VIPER results, we analyzed a subset of samples for which scRNASeq data was available using high-dimensional flow cytometry with a 19-marker lymphoid panel and 19-marker myeloid panel. Manual gating of specific marker pairs broadly recapitulated the populations identified by VIPER-based cluster analysis. For example, flow cytometry identified a population of CD8+/PD1+/CD39+ CD8 T cells, with numerically higher normalized counts in tumor vs. adjacent non-tumor samples ($p = 0.057$), consistent with the cluster of tumor-enriched exhausted CD8 T cells identified by VIPER (**Figure 16**). Cytometry also identified a population of CD4+/CD127^{low}/FOXP3+ Tregs, with higher representation in the majority of tumor vs. adjacent normal samples ($p = 0.072$). Similarly, these flow studies confirmed the existence of two distinct NK cell subpopulations (i.e., CD56^{high}/CD16^{low} vs CD56^{low}/CD16^{high}), a CD11C+/CD163+ macrophage population with higher representation in tumor vs. adjacent normal ($p = 0.076$), and three distinct monocyte subpopulations (CD14+/CD16+, vs. CD14+/CD16-, vs CD14-/CD16+).

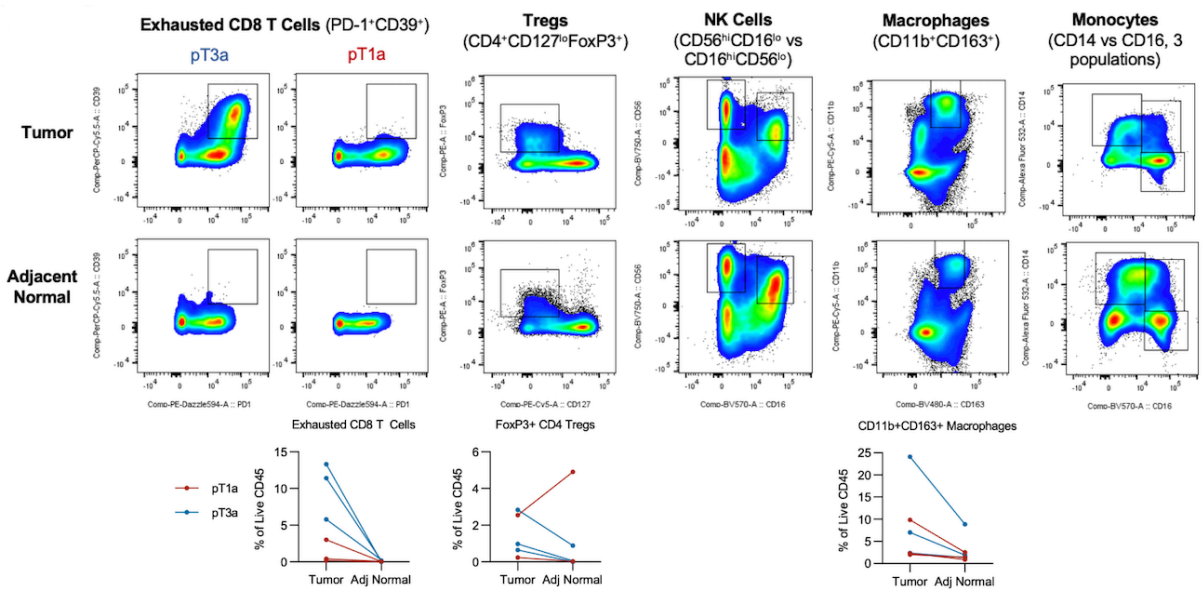


Figure 16: Characterization of Immune Infiltrate by Manual Flow Cytometry Gating

Representative Flow Cytometry Gating in Tumor and Adjacent Normal and frequency plots in tumor and adjacent normal for all manually gated populations. Populations of PD1+CD39+ exhausted CD8 cells, Tregs, and CD11B+CD163+ Macrophages are of higher frequency in Tumor than Adjacent Normal. Representative plots showing two distinct NK cell phenotypes and three monocytic sub-phenotypes, consistent with Figure 13B.

While these populations were validated by manual gating of specific proteins, our analysis shows that they could not have been inferred directly from the high-dimensional flow cytometry data (**Figure 17**). To test this, we performed unsupervised clustering of the flow cytometry dataset, using the Resolution-Optimized Louvain cluster analysis algorithm by which we infer expression and activity-based clusters. Based on the lymphoid panel, the analysis identified four distinct clusters: CD4 and CD8 T cells, Myeloid cells, and B cells (**Figure 17A**) while the myeloid panel yielded 6 clusters: Lymphoid cells, B cells, 3 monocyte cell types characterized as CD14+/CD16+, CD14+/CD16-, and CD14-/CD16+ respectively, and a macrophage cluster only

represented in the tumor compartment, with relative over-expression of CD86, CD1D, CD16, CD163, CD169, CD56, CXCR2, CD14, and CD33 proteins (**Figure 17D**).

To assess whether scRNASeq data could recapitulate these findings, we restricted unsupervised cluster analysis to genes encoding for proteins represented in the flow cytometry panels (**Figure 17B,E**). This analysis failed to reveal biologically relevant clusters due to high gene dropout rates, even on markers that should be highly expressed. Indeed, expression of the genes encoding for the 19 lymphoid and 19 myeloid proteins was too noisy and sparse to support cluster inference consistent with established cell types. This result emphasizes the extremely noisy nature of scRNASeq measurements when restricted to specific genes of interest.

We next tested whether analysis of VIPER-inferred activity for lymphoid and myeloid markers could recapitulate meaningful cell types. As shown in Figure 3D, the vast majority of proteins in the lymphoid flow cytometry panel were well-resolved by VIPER, with the exceptions of PTGDR2, FCGR3B, and NT5E. Similarly, the majority of the proteins in the myeloid flow panel were also well-resolved by VIPER, with the exception of FCGR3B and CD33 (**Figure 17F**). Taken together, 34 of 39 proteins (77%) were well-represented by VIPER, consistent with the 70-80% previously reported recovery in protein activity measurements [21]. This was even more remarkable because the panel analyzed here included mostly surface markers not directly involved in transcriptional regulation, with a few exceptions (e.g. FOXP3).

Protein activity analysis restricted to the lymphoid panel was effective in recovering lymphoid cell diversity, and was able to distinguish Monocytes from Macrophages despite limited profiling

of macrophage lineage markers by the lymphoid panel (**Figure 17C**), thus comparing favorably with analysis of flow cytometry data, which only identified 4 of these 7 cell types. For instance, protein activity analysis was effective in identifying Tregs, which were missed by flow-based clustering due to low intensity of FOXP3 staining. Comparing protein abundance and activity in matched flow-cytometry and VIPER clusters (e.g., CD4 T cells), the reproducibility of activity data was approximately 2-fold higher, on average, based on coefficient of variation (CV) analysis, defined as the ratio of the standard deviation over the mean. When averaged over the top proteins differentially represented in the CD4 and CD8 T cell clusters, flow-based analysis produced $CV_{CD4} = 0.206$ and $CV_{CD8} = 0.209$, while activity-based analysis yielded $CV_{CD4} = 0.151$ and $CV_{CD8} = 0.124$, reflecting higher noise in antibody-based measurements.

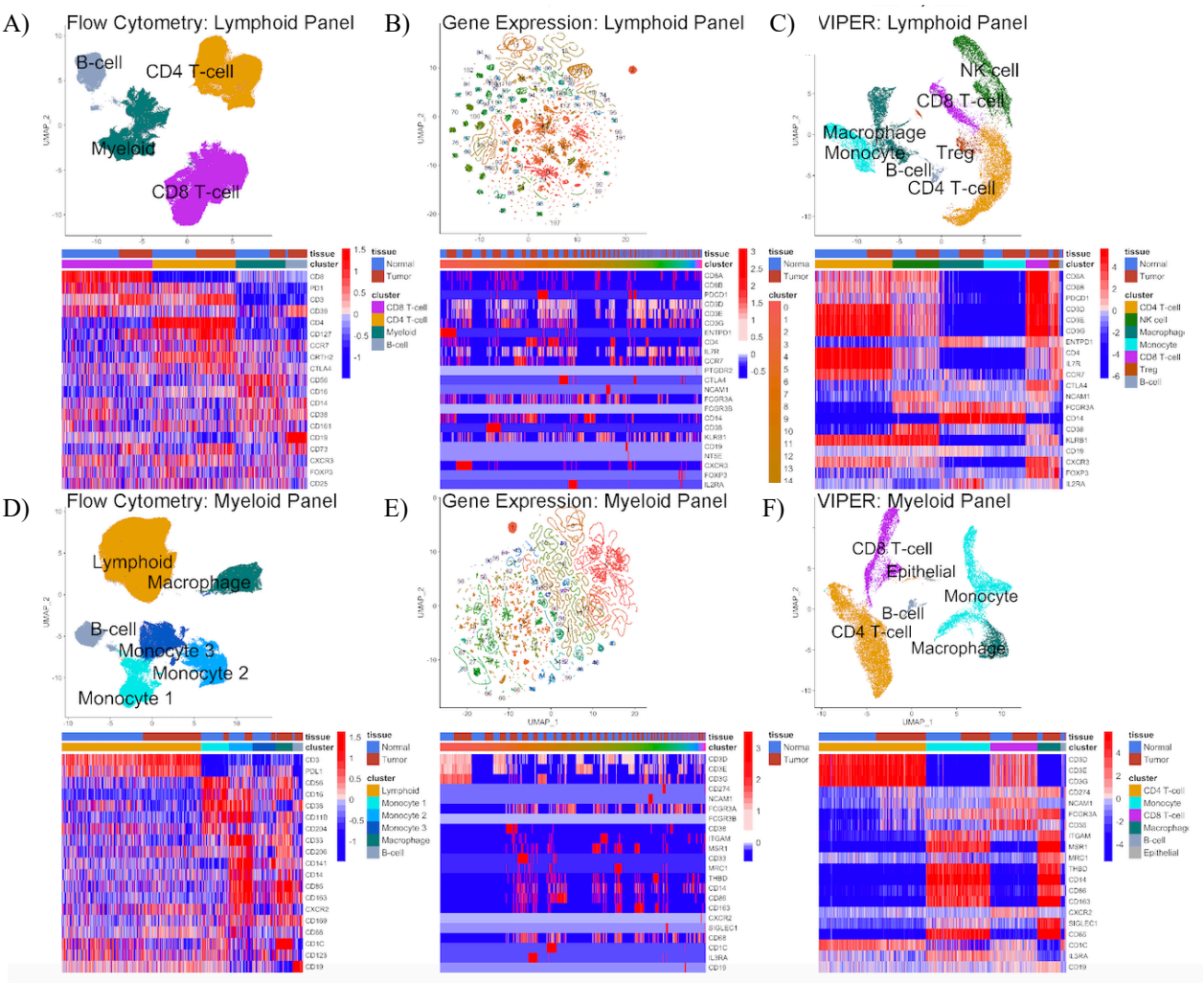


Figure 17: Flow Cytometry is Better Recapitulated by Protein Activity Than by Gene Expression

A) UMAP projection, clustering, and heatmap by flow cytometry proteins profiled in CyTEK Lymphoid Panel. B) UMAP and clustering by scRNASeq gene expression subset to the proteins profiled in A, showing noise-induced decrease in clustering resolution. C) UMAP and clustering by scRNASeq VIPER inference subset to the proteins profiled in A. D) UMAP and clustering by flow cytometry proteins profiled in CyTEK myeloid panel. E) UMAP and clustering by scRNA-Seq gene expression, subset to the proteins profiled in D. F) UMAP and clustering by scRNA-Seq VIPER inference, subset to the proteins profiled in D.

These findings were validated in a public CITE-Seq dataset profiling antibody staining and gene expression simultaneously in single cells derived from cord blood [46]. Here, clustering by antibody profile was lost by single-cell gene expression but was completely recovered by VIPER analysis. Moreover, the cell-matched coefficient of variation within each cluster was significantly lower for VIPER-inferred protein activity as compared to gene expression (p=0.0004) and compared favorably with antibody-based measurements (p=0.0083). As above, VIPER-measured protein activity correlated more closely with antibody staining than did gene expression. Taken together, these results show that protein activity-based analyses can recapitulate cell-type identification based on protein-level data derived by flow cytometry even from a relatively restricted set of marker proteins, suggesting that such analyses effectively mitigate the gene dropout inherent in scRNASeq.

Protein Activity in CD45- Cells Distinguishes Tumor cells from Adjacent Normal Epithelium*:* We next compared expression and activity-based clustering of non-hematopoietic (CD45-) ccRCC TME cells. Expression-based clustering of CD45- cells from all patients (including tumor and adjacent normal) revealed four predominant cell types: epithelial cells, endothelial cells, fibroblasts, and M2 macrophages (**Figures 18A**). These populations showed differential representation in tumor vs. adjacent normal, with M2 macrophages predominant in adjacent normal (p_{M2} = 0.007), and fibroblasts and epithelial cells over-represented in tumor tissue (p_{FB} = 0.009, p_{Epi} = 0.0005). The epithelial cluster, mostly comprised of tumor cells, showed tumor compartment specificity as compared to adjacent normal. As expected, it was more highly represented in patients with pT3a compared to pT1a disease ($p_{Epi|pT3\text{-}pT1}$ = 0.011) (**Figure 18E**). The 5 most upregulated genes for each cluster are shown in in **Figure 18C**. Of note, epithelial

cells overexpressed SERPINA1, a protease inhibitor upregulated in multiple cancer types, as well as CD24, recently described as a macrophage immune checkpoint protein [67], whose expression is associated with worse prognosis in ccRCC [68]. Additionally, the epithelial cluster showed upregulation of keratin 16 (*KRT16*) and 8 (*KRT8*), which have also been associated with poor prognosis in ccRCC [69].

VIPER was equally successful in identifying fibroblast, endothelial, and M2 macrophage clusters, while revealing a deeper level of heterogeneity. Specifically, epithelial cells stratified into four distinct clusters (**Figure 18B**). While clusters E1, E3, and E4 were more represented in the tumor as compared to adjacent non-tumor ($p_{E1} = 0.001$, $p_{E3} = 0.056$, $p_{E4} = 0.028$), cluster E2 was more represented in adjacent normal ($p_{E2} = 0.312$) (**Figure 18E**). This population represents normal epithelial cells, whose gene expression was not sufficiently distinct from tumor cells to be effectively stratified without VIPER. For visualization purposes, we show the 5 most differentially active proteins for each cluster (**Figure 18D**).

Further analyses showed that Cluster E1—the most prevalent among the four epithelial clusters—was significantly over-represented in stage 3 tumors as compared to stage 1 tumors ($p_{E1|S3:S1} = 0.018$), while lower-frequency populations E2, E3, and E4 were represented in both stage 1 and stage 3 patients. This analysis suggests that differential frequency of a dominant epithelial cell population whose transcriptional state is virtually identical across patients effectively stratifies Stage 1 vs. Stage 3 tumors (**Figures 18E and 18F**). These data are relevant because transcriptionally distinct tumor cell subpopulations may have differential drug sensitivity and because protein activity-based analysis but not gene expression allowed distinct identification of normal vs. tumor-related cells.

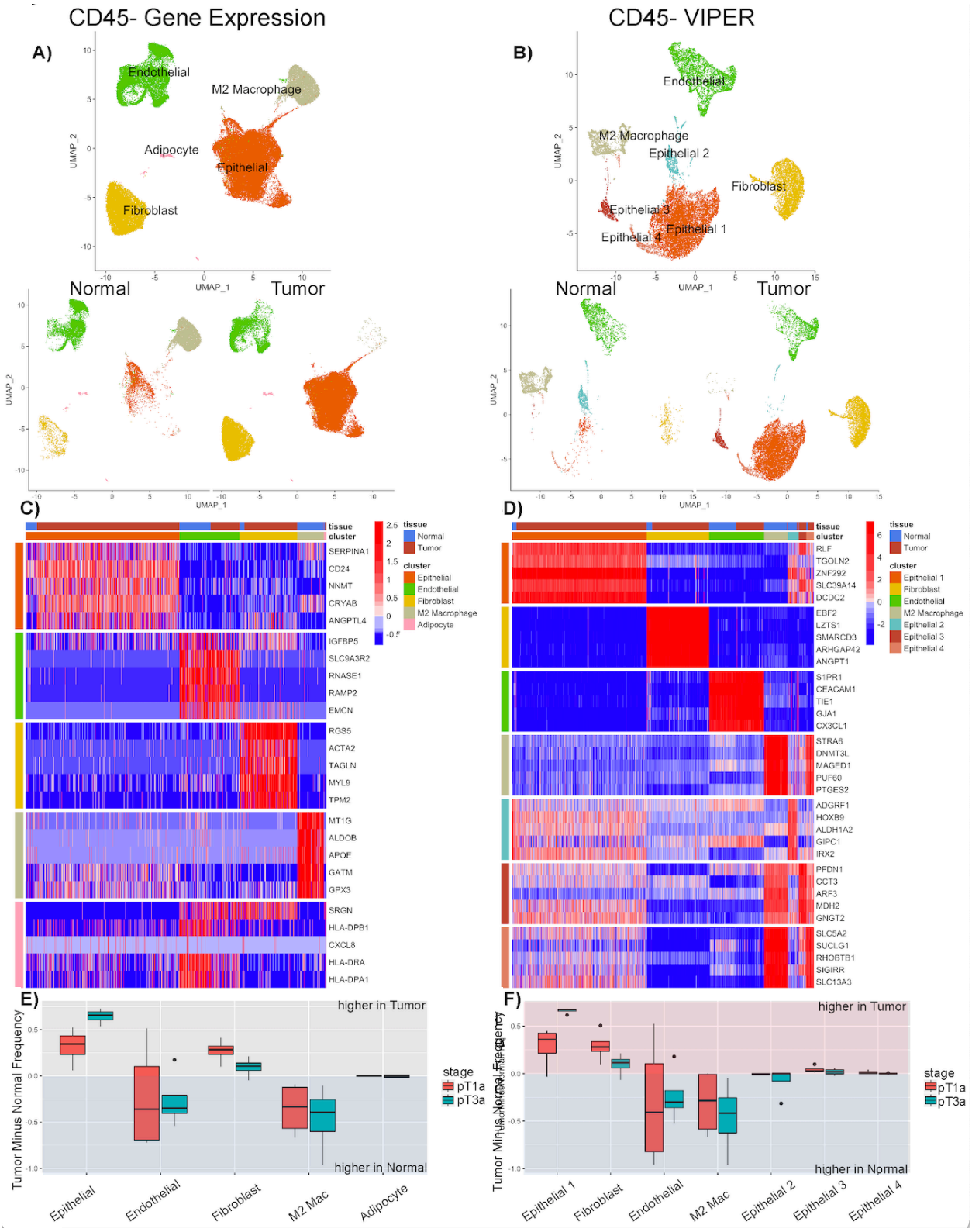
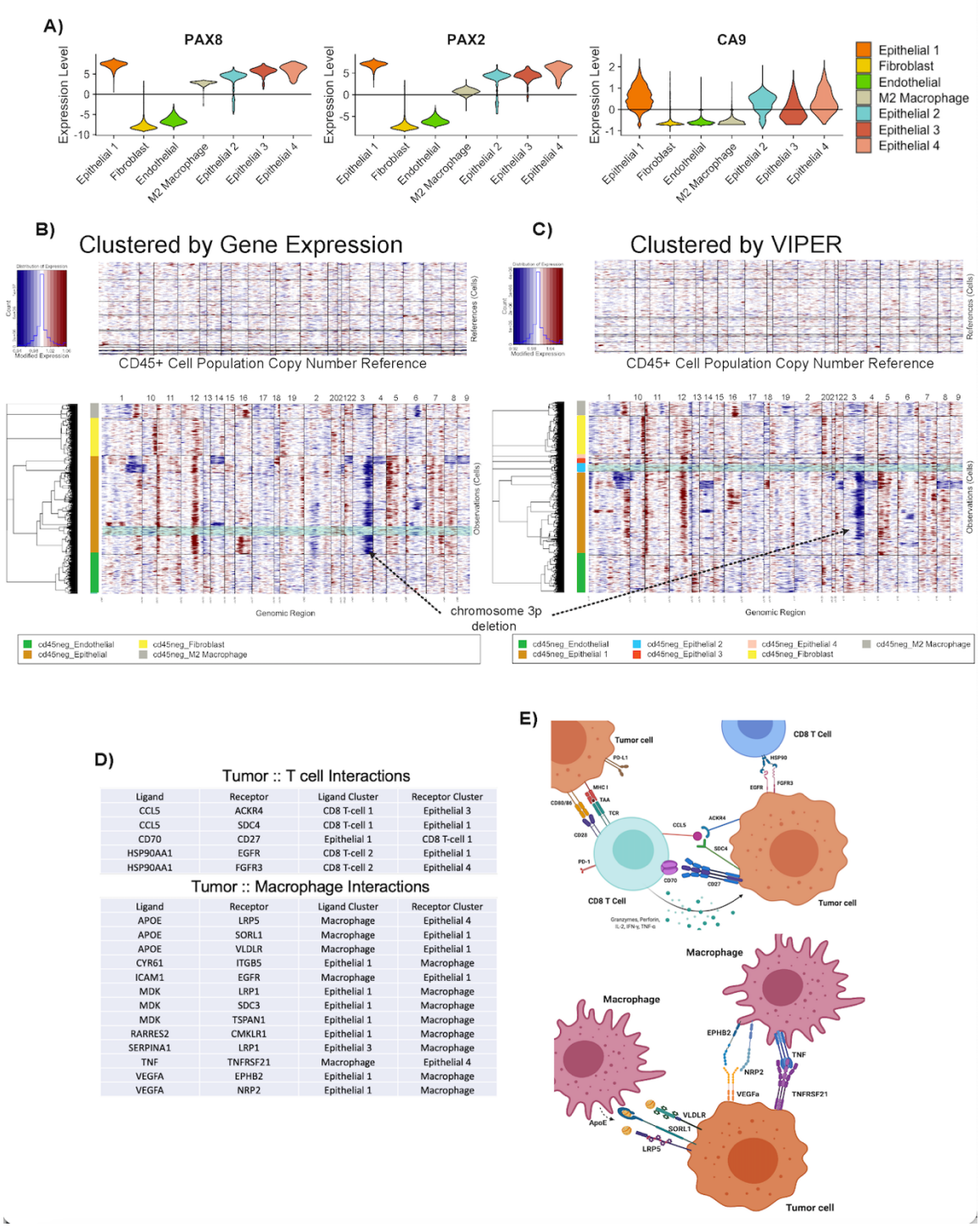


Figure 18: Deep Profiling of CD45- Cells by Gene Expression and Protein Activity Distinguishes Tumor Cells from Normal Epithelium

A) UMAP of single-cell gene expression pooled across all CD45- samples, clusters labelled by cell type. Bottom plot is split by Tumor vs Adjacent Normal label B) UMAP of VIPER-inferred protein activity pooled across all CD45- samples, clusters labelled by cell type. Bottom plot is split by Tumor vs Adjacent Normal. C) Heatmap of top5 differentially upregulated genes for each cluster by expression; each row represents a gene and each column represents a cell. Legend shows cluster identity with cell type inferred by SingleR and Tumor (red) or Adjacent Normal (blue). D) Heatmap of top5 differentially upregulated proteins for each cluster by VIPER-inferred activity. Legend as in C. E) Bar plots of patient-by-patient cluster frequency in Tumor minus frequency in Adjacent Normal for each Gene Expression cluster, grouped by stage; values < 0 (blue) indicate higher frequency in Adjacent Normal, values greater < 0 (red) indicate higher frequency in Tumor. F) Bar plots of patient-by-patient cluster frequency in Tumor minus frequency in Adjacent Normal for each VIPER cluster, grouped by stage, as in E.

Inferred Tumor-Cell Copy Number Alterations are Characteristic of Clear Cell Renal Carcinoma: To further understand the epithelial cell clusters, we assessed VIPER-inferred activity of PAX8, PAX2, and CAIX, proteins expressed in renal epithelium; these are upregulated in malignancy and commonly used as markers for ccRCC [70]. This analysis confirmed increased activity of these markers in epithelial clusters E1, E3, and E4 (**Figure 19A**). To more precisely determine which epithelial clusters represent tumor cells, we performed Copy Number Alteration (CNA) inference clustered by expression-based (**Figure 19B**) or activity-based (**Figure 19C**) analysis. We inferred CNAs for each CD45 negative cell, using CD45 positive cells as normal ploidy controls. The results (**Figure 19C**) showed that aberrant CNA regions are present in epithelial clusters E1, E3, and E4 but not E2, including recurrent 3p chromosomal deletions not detected in any other cell type. Of note, chromosome 3p deletions occur in >96% of all ccRCC patients, as that region contains the VHL tumor suppressor locus [71]. The epithelial cluster inferred by expression-based analysis included cells lacking gross

copy number alterations (highlighted in **Figure 19B**), corresponding exactly to activity-based cluster E2 (**Figure 19C**).



D)

Tumor :: T cell Interactions

Ligand	Receptor	Ligand Cluster	Receptor Cluster
CCL5	ACKR4	CD8 T-cell 1	Epithelial 3
CCL5	SDC4	CD8 T-cell 1	Epithelial 1
CD70	CD27	Epithelial 1	CD8 T-cell 1
HSP90AA1	EGFR	CD8 T-cell 2	Epithelial 1
HSP90AA1	FGFR3	CD8 T-cell 2	Epithelial 4

Tumor :: Macrophage Interactions

Ligand	Receptor	Ligand Cluster	Receptor Cluster
APOE	LRP5	Macrophage	Epithelial 4
APOE	SORL1	Macrophage	Epithelial 1
APOE	VLDLR	Macrophage	Epithelial 1
CYR61	ITGB5	Epithelial 1	Macrophage
ICAM1	EGFR	Macrophage	Epithelial 1
MDK	LRP1	Epithelial 1	Macrophage
MDK	SDC3	Epithelial 1	Macrophage
MDK	TSPAN1	Epithelial 1	Macrophage
RARRES2	CMKLR1	Epithelial 1	Macrophage
SERPINA1	LRP1	Epithelial 3	Macrophage
TNF	TNFRSF21	Macrophage	Epithelial 4
VEGFA	EPHB2	Epithelial 1	Macrophage
VEGFA	NRP2	Epithelial 1	Macrophage

Figure 19: Tumor Cell Labeling is Validated by Copy Number Inference and Tumor Marker Expression

A) ViolinPlots of VIPER-Inferred Activity for ccRCC tumor markers PAX2, PAX8, and CA9. Plots grouped by CD45- cluster label revealing increased expression in epithelial cells. B) CNA Inference for all CD45- populations, using CD45+ cells as reference. Columns represent chromosomal regions and rows represent cells, grouped by Gene Expression cluster, with a subset of copy-number-normal epithelial cells highlighted in green. C) CNA Inference re-grouped by VIPER cluster. Epithelial cell clusters 1, 3 and 4 contain consistent chromosome 3p deletions characteristic of ccRCC, while Epithelial cluster 2, highlighted in green, is grossly Copy-Number normal. D) Table of known receptor-ligand interaction pairs in which ligand is significantly upregulated by Gene Expression in one cluster and receptor is significantly upregulated by VIPER in another. Subset to interactions inferred between Tumor cells and T-cells, or between APOE+/TREM2+/C1Q+ Tumor Macrophages and Tumor cells. E) Visualization of receptor-ligand interaction pairs shown in D.

Dissecting Receptor-Ligand Interactions in ccRCC: A critical challenge that may benefit from VIPER-based protein activity measurements is the elucidation of cross-compartment interactions that may modulate tumor homeostasis. The extensive scRNAseq dataset generated by these studies, which included data from both hematopoietic and non-hematopoietic cells, supported *in silico* interrogation of putative receptor/ligand interactions between cell types. To that end, we identified overexpressed genes encoding for secreted ligands and differential VIPER activity of their cognate binding receptors between all possible subpopulation pairs and across patients. Referencing a curated public database of 2,557 known receptor-ligand interaction pairs [72], we identified interactions supported by significant overexpression of the ligand in any cell population and concomitant activation of the cognate receptor by VIPER in any patient-matched subpopulation. A total of 276 candidate receptor-ligand pairs were identified. Of these, several had been previously established in ccRCC. For example, we identified receptor/ligand pair KDR/VEGFA in tumor cells and endothelial cells, respectively. **Figure 19D** shows a curated

subset of predicted receptor/ligand interactions between tumor cells and tumor-enriched immune populations (T cells and Macrophages). Of these, the potential interaction between CD70 and its cognate receptor CD27, in tumor and CD8 T cells respectively, is of interest, given ongoing investigation of CD70 as a therapeutic target [73]. The majority of interactions thus identified were not previously reported, providing a systematic resource for future studies.

A Tumor-Specific Macrophage Signature is Associated with Disease Recurrence: As above, protein activity analysis identified a novel, tumor-specific sub-population of macrophages and their top differentially active proteins (i.e., LILRB5, APOE, and TREM2) and differentially expressed genes (i.e., C1QA-C, APOE, and TREM2). To assess the clinical significance of this population, we leveraged single-cell ARACNe networks to transform bulk RNA Sequencing data from two independent cohorts (n=8 and n=157) using VIPER. Here, gene expression of each cohort was scaled by the mean and standard deviation of each gene and VIPER was applied. We defined a set of statistically upregulated proteins ($p<0.05$) in the Tumor-Specific Macrophage population and computed Normalized Enrichment Score (NES) of this gene set in the ranked differential protein activity signature of patients with post-surgical disease recurrence compared to those without recurrence. This analysis was first performed in a small (n=8), well clinically-annotated cohort of bulk RNASeq samples from untreated ccRCC surgical resections; here we found a significant enrichment of tumor macrophage signature in 4 patients with recurrence compared to 4 age- and stage-matched controls (**Figure 20A**) (Normalized Enrichment Score NES = 4.08, $p = 4.5\times10^{-5}$). We found that the leading-edge proteins included marker proteins APOE and TREM2, as well as other macrophage-associated proteins of potential clinical interest, such as LILRB5, MERTK, and IGF1 (**Figure 20C**). Sample-by-sample NES of the

Tumor Macrophage Gene Set was computed directly on the ranked VIPER activity of proteins in each bulk RNA-Seq sample, and was consistent with the group-wise analysis in Figure 6A, such that all non-recurrent patients had significant depletion of tumor macrophage markers and recurrent patients had strong enrichment, with the sole exception of a single patient who recurred late (82 months post initial surgery) (**Figure 20D**). To further explore this clinical association, we performed Univariate Cox regression of normalized enrichment scores (NES) versus time to recurrence (TTR) on a patient-by-patient basis (p=0.057). Binary log-rank test of macrophage enrichment, with NES > 0 = "high" and NES ≤ 0 = "low", showed a strong statistically significant association between signature enrichment and shorter time-to-recurrence ($p = 6.7x10^{-3}$) despite a relatively small sample size, suggesting a strong effect (**Figure 20B**). To validate the association of markers representative of this rare population with recurrence, we next measured their enrichment in a larger cohort of bulk RNA-Seq samples from 157 treatment-naïve ccRCC surgical resections, annotated with time to post-surgical recurrence. This validation cohort showed a consistent up-regulation of tumor macrophage markers in patients with post-surgical recurrence (**Figure 20E**) (NES = 4.33, $p = 1.5x10^{-5}$), with a significant cox regression p-value of 0.012 and binarized log-rank p-value of 0.0029 (**Figure 20F**).

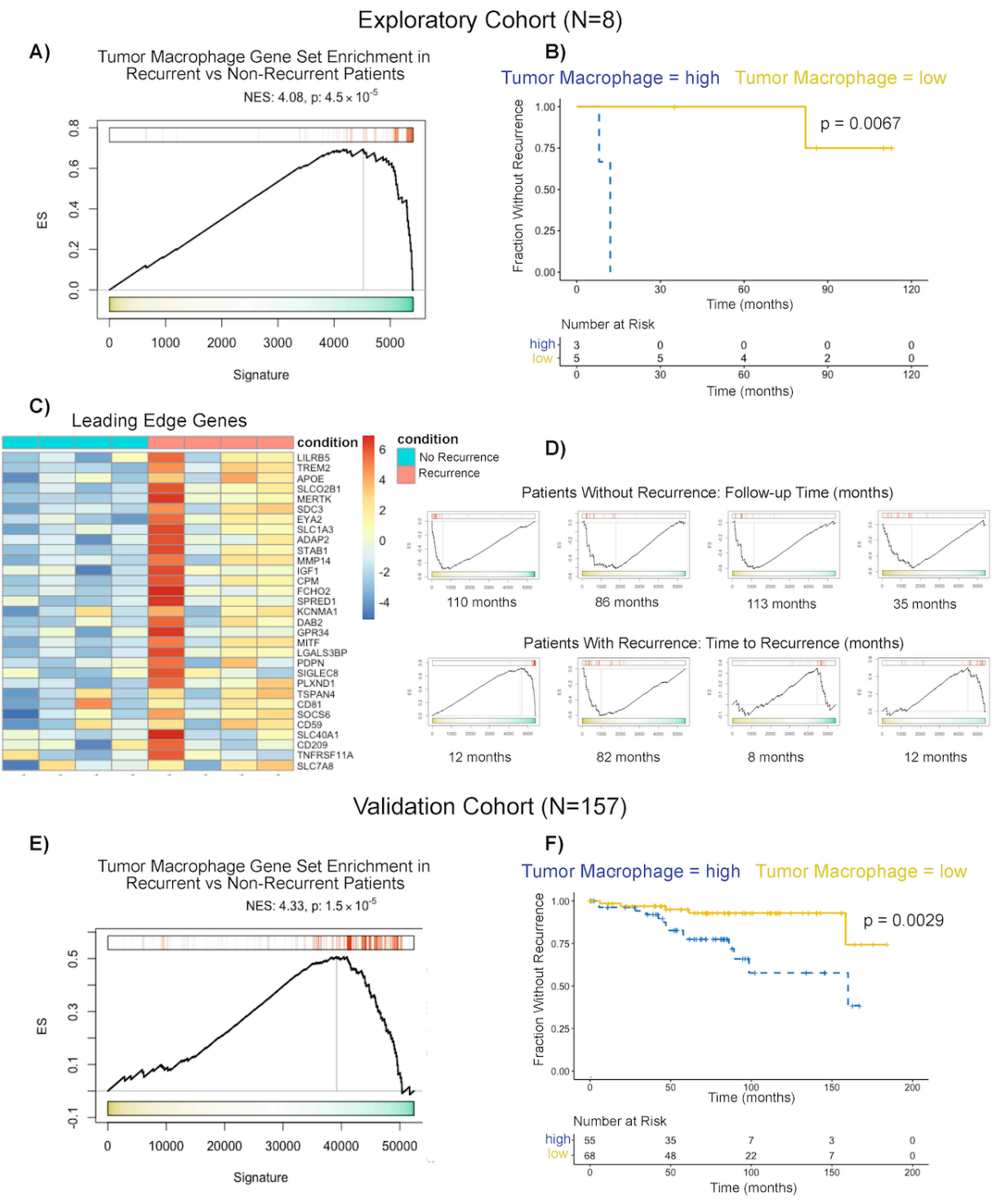


Figure 20: Enrichment of Tumor-Specific Macrophage Markers Defined from Single-cell RNASeq in Bulk RNASeq Data is Associated with Shorter Time-to-Recurrence

A) Gene Set Enrichment Analysis (GSEA) of tumor-specific macrophage marker proteins in VIPER-transformed bulkRNASeq data from 4 patients with post-surgical recurrence vs 4 without. Proteins ranked by fold-change in recurrence versus no recurrence, p-value computed by GSEA vs gene shuffling null model with 1000 permutations. Note enrichment in patients with recurrence (NES=4.08, $p=4.5*10^{-5}$). B) Kaplan-Meier curve of sample-by-sample tumor-specific macrophage GSEA associated with time to recurrence, yellow line indicates patients with low enrichment, blue line indicates patients with high enrichment. Log-rank p-value = 0.0067. C) Heatmap of leading-edge protein set from A. D) Sample-by-sample tumor macrophage GSEA, annotated with each sample's recurrence status and time to recurrence or total observation time. Proteins ranked by inferred activity. E) Macrophage signature GSEA in recurrence vs. no recurrence in validation cohort (N=157). F) Kaplan-Meier curve of sample-by-sample GSEA in association with time to recurrence in the validation cohort, log-rank p-value = 0.0029.

***C1Q/TREM2 Expressing Macrophages are Tumor-Restricted and Associated with Post-Surgical Recurrence*:** We next queried whether markers of the macrophage population associated with poor outcome were co-expressed in cells by Immuno-fluorescence staining and analyzed their spatial localization to determine if these markers were tumor, T cell or macrophage-related. We specifically interrogated C1Q, APOE, and TREM2, as the latter two were identified as protein activity markers strongly associated with clinical outcome (**Figure 20C**) and C1Q was highly overexpressed but could only be identified following VIPER clustering of the single-cell data. For these studies, we developed a fluorescence-based panel that included the three markers, as well as subpopulation specific markers CA9 (tumor cells), CD3 (T cells), and CD69/CD163 (pan-macrophage) (**Figure 21A**). Using this panel, we stained each of the 11 samples interrogated by scRNASeq and quantified expression in multiple segments of tumor and adjacent non-tumor tissue. Both C1Q and TREM2 were strongly enriched in macrophages across all samples (**Figure 21B**).

To determine which macrophage populations were comparatively tumor-restricted, we tested the representation of cells co-staining for C1Q, TREM2, and CD68/CD163 within tumor stroma vs. adjacent normal. In contrast to total CD68/CD163 macrophages, and as predicted by VIPER analysis, C1Q+ and TREM2+ macrophages were significantly tumor restricted, whilst the double positive population (C1Q+, TREM2+) appeared to be almost completely exclusive to tumor tissue (**Figure 21C**). Furthermore, these C1Q+/TREM2+/APOE+ macrophages were localized more closely to tumor cells than control macrophages (C1Q-TREM2-APOE-) with a relative distance to the nearest CA9+ cell of 15.25 um vs 23.28 um, $p = 1.7x10^{-14}$, respectively. The strong tumor restriction of this population did not appear to correlate with tumor stage (**Figure 21D**). To further assess for association with disease recurrence, we stained the samples from the same dataset we had used for initial bulk RNAseq studies (**Figure 20A-D**). These samples showed that both TREM2 and C1Q were significantly enriched in the tumor stroma of patients with disease recurrence as compared to patients without recurrence (**Figure 21E**) ($p_{C1Q} = 0.047$, $p_{TREM2} = 0.038$, $p_{C1Q/TREM2} = 0.009$). C1Q+ macrophages, in particular, were significantly associated with disease recurrence ($p = 0.028$). These data suggest that assessment of intra-tumoral C1Q+ macrophage density by IF (Figure S5) may provide a useful prognostic biomarker for recurrence. We explored this hypothesis by first calculating a cutoff for C1Q macrophage frequency that maximized the log-rank statistic, and next performing log-rank regression (**Figure 21F**). A C1Q+ macrophage frequency threshold of 0.01 significantly separated patients with post-surgical recurrence from those without recurrence, with a log-rank *p*-value of $6.7x10^{-3}$ and Area under the Curve (AUC) of 0.9375. These data recapitulated the disease recurrence Kaplan-Meier curve defined by GSEA analysis (**Figure 20B**), and independently support the association of tumor-infiltrating macrophage density with post-

surgical recurrence, highlighting a strong consistency between IF staining and scRNASeq analysis.

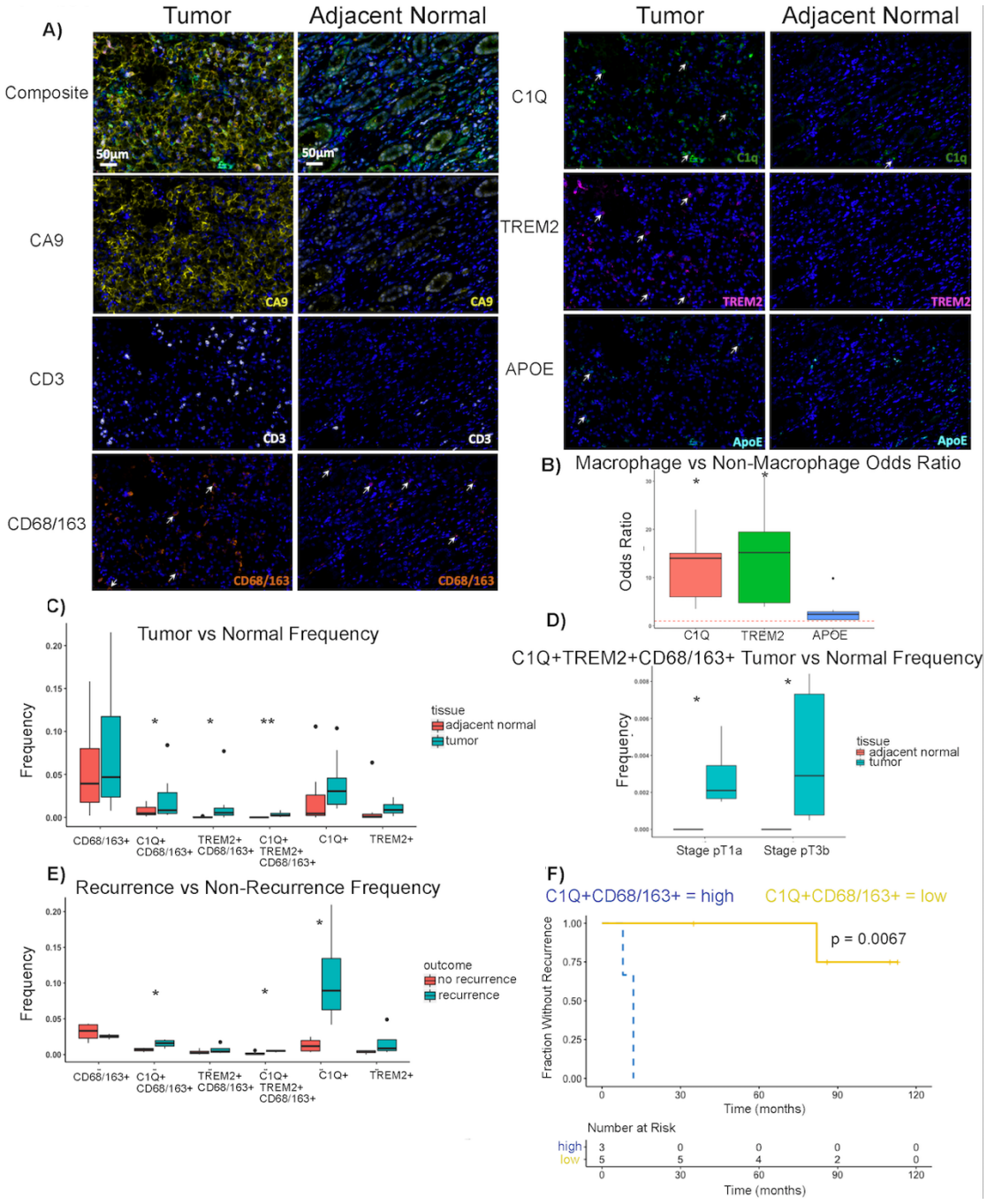


Figure 21: A Novel Population of C1Q/TREM2+ Macrophages are Tumor-Specific and Associated with shorter time-to-recurrence by Immunohistochemistry (IHC)

A) Representative IHC images for each marker in tumor stroma vs adjacent normal. Note high C1Q/TREM2/APOE staining within CA9+ tumor as compared to tumor-adjacent (CA9-) regions. B) Odds ratios (OR) across samples of tumor-specific macrophage markers C1Q, TREM2, and APOE co-staining with CD68/CD163+ macrophage cells vs CD68/CD163- non-macrophage cells, note association of C1Q and TREM2 with macrophage markers. Dotted red line represents OR=1. Individual OR for C1Q and TREM2 co-staining with CD68/CD163 is statistically significant by Fisher's exact test ($p<0.01$). C) Frequency by IHC of C1Q+ or TREM2+ macrophages in tumor stroma vs adjacent normal across the 11 patient samples profiled by scRNASeq. Enrichment in tumor compared to adjacent normal assessed by paired Wilcox test, *<0.05, **<0.01. D) Frequency of C1Q+TREM2+CD68/CD163+ macrophages in tumor vs adjacent normal, plotted by stage (pT1a vs pT3b). No C1Q+TREM2+CD68/CD163+ cells were present in adjacent normal. E) Frequency of C1Q+ or TREM2+ macrophages in tumor stroma of patients with or without post-surgical recurrence, from the cohort profiled by bulkRNASeq in Figure 20A-20D. Higher frequency in patients with recurrence assessed by unpaired Wilcox test, *<0.05. F) Kaplan-Meier plot of C1Q+CD68/CD163+ frequency in association with time to recurrence. Log-rank p-value = 0.0067, with sample-by-sample frequency binarized by log-rank maximization to >0.01 = "high" and <0.01 = "low."

2.4 Discussion

We report a systematic single-cell analysis of the cell populations that comprise the immune and non-immune compartments of clear cell renal cell carcinoma (ccRCC), from >200,000 cells representing tumor and adjacent normal tissue from 11 patients, with either stage 1 or stage 3 disease. By incorporating both transcriptomic and VIPER-based proteomic data, our analysis characterizes sub-populations, key regulatory proteins, and candidate ligand/receptor-mediated interactions, providing a previously unavailable window into the microenvironment of ccRCC.

These studies provide insight that could only be gleaned using our comprehensive VIPER-based scRNASeq protein activity analysis pipeline. In particular, key tumor-specific populations,

comprising both immune and non-immune cells, and their established lineage markers were missed by expression-based cluster analysis and by flow cytometry, due to significant gene dropout effects, a limited set of antibodies, and measurement reproducibility. By contrast, activity-based analyses provided high-resolution sub-structure and revealed a novel tumor-specific macrophage population prognostic for recurrence.

To confirm that activity-based analysis tracked protein expression (quantified using flow cytometry and IF), we performed a comprehensive validation of VIPER results using high-parameter spectral flow cytometry. Our results show that VIPER-based analyses may potentially outperform antibody-based measurements in terms of both detection and reproducibility, while providing quantitative activity assessment for >6,000 proteins in a single experiment. By contrast, gene expression-based analyses of scRNASeq data could not recapitulate flow cytometry results, due to significant gene dropout effects. Thus, a key novel finding of this study is feasibility, accuracy, and reproducibility of network-based protein activity inference from single cell gene expression profile data.

Importantly, activity-based analyses identified several known immune checkpoint and master regulatory proteins missed by gene expression analysis alone. In exhausted CD8 T-cells, for instance, these included LAG-3, PD-1, and CTLA-4, while in Tregs, they included FOXP3 and CTLA-4. Thus, the full set of differentially active regulatory proteins reported represents a previously unavailable resource for the study of these cell types. Analysis of both hematopoietic and non-hematopoietic cells allowed us to study the interaction between tumor-related cells and immune subpopulations in the TME, especially with respect to tumor-infiltrating macrophages.

Inference of interactions among established receptor-ligand pairs [72] identified >200 ligand/receptor-mediated cell-cell interactions, which were consistently yet independently detected across all patients**.** Among multiple potentially significant interactions identified, we highlight CD70, localized to the surface of tumor cells, interacting with its cognate ligand, CD27, on tumor-infiltrating T cells. CD70 is expressed in many solid tumors, including RCC [74], and may facilitate tumor cell escape by inducing tumor cell proliferation and survival [73] - these data establish it as a potential therapeutic target in ccRCC.

In terms of potential clinical relevance, activity-based analysis identified a tumor-specific macrophage subpopulation characterized by upregulation of C1Q, APOE, and TREM2 and high activity of the LILRB5 protein (**Figure 13, Figure 14**). This subpopulation was consistently detected in all tumors, and GSEA analysis of its single-cell RNASeq protein signature in independent bulk RNASeq profiles revealed its significant association with shorter time to post-surgical recurrence. These findings were confirmed in a validation cohort of 157 patients (**Figure 20**). Of note, the VIPER based test to measure activity of these proteins in patients, based on their tumor's mRNA profile (OncoTarget) [44], recently received CLIA-certification by the NY and CA departments of health.

Protein-level qmIF confirmed the clinical significance of C1Q+ tumor-specific macrophages and recapitulated association with shorter time to post-surgical recurrence identified at the transcriptional level (**Figure 21F**). Though the functional role of these tumor-specific macrophages is currently unknown, a recent study in primary renal tumors also found that high density of C1Q-expressing cells correlates with poor prognosis [75] although that study did not

characterize these findings at the single-cell level or report their tumor-specificity and interaction with tumor cells. Functionally, complement-deficient mice developed high densities of C1Q expressing macrophages with concurrent upregulation of immune checkpoints PD-1, LAG-3, and PD-L1. These orthogonal data are consistent with our observations in human single-cell data, as we documented a C1Q-expressing tumor-specific macrophage population and a high frequency of likely exhausted LAG-3+ PD-1+ T cells in the ccRCC TME.

Our studies also highlighted APOE and TREM2 - a member of the immunoglobulin superfamily that plays an important immunomodulatory role in the regulation of inflammatory processes [76] [77] and enhances tumor proliferation [78] [79]. The role of the TREM2-ApoE pathway in RCC tumor biology has not been fully explored. A recent study profiling a murine ccRCC model using scRNASeq in conjunction with intracellular proteomic staining identified a population of TREM2+ tumor-infiltrating macrophages which appears to be phenotypically similar to the population we discovered in patients [80]. Functional studies showed that co-culture of CD8 T-cells with these macrophages significantly impeded T-cell proliferation, and that TREM2 knockdown led to favorable pre-clinical outcomes.

In conclusion, we report the development and application of a novel and broadly generalizable scRNASeq analytic pipeline which complements gene expression with inferred protein activity to comprehensively dissect the repertoire of subpopulations in the TME. While our analysis focused on treatment-naïve clear cell renal carcinoma, our validation with proteins concurrently profiled by flow cytometry suggests that this approach could be effectively applied to any tumor of interest, and potentially to other tissue-based studies. Our scRNASeq data are limited by the

relatively small number of cases, but it should be noted that we cumulatively profiled >200,000 cells with high data quality and that the populations identified were remarkably consistent across patients (**Figure 12**), suggesting that additional patients would not dramatically affect the conclusions of the study. Enrichment of single-cell signatures in bulk data and qmIF studies showed strong and statistically significant association between tumor infiltration by a C1Q-expressing macrophage subpopulation and disease recurrence. One implication of these findings is that ccRCC patients with an increased density of C1Q-expressing macrophages in the tumor stroma at baseline might be at increased risk of post-surgical disease recurrence, and thus may be suitable candidates for adjuvant therapy or more aggressive neoadjuvant approaches in the context of clinical trials. A more intriguing possibility is that these cells could be causal of (rather than associated with) recurrence; hence targeting their top master regulators and/or proteomic markers could be of clinical value in ccRCC.

Limitations of the Study:

Consistent with prior results, we found that recovery rates for protein activity inference using this analysis pipeline were in the 70-80% range, i.e., 20-30% of differentially active proteins may be missed. Although this compares favorably with gene expression, where >80%-90% of genes may be undetected, we expect that future studies aimed at improving the population-specific reporter assays used to infer protein activity by VIPER will address these limitations.

2.5 Methods

Data and Code Availability:

Data files and specific code used to perform all analyses in this manuscript are available at https://github.com/Aleksobrad/single-cell-rcc-pipeline. General pipeline for VIPER analysis of scRNASeq data is available as an actively maintained and updated R package at https://github.com/califano-lab/PISCES. Source data for all analysis in this study will also be publicly hosted on Mendeley at http://dx.doi.org/10.17632/nc9bc8dn4m.1

Experimental Model and Subject Details:

Fresh Surgical Tumor Tissue and Matched Adjacent Normal Tissue were obtained from 11 adult patients with clear cell renal carcinoma undergoing nephrectomy to treat primary, non-metastatic disease, ranging from grade 1 to grade 4 and including 6 patients with stage pT1a disease and 5 patients with pT3a disease. These were dissociated immediately for Single-cell RNASeq and flow Cytometry analysis. Formalin-Fixed Paraffin-Embedded (FFPE) Pathology samples of the same patient tumors were obtained for follow-up Immunohistochemistry analysis. A separate cohort of FFPE tissue from primary tumor of 11 adult patients also treated for clear cell renal carcinoma by surgical nephrectomy was curated retrospectively to identify patients who recurred after surgery and match them by age and tumor stage to those who didn't. This cohort was tracked over a period of 5-113 months, during which time 6 patients experienced disease recurrence (between 5-82 months after surgery) and 5 patients had no recurrence (between 35-113 months after surgery). We used this cohort for validation of Immunohistochemical analysis, as well as for exploratory profiling by bulk RNA sequencing and association of markers identified from single-cell profiling with time-to-recurrence. A larger validation cohort was curated from the Vanderbilt tissue bank, identifying 157 treatment-naïve patients with varying follow-up time and annotation of post-surgical disease recurrence and profling them by bulk

RNA sequencing. The studies were conducted in accordance with the guidelines approved by the Institutional Review Board (IRB) protocols, AAAO5706 and AAAA9967, respectively.

Tissue Dissociation:

Fresh Tumor or Adjacent Normal tissue were minced to 2-4 mm sized pieces in separate 6-cm dishes and digested to single cell suspension using Multi Tissue Human Tumor Dissociation Kit 1 (Miltenyi Biotec) and a gentleMACS OctoDissociator (Miltenyi Biotec) according to the manufacturer's instructions. Dissociated cells from both Tumor and Adjacent Normal tissue were aliquoted for Flow Cytometry Analysis and single-cell sequencing, with 2-3x10^6 cells allocated for flow cytometry and the remainder used for single-cell sequencing. Cells aliquoted for sequencing were stained for Live/Dead (eBioscience, cat#50-112-9035) and CD45 (BioLegend, cat#368524) and then fluorescence-activated cell sorted (FACS) using BD Influx™ cell sorter into a Live CD45 positive population and a Live CD45 negative population, each of which were separately loaded for single-cell RNA sequencing. Boundaries between positive and negative cell fractions were drawn based on single-color stain. An example gating strategy is shown in Figure S1. For the first set of three patients processed (Patients A-C), only the sorted CD45-positive population was further processed for single-cell RNASequencing, and for a second set of eight patients (patients 1-8), both CD45-positive and CD45-negative cells were processed for single-cell RNASequencing.

Single-Cell RNA Sequencing:

Sorted CD45-positive and CD45-negative samples were processed for single-cell gene expression capture (scRNASeq) using the 10X Chromium 3' Library and Gel Bead Kit (10x

Genomics), following the manufacturer's user guide at the Columbia University Human Immune Monitoring Core (HIMC). After GelBead in-Emulsion reverse transcription (GEM-RT) reaction, 12-15 cycles of polymerase chain reaction (PCR) amplification were performed to obtain cDNAs used for RNAseq library generation. Libraries were prepared following the manufacturer's user guide and sequenced on Illumina NovaSeq 6000 Sequencing System. Single-cell RNASeq data were processed with Cell Ranger software at the Columbia University Single Cell Analysis Core. Illumina base call files were converted to FASTQ files with the command "cellranger mkfastq." Expression data were processed with "cellranger count" on the pre-built human reference set of 30,727 genes. Cell Ranger performed default filtering for quality control, and produced for each sample a barcodes.tsv, genes.tsv, and matrix.mts file containing counts of transcripts for each sample, such that expression of each gene is in terms of the number of unique molecular identifiers (UMIs) tagged to cDNA molecules corresponding to that gene. These data were loaded into the R version 3.6.1 programming environment, where the publicly available Seurat package was used to further quality-control filter cells to those with fewer than 10% mitochondrial RNA content, more than 1,500 unique UMI counts, and fewer than 15,000 unique UMI counts. Pooled distribution across all samples of UMI counts, unique gene counts, and percentage of mitochondrial DNA after QC-filtering is shown in Figure S1, with total post-filtering cell counts and median UMIs/cell shown for each individual sample in Table S1.

Single-cell RNA-Seq Gene Expression Processing:

Gene Expression UMI count matrices for each sample were processed in R using the Seurat SCTransform command to perform a regularized negative binomial regression based on the 3000 most variable genes. Each sample was then individually clustered by the Resolution-Optimized

Louvain Clustering Algorithm described below, and within each cluster metaCells were computed for downstream regulatory network inference by summing SCTransform-corrected template counts for the 10 nearest neighbors of each cell by Pearson correlation distance. Normalized datasets for both Tumor and Adjacent Normal tissue across all patients were combined separately for CD45-positive and CD45-negative samples using the FindIntegrationAnchors and IntegrateData functions in Seurat, with the default parameters. The resulting datasets of 102,509 CD45-positive cells from 11 patients and 61,423 CD45-negative cells from 8 patients were projected into their first 50 principal components using the RunPCA function in Seurat, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Differential Gene Expression between clusters was computed by the MAST hurdle model for single-cell gene expression modeling, as implemented in the Seurat FindAllMarkers command, with log fold change threshold of 0.5 and minimum fractional expression threshold of 0.25, indicating that the resulting gene markers for each cluster are restricted to those with log fold change greater than 0 and non-zero expression in at least 25% of the cells in the cluster.

Resolution-Optimized Louvain Clustering Algorithm:

For each clustering step in the analysis, clustering was performed in two steps. The Louvain algorithm as implemented in Seurat uses the FindNeighbors and FindClusters functions, such that the FindClusters function includes a resolution parameter that allows selection of a progressively higher number of clusters as the parameter is increased, which does not control for over-clustering or allow for objective evaluation of cluster purity. Therefore, clustering was

performed with resolution values ranging from 0.01 to 1.0 at intervals of 0.01, and cluster quality was evaluated at each resolution value to select an optimum in this range. For each resolution value, the clustered cells were subsampled to 1000, and silhouette score was computed for these 1000 cells and their cluster labels. For gene expression data, Pearson correlation was used as the distance metric in computation of silhouette score, and for VIPER-inferred protein activity data ViperSimilarity as implemented in the VIPER package was used as the distance metric. This procedure was repeated for 100 random samples of 1000 cells to compute a mean and standard deviation of average silhouette score at each resolution value. The highest resolution value that maximizes mean silhouette score was selected as the optimal resolution at which to cluster the data without over-clustering.

Semi-Supervised Cell Type Calling:

For each single cell gene expression sample, cell-by-cell identification of cell types was performed using the SingleR package and the preloaded Blueprint-ENCODE reference, which includes normalized expression values for 259 bulk RNASeq samples generated by Blueprint and ENCODE from 43 distinct cell types representing pure populations of stroma and immune cells [49] [50]. The SingleR algorithm computer correlation between each individual cell and each of the 259 reference samples, and then assigns both a label of the cell type with highest average correlation to the individual cell and a p-value computed by wilcox test of correlation to that cell type compared to all other cell types. Projection of cell-by-cell SingleR labels with $p<0.05$ onto the Gene Expression UMAP space is shown in Obradovic et al. (Supplemental Figure S3), such that localization of SingleR labels is highly concordant with the unsupervised clustering. Unsupervised Clusters determined by the resolution-optimized Louvain algorithm are

labelled as a particular cell type based on the most-represented SingleR cell type label within that cluster.

Regulatory Network Inference:

From each sample, metaCells were computed within each gene expression cluster by summing SCTransform-corrected template counts for the 10 nearest neighbors of each cell by Pearson correlation distance. 200 metaCells per cluster were sampled to compute a regulatory network from each cluster in each patient. All regulatory networks were reverse engineered by the ARACNe algorithm. ARACNe was run with 100 bootstrap iterations using 1785 transcription factors (genes annotated in gene ontology molecular function database as GO:0003700, "transcription factor activity", or as GO:0003677, "DNA binding" and GO:0030528, "transcription regulator activity", or as GO:0003677 and GO:0045449, "regulation of transcription"), 668 transcriptional cofactors (a manually curated list, not overlapping with the transcription factor list, built upon genes annotated as GO:0003712, "transcription cofactor activity", or GO:0030528 or GO:0045449), 3455 signaling pathway related genes (annotated in GO biological process database as GO:0007165, "signal transduction" and in GO cellular component database as GO:0005622, "intracellular" or GO:0005886, "plasma membrane"), and 3620 surface markers (annotated as GO:0005886 or as GO:0009986, "cell surface"). ARACNe is only run on these gene sets so as to limit protein activity inference to proteins with biologically meaningful downstream regulatory targets, and we do not apply ARACNe to infer regulatory networks for proteins with no known signaling or transcriptional activity for which protein activity may be difficult to biologically interpret. Parameters were set to zero DPI (Data Processing Inequality) tolerance and MI (Mutual Information) p-value threshold of 10^{-8},

computed by permuting the original dataset as a null model. Each gene list used to run ARACNe is available on github, along with the generated patient-by-patient ARACNe tables organized into CD45-positive and CD45-negative clusters.

Protein Activity Inference:

Protein activity was inferred for CD45-positive cells from each patient by running the metaVIPER algorithm with all CD45-positive ARACNe networks across all patients on the SCTransform-scaled and Anchor-Integrated gene expression signature of single cells from each patient. Because the SCTransform-scaled gene expression signature is already normalized, VIPER normalization parameter was set to "none." The resulting patient-by-patient VIPER matrices were combined by sub-setting to the VIPER proteins for which activity was inferred in each patient sample, resulting in 2,562 proteins with successfully inferred activity across all CD45-positive patient samples. For CD45-negative single cells, protein activity was inferred by running the metaVIPER algorithm with all CD45-negative ARACNe networks across all patients in the same way, and then taking the intersection of 2,667 proteins with successfully inferred activity across all CD45-negative patient samples. VIPER-Inferred Protein Activity matrices were loaded into a Seurat Object with CreateSeuratObject, then projected into their first 50 principal components using the RunPCA function in Seurat, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Differential Gene Expression between clusters identified by resolution-optimized Louvain was computed using bootstrapped t-test, run with 100 bootstraps, and top proteins for each cluster were ranked by p-value.

Copy Number Inference:

Copy Number Alteration (CNA) across CD45-negative cells was inferred from gene expression counts at the single cell level using the InferCNA package. Cells were clustered according to their unsupervised clustering label by either gene expression or VIPER. The entire set of CD45-positive cells was taken as a reference set to infer CNAs shown for each of the CD45-negative populations.

Flow Cytometry Data Acquisition:

From each of the 8 patient samples profiled by single-cell sequencing of both the CD45-positive and CD45-negative cells, an aliquot of roughly $2x10^6$ cells was taken for staining and high-throughput flow cytometry on CyTEK Aurora flow cytometer. Cells were stained for 10 minutes with Zombie NIR dye (1:1000 concentration), then stained with surface antibodies for 30 minutes on ice protected from light. After washing, cells stained with myeloid panel antibodies were run fresh on the cytometer. Cells stained with the lymphoid panel were fixed with the FoxP3 Fix/Perm kit (ThermoFisher) for at least 30 minutes, then stained with intracellular markers for 30 minutes on ice protected from light. All antibodies used can be found in Supplementary Table X. For both panels, single stain reference controls were created using UltraComp eBeads (ThermoFisher). Due to poor staining quality in Patient 8, samples from this patient were excluded in downstream analysis of flow cytometry data. Data was evaluated by multi-dimensional analysis in R, and follow-up manual gating was performed as shown in Figure 2 using FlowJo v10.5.3.

Multi-Dimensional Analysis of Flow Cytometry Data:

Flow cytometry samples from all samples were combined and gated on Live CD45-positive, then the gated .fcs files for both lymphoid and myeloid antibody panel were separately exported from FlowJo software and analyzed in R with the flowCore and ggcyto packages. For each panel, raw fluorescence data were normalized with the estimateLogicle and transform functions, and in order to reduce computational burden of downstream analysis a sampled set of normalized fluorescence data from 250,000 cells were then loaded into a Seurat object with CreateSeuratObject. Two-dimensional representation of these data was computed by RunUMAP and resolution-optimized Louvain clustering was performed. Fluorescence of all markers was visualized in a heatmap, with cells grouped by cluster. For single-cell sequencing data, the CD45-positive gene expression matrix and VIPER-inferred protein activity matrix were each subset to genes corresponding to the proteins profiled by flow cytometry, and re-clustered by the resolution-optimized Louvain algorithm. Side by side comparison of the clustering and heatmaps for flow cytometry protein expression, gene expression, and inferred protein activity are shown in Figure 17.

Receptor-Ligand Interaction Inference:

A curated database of 2,557 known receptor-ligand interaction pairs was downloaded from the RIKEN FANTOM5 database. This list of receptor-ligand pairs was subset to pairs for which the ligand is significantly upregulated by gene expression in at least one VIPER cluster across patients and the receptor is significantly upregulated by protein activity in at least one VIPER cluster across patients. This reduced the total set of receptor-ligand pairs detected in our dataset to 276. For each pair we annotate a ligand cell type with highest median gene expression and a receptor cell type with highest median protein activity. Filtering to interactions involving the

Tumor macrophage cluster and any T cell cluster returns 5 interaction pairs, filtering to interactions involving any Tumor cell clusters and any T cell cluster returns 5 interaction pairs, and filtering to interactions involving any Tumor cell and the Tumor macrophage cluster returns 13 interaction pairs.

Staining for Multiplex Immunohistochemistry:

After consulting with a pathologist, patient FFPE tissue blocks with at least 50% tumor were chosen for sectioning on to Superfrost™ slides. Representative full section 4 µm slides of tissue specimens were stained for H&E and viewed by the pathologist to determine areas of tumor, stroma, regression, and immune infiltrates. Each patient's tissue specimen was then stained using Opal™ 7-color multiplex IHC kit, according to the manufacturer's protocol (Akoya Biosciences) with minor modifications. Briefly, the slides were baked at 60°C for approximately 2 hrs before de-paraffinization and retrieval of antigen at pH 9. The slides were then blocked using 3% hydrogen peroxide (in 1X Tris Buffer with 0.05%Tween20), followed by an additional block using the antibody diluent, before staining with the primary antibodies, which include (in the order of staining) TREM2 (clone-D8I4C, Cell Signaling, cat# 91068S, 1:400, AR9), C1q (clone-C1QA/2956, AbCam, cat#ab268120, 1:100, AR6), CD3 (clone-LN10, Leica, cat#NCL-L-CD3-565,1:100, AR6), ApoE (clone – D17N, Cell Signaling, cat#13366S, 1:300, AR6), CA9 (polyclonal, AbCam, cat#ab15086, 1:1000, AR9) and CD68 (clone – KP1, BioGeneX, cat#AM416-5M, RTU, AR6) along with CD163(clone – 10D6, AbCam, cat#ab74604, ready-to-use (RTU), AR6). For each staining cycle, the slides were first incubated with primary antibody, followed by the secondary HRP-polymerization, and signal amplification using Tyramide conjugated to an Opal fluorophore and microwave treated in the AR6 or AR9 buffer as required

by the next round of primary antibody staining. Single color controls for each fluorochrome and an unstained slide were processed in the same batch and used to create the library for spectral unmixing.

Multispectral Imaging:

For each patient specimen, slides were scanned using Vectra 3 (PerkinElmer), with nine representative areas chosen for multispectral imaging – (i) areas with 50% tumor and 50% stroma, (ii) areas with >90% tumor, and (iii) one area with >90% adjacent normal, wherever possible. These images were factored equally for each patient during analysis using InForm™ software (PerkinElmer). Single stained slides and unstained slides were used for building the spectral library and for unmixing, taking autofluorescence spectrum of patient tissue into account.

Image analysis:

Using the spectral library, the nine representative regions for each patient were spectrally unmixed before manual tissue segmentation using inForm software (Version 2.6, PerkinElmer). Tissue segmentation included highlighting examples of CA9+ renal tumor tissue, classifying the CA9- highly cellular regions as stroma; and the spatially distant CA9- tubular regions as adjacent normal. This trained the InForm algorithm to characterize each of the three tissue types and segment all the corresponding regions for each patient sample. Cellular components were then identified for each cell using the DAPI nuclear counter stain to define the nucleus; and CD3 and C1q stains to detect the associated membrane and cytoplasm, respectively. Using DAPI, we adjusted the nuclear splitting intensity to prevent incorrect identification and quantification of

cells due to clumping. Individual cells were then phenotyped manually on the basis of their staining as Tumor cells (CA9+, yellow), macrophages (CD68+ or CD163+, orange), T cells (CD3+, white), ApoE+ (aqua), TREM2+ (magenta) and C1q+ (green). A training set was defined for the InForm algorithm of around 30 cells for each phenotype, from which we were able to distinguish between the cell densities across all tissue types within the nine fields for each patient. The cells were then scored for the staining intensities of each individual marker, including co-expression across the three tissue types – tumor, stroma and adjacent normal, and threshold fluorescence value of positive staining vs background was computed for each marker by the InForm software. The data from each field was compiled to summarize the position, phenotype and density of cells for each patient. Data were further analyzed in R version 3.6.1 using the phenoptr package, such that all fields for each patient sample were combined into a single data frame with cell-by-cell annotation of classified tissue context (tumor, tumor stroma, or adjacent normal), and fluorescence intensity of all markers.

Co-staining of C1Q, TREM2, and APOE with known macrophage markers was determined by generating contingency tables of C1Q, TREM2, or APOE positive cells with CD68/CD163 positive cells, and testing for statistical over-representation of C1Q/TREM2/APOE on macrophages by Fisher's Exact Test. Odds ratios of co-staining with CD68/CD163+ vs CD68/CD163- cells were computed across all 11 patients in the cohort profiled by single-cell RNASeq, shown as a boxplot in Figure 21B. Cell counts were computed and normalized in the tumor stromal and adjacent normal tissue contexts for each combination of C1Q+/TREM2+/APOE+ Macrophages, defined by positive staining for DAPI and CD68/CD163 and negative staining for CD3 and CA9. Frequencies of each cell population in tumor stroma vs

adjacent normal tissue were compared by paired Wilcox test for the 7 patients in which regions of both tumor stroma and adjacent normal tissue were identified on the same stained tissue slices (Figure 21C, 21D).

In the separate validation cohort of 8 patients for which 4 experienced early post-surgical recurrence and 4 did not, frequency of each combination of C1Q+/TREM2+/APOE+ Macrophage cells was computed in the same way as described above, and frequencies of each population in tumor stroma of recurrent vs non-recurrent patients were compared by unpaired Wilcox test (Figure 21E). Since C1Q+ cells and C1Q+CD68/CD163+ cells were significantly enriched in tumor stroma of patients with early recurrence, fraction of cells staining for these markers was tested for association with time-to-recurrence. Threshold for defining high vs low fraction of cells positive for these markers was determined by maximization of the log-rank statistic, such that frequency of C1Q+ cells > 0.02 was determined to be high C1Q+ and frequency of C1Q+CD68/CD163+ cells >0.01 was determined to be high C1Q+CD68/CD163+. Kaplan-Meier curve was plotted for each population, with statistical significance assessed by log-rank test.

Association of Tumor Macrophage Signature with Clinical Recurrence:

A protein signature for the Tumor-Specific Macrophage cluster was defined based on proteins differentially upregulated in the VIPER macrophage cluster (see Supplemental Table S2 in Obradovic et al. for gene and protein marker lists defining each VIPER cluster). In the dataset of FFPE samples profiled by bulkRNASeq that had been followed for time-to-recurrence after nephrectomy, outlier samples with low total read-counts were filtered out, and signature of

remaining patients with recurrence (time-to-recurrence 8 months, 12 months, 12 months, and 82 months) vs patients without recurrence (observation period 35 months, 86 months, 110 months, and 113 months), was computed by z-score scaling of log10(TPM) normalized counts. Protein activity was computed from gene signature by VIPER using the CD45+ ARACNe networks inferred from single-cell data. Enrichment of the Tumor-Specific Macrophage protein marker set in the VIPER-transformed signature of recurrence vs no recurrence from bulkRNASeq was computed by Gene Set Enrichment Analysis (GSEA), with normalized enrichment score and p-value determined by 1000 random permutations of gene labels. Activity of proteins in the leading edge of the enrichment was plotted sample-by-sample in a gene expression heatmap. Sample-by-Sample Normalized Enrichment Scores were also computed by ranking proteins in each sample according to decreasing activity. Cox regression of the raw normalized enrichment scores against time to disease recurrence was performed. Normalized enrichment scores for each sample were then binarized to less than zero (low) or greater than zero (high), and Kaplan-Meier curve showing association with time to recurrence was plotted along with the binarized log-rank p-value.

Results were further validated by repeating the sample-by-sample gene set enrichment of VIPER macrophage markers in a larger cohort of 157 patients profiled by bulk-RNASeq, where enrichment of macrophage signature was associated with shorter time to post-surgical disease recurrence with log-rank p-value of 0.0029. This analysis was performed using the ggsurvplot and survminer packages in R and is shown in Figure 20. Validation of tumor-specific macrophage association with time-to-recurrence was also performed by immunohistochemical staining of FFPE tissue from the same 8 patients analyzed by RNA sequencing.

Immunohistochemical staining and fluorescence thresholding was performed as described above, and proportion of C1Q, TREM2, APOE, and CD68/CD163 positive cells in the tumor stroma was compared in recurrent vs non-recurrent patient samples, along with the proportions of cell co-staining for every combination of those markers. Significance of the difference in frequency between recurrence and non-recurrence samples was assessed by unpaired Wilcox test. Cell populations with significant difference in staining between the two groups were further assessed by log-rank regression against time-to-recurrence. Frequency threshold for high vs low level of staining was determined by maximizing the log-rank statistic, and Kaplan-Meier curve associating IHC staining with time-to-recurrence was generated, shown in Figure 21F.

Quantification and Statistical Analysis:

All quantitative and statistical analyses were performed using the R computational environment and packages described above. Differential gene expression was assessed at the single-cell level by the MAST single-cell statistical framework as implemented in Seurat v3 [32], and differential VIPER activity was assessed by t-test, each with Benjamini-Hochberg multiple-testing correction. Comparisons of cell frequencies were performed by non-parametric Wilcox rank-sum test, and survival analyses were performed by log-rank test. In all cases, statistical significance was defined as an adjusted p-value less than 0.05. Details of all statistical tests used can be found in the corresponding figure legends.

Chapter 3: Single-Cell Protein Activity Enables Personalized Drug Sensitivity Inference Targeting Individual Tumor Cells

3.1 Summary

Cholangiocarcinoma is an aggressive tumor type with extremely limited treatment options. Furthermore, since cholangiocarcinoma is a heterogeneous tumor with significant stromal involvement, it is difficult to isolate the tumor cell transcriptional profile for personalized treatment by traditional RNA-Sequencing approaches. Here, we present the first comprehensive single-cell RNA-Sequencing profile of a cholangiocarcinoma patient from which we have developed and applied a novel framework for single-cell drug sensitivity prediction, building on CLIA-approved OncoTreat and OncoTarget algorithms. We have found the tumor micro-environment to be heavily infiltrated by immune cells, with T-cells comprising nearly half of all cells and tumor cells representing fewer than 10%. However, isolation of tumor cells reveals three distinct sub-populations, with a set of five candidate drugs predicted to target them all to varying extents. Validation in a patient-derived xenograft model identified Plicamycin and Dacinostat as drugs able to effectively control tumor growth rate in vivo. This work identifies Plicamycin and Dacinostat as promising candidates for follow-up trials in cholangiocarcinoma, alone or in combination with anti-PD1 immunotherapy and current standard-of-care chemotherapies. Furthermore, we present a flexible pipeline for prediction of drug sensitivities at single-cell-resolution, with potential application in precision medicine across a broad range of tumor types.

3.2 Introduction

Cholangiocarcinoma (CCA) is an aggressive biliary adenocarcinoma with both intrahepatic (iCCA) and extrahepatic subtypes, and the extrahepatic subtype is further characterized as either perihilar (pCCA) or distal (dCCA). Collectively, they account for up to 20% of newly diagnosed primary hepatic tumors each year, making CCA the second most common hepatic malignancy after hepatocellular carcinoma (HCC) [81].

A diagnosis of CCA carries with it a very poor prognosis, with a median survival of only 12-37.4 months [82]. The American Cancer Society reports that even localized disease is associated with high mortality, given 5-year survival rates of 15% and 24% for extra- and intrahepatic CCA, respectively, and metastatic CCA offers only a 2% 5-year survival rate [83].

For patients with non-resectable disease, management differs depending on anatomic subtype, though no treatment option is considered curative. Localized, unresectable iCCA is sometimes treated with locoregional therapies, such as trans-arterial chemoembolization (TACE), yet median overall survival of these patients is still only 12-15 months. Certain patients diagnosed with pCCA may receive liver transplantation following neoadjuvant chemotherapy, with 5-year disease-free survival rates reportedly as high as 65%, though very few patients meet the criteria for transplantation. However, for the majority of patients with non-resectable or advanced cholangiocarcinoma, cytotoxic chemotherapy with gemcitabine-cisplatin is first-line, and multiple studies report a median survival of only about 11 months with this regimen [82].

Despite rising incidence and lack of effective therapies, studies of the cholangiocarcinoma micro-environment and intratumoral heterogeneity have been limited to immunohistochemical characterization of individual components of the stroma, especially fibroblasts, which produce an extracellular matrix so prominent that it outweighs the tumor component [84]. The extent of infiltration by diverse stromal cells further presents an obstacle to transcriptional profiling of patient tumors, which are likely to be heavily influenced by the extent of infiltration by various stromal components rather than reflective of the transcriptional state of the tumor itself.

High-throughput droplet-based single-cell RNA Sequencing (scRNASeq) has recently emerged as a valuable tool to characterize the diverse cellular subpopulations that comprise the tumor microenvironment, with the ability to identify representative gene expression signatures from thousands of individual cells in a single sample [32] [16]. In contrast to traditional bulk RNASequencing, scRNASeq can provide a rough characterization of the transcriptional state of individual cell types contributing to emergence of specific tumor phenotypes, thus potentially highlighting the role of rare populations, whose gene expression signature would be diluted below the limits of detection in bulk samples [20]. Furthermore, single-cell RNASequencing enables the characterization of transcriptional heterogeneity among tumor cells themselves, whereas bulk sequencing even on purified tumor cells captures only the average gene expression of the entire tumor. In contrast to antibody-staining approaches, scRNASeq generates a transcriptome-wide profile of each individual cell, without manual selection of predefined proteomic markers. The value of scRNASeq has been demonstrated in recent studies of melanoma [18] [17] and breast cancer [34]. However, no single-cell RNA Sequencing studies have been so far performed in cholangiocarcinoma.

One significant limitation of scRNASeq is that the physical limits of RNA capture efficiency and total RNA molecules per cell dramatically reduce the number of unique genes detected in any single cell, especially from patient-derived tumor tissue. As a result, single-cell gene expression matrices are extremely sparse, with ~80% - 90% of genes undetected by even one read, on average. We have successfully addressed this issue by applying an algorithm (VIPER), which leverages highly-multiplexed, fully tissue-specific gene-reporter assays to accurately measure the activity of up to ~6,500 regulatory proteins, including transcription factors (TFs), co-factors (co-TFs), signaling proteins (SPs), and surface markers (SMs), based on the expression of their downstream regulatory targets (*regulon*) [21]. Tissue specific regulons are inferred directly from analysis of single-cell populations, using ARACNe [23], an information theoretic algorithm that has been experimentally validated in dozens of tissue contexts, with a >70% accuracy in target identification [40]. By integrating the expression of ~100 downstream targets per regulon, on average, VIPER can effectively measure even the activity of proteins whose encoding gene is undetected in scRNASeq data, thus virtually eliminating gene dropouts [26]. The VIPER approach to single-cell data analysis has been shown in clear cell renal carcinoma (ccRCC) to improve resolution of both tumor and immune cell phenotypes with dramatic improvement in signal-to-noise for individual regulatory genes of interest [35].

Successful application of VIPER to single-cell data also allows the adaptation of two algorithms previously developed for application on bulk RNA-Sequencing data, OncoTarget and OncoTreat, to single-cell RNA-Sequencing [24]. OncoTarget infers druggable proteins with increased VIPER activity relative to a standardized external reference, and OncoTreat leverages a database

of large-scale drug perturbation RNA Sequencing assays on tumor cell lines to infer drugs that invert overall VIPER activity profile of an individual patient tumor relative to external reference. These are both CLIA-certified algorithms for personalized drug prediction, now adapted for the first time to the level of individual tumor cells. The value of extending these algorithms to the single-cell level is that single-cell RNA-Seq allows the purification of tumor cells among immune and stromal cell infiltrate, significantly reducing transcriptional noise, and enabling identification of tumor cell heterogeneity and prediction of drugs targeting the largest number of patient-specific tumor cell sub-populations comprising a tumor mass.

Here we present a case study of the first comprehensive profiling of cholangiocarcinoma tumor micro-environment at the resolution of single-cell RNA-Sequencing, and the first application of a unique OncoTarget and OncoTreat approach to assess potentially actionable drug targets at the single-cell level. This has significant implications for the application of single-cell RNA-Sequencing in precision medicine for personalized treatment of cholangiocarcinoma and other highly treatment-resistant malignancies, as well as for improving the understanding of treatment resistance mechanisms in the cholangiocarcinoma micro-environment.

3.3 Results

Clinical Presentation:

The patient first presented to the Emergency Department with signs and symptoms of obstructive jaundice and a 6.8cm lesion in the gall bladder fossa. Fine-needle aspiration was positive for adenocarcinoma and PET CT showed one PET avid lesion in the hilum and non-avid sub-centimeter lung nodules. Two months after initial presentation, the patient underwent a central

hepatectomy with cholecystectomy and excision of bile duct tumor, which was collected for dissociation and single-cell RNA Sequencing. Pathology at that time revealed adenocarcinoma, biliary type, moderately to poorly differentiated (grades 2-3), with contiguous involvement of multiple segments of the biliary tree, including common bile duct, gall bladder, cystic duct, left and right hepatic bile ducts, liver, and perihilar soft tissue. There was also extensive lymphovascular and perineural invasion. Tumor staging was pT3N1M1, given 2/5 regional lymph nodes were positive for adenocarcinoma, and there were distant metastases noted on the falciform ligament. Immunohistochemistry was negative for HER2 overexpression, and PD-L1 combined positive score was 10. Next-generation sequencing of 467 cancer-associated genes showed a TP53 mutation as well as multiple variants of uncertain significance (EPHA5, STAT3, FAT1), with an intermediate tumor mutational burden (3.15 mutations/Mb). The tumor was microsatellite stable, and CA19-9 tumor marker levels were regularly measured to track tumor progression. The patient was initiated on combination gemcitabine, cisplatin, and paclitaxel, which were continued for seven 21-day cycles.

Cholangiocarcinoma microenvironment is highly immune-infiltrated:

Single-cell RNA-Sequencing revealed significant tumor heterogeneity at the gene expression level, identifying 2,738 cells of high data quality grouped into 8 major clusters (Figure 22A-B).

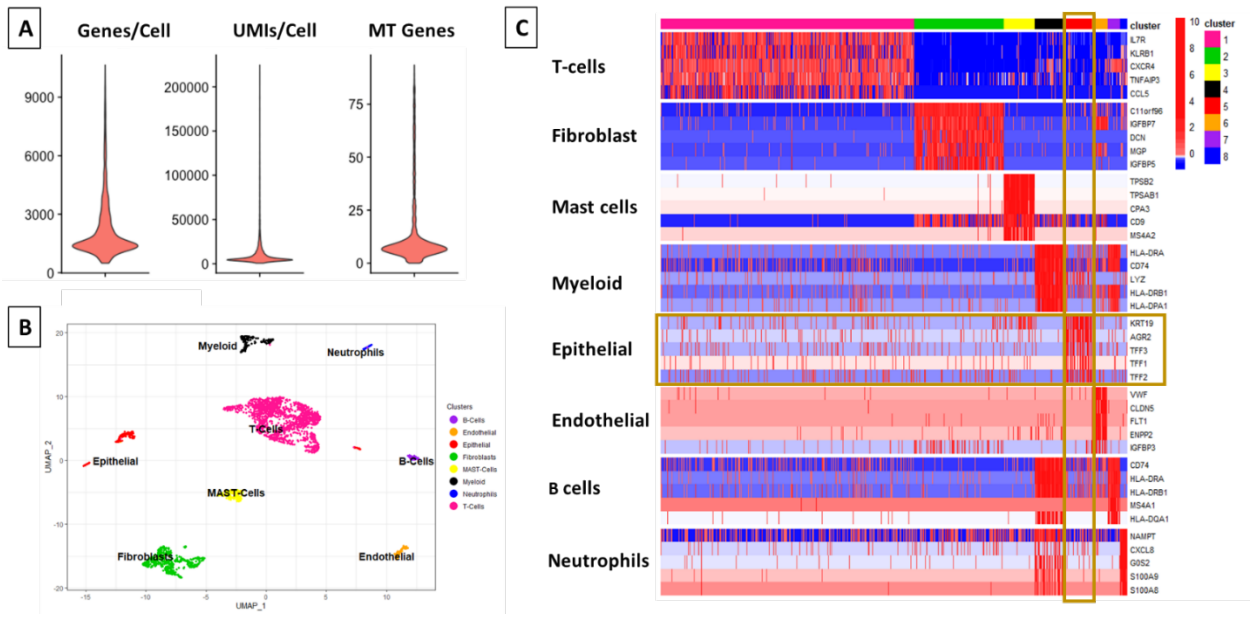


Figure 22: Data Quality and Cell Types Represented in Cholangiocarcinoma Tumor Micro-Environment by Gene Expression

(A) Violin plots showing the distributions of the number of genes detected in each cell, the number of UMIs per cell and the percentage of UMIs in MT genes. These are the QC metrics used to filter out low quality cells from the dataset. **(B)** UMAP projection showing the results of the unsupervised clustering analysis performed at gene expression level (Louvain algorithm). Cell types are inferred for each cluster by SingleR. **(C)** Heatmap containing top-5 differentially upregulated genes in each cluster. The markers of each cluster are identified comparing the average expression of all the genes in one cluster versus the average expression of the same genes in the rest of the cells (MAST test).

Differentially upregulated genes in each cluster are shown in Figure 22C, with cell type inferred by SingleR, which correlates cell-by-cell expression with a sorted bulkRNA-Seq reference database. Most strikingly, the largest cluster by cell counts consisted of Tumor-Infiltrating T-cells, representing over half of all cells profiled by single-cell RNASeq. By VIPER, we are able to further sub-cluster these T-cells, identifying cytotoxic CD8 T-cells with high inferred activity of GZMB as the predominant population, as well as a cluster of CD4 T-cells with high activity of

CCR7, consistent with a central memory phenotype, as well as VIM, a marker of tissue residency. Finally, there is a population of activated CD4 T-cells with high STAT4 signaling (Figure 23A-B).

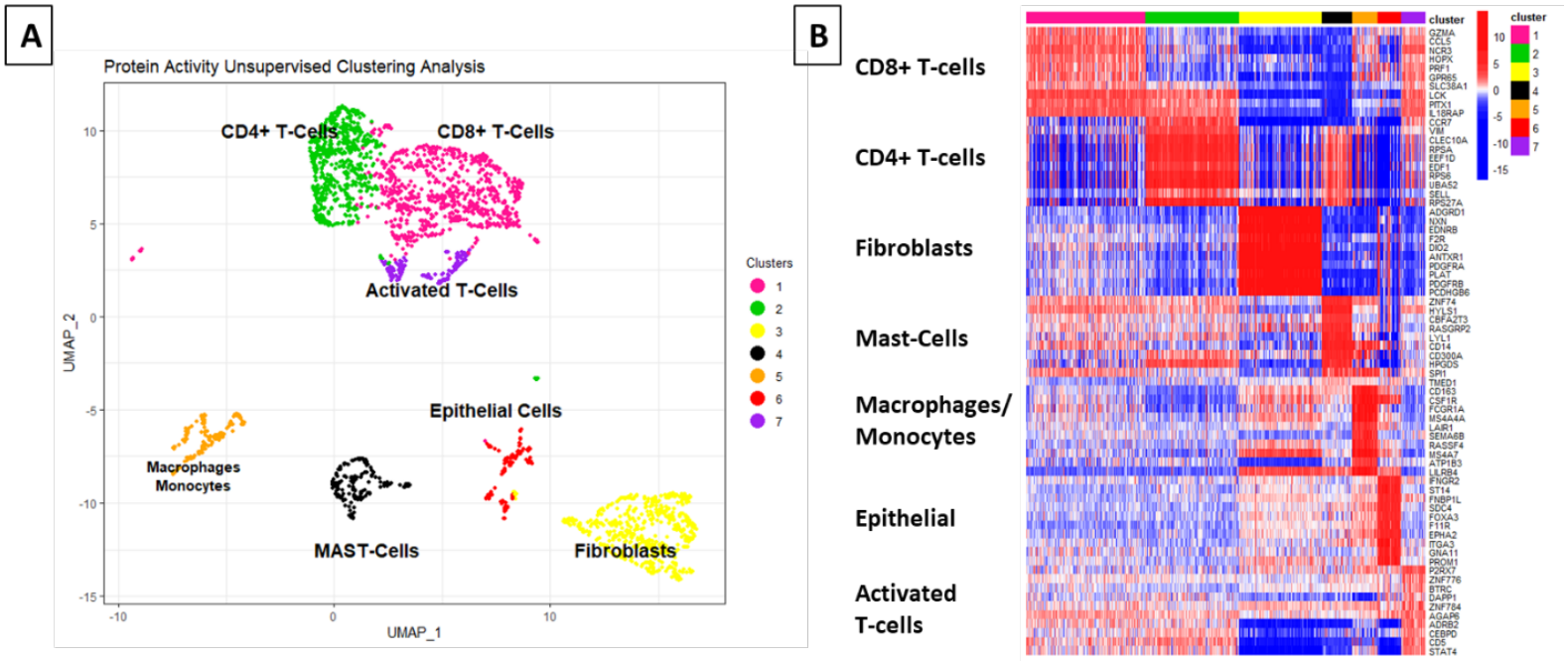


Figure 23: Cell Types Represented in Cholangiocarcinoma Tumor Micro-Environment by Protein Activity

(A) Results of unsupervised clustering analysis performed on the VIPER-inferred protein activity signatures of the single-cells (Louvain algorithm). UMAP projections are used to plot the cells. **(B)** Heatmap showing the top-10 differentially activated proteins of each cluster identified with a t-Test comparing the average activity of each VIPER-inferred regulator in one cluster versus the activity of the same regulators in all the other clusters.

Other populations highly represented in the Tumor micro-environment included Fibroblasts, myeloid cells, mast cells, endothelial cells, B-cells, and neutrophils. Of these, several have previously been associated with immunosuppression and tumor immune evasion, including neutrophil myeloid-derived suppressor cells [85] and fibroblasts [86]. A relatively small population of 140 cells were identified as epithelial in origin, and expressed KRT19, a marker of

cholangiocarcinoma [78], as the top most up-regulated gene. These cells were further validated as tumor cells by copy-number variation inference (Figure 24A).

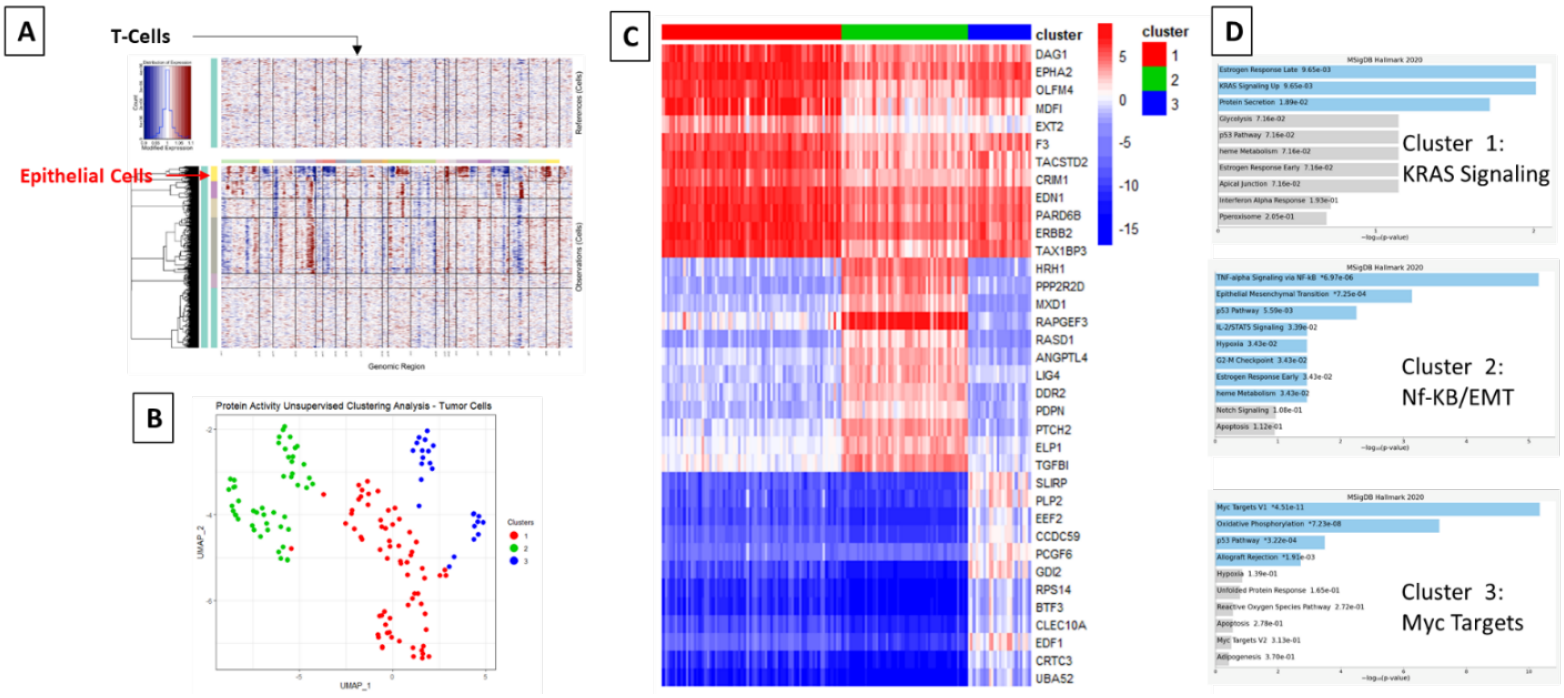


Figure 24: Cholangiocarcinoma Tumor Cell Sub-Clustering and Phenotyping

(A) InferCNV analysis was performed to confirm that the cluster of epithelial cells contains tumor cells, as they show substantially more copy number alterations across most of the chromosomes. T-Cells were chosen as a copy-number normal reference to infer copy number variations. **(B)** UMAP projection of the 3 clusters identified from the sub-clustering analysis of tumor cells. **(C-D)** Heatmap of the top MRs for each one of the tumor sub-clusters and the corresponding most statistically significant enriched pathways from Hallmarks of Cancer.

Tumor cells are heterogeneous by inferred protein activity:

From application of VIPER to infer protein activity of tumor cells within the cholangiocarcinoma sample, three strikingly distinct sub-populations could be identified (Figure 24B). These sub-populations were only revealed following VIPER analysis as the initial gene expression data are too noisy to further sub-cluster and suffer from gene dropout. In contrast,

VIPER profiling was able to reliably recover activity of 1602 regulatory proteins in each individual tumor cell and demonstrated high within-cluster consistency of protein activity profiles. Top differentially active proteins within each cluster are shown in Figure 24C, with the predominant cluster defined by upregulation of KRAS signaling and estrogen response pathway, a second cluster defined by upregulation of TNFa signaling via NF-kB and epithelial-mesenchymal transition, and the smallest cluster defined by upregulation of Myc targets and oxidative phosphorylation pathway (Figure 24D).

Cluster-specific druggable protein targets are inferred by single-cell OncoTarget

We applied VIPER using TCGA [87] as an external reference to perform single-cell OncoTarget and OncoTreat analysis, enabling both inference of druggable proteins with aberrant activity and drug candidates with transcriptional effects complementary to the overall protein activity profile of each tumor cell (Figure 25).

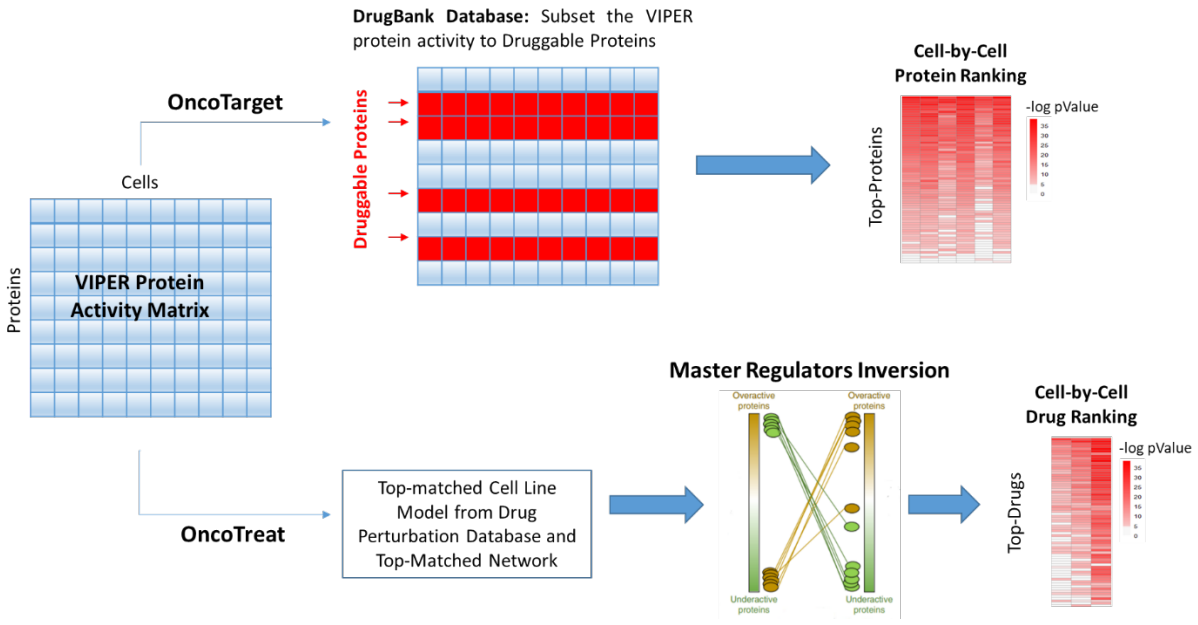


Figure 25: Flowchart of the Single-Cell OncoTarget/OncoTreat Drug Prediction Pipeline.

OncoTarget: Druggable proteins are first selected from the VIPER-inferred protein activity matrix using the DrugBank database and are then ranked according to the integrated p-Value across all the cells. OncoTreat: VIPER analysis using multiple ARACNe networks reconstructed from different TCGA cohorts and multiple tumor cell lines was performed to identify the best model to recapitulate the Master Regulators of tumor cells. Best-matched networks/cell lines were used to prioritize a list of drugs based on their ability to invert the activity of master regulator proteins cell-by-cell.

Applying single-cell OncoTarget to infer only druggable protein proteins identified significant differences between tumor cell sub-clusters in predicted drug sensitivity, with very high consistency within-clusters, and between single-cell resolution and OncoTarget inference on artificial bulk of each cluster (Figure 26B-D).

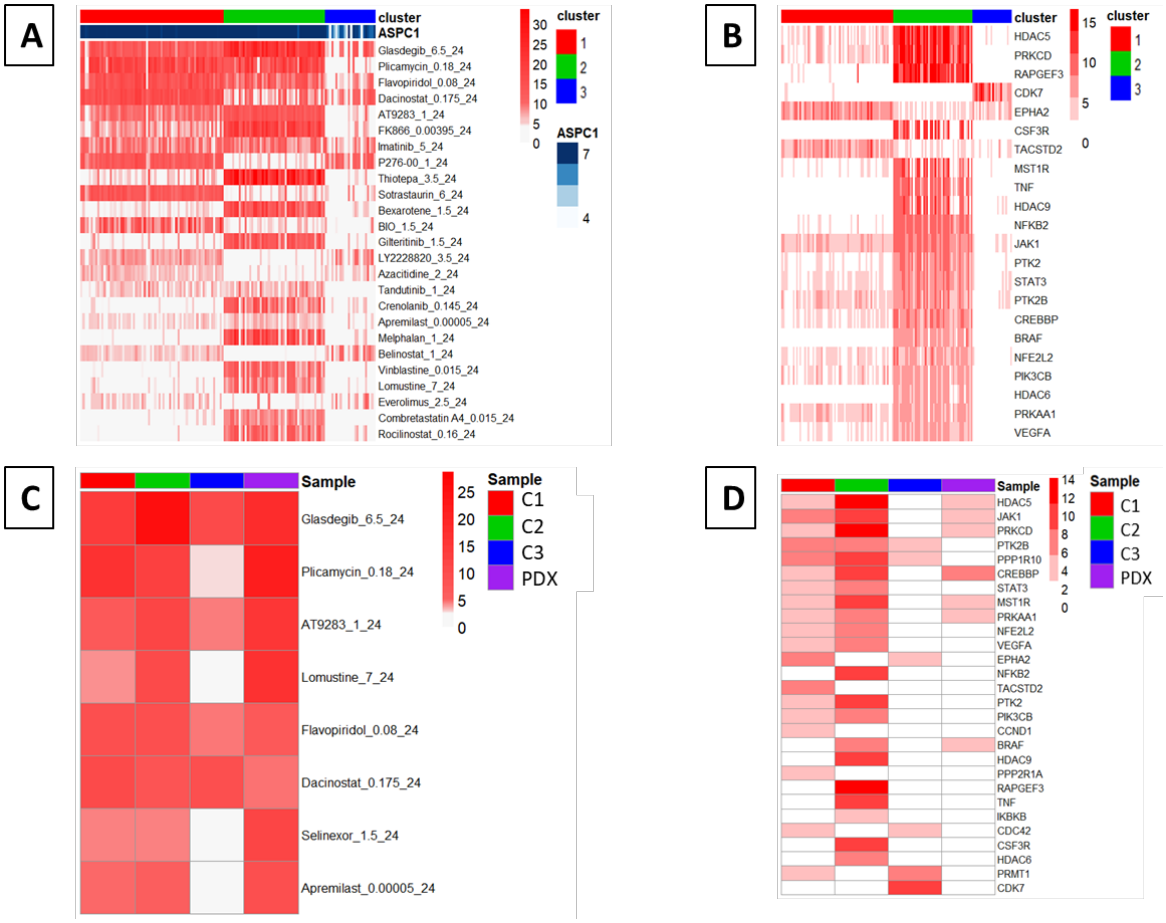


Figure 26: Results of OncoTreat and OncoTarget drug predictions.

(A) Negative log10 pValues of the top-drugs predicted by OncoTreat from the single-cells protein activity profiles of the three tumor sub-clusters. (B) Negative log10 pValues of the most statistically significant activated proteins in the DrugBank database, i.e. OncoTarget predictions. (C) OncoTreat drug predictions performed on synthetic bulk samples generated from the three tumor sub-clusters and on the bulk protein activity profile of the PDX model. (D) Top-activated proteins predicted by OncoTarget from synthetic bulk samples generated from the three tumor sub-clusters and from the bulk of the PDX.

Despite the extremely noisy nature of single-cell gene expression profiles, with >90% of genes undetected in any given cell, VIPER activity and drug predictions were dramatically conserved across single cells, confirming that the transcriptional states of the single cell populations that

comprise a tumor mass are remarkably stable. Twenty-two druggable targets were inferred by OncoTarget, including several HDAC proteins and other proteins of biological interest druggable by known FDA-approved or investigational compounds.

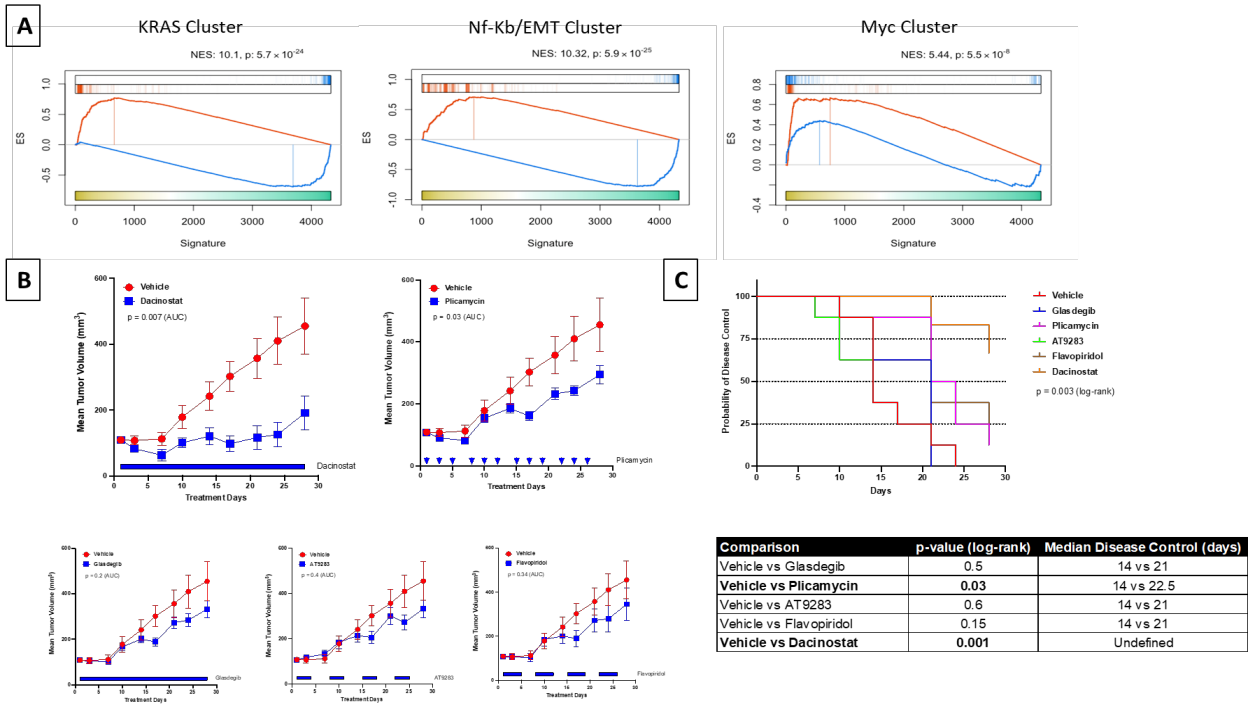
Drug predictions are consistent by single-cell OncoTreat and OncoTarget:

Drug predictions by single-cell OncoTreat are shown in Figure 26A-C, and are highly consistent with drug target predictions by single-cell OncoTarget, with six drugs in the top-25 predicted by OncoTreat known to target proteins predicted by OncoTarget. These included Dacinostat and Belinostat (which target HDAC5, HDAC9, and HDAC6), Rocilinostat (which targets HDAC6), flavopiridol (which targets CDK7), Sotrastaurin (which targets PTK2B), and Bardoxolne methyl (which targets NFE2L2 and NFKB2). The top five drugs predicted to target the greatest number of tumor cells were: Glasdegib, Plicamycin, Flavopiridol, Dacinostat, and AT9283.

Patient-Derived Xenograft Model Recapitulates Tumor Cell Population Signatures and Demonstrates Sensitivity to OncoTreat- Predicted Drugs:

We successfully engrafted and propagated a patient-derived xenograft model from resected tumor tissue at time of biopsy. Bulk RNA-Sequencing of the engrafted model demonstrates significant enrichment for Master Regulator Proteins of all three single-cell tumor populations observed in the initial biopsy (Figure 27A), with greatest enrichment for the two largest clusters. OncoTarget predictions on the PDX model include only proteins also observed in OncoTarget for at least one of the single-cell tumor clusters (Figure 26D), and OncoTreat predictions overlap entirely between PDX and the two largest clusters, with partial overlap in the third cluster of tumor cells (Figure 26C). In this analysis, Glasdegib, Plicamycin, Flavopiridol, AT9283, and

Dacinostat, which were ranked as the top 5 drugs with best overall tumor cell coverage in single-cell OncoTreat (Figure 26A), were also identified at the level of artificial bulk for each cluster and for the PDX (Figure 26C). Therefore, we administered these five drug candidates as well as vehicle control to a cohort of 8 mice per treatment arm, in order to evaluate their effectiveness in vivo for control of tumor growth rate and extension of time-to-disease-control. Dacinostat and Plicamycin were found to significantly reduce tumor growth rate (p=0.007 and p=0.03, respectively), with Dacinostat demonstrating stable tumor size over 28 days of treatment (Figure 27B). Both of these drugs were also found to significantly survival time by Kaplan-Meier regression (Dacinostat p=0.001 with median survival time exceeding 28 days vs 14 days by vehicle alone, Plicamycin p=0.03, with median survival time 22.5 days vs 14 days by vehicle alone) (Figure 27C).



Comparison	p-value (log-rank)	Median Disease Control (days)
Vehicle vs Glasdegib	0.5	14 vs 21
Vehicle vs Plicamycin	**0.03**	14 vs 22.5
Vehicle vs AT9283	0.6	14 vs 21
Vehicle vs Flavopiridol	0.15	14 vs 21
Vehicle vs Dacinostat	**0.001**	Undefined

Figure 27: Dacinostat and Plicamycin Significantly Inhibit Growth Rate of Cholangiocarcinoma Patient-Derived Xenograft Model

(A) Gene Set Enrichment Analysis of the top and bottom MRs of each tumor sub-cluster in the protein activity signature of the PDX model. **(B)** Tumor growth curves comparing the volume of the tumor in PDX models treated with the 5 selected drugs vs Vehicle control. Each drug was tested in N=8 mice, except for Dacinostat for which we have N=6. **(C)** Kaplan-Meier curve analysis comparing the time to failure of Disease Control achieved by each of the treatment conditions.

3.4 Discussion

We present in this case study the first single-cell RNA-Sequencing profile of a cholangiocarcinoma patient, to our knowledge. This profiling has clarified the composition of the cholangiocarcinoma microenvironment in this patient, with significant therapeutic implications for other patients. Cholangiocarcinoma has been previously described as having a significant stromal component, particularly with respect to fibroblasts producing a dense extracellular matrix [84]. However, in addition to a significant fibroblast infiltrate, here we observe T-cells comprising nearly half of the patient's tumor. This extent of lymphocytic infiltration suggests checkpoint immunotherapies as a potentially valuable adjunct to traditional standard-of-care treatments. Additional immune populations of interest include sizeable clusters of mast cells, myeloid cells, B-cells, and a cluster of neutrophils actively expressing IL8, which has been observed in studies of melanoma [88] and prostate cancer [89] to correlate with poor clinical outcomes via accumulation of myeloid-derived suppressor cells, and for which several inhibitor compounds are currently in clinical trials [90].

Furthermore, the low overall representation of tumor cells (<10% of total cell count) demonstrates the critical value of single-cell sequencing, as their transcriptional profile in bulk RNASeq would be overwhelmed by noise from the transcriptional profiles of the more highly

represented non-tumor cell types. This level of resolution combined with our protein activity inference tools has enabled us to profile the phenotypes of individual tumor cells in this patient, identifying three distinct subpopulations characterized by upregulation of KRAS signaling, TNFa signaling via NF-kB with epithelial-mesenchymal transition, and Myc signaling, respectively. This level of phenotypic heterogeneity of tumor cells has never previously been described in cholangiocarcinoma studies, and represents an important characteristic of the disease, as effective treatment must target all three of the populations identified in this patient. It remains to be determined in future work which of these phenotypes may generalize across patients, but it is reasonable to expect some degree of intratumoral heterogeneity requiring a precision medicine approach in this highly treatment-resistant disease.

We have extended our OncoTreat and OncoTarget personalized drug prediction algorithms to the single-cell level for the first time in analysis of this patient's tumor, identifying five candidate drugs with predicted activity against all of the observed tumor cell states. Indeed, our drug sensitivity and drug target predictions at the single-cell level precisely matched the predictions from synthetic bulk of each tumor cell cluster. Because we were able to establish a patient-derived xenograft model and confirm that it represents a good match for the tumor cell states observed in the patient both by overall gene set enrichment and OncoTreat/OncoTarget profile, we were able to validate our drug predictions in vivo. Of the predicted drugs, Dacinostat and Plicamycin were found to significantly decrease tumor growth rate and improve disease control rate vs vehicle control in a cohort of 8 mice per treatment arm. Plicamycin, which was predicted to strongly inhibit the first two tumor cell clusters but weakly inhibit the third, Myc-pathway-expressing, cluster, resulted in a slowed but continued tumor growth rate, with only one out of

eight mice still demonstrating disease control at 28 days. Dacinostat, however, was predicted to target all three cell clusters with similar effectiveness by OncoTreat and demonstrated stable tumor size with no significant growth from baseline for the majority of treated mice by 28 days.

In a prior study by Li et al. of in vitro manual drug screening on 27 organoids derived from 3 cholangiocarcinoma and 2 hepatocellular carcinoma patients, Plicamycin was found to be pan-effective across a broad range of organoids [91]. Out of 129 drugs screened in vitro, the authors found 9 drugs with at least 90% killing across all organoids, belonging to 5 classes of antineoplastic agents (histone deacetylase [HDAC] inhibitors, proteasome inhibitors, DNA topoisomerase II inhibitors, protein translation inhibitors, and RNA synthesis inhibitors). The authors found Plicamycin to be effective across organoids in vitro, consistent with our observed effectiveness of Plicamycin in slowing tumor growth rate in vivo. This prior in vitro evidence across a large number of organoids points toward the potential for generalized utility of Plicamycin across cholangiocarcinoma patients, with important implications for future clinical trials. In contrast to [91] Pant, K, we have identified Plicamycin here as a candidate adjunct to treatment for cholangiocarcinoma from a personalized drug prediction pipeline which narrowed down the space of all cancer drugs in a systematic way from several hundred to only five candidates, two of which were successfully validated in vivo. This is a much-improved success rate when compared to 9 out of 129 drugs identified as potentially effective by manual screening and allows for rapid and feasible validation of personalized drug predictions in vivo.

Dacinostat, which we identify to be even more effective than Plicamycin, was not included in the drug panel assessed by [91], and represents a novel candidate for cholangiocarcinoma therapy.

CG200745, an HDAC inhibitor, was previously found to induce anti-tumor effects in cholangiocarcinoma cell lines via miRNAs targeting the Hippo pathway [92]. HDACs are also known to have a role in cholangiocarcinoma carcinogenesis and are being actively explored as therapeutic candidates, but each HDAC inhibitor has highly variable binding affinities for the different HDAC proteins [93], with potential for very different therapeutic efficacy based on broad downstream transcriptional effects, which we directly assess by OncoTreat. To our knowledge, we are the first to identify Dacinostat as an HDAC inhibitor effective in treatment of cholangiocarcinoma. These findings are limited to the patient assessed in the case study due to difficulty of cholangiocarcinoma PDX-engraftment, but may generalize across patients similarly to Plicamycin, meriting further clinical follow-up alone or in combination with current treatments, particularly given the dismal treatment outcomes and lack of response to current standard-of-care therapy.

In addition to translating the findings from this study to trials in combination of immunotherapy, Dacinostat, and Plicamycin with traditional approaches for cholangiocarcinoma treatment, we hope in the future to extend our single-cell OncoTreat/OncoTarget approach to personalized drug prediction for additional cholangiocarcinoma patients and for a broad range of aggressive tumor types with limited treatment options. This would enable rapid reporting of plausible drug candidates for inhibition of tumor cells even in highly heterogeneous or stromally-infiltrated tumors, representing a true precision medicine approach to treatment-resistant patients.

3.5 Methods

Human Research Participation and Clinical Course of Treatment

Fresh surgical tissue was obtained with patient's consent from central hepatectomy performed at tumor stage pT3N1M1. Tissue was dissociated immediately for Single-cell RNA Sequencing. Following data analysis, tumor composition and OncoTreat/OncoTarget drug predictions were communicated to the medical team. Treatment with combination gemcitabine, cisplatin, and paclitaxel was initiated on a 21-day schedule one month after surgery, with dosing at day1 and day8, and CA19-9 tumor marker levels were drawn to monitor progression. With the patient's consent a trial of nivolumab was initiated following non-response to triple-combination chemotherapy and discovery of significant tumor T-cell infiltration by single-cell RNA Sequencing. Following interval improvement in CA19-9, patient was enrolled in an ongoing trial of TP-1287 (NCT03604783). Research was conducted in accordance with the Declaration of Helsinki.

Tissue Dissociation

Fresh tumor tissue was minced to 2-4 mm sized pieces in a 6-cm dish and subsequently digested to single cell suspension using Multi Tissue Human Tumor Dissociation Kit 1 (Miltenyi Biotec) and a gentleMACS OctoDissociator (Miltenyi Biotec) according to the manufacturer's instructions. Dissociated cells were aliquoted for single-cell sequencing.

Single-Cell RNA-Sequencing

Dissociated sample was processed for single-cell gene expression capture (scRNASeq) using the 10X Chromium 3' Library and Gel Bead Kit (10x Genomics), following the manufacturer's user

guide at the Columbia University Genome Center. After GelBead in-Emulsion reverse transcription (GEM-RT) reaction, 12-15 cycles of polymerase chain reaction (PCR) amplification were performed to obtain cDNAs used for RNAseq library generation. Libraries were prepared following the manufacturer's user guide and sequenced on Illumina NovaSeq 6000 Sequencing System. Single-cell RNASeq data were processed with Cell Ranger software at the Columbia University Single Cell Analysis Core. Illumina base call files were converted to FASTQ files with the command "cellranger mkfastq." Expression data were processed with "cellranger count" on pre-built human reference set of 30,727 genes. Cell Ranger performed default filtering for quality control, and produced a barcodes.tsv, genes.tsv, and matrix.mts file containing transcript counts for each cell, such that expression of each gene is in terms of the number of unique molecular identifiers (UMIs) tagged to cDNA molecules corresponding to that gene. These data were loaded into the R version 3.6.1 programming environment, where the publicly available Seurat package was used to further quality-control filter cells to those with fewer than 25% mitochondrial RNA content, more than 1,000 unique UMI counts, and fewer than 25,000 unique UMI counts. Pooled distribution of UMI counts, unique gene counts, and percentage of mitochondrial DNA after QC-filtering is shown in Figure 22A.

Single-cell RNASeq Gene Expression Processing

Gene Expression UMI count matrix was processed in R using the Seurat SCTransform command to perform a regularized negative binomial regression based on the 3000 most variable genes. The sample was then clustered on gene expression by a Resolution-Optimized Louvain Algorithm leveraging mean silhouette score of sub-sampled data to select optimal resolution without over-clustering [35]. Within each cluster metaCells were computed for downstream

regulatory network inference by summing SCTransform-corrected template counts for the 10 nearest neighbors of each cell by Pearson correlation distance. The resulting dataset of 2738 cells was projected into its first 50 principal components using the RunPCA function in Seurat, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Differential Gene Expression between clusters was computed by the MAST hurdle model for single-cell gene expression, as implemented in the Seurat FindAllMarkers command, with log fold change threshold of 0.5 and minimum fractional expression threshold of 0.25, indicating that the resulting gene markers for each cluster are restricted to those with log fold change greater than 0 and non-zero expression in at least 25% of the cells in the cluster.

Semi-Supervised Cell Type Calling

Unbiased inference of cell types was performed using the SingleR package and the preloaded Blueprint-ENCODE reference, which includes normalized expression values for 259 bulk RNA-seq samples generated by Blueprint and ENCODE from 43 distinct cell types representing pure populations of stromal and immune cells [49]. The SingleR algorithm computes correlation between each individual cell and each of the 259 reference samples, and then assigns a label of the cell type with highest average correlation to the individual cell and a p-value computed by wilcox test of correlation to that cell type compared to all other cell types. Unsupervised Clusters determined by the resolution-optimized Louvain algorithm are labelled as a particular cell type based on the most highly represented SingleR cell type label within that cluster among labels with p-value < 0.05.

Copy Number Inference

Copy Number Variation (CNV) was inferred from gene expression counts at the single cell level using the InferCNV package. Cells were clustered according to their unsupervised clustering label by gene expression, and the large T-cell cluster was used as a Copy-Number-Normal reference set. The cluster labelled by SingleR as epithelial cells was confirmed to exhibit significant copy-number alteration across the entire genome relative to other cell types, confirming this 140-cell cluster as Tumor cells.

Regulatory Network Inference

Within each gene expression cluster, metaCells were computed by summing SCTransform-corrected template counts for the 10 nearest neighbors of each cell by Pearson correlation distance. For clusters exceeding 200 cells, metaCells were sub-sampled to 200, and for each cluster a transcriptional regulatory network was inferred. All regulatory networks were reverse engineered by the ARACNe algorithm. ARACNe was run with 200 bootstrap iterations using 1785 transcription factors (genes annotated in gene ontology molecular function database as GO:0003700, “transcription factor activity”, or as GO:0003677, “DNA binding” and GO:0030528, “transcription regulator activity”, or as GO:0003677 and GO:0045449, “regulation of transcription”), 668 transcriptional cofactors (a manually curated list, not overlapping with the transcription factor list, built upon genes annotated as GO:0003712, “transcription cofactor activity”, or GO:0030528 or GO:0045449), 3455 signaling pathway related genes (annotated in GO biological process database as GO:0007165, “signal transduction” and in GO cellular component database as GO:0005622, “intracellular” or GO:0005886, “plasma membrane”), and 3620 surface markers (annotated as GO:0005886 or as GO:0009986, “cell surface”). Parameters

were set to zero DPI (Data Processing Inequality) tolerance and MI (Mutual Information) p-value threshold of 10^{-8}, computed by permuting the original dataset as a null model. Each gene list used to run ARACNe is available on github, along with the generated ARACNe tables.

Protein Activity Inference

Protein activity was first inferred for all cells by running the metaVIPER algorithm with all ARACNe networks on the SCTransform-scaled single-cell gene expression signature. Because the SCTransform-scaled gene expression signature is already normalized, VIPER normalization option was set to "none." The resulting VIPER matrix included 1602 proteins with successfully inferred activity across all 2738 cells. Subsetting to tumor cells only, this protein activity matrix was loaded into a Seurat Object with CreateSeuratObject, then projected into its first 50 principal components using the RunPCA function in Seurat, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Differential Protein Activity between clusters identified by resolution-optimized Louvain was computed using bootstrapped t-test, run with 100 bootstraps, and top proteins for each cluster were ranked by p-value. This analysis identified three phenotypically distinct clusters of tumor cells by VIPER which were not previously identifiable by raw gene expression.

Single-cell OncoTarget

In order to highlight potentially druggable proteins with aberrant activity in the three tumor cell sub-clusters identified in this patient, we generated single-cells differential gene expression signatures by scaling the log-Normalized Counts-Per-Million (CPM) single-cells data

(log10[CPM-UMI count +1]) to the log-Normalized Transcripts-Per-Million (TPM) data (log10[TPM+1]) of the entire publicly available TCGA, as an external reference to enable identification of any proteins concurrently upregulated in all tumor cells in addition to proteins differentially upregulated across tumor cell sub-clusters. We ran VIPER on this gene expression signature using the single-cell tumor cluster ARACNe network. Resulting matrix was sub-set to proteins targeted by known FDA-approved or investigational drug compounds in DrugBank. Enrichment scores for each protein in each cell were then converted to Bonferroni-corrected p-values and subset to proteins with median p-value $< 10^{-5}$ in any tumor cell sub-cluster. To assess robustness of OncoTarget predictions, the same analysis was performed on artificial bulk of each tumor cell sub-cluster, increasing sample depth at the cost of single-cell resolution by summing UMI counts across all cells in each cluster. Notably, this resulted in the exact same druggable protein predictions as predicted at the single-cell level (Figure 26D).

Single-cell OncoTreat

In order to incorporate off-target and downstream drug effects on the entire transcriptional profile to improve drug prediction, we further leveraged a database of drug perturbation RNA Sequencing on multiple tumor cell lines generated in Califano Lab (BT20, HCC1143, GISTT1, GIST430, HSTS, KRJ1, IOMM, U87, HF2597, H1793, ASPC1, PANC1, LNCAP, DU145, TCCSUP, EFO21, ASPC1, PANC1). Analogously to OncoTarget analysis, we computed single-cell differential gene expression signatures with respect to TCGA. In order to identify the top-matched cell lines able to recapitulate the Master Regulators of the tumor sub-clusters and the best regulatory model, we first applied VIPER using ARACNe networks derived from distinct TCGA tumor cohorts. The Normalized Enrichment Scores produced by VIPER were converted

into bonferroni-corrected p-values and for each single-cell we measured the ratio between the AUC of the cumulative distribution of the regulators with a p-value less than 0.01 and the AUC of the null model. The pancreatic adenocarcinoma ARACNe network showed the highest median score among the tumor types for which we had drug perturbational data (Figure 28A).

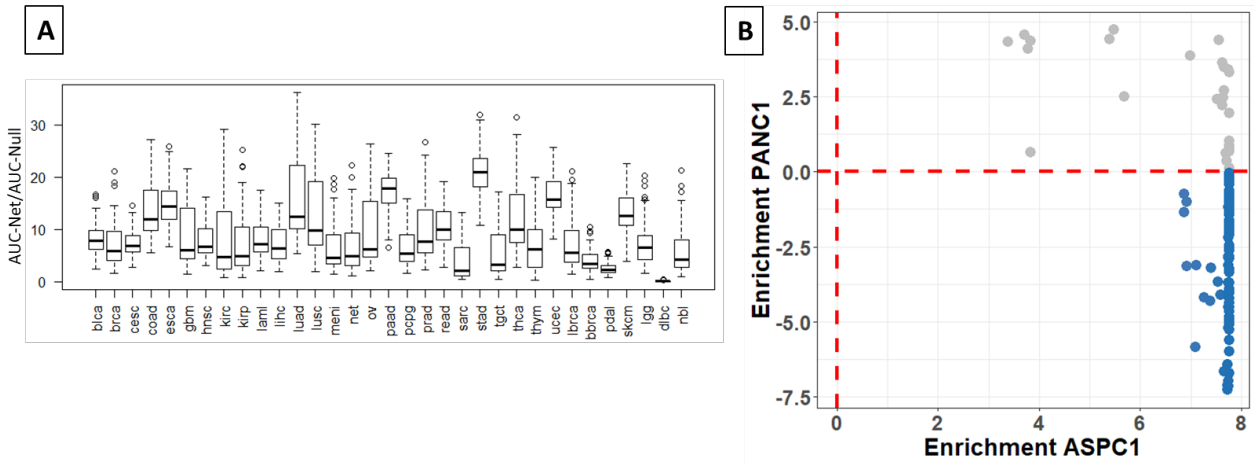


Figure 28: Selection of Best-Matched Cell Line From Drug Perturbational Database

(A) Results of the ARACNe network-matching procedure applied to determine the best regulatory model for the OncoTreat analysis of tumor cells. The top-matched network for which we have perturbational data is the Pancreatic Adenocarcinoma Interactome (second one in the overall ranking) (B) Results of the Gene Set Enrichment Analysis of the top/bottom 50 MRs of each single-cells in the protein activity signature of the two pancreatic cell lines ASPC1 and PANC1.

Consequently, we tested the enrichment of the top/bottom 50 MRs of each single cell in the protein activity signature of the two available pancreatic cell lines in the Califano Lab database, ASPC1 and PANC1 (Figure 28B). We subsequently applied OncoTreat using the perturbational data of ASPC1, since all of the tumor cells have a strong statistically significant enrichment for the signature of this cell line. From the drug perturbation data on ASPC1, we matched each

individual cell to the drug that best inverted the protein activity profile of that cell, using the standard CLIA-certified OncoTreat algorithm to estimate p-value for each drug's ability to invert the tumor cell profile. Drug prioritization was weighted by p-values of each drug in each cell by fisher integration of p-values, such that drugs predicted to invert the profile of a larger number of cells were higher-ranked. In Figure 26A we show p-values across all tumor cells for the top-25 drugs predicted from ASPC1 drug perturbation data. As we did for OncoTarget, in order to assess the robustness of the top predicted drugs from single cells, the same analysis was performed on artificial bulk of each tumor cell sub-cluster computed by summing the UMI counts by gene.

OncoTarget/OncoTreat analysis of PDX bulk RNA-Seq data

Bulk RNA-Seq data of the PDX model was analyzed using the same pipeline conceived for single-cells and artificial bulk samples. Consequently, a differential gene expression signature was computed scaling the log-transformed TPM data to the average of TCGA and then processed through the VIPER algorithm. The enrichment of the MRs of the three tumor sub-clusters in the protein activity signature of the PDX was computed with Gene Set Enrichment Analysis (GSEA). Results in Figure 6A show that the PDX is able to recapitulate the MRs of the tumor cell sub-clusters, with particular enrichment of genes marking the two largest cluster defined by upregulation of KRAS and Nf-kb/EMT signaling. Finally, the OncoTarget/OncoTreat analysis was performed as described in the previous sections, using the same regulatory models and cell lines used for single-cells and synthetic bulk samples.

Patient-Derived Xenograft Model & Treatment

All mice were maintained under barrier conditions and experiments were conducted using protocols and conditions approved by the Memorial Sloan Kettering Cancer Center (MSKCC) Institutional Animal Care and Use Committee (IACUC) under protocol 16-08-011. Patient derived tumor tissue to generate PDX models were obtained under the MSKCC Institutional Review Board (IRB)-approved protocols #17-387 and #06-107. PDX mouse models were established by implanting tumor cells subcutaneously into non-obese diabetic/severe combined immunodeficiency interleukin-2R gamma null, HPRT null (NSGH) mice (Jackson Labs, IMSR Cat# JAX:012480, RRID: IMSR_JAX:012480). Mice were treated with vehicle or one of five treatment groups for four weeks: (1) Vehicle control, (2) Glasdegib 100mg/kg PO daily, (3) Plicamycin 0.2mg/kg IP 3 times/week, (4) AT9283 15mg/kg daily, IP 4 days on 3 days off, (5) Flavopiridol 15mg/kg PO daily, 5 days on 2 days off, (6) Dacinostat 25mg/kg IP daily. Fifty-four mice were implanted in total. Once the tumor reached a volume of 80-120 mm3, mice were assigned to treatment groups using block randomization with eight animals per group. Tumors were measured by caliper measurement twice weekly and tumor volume (TV) was calculated as follows: TV = width^2 X ½ length. Treatment failure was defined as >100% increase in tumor volume relative to baseline in each respective mouse. For in vivo statistical analysis, the Mann-Whitney-Wilcoxon method was used to evaluate differences in distribution of tumor volume between treatment groups. Vardi's test was used to evaluate difference in area under the curve (AUC) between treatment groups. Event-free survival (EFS) was defined as the percentage of mice that survived at any given time point without treatment failure. Kaplan-Meier survival curves were compared using the log-rank test. Statistical analysis was performed using R software (v3.5.0). Waterfall plots and tumor volume curves for in vivo analysis were generated with GraphPad Prism (v8.4.1). Statistical significance was defined as p values < 0.05.

Chapter 4: Cancer-Associated Fibroblast Sub-Populations Mediate Clinical Immunotherapy Response in Head and Neck Squamous Cell Carcinoma

4.1 Summary

The heterogeneity of cancer associated fibroblasts (CAF) has precluded rigorous understanding of their function in the tumor microenvironment. Using an inferred protein network-based methodology to generate CAF atlas from single-cell transcriptomic profiles of human head and neck carcinoma, pre and post nivolumab treatment, we resolved 5 unique CAF subtypes. The head and neck CAF (HNCAF) protein activity profiles, derived from a nested cohort of paired single-cell RNA sequencing profiles, were then used to perform protein activity enrichment analysis on the 36-patient parental cohort of clinically annotated bulk transcriptomic profiles. Among these subtypes, HNCAF-0/3 emerged as predictive of nivolumab response, while HNCAF-1 was associated with immunosuppression. Functionally, HNCAF-0/3 were found to reduce TGFβ dependent PD-1^{+}TIM-3^{+} exhaustion of T cells and increase CD103^{+}NKG2A^{+} resident memory phenotype to enhance the overall cytolytic profile of T cells. Our findings implicate distinct HNCAF subsets as clinically actionable modulators of human TIL repertoire.

Significance: Our utilization of systems biology-based master regulator analysis of single-cell transcriptomic profiles unveiled unique CAF subtypes that can predict clinical responses to aPD-1 blocking antibodies to warrant biomarker validation studies. Furthermore, our CAF atlas will open and inform the understudied immunobiology of the stroma in head & neck cancer.

4.2 Introduction

Anti-PD-1 immune checkpoint inhibitors (ICI) are currently the first line therapy for recurrent/metastatic head and neck squamous cell carcinoma [94] [95]. Yet, overall response rates can be as low as 20%, with increased responses in tumors with elevated PD-L1 expression and tumor infiltrating T cells [94] [96] [97]. While this is partly a T cell dependent mechanism, there may be additional cellular subpopulations in the tumor microenvironment (TME) responsible for mediating response to ICI (5). Recent studies have suggested that cancer associated fibroblasts (CAF) are associated with this resistance, but their role in T cell immunomodulation is still unclear in the human TME [98] [99] [100].

In human breast cancer, four CAF subtypes (CAF-S1 to S4), were identified based on the expression of six fibroblast markers — fibroblast activation protein (FAP), integrin β1 (CD29), α-smooth muscle actin (α-SMA), fibroblast-specific protein-1 (FSP-1), platelet-derived growth factor receptor β (PDGFRβ), and caveolin-1 (CAV1) [101]. In their follow-up study, single-cell RNA sequencing further divided CAF-S1 into eight subtypes with the majority of these subtypes linked to immunosuppression and resistance to immunotherapy [100]. In contrast, only three molecularly and phenotypically distinct CAF subpopulations were identified in pancreatic cancer, based on spatial location and imputed function, as defined by cytokine and surface marker expression. These include inflammatory CAF (iCAF), myofibroblastic CAF (myCAF) and antigen-presenting CAF (apCAF) [33] [102]. In head and neck, three CAF types were previously identified by single-cell RNA sequencing corresponding to myCAF and two undefined CAF subtypes (CAF1 and CAF2), but the functionality of these subtypes and their association with immunotherapy response remains unknown [103]. In general, these various

orthogonal approaches to cancer specific CAF characterizations have not provided a concordant classification of this important stromal host cell type.

To assess whether CAF-related or other TME subpopulations can regulate clinical responses to nivolumab, we leveraged single-cell RNA-sequencing (scRNA-Seq) to longitudinally profile pre and post treatment human head and neck squamous cell carcinoma to generate a dynamic atlas of the human HNSCC TME. Our novel bioinformatic approach uses the VIPER algorithm [21] [26] to address limitations imposed by high noise and significant gene dropout rates in most scRNA-Seq analysis platforms. Specifically, VIPER leverages knowledge of regulatory networks to allow full quantitative characterization of protein activity by assessing the enrichment of their transcriptional targets in differentially expressed genes. On average, the resulting protein activity profiles outperform antibody-based measurements and dramatically outperform gene expression-based analyses in terms of identifying and characterizing molecularly distinct TME subpopulations [35], thus enabling mechanistic dissection of the HNSCC microenvironment at hitherto unattained resolution. We present the results of these protein activity-based analyses on clinical biospecimens to generate a high-resolution atlas of the human HNSCC immune and stromal micro-environment under ICI pressures.

4.3 Results

Proteomic Master Regulatory Network Analysis of Longitudinal Single-Cell Transcriptomic Profile Identifies Functionally Unique CAF Populations in The HNSCC Microenvironment.

Longitudinal scRNA-Seq of patient tumors, pre- and post-treatment with nivolumab, and gene expression clustering with Seurat revealed 12 broadly distinct cellular populations, consistently expressed across all the tumors sequenced (Fig. 29A-B).

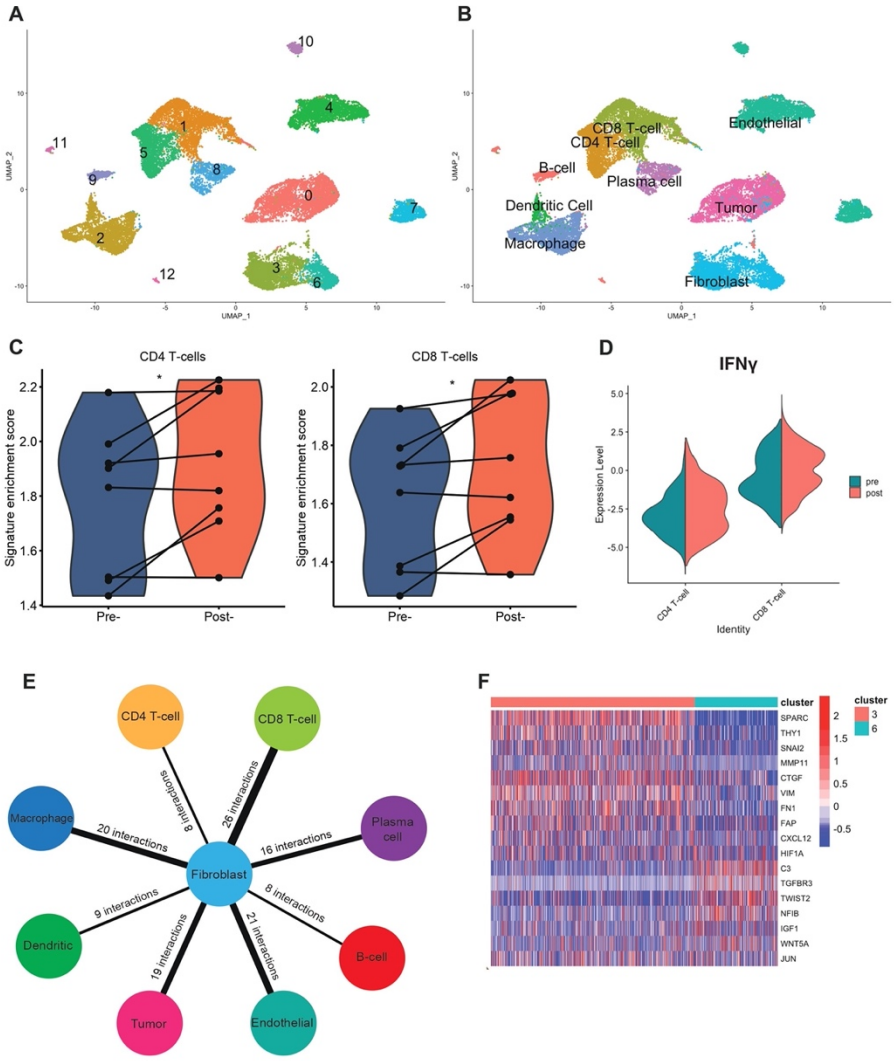


Figure 29: Single-Cell RNA-Sequencing and VIPER Inference Shows Increased T-cell Activity Induced by Nivolumab

A) 2-dimensional UMAP projection of single-cell RNA-Seq gene expression data before VIPER is applied, colored by unsupervised cluster grouping. B) 2-dimensional UMAP projection of gene expression data from A, colored by cell type inferred from SingleR, as in Figure 1D. C) Validation of CD4 and CD8 T cell gene signatures in responders of HNSCC patients during

αPD-1 treatment with bulk RNAseq data. D) Interferon Gamma VIPER-inferred protein activity among T-cell populations pre vs post immunotherapy. E) Plot of inferred receptor-ligand interactions between cell types, such that receptor-ligand pairs with known interaction have significant upregulation of ligand gene expression among fibroblasts, and significant upregulation of corresponding receptor protein activity by VIPER in another cell type. Width of lines is weighted by the number of inferred interactions between fibroblasts (in the middle), and each other cell type. F) Heatmap of differentially expressed CAF1/CAF2 genes defined by Puram et al. for each gene expression-based CAF cluster identified in A (Clusters 3 and 6).

To achieve higher resolution of cellular subpopulation characterization, we used scRNA-Seq profiles from each cluster to infer subpopulation-specific gene regulatory networks with the ARACNe algorithm [40], followed by protein activity analysis using the Virtual Inference of Protein-activity by Enriched Regulon (VIPER). Protein activity-based re-clustering identified two additional, previously undetected clusters for a total of 14 distinct cellular populations which were also consistently expressed across all four patients (Fig. 30A). To visualize key differences between these cellular populations, we generated a heatmap showing the activity of the five proteins most differentially activated in each cluster (Fig. 30B). We first assessed the ability of these VIPER generated populations to accurately respond to expected treatment-induced changes. As expected, both gene expression and protein activity analyses revealed increased T cells and interferon-gamma protein activity following nivolumab treatment (Fig. 30C (cluster 8), Fig. 29C-29D). When we interrogated changes in the abundance of other cell populations, VIPER clustering revealed heterogeneity among fibroblast cells not discoverable from gene expression clustering alone, with two clusters (clusters 4 and 9) presenting highly statistically significant post-treatment cellular fraction increase (Fig. 30C). Cell lineage inference, using SingleR [48], identified both clusters as fibroblasts, suggesting that PD-1 targeted

immunotherapy in head and neck cancer was associated with CAF upregulation (Fig. 30D). Accordingly, imputed receptor-ligand interactions between cell types [35] suggested strong interplay between CAF and CD8 T cells (Fig. 29E). Furthermore, two additional clusters (clusters 6 and 7), also characterized as fibroblasts by SingleR (Fig. 30D), did not show significant fractional representation differences following immunotherapy (Fig. 30C), thus suggesting the existence of functionally distinct CAF subpopulations within the HNSCC TME.

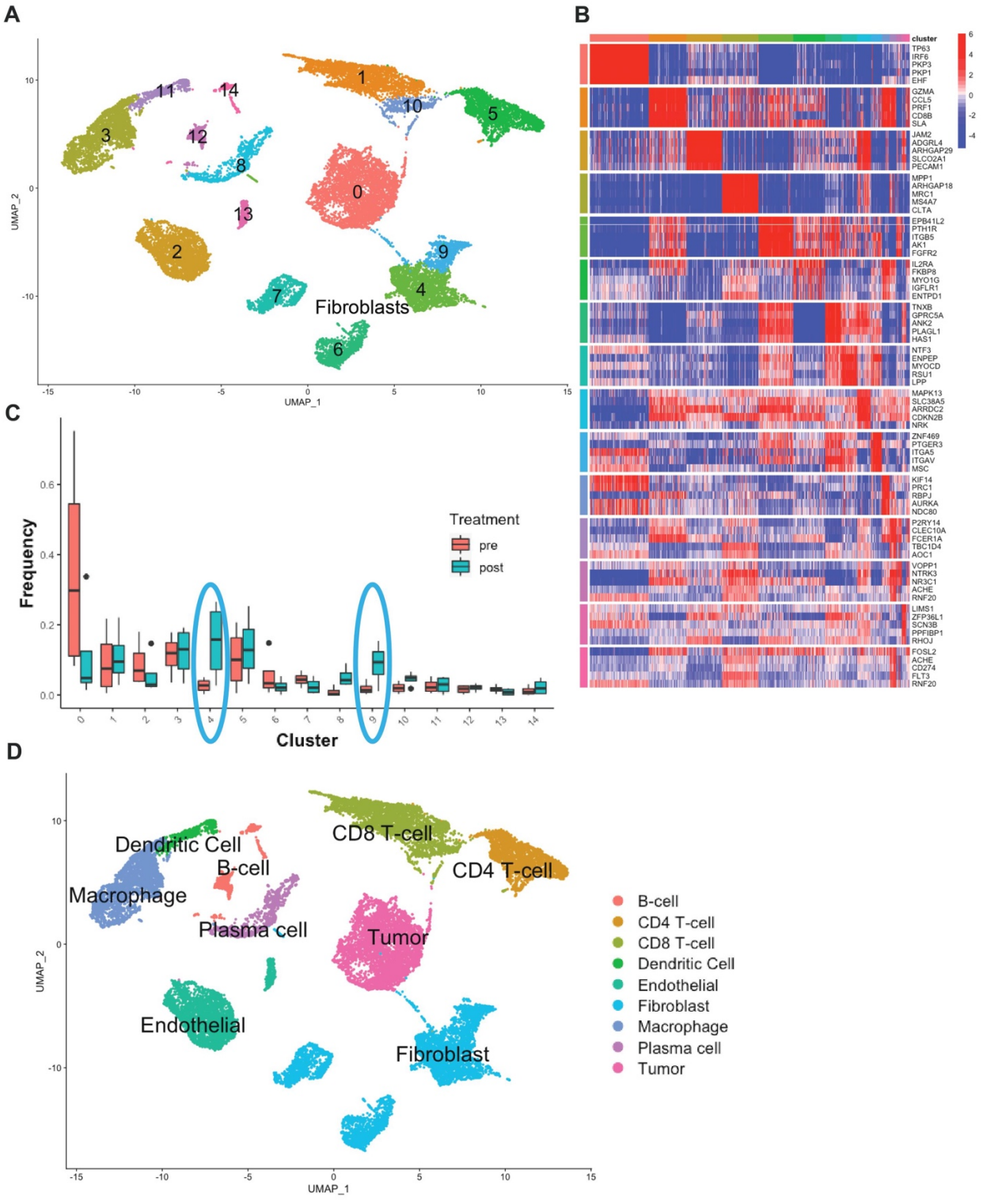


Figure 30: VIPER analysis of longitudinal single-cell transcriptomic profiles of HNSCC show CAF changes associated with immunotherapy.

A) 2-dimensional UMAP projection of cells across all samples, processed by VIPER and clustered by resolution-optimized Louvain. Cells are colored by unsupervised cluster number, with fibroblast clusters (4,6,7,9) further labelled by cell type. **B)** Heatmap of top 5 most differentially upregulated proteins per cluster for each cell population in A. **C)** Boxplot of population frequency at baseline and following αPD-1 immunotherapy for each cell type cluster in A. CAF subtypes increasing in response to immunotherapy ($p<0.01$) are circled in blue. **D)** SingleR cell type inference projected on UMAP plot. Each cluster is assigned a lineage cell type based on its majority SingleR-inferred label.

VIPER Fibroblast Clustering Identifies Unique Sub-Populations Associated with Response and Resistance to Immunotherapy

To further evaluate functional differences between the distinct CAF sub-populations in the HNSCC TME, we performed protein activity-based sub-clustering focusing only on fibroblast cells using ARACNe and VIPER. The analysis identified five molecularly-distinct CAF clusters, preliminarily termed HNCAF-0 – HNCAF-4 (Fig. 31A). Importantly, gene expression-based sub-clustering of fibroblasts only identified two distinct CAF populations corresponding to the two fibroblast populations seen in Fig. 29A (Clusters 3 and 6). As expected, these two clusters phenotypically match the two CAF populations previously identified in HNSCC such that Cluster 3 corresponds to CAF1 and Cluster 6 corresponds to CAF2 (Fig. 29F) [103]. Among the five HNCAF populations identified by protein activity-based clustering, cell fractional representation increased for HNCAF-0 and HNCAF-3, decreased for HNCAF-1 and HNCAF-2, and was unaffected for HNCAF-4 (Fig. 31B). The top ten most differentially active proteins, presented as a ranked list of differentially active transcription factors and signaling molecules, in each of the five clusters highlight their potential functional properties (Fig. 31C) to help define the molecular biology of each HNCAF phenotype. To assess the associations of each HNCAF

subpopulation with clinical response to αPD-1 immunotherapy, we used the HNCAF molecular signatures to analyze the bulk RNA sequencing profiles from the 36-patient parental cohort annotated with clinical response to nivolumab. For this purpose, we first used VIPER to generate protein activity profiles from each bulk profile, using fibroblast specific regulatory networks generated at the single-cell level, and then evaluated the enrichment of the most differentially active proteins in each HNCAF subpopulation (marker protein sets) among proteins differentially active in responders vs non-responders. The analysis revealed statistically significant enrichment of HNCAF-0 and HNCAF-3 marker genes in pre-treatment samples of patients who subsequently responded to αPD-1 immunotherapy (Fig. 31D). This result indicates that the HNCAF-0 and HNCAF-3 populations, which also expand following nivolumab treatment, may be predictive of clinical response in human HNSCC patients. By contrast, HNCAF-1, HNCAF-2, and HNCAF-4 cells did not expand following therapy and their markers were not significantly enriched in responders vs non-responders.

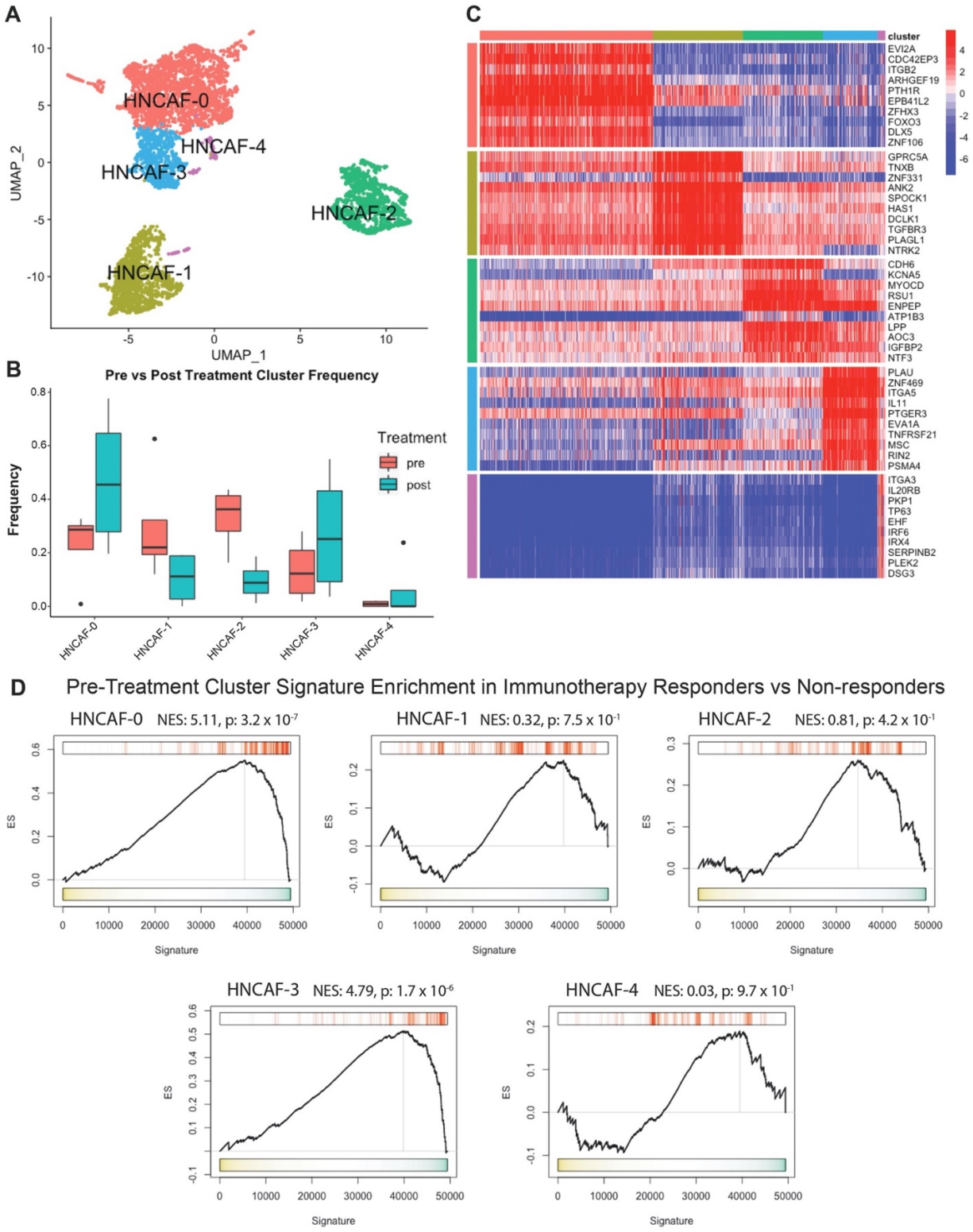


Figure 31: Fibroblast sub-clustering reveals distinct populations associated with differential responses to αPD-1-based immunotherapy

A) 2-dimensional UMAP projection of CAF across all samples, re-clustered by resolution-optimized Louvain and colored by cluster identity. **B)** Boxplot of cluster frequencies pre vs post nivolumab therapy, such that HNCAF-0 and HNCAF-3 show statistically significant increase in frequency ($p<0.01$) while HNCAF-1 and HNCAF-2 show significant decrease ($p<0.01$). **C)** Heatmap of top 10 most differentially upregulated proteins per cluster for each CAF population. **D)** Protein Activity Profile Enrichment plots of single-cell protein population markers for each HNCAF cluster (Supplemental Table 1) in bulk-RNASeq signature of immunotherapy responders (18 patients) vs non-responders (18 patients), profiled pre-treatment. HNCAF-0 and HNCAF-3 profiles are significantly enriched in treatment responders ($p=3.2 \times 10^{-7}$ and $p=1.7 \times 10^{-6}$, respectively).

Fibroblast Subtype Analysis Reveals Novel Classification Paradigm in HNSCC

Due to scant literature on CAF in human HNSCC, we next quantified the extent of CAF infiltration from surgical HNSCC specimens using flow cytometry. CAF abundance — as defined by $CD45^-$ $EpCAM^-$ $CD31^-$— ranged between 12% and 58% of the total live cells (Fig. 33A). Having confirmed significant abundance of CAF in human HNSCC, we proceeded to assess the presence of novel HNCAF subpopulations predicted by the VIPER analysis. Distinct CAF subpopulations termed CAF-S1, CAF-S2, CAF-S3, and CAF-S4 have been previously identified in breast cancer based on the expression of CD29 and fibroblast activation protein (FAP) by flow cytometry [101]. Kieffer et al. showed that CAF-S1 can be found in HNSCC but the presence of other CAF-S populations in HNSCC remains unknown [100]. Hence this protein-based framework was initially used to assess how the VIPER imputed HNCAF align with breast cancer CAF scheme. Following the same gating strategy employed by Costa et al. (Fig. 32), we confirmed the presence of all four breast cancer CAF-S populations in HNSCC (Fig. 33B). Interestingly, CAF-S1 and CAF-S2 were most abundant, while CAF-S3 and CAF-S4 abundance was quite minimal (Fig. 33C). We then sorted CAF-S1 – S4 cells from HNSCC tumors (Fig. 32)

and performed bulk RNA sequencing of each subpopulation to assign these sorted cells to the VIPER generated HNCAF populations. Pairwise gene set enrichment analysis of the HNCAF protein activity signatures in the bulk transcriptome indicated that the gene sets representative of HNCAF-0, HNCAF-1, and HNCAF-3 were all enriched in the same breast subtype (CAF-S1), while HNCAF-2 and HNCAF-4 were both enriched in CAF-S4 with HNCAF-4 also enriched in CAF-S3 (Fig. 34A). CAF-S2—primarily defined as double-negative for FAP and CD29—did not significantly align with any HNCAF. These analyses importantly showed that VIPER-clustered HNCAF provide much greater resolution of functionally distinct CAF phenotypes compared to the flow-based CAF-S1/S2/S3/S4 framework. Specifically, the CAF-S1 subtype matched three distinct HNCAF subtypes, which have opposing association with clinical responses to immunotherapy. Additionally, the HNCAF subtypes did not clearly correlate with any of the CAF-S1 subclusters later identified by gene expression in breast cancer apart from ecm-myCAF (Fig. 34B). HNCAF-0,1,2 and 3 were all significantly enriched for the ecm-myCAF signature, with HNCAF-0 and HNCAF-3 being more enriched than HNCAF-1 and HNCAF-2. However, ecm-myCAF are associated with immunosuppression and resistance to immunotherapy leading us to believe that HNCAF-0 and HNCAF-3 differ from ecm-myCAF in terms of functionality [100].

We also tested for concordance of our classification schema with previously defined gene set markers of inflammatory CAF (iCAF) and myofibroblastic CAF (myCAF), as first described in pancreatic cancer [33]. Cell-by-cell enrichment of iCAF and myCAF gene signatures revealed an enrichment for the iCAF signature in HNCAF-1 cells and for myCAF in HNCAF-2 cells (Fig. 34C-D). Correlations between our HNCAF populations and the breast or pancreatic CAF

phenotypes is summarized in the classification scheme shown in Fig. 33D. While presenting some similarity to CAF-S1 cells, we conclude that HNCAF-0 and HNCAF-3 represent novel, molecularly distinct fibroblast subpopulations, potentially unique to head and neck cancer and predictive of patient outcome (Fig. 31D), which do not completely match the iCAF/myCAF classification. Furthermore, in contrast to the HNCAF-0 and HNCAF-3 subtypes, iCAF/myCAF and CAF-S1/S2/S3/S4 are not significantly enriched in responder cohorts, suggesting that these classification schemes do not accurately depict CAF function in HNSCC (Fig. 35). Cumulatively, these data underscore the greater resolution of HNCAF from our VIPER analysis compared to previous efforts, and more critically, the HNCAF populations identified through VIPER allow prognostic correlations of CAF cells in HNSCC.

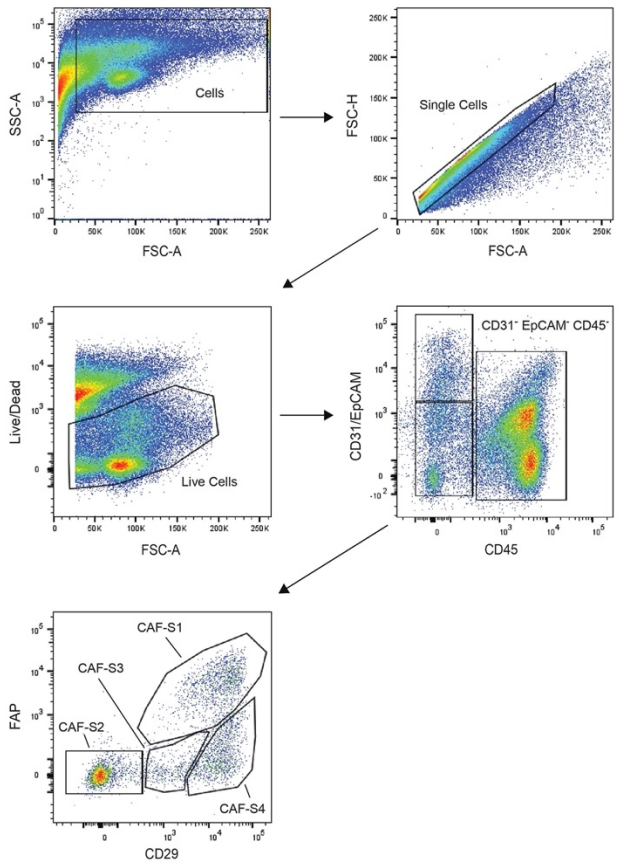


Figure 32: Flow cytometry gating strategy for sorting of CAF-S1 through CAF-S4 populations.

CAF-S1 – S4 gating from representative patient HNSCC tumor used for fluorescence-activated cell sorting.

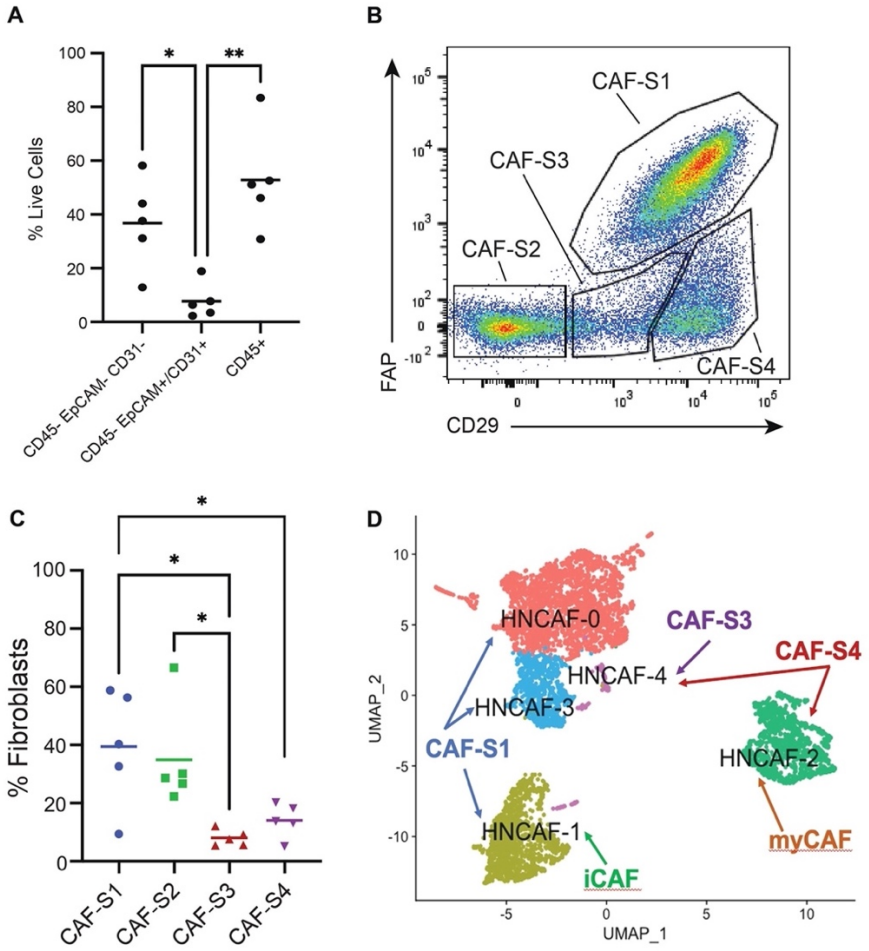


Figure 33: Prognostically associated HNCAF sub-populations as defined by scRNA-Seq provide greater resolution than CAF phenotypes previously characterized.

A) Relative frequencies across 5 HNSCC patients of stromal (CD45-Epcam-CD31-), epithelial/endothelial (CD45-Epcam+/CD31+) and immune (CD45+) cells components quantified by flow cytometry. **B)** Flow cytometry gating strategy to isolate CAF phenotypes previously described in the literature, implemented as described in [101]. **C)** Relative frequency for each patient of CAF subtypes from B among total CAF quantified by flow cytometry. **D)** Phenotype-matching between unsupervised clusters from single-cell RNA-Seq and bulk-RNASeq of sorted populations CAF-S1 to CAF-S4, as well as iCAF and myCAF, from [33]. Each single-cell population is labelled as the sorted population with highest gene set enrichment, as shown in Figure S3. The data in A and C were analyzed using one-way ANOVA and * indicates $p<0.05$, ** indicates $p<0.01$.

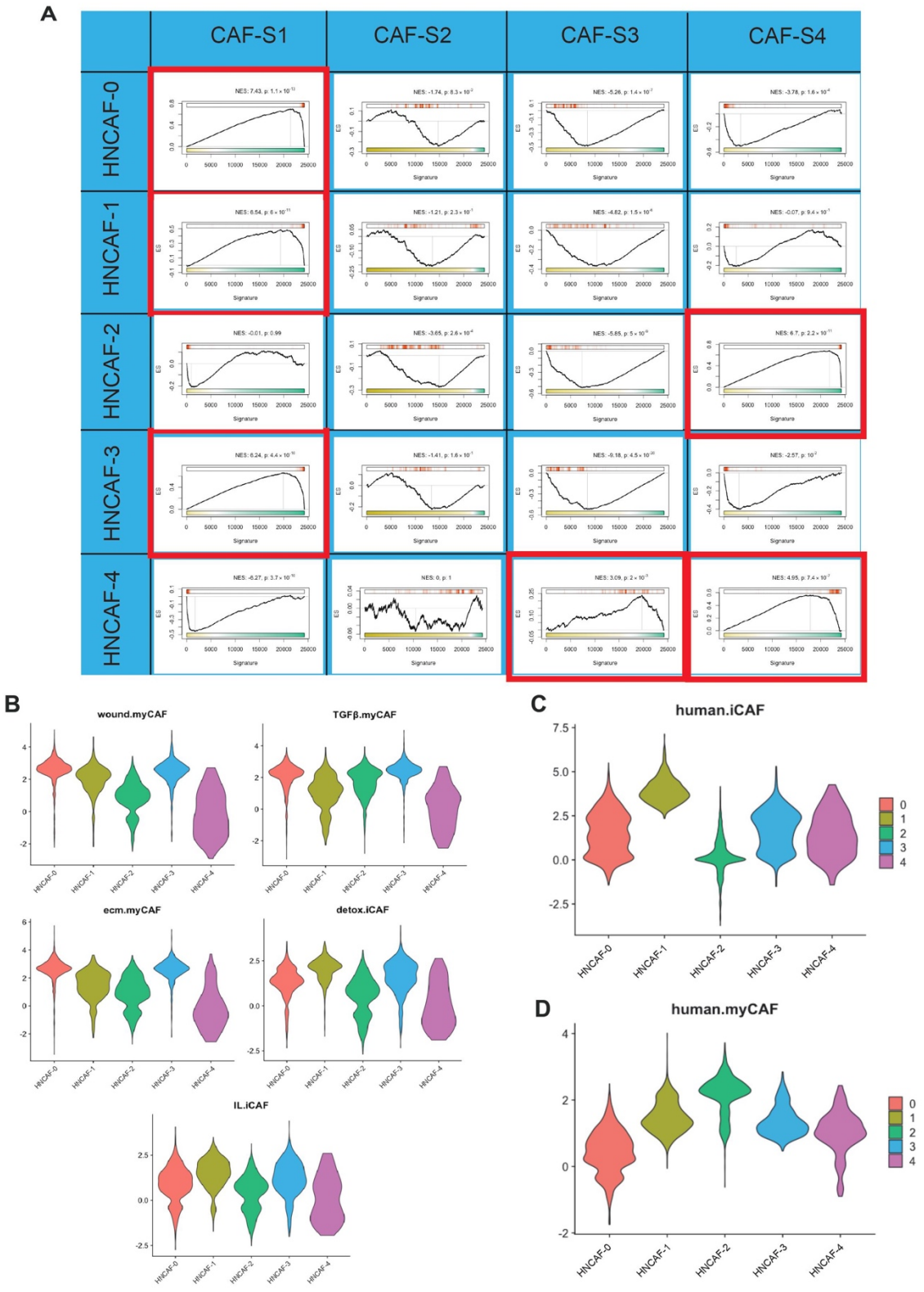


Figure 34: Phenotypic Matching of Single-Cell HNCAF Populations to Flow-Sorted Populations.

A) Pairwise Protein Activity Gene Set Enrichment of single-cell HNCAF gene sets (rows) among bulk RNA-Seq of sorted CAF populations, CAF-S1, CAF-S2, CAF-S3, and CAF-S4, as defined by [101]. Best match by enrichment for each HNCAF cluster is outlined in red. B) Cell-by-Cell enrichment of published CAF-S1 subcluster gene sets [100] in our single-cell HNCAF dataset, grouped by cluster. C-D) Cell-by-Cell enrichment of published iCAF and myCAF protein activity gene sets (Elyada at. al, 2019) in our single-cell HNCAF dataset, grouped by cluster, such that HNCAF-1 is enriched for iCAF gene set, and HNCAF-2 is enriched for myCAF gene set.

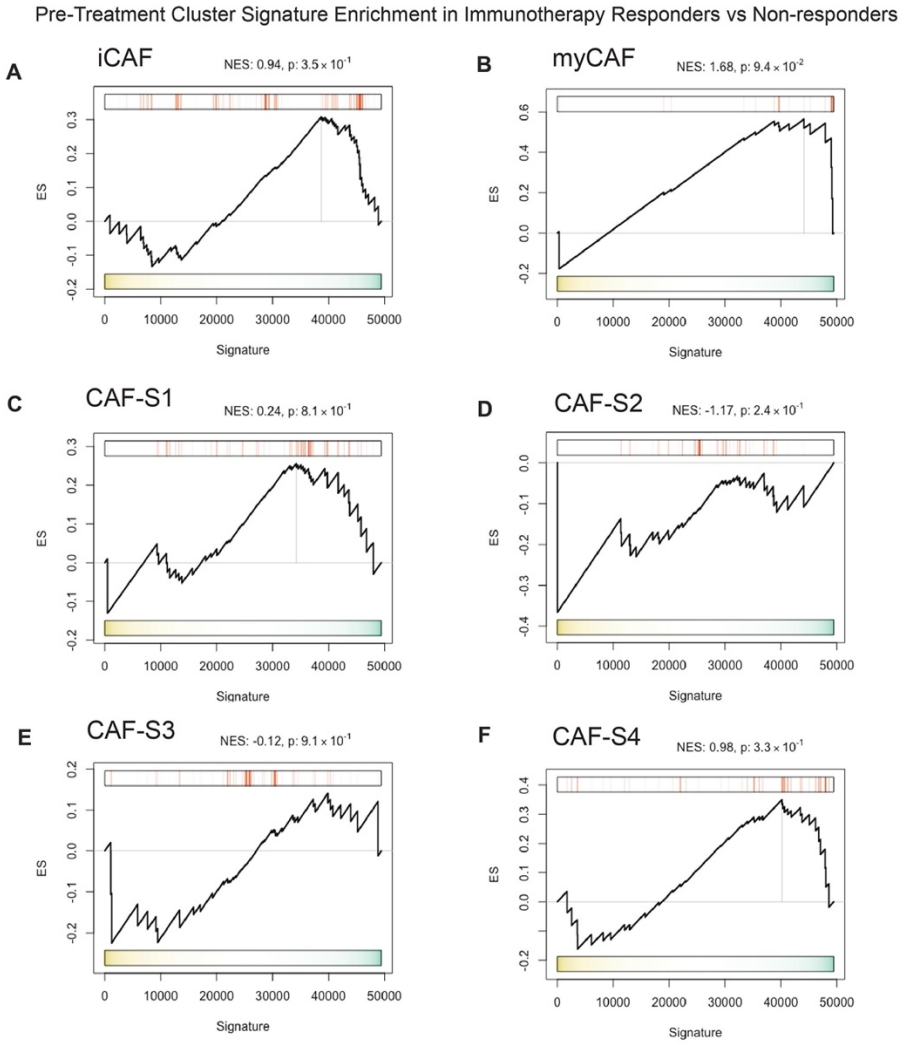


Figure 35: iCAF/myCAF and CAF-S1/S2/S3/S4 Population Markers are Not Significantly Enriched in Immunotherapy Responders vs Non-Responders Pre-Treatment.

A) Gene Set Enrichment plot of iCAF protein activity markers [33] in bulkRNA-Seq signature of immunotherapy responders vs non-responders, profiled pre-treatment. B) Gene Set Enrichment plot of myCAF protein activity markers [33] in bulkRNA-Seq signature of immunotherapy responders vs non-responders, profiled pre-treatment. C) Gene Set Enrichment of Top50 most-upregulated proteins from bulk-RNA-Seq of CAF-S1 vs CAF-S1/S2/S3/S4 in bulk-RNA-Seq signature of immunotherapy responders vs non-responders, profiled pre-treatment. D) Gene Set Enrichment of Top50 most-upregulated proteins from bulk-RNA-Seq of CAF-S2 vs CAF-S1/S2/S3/S4 in bulk-RNA-Seq signature of immunotherapy responders vs non-responders, profiled pre-treatment. E) Gene Set Enrichment of Top50 most-upregulated proteins from bulk-RNA-Seq of CAF-S3 vs CAF-S1/S2/S3/S4 in bulk-RNA-Seq signature of immunotherapy responders vs non-responders, profiled pre-treatment. F) Gene Set Enrichment of Top50 most-upregulated proteins from bulk-RNA-Seq of CAF-S4 vs CAF-S1/S2/S3/S4 in bulk-RNA-Seq signature of immunotherapy responders vs non-responders, profiled pre-treatment.

HNCAF-0 Predicts Favorable Disease Course in TCGA, in Contrast to HNCAF-1.

To evaluate the prognostic relevance of the CAF populations identified in a setting free from external immunotherapeutic pressures, we quantified the enrichment of HNCAF protein activity signatures in The Cancer Genome Atlas (TCGA) HNSCC cohort (Fig. 36, Fig. 37). Gene set enrichment (GSEA) analysis [104], on a patient-by-patient basis, revealed significant enrichment of the HNCAF-0 signature in patients with better overall survival in TCGA (Fig. 36A), suggesting that these cells may not only be important regulators of immunotherapy response but may also play a key role in mounting clinically relevant, endogenous immune responses against HNSCC. In contrast, the HNCAF-1 protein activity signature was associated with early worse overall survival when there were sufficient number of patients (Fig. 36B). Prognosis in HNSCC has been associated with higher TIL level in the TCGA cohort [105], and these results were consistent with the differential immunomodulatory functional roles of distinct VIPER derived HNCAF cells.

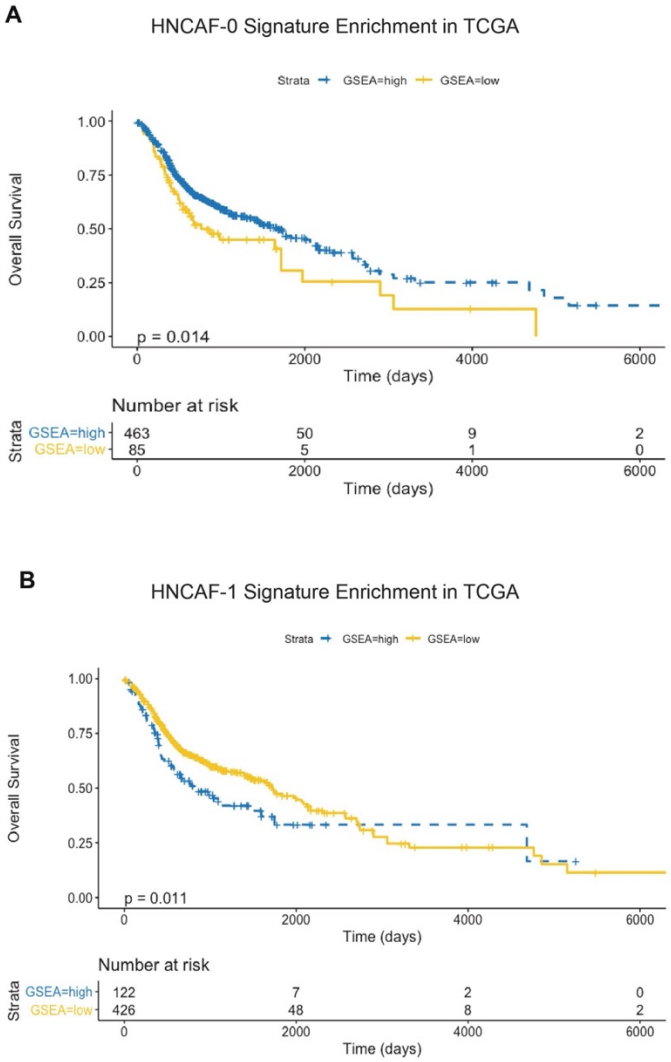


Figure 36: HNCAF-0 and HNCAF-1 have contrasting prognostic associations.

A) Kaplan-Meier plot of HNCAF-0 Protein Activity Gene Set Enrichment among TCGA dataset of head and neck squamous cell carcinoma patients in association with overall survival time. Enrichment scores were binarized by log-rank maximization to "high HNCAF-0" and "low HNCAF-0" and showed significant association with improved survival (p=0.014, median survival time = 602 days vs 1671 days). **B)** Kaplan-Meier plot of HNCAF-1 Protein Activity Gene Set Enrichment among TCGA head and neck squamous cell carcinoma patients in association with overall survival time, as in A. HNCAF-1 enrichment shows significant association with worse survival (p=0.011, median survival time = 1718 days vs 773 days). The crossover occurs at a point with very low number of patients.

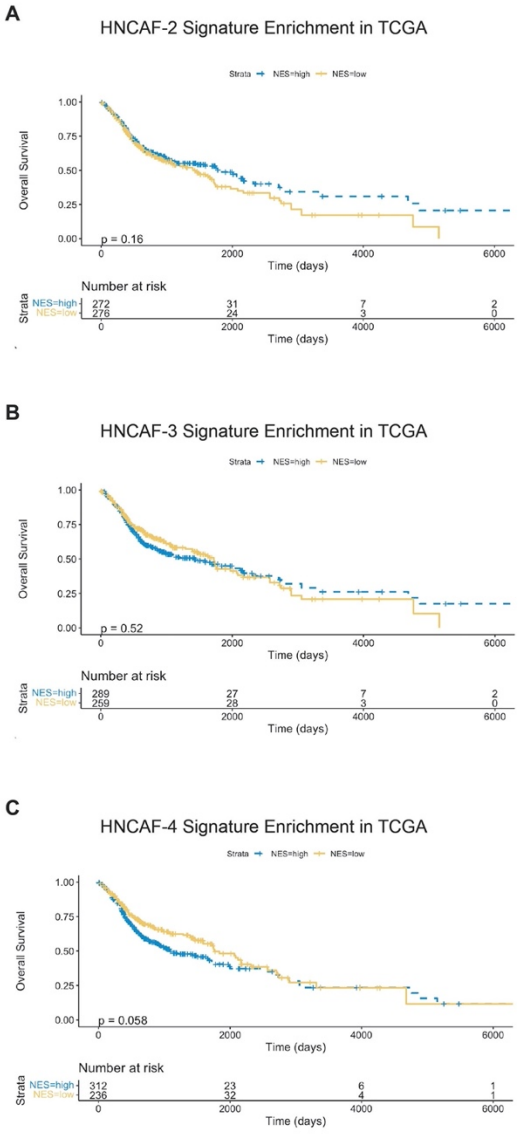


Figure 37: HNCAF-2, HNCAF-3, and HNCAF-4 are not associated with overall survival in TCGA.

Kaplan-Meier plot of HNCAF-2 (A), HNCAF-3 (B), and HNCAF-4 (C) Protein Activity Gene Set Enrichment among TCGA head and neck squamous cell carcinoma patients in association with overall survival time. Enrichment scores were binarized by log-rank maximization to "high HNCAF" and "low HNCAF" and show insignificant association with improved survival.

HNCAF-0/3 cells inhibit TGFβ dependent T-cell exhaustion in functional co-culture experiments.

Prompted by these intriguing clinical findings (Fig. 31D), we studied the potential interactions of HNCAF-0 and HNCAF-3 cells with other TME subpopulations. Interactome analysis showed that HNCAF have more receptor-ligand interactions with CD8 T cells than with any other cell subtype in the TME (see methods) (Fig. 29E). Therefore, we next interrogated the relationship between HNCAF-0 and human CD8 T cells *in situ* and *in vitro.* Digital spatial profiling (DSP, NanoString) of immune-related transcripts and protein markers was performed on HNSCC tissue from patients prior to nivolumab treatment. We first analyzed global CAF infiltration pattern in HNSCC tissue using aSMA, aCD8, and acytokeratin antibodies. These multiplexed immunofluorescent images exhibited colocalization of CAF with CD8+ T cells in the stromal compartment (Fig. 38A, white arrow). Reliable validated antibodies for each of the VIPER derived HNCAF subpopulations are currently unavailable to prevent histological analysis of each HNCAF population at this time.

To test whether HNCAF-0 cells can directly affect the biology of the CD8 T cells, we performed *in vitro* co-culture assays with HNCAF harvested from surgical resection mixed with either naïve T cells or tumor-infiltrating T cells. Primary fibroblasts enriched for HNCAF-0/3 were isolated from human HNSCC samples and their transcriptional identity was verified by RNA-Sequencing and protein activity analysis (Fig. 39). Due to the phenotypic similarity between HNCAF-0 and HNCAF-3, we were unable to enrich solely for HNCAF-0 and proceeded with a heterogeneous population enriched for both HNCAF-0 and HNCAF-3. When co-cultured with T cells isolated from peripheral blood mononuclear cells of healthy donors, HNCAF-0/3 cells reduced the PD-

1+TIM-3+ exhaustion phenotype among exogenously activated CD8 T cells and increased the CD103+NKG2A+ tissue resident memory phenotype, as well as their cytolytic function, based on Perforin and Granzyme B expression (Fig. 38B). It is important to note that reduced exhaustion and increased tissue resident memory phenotypes caused by HNCAF-0/3 was not due to PD-1:PD-L1 signaling (Fig. 40A-40B). Transwell co-culture assays revealed that HNCAF-0/3-mediated T cell activation increase and induction of tissue resident memory phenotypes—but not T cell exhaustion phenotype mitigation—depends on cell-to-cell contact (Fig. 38C). Additionally, coculture of HNCAF-0/3 cells with tumor-infiltrating T cells isolated directly from human HNSCC specimens resulted in an equivalent increase in tissue resident memory phenotype and cytotoxicity among CD8 T cells. HNCAF-0/3 cells could not rescue the exhaustion phenotype of terminally exhausted tumor-infiltrating T cells (Fig. 38D). Regardless, HNCAF-0/3 cells strongly promoted production of the activation markers, Perforin, Granzyme B, and IFNɣ, in tumor-infiltrating T cells (Fig. 38D-E), suggesting that HNCAF-0/3 may prevent terminal exhaustion in early activated T cells, while not be able to reverse the phenotype of already exhausted T cells from the TME. Notably, we found that HNCAF-0/3 completely rescued TGFβ-mediated PD-1/TIM-3 induction in culture, without inhibiting total TGFβ signaling, as defined by CD103 induction (Fig. 38F).

To evaluate CAF influences on T cell exhaustion *in situ* without a validated antibody for each HNCAF populations, we leveraged the digital spatial profiling data to evaluate colocalization of HNCAF-0 and HNCAF-1 protein activity signatures in regions enriched for the T-cell functional exhaustion signature. Indeed, the HNCAF-1 signature enrichment significantly correlated with increased T cell exhaustion signature enrichment ($r = 0.94$, $p = 0.0014$). In sharp contrast, the

HNCAF-0 signature was not significantly associated with the T-cell exhaustion signature in the TME region of interest (Fig. 38G-H). Given the association of HNCAF-1 cells with an immunosuppressive environment, we aimed to evaluate the direct impacts of isolated HNCAF-1 cells on T cell phenotypes in co-culture, as performed for HNCAF-0/3 cells. However, despite repeated experiments, T cells co-cultured with HNCAF-1 rapidly died, leaving an insufficient number of viable cells for further analyses (Fig. 38I, Fig. 40C-E). T cell death induction was not observed when T cells were cultured in isolation or with HNCAF-0/3 cells, suggesting HNCAF-1-mediated accelerated T cell apoptosis *ex vivo*.

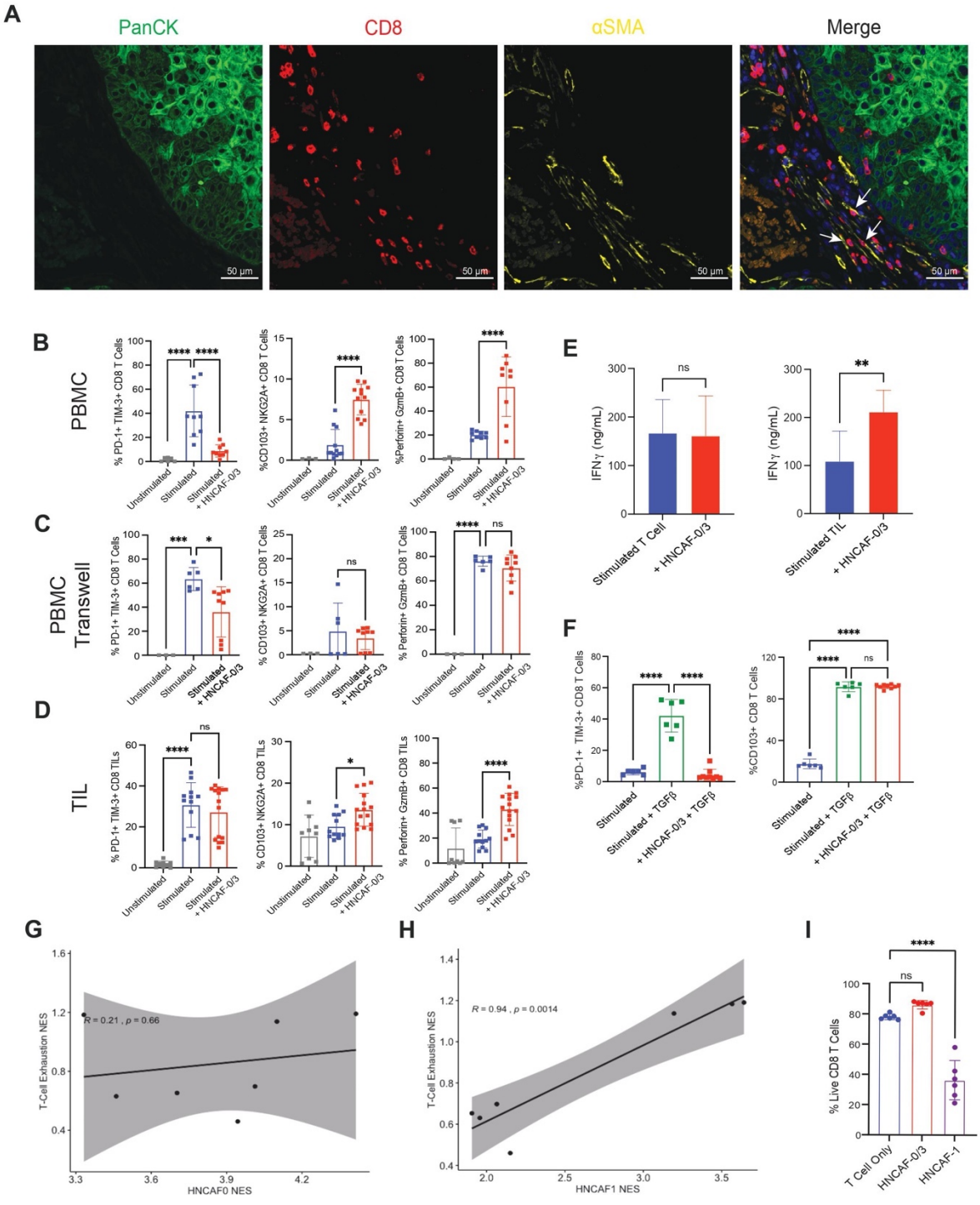
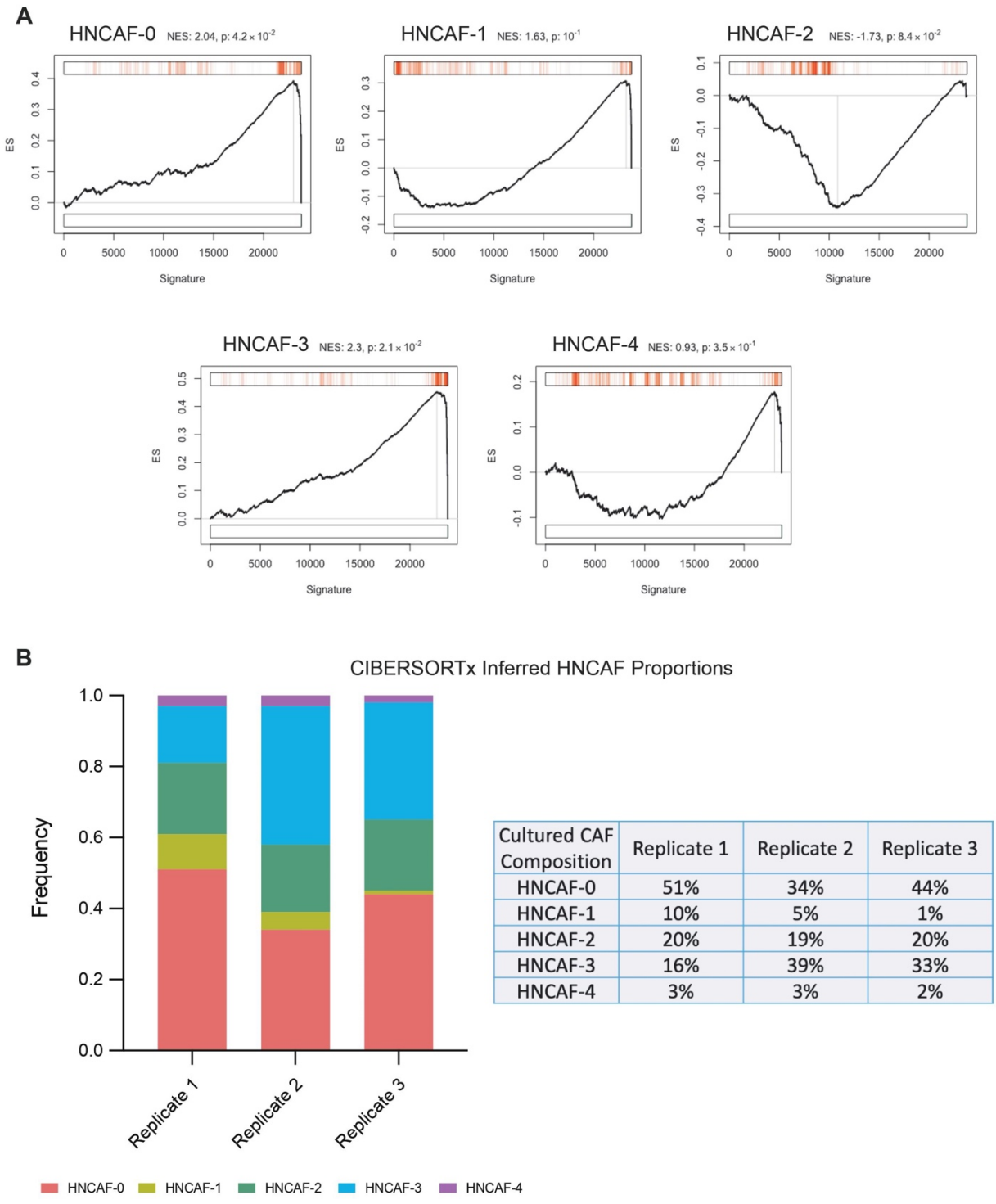


Figure 38: HNCAF spatially co-localizes with CD8 T-cells and HNCAF-0/3 functionally decrease TGFb dependent T-cell exhaustion in vitro.

A) Pre-treatment DSP immunofluorescence imaging from representative patient treated with αPD-1 immunotherapy, such that tumor cell localization is indicated by panCK staining (green),

CD8 T-cell localization, CD8 staining (4), fibroblast localization, αSMA staining (yellow), and nucleated cells, DAPI staining (blue). Arrows indicate interactions between αSMA+ fibroblasts and CD8+ T cells. **B)** Co-culture experiment of naïve T cells derived from peripheral blood mononuclear cells (PBMC) with CD3/CD28 stimulation and isolated HNCAF-0/3 cells, showing reduced T cell exhaustion (%PD-1+ TIM-3+ cells), increased tissue residency markers (%CD103+ NKG2A+ cells), and increased cytotoxicity (%Perforin+ GzmB+ cells). **C)** Co-culture experiment of T cells with HNCAF-0/3 cells as in B in contact-isolating transwell culture, showing reduced T cell exhaustion, but no significant difference in tissue residency markers or cytotoxicity. **D)** Co-culture experiment of Tumor-Infiltrating Lymphocytes (TIL) with CD3/CD28 stimulation and isolated HNCAF-0/3 cells, showing increased tissue residency markers and increased cytotoxicity. **E)** Interferon gamma levels in co-culture of naïve T cells derived from PBMC and TIL with HNCAF-0/3 cells determined by ELISA, showing significant increase in co-culture with TIL but not naïve T cells. **F)** Rescue experiment of T cells with CD3/CD28 stimulation and TGFβ with or without HNCAF-0/3. T-cell exhaustion markers on the left plot (%PD-1+ TIM-3+ cells) are rescued by HNCAF-0/3, and tissue localization markers on the right plot (%CD103+ cells) increase with TGFβ but are unaffected by addition of HNCAF-0/3. **G)** Spatial enrichment of HNCAF-0/3 gene set vs enrichment of T-cell exhaustion signature in Nanostring DSP of tissue slices across patients. No statistically significant association in spatial co-enrichment. **H)** Spatial enrichment of HNCAF-1 gene set vs enrichment of T-cell exhaustion signature in Nanostring DSP of tissue slices across patients. Signatures are positively correlated with respect to spatial co-localization (correlation coefficient = 0.94, $p = 0.0014$). **I)** Quantitation of live cells out of total CD8 T cells determined by flow cytometry from co-culture with HNCAF-0/3 or HNCAF-1. **B-D, F)** Percentages were quantified by flow cytometry. Results are shown as mean ± SD and are representative of at least three independent experiments. The data were analyzed using one-way ANOVA (B-D, F, I) or the Student *t* test (E) and * indicates $p<0.05$, ** indicates $p<0.01$, *** indicates $p<0.001$, and **** indicates $p<0.0001$.



Cultured CAF Composition	Replicate 1	Replicate 2	Replicate 3
HNCAF-0	51%	34%	44%
HNCAF-1	10%	5%	1%
HNCAF-2	20%	19%	20%
HNCAF-3	16%	39%	33%
HNCAF-4	3%	3%	2%

Figure 39: Isolated CAF in co-culture experiments are strongly enriched for HNCAF-0 and HNCAF-3.

A) Gene Set Enrichment plots of single-cell protein activity population markers for each HNCAF cluster (Supplemental Table 1) in bulk-RNASeq of isolated CAF used for co-culture

experiments. HNCAF-0 and HNCAF-3 gene sets are significantly enriched in cultured CAF ($p=4.2 \times 10^{-2}$ and $p=2.1 \times 10^{-2}$, respectively), and no other HNCAF signatures show statistically significant gene set enrichment. B) CIBERSORTx inference of cell type abundances in bulk-RNASeq of sorted CAF used for co-culture experiments, across three technical replicates, such that HNCAF-0 and HNCAF-3 constitute the majority of inferred cell frequency. For reference, in single-cell RNASeq (Figure 2) HNCAF-0 represents 43% of overall CAF frequency, HNCAF-1 represents 22%, HNCAF-2 represents 20%, HNCAF-3 represents 13%, and HNCAF-4 represents 2%.

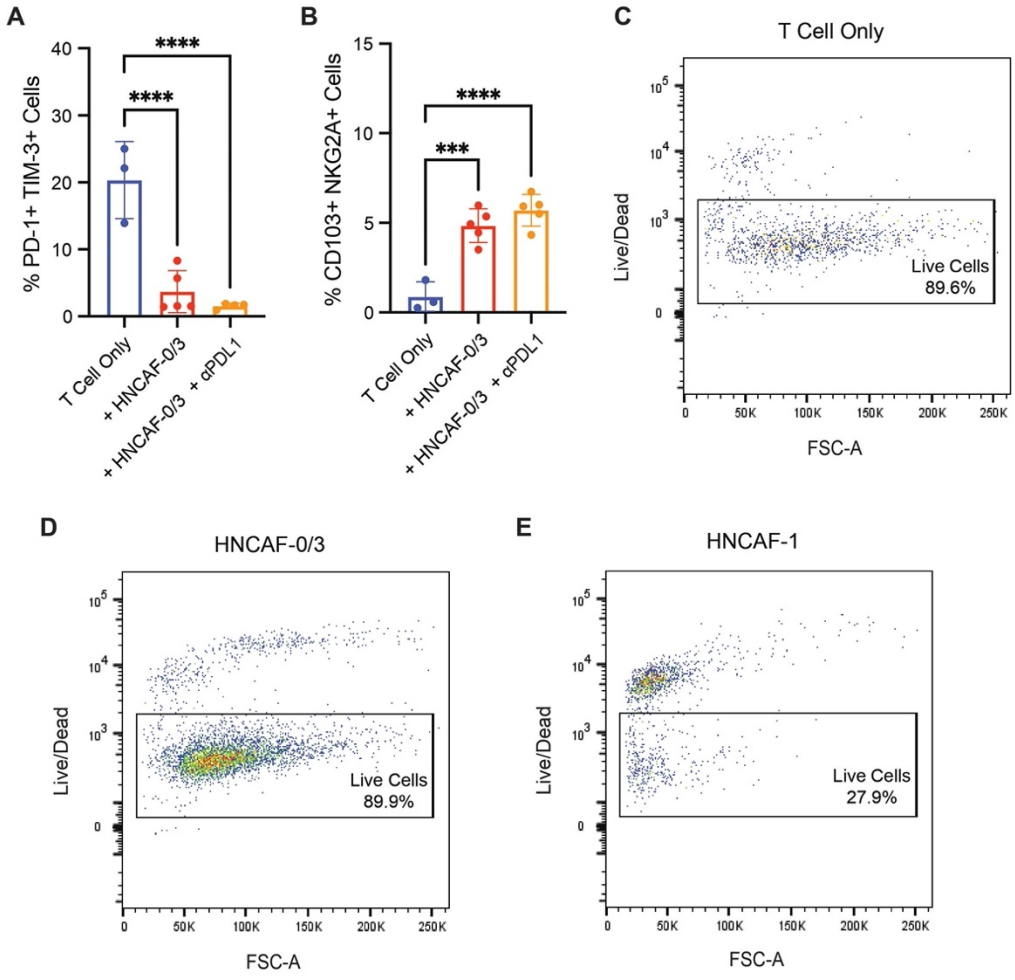


Figure 40: PD-L1 blockade does not affect HNCAF-0/3-induced T cell phenotypes and HNCAF-1 fibroblasts induce T cell death in co-culture experiments.

A-B) Co-culture experiment of Peripheral Blood Mononuclear Cells (PBMCs) with isolated HNCAF-0/3 cells with or without PD-L1 blocking antibody, showing reduced PD-1/TIM-3 (A) and increased CD103/NKG2A (B). *** indicates $p<0.001$ and **** indicates $p<0.0001$. C-E) Representative flow plots of T cells co-cultured without CAF (C) with HNCAF-0/3 (D) or with HNCAF-1 (E) for 5 days, gated on CD8+ T cells.

HNCAF-0 can predict clinical outcome to aPD-1 blocking antibodies.

To test for the potential generalizability of these HNCAF subpopulations, we next performed enrichment of HNCAF protein activity signatures across TCGA, by tumor type, focusing on tumors with high stromal cell content. Enrichment analyses revealed that HNCAF-0 enrichment is relatively specific to HNSCC while HNCAF-1 enrichment is more broadly observed (Fig. 41A-B). Intriguingly, HNCAF-1 enrichment is highest in pancreatic adenocarcinoma, which is known to be unresponsive to PD-1 based immunotherapy (Fig. 41B). HNCAF-1 phenotypically matches the previously defined iCAF population from pancreatic cancer (Fig. 34D). To externally validate our HNCAF-0/3's potential for clinical response prediction, we tested for enrichment of protein activity signatures in another cohort of HNSCC patients treated with αPD-1 immunotherapy, pembrolizumab [106]. The analysis revealed statistically significant enrichment of HNCAF-0 and HNCAF-3 marker genes in pre-treatment samples of patients who subsequently responded to pembrolizumab, validating the association of HNCAF-0/3 with immunotherapy response in an independent cohort (Fig. 42). Although not as enriched as HNCAF-0 and HNCAF-3, analysis of this cohort also revealed a significant enrichment of HNCAF-2 in responders pre-treatment (Fig. 42).

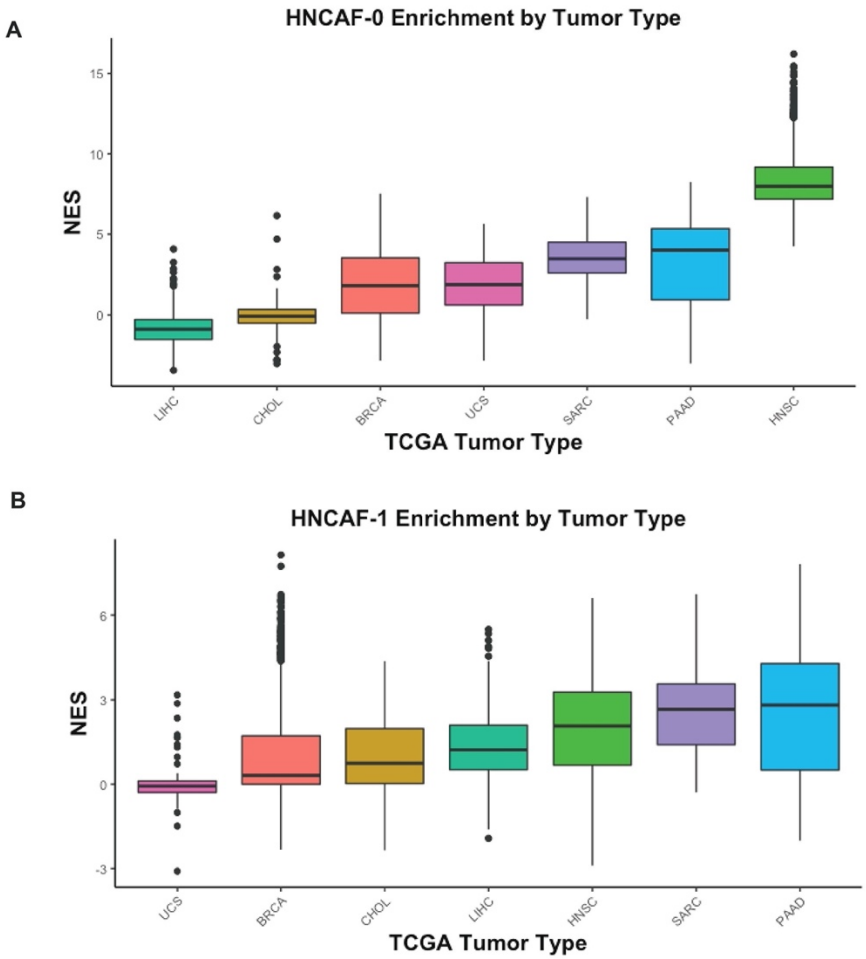


Figure 41: HNCAF-0 enrichment is highly specific to Head and Neck Squamous Cell Carcinoma.

A) Boxplot of HNCAF-0 protein activity gene set enrichment among TCGA tumor types with high stromal involvement. B) Boxplot of HNCAF-1 protein activity gene set enrichment among TCGA tumor types with high stromal involvement. LIHC: Liver Hepatocellular Carcinoma, CHOL: Cholangiocarcinoma, BRCA: Breast Cancer, UCS: Uterine Carcinosarcoma, SARC: Sarcoma, PAAD: Pancreatic Adenocarcinoma, HNSC: Head and Neck Squamous Cell Carcinoma

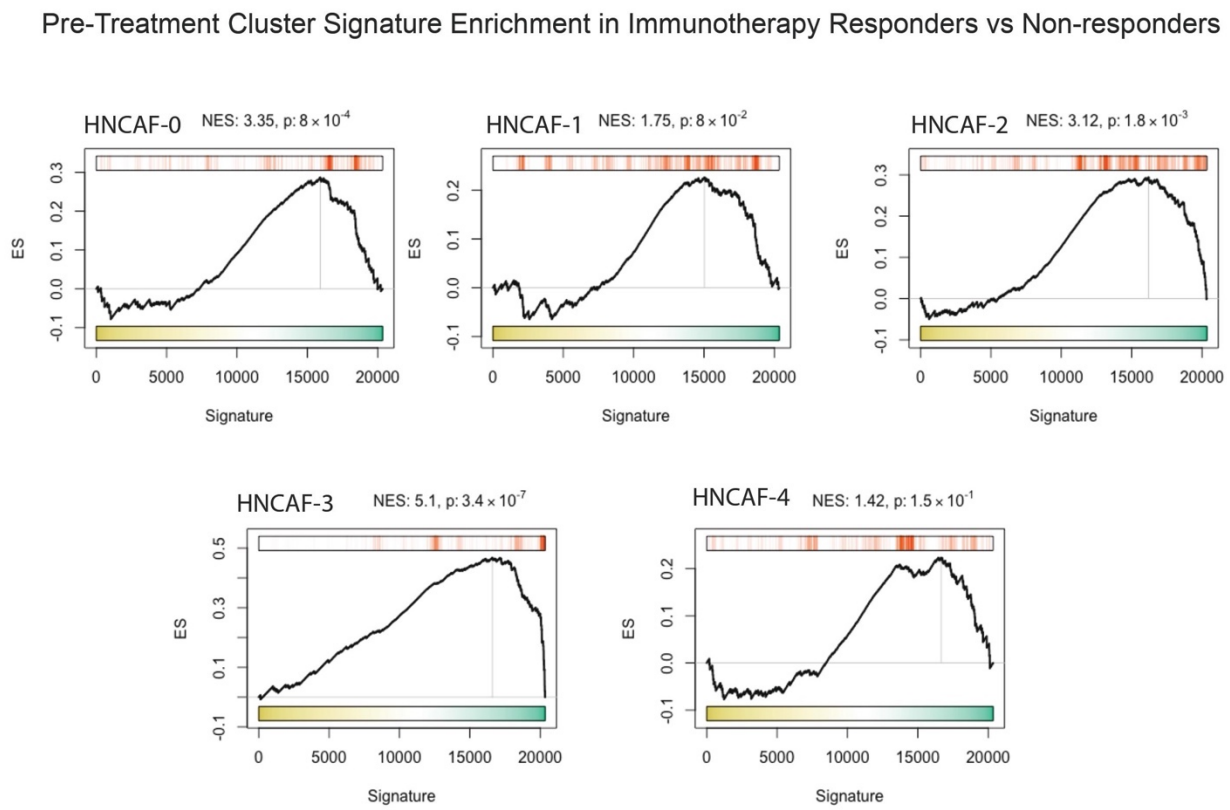


Figure 42: HNCAF-0 and HNCAF-3 are also predictive of favorable responses to pembrolizumab.

Protein Activity Profile Enrichment plots of single-cell protein population markers for each HNCAF cluster in bulk-RNASeq signature of pembrolizumab immunotherapy responders (5 patients) vs non-responders (15 patients), profiled pre-treatment from (Uppaluri et al, 2020). HNCAF-0, HNCAF-2, and HNCAF-3 profiles are significantly enriched in treatment responders ($p=8.0 \times 10^{-4}$, $p=1.8 \times 10^{-3}$, and $p=3.4 \times 10^{-7}$, respectively).

4.4 Discussion

In this study, we used protein activity profiles, as measured by the VIPER algorithm analysis of a longitudinal single-cell transcriptomics HNSCC dataset, to identify five molecularly distinct CAF subtypes. We took advantage of the longitudinally harvested biospecimens from a neoadjuvant clinical trial to show that two subtypes, HNCAF-0 and HNCAF-3, are predictive of favorable clinical responses to PD-1 checkpoint blockade therapy. Moreover, we discovered

HNCAF-0/3 cells have an immunostimulatory effect on CD8 T cells while HNCAF-1 cells are associated with immunosuppression. From a functional perspective, we have shown that HNCAF-0/3 fibroblasts prevent induction of an exhausted T cell phenotype and are associated with increased CD8 T cell cytotoxicity. HNCAF-1 fibroblasts, in contrast, correlate with increased T cell exhaustion *in situ* and T cell apoptosis *in vitro*, suggesting contrasting roles for these CAF subtypes. Functional significance of HNCAF-2 and HNCAF-4 subtypes have yet to be defined.

Protein activity-based clustering with VIPER has previously been shown to outperform gene expression-based clustering and we have corroborated this with the identification of five HNCAF subtypes compared to two with gene expression-based clustering [35]. Although we are not the first to identify CAF in HNSCC by single-cell sequencing, we were able to achieve greater resolution with VIPER allowing for a more granular picture of the CAF composition in HNSCC [100] [103]. We were also able to integrate the major subclasses of CAF identified in breast cancer and pancreatic cancer into our HNSCC CAF data. We showed through GSEA that CAF-S1, identified in breast cancer and in our HNSCC samples as $CD29^{+}FAP^{+}$ fibroblasts, correspond to three of the HNCAF groups we identified, HNCAF-0, HNCAF-1 and HNCAF-3 [101]. Kieffer et al. further dissected CAF-S1 in breast cancer and identified 8 subgroups but only focused on the 5 most abundant groups [100]. We performed GSEA of these 5 subgroups but were only able to identify one with significant enrichment in our HNCAF subpopulations - ecm-myCAF. We did find that the ecm-myCAF signature was significantly enriched on 4 of our 5 HNCAF groups. The discrepancies between our sub-clustering of CAF-S1 and Kieffer et al. may be due to tumor specific differences in the CAF heterogeneity between breast and HNSCC as

well as the use of a protein activity-based algorithm in contrast to gene expression. Additionally, we have found a CAF population (HNCAF-0) that is relatively specific to HNSCC amidst the other CAF types seen across different cancers. At this time, we hypothesize that tumor intrinsic factors may be influencing the differentiation of tumor specific CAF subpopulations (HNCAF-0) while nonspecific CAF subpopulations like HNCAF-1 could be derived from mesenchymal stem cells [107] [108] [109] [110] [111].

Despite the incomplete biological understanding of the HNCAF, our work introduces a novel clinically actionable biomarker for HNSCC. Immune checkpoint inhibitors (ICI) have revolutionized the field of cancer immunotherapy with monoclonal antibodies targeted against CTLA-4, PD-1, and PD-L1 being recently approved for use as frontline therapies for HNSCC and other cancer types [94] [96] [112]. The factors driving resistance to ICI remain largely unknown, making it difficult to select those who will respond and who will not. Accordingly, there remains an unmet need for reliable biomarkers predictive of response to guide patient selection and optimization of ICI treatment. In recent studies, CAF have been implicated to influence resistance to checkpoint immunotherapy. A preclinical model of pancreatic ductal adenocarcinoma showed that depletion of CAF expressing high levels of FAP improves response to αPD-L1 blockade [98]. Similarly, single-cell RNA sequencing revealed a LRRC15+ CAF population associated with worse response to αPD-L1 immunotherapy in a clinical trial for bladder cancer [99].

Furthermore, distinct CAF populations identified in breast cancer were also shown to be associated with poor αPD-1 immunotherapy response in melanoma and lung cancer [100]. All

these studies have implicated CAF primarily as contributors to resistance; however, the precise nature of molecularly distinct CAF subtypes and their role in mediating the effect of immunotherapy remains poorly investigated. With our ability to provide a higher resolution CAF repertoire, we show that the presence of two unique HNSCC-specific CAF subtypes is predictive of clinical response to immunotherapy in HNSCC. In particular, our functional findings suggest that HNCAF-0/3 fibroblasts are active participants in the immune response elicited by PD-1 directed immunotherapy through T cell modulation. For HNSCC, we are currently evaluating whether these CAF subtypes behave differently in virally associated HPV+ HNSCC vs. HPV- tumors.

Previous studies have typically shown CAF as promoters of immunosuppression. CAF have been shown to prevent T cell infiltration and to kill T cells in an antigen-dependent manner, via PD-L2 and FasL [98] [113]. CAF have also been shown to suppress T cells through upregulation of PD-L1 and PD-L2 and through recruitment of regulatory T cells [101] [114]. While confirming the immunosuppressive role of some HNCAF subtypes, our work has also established a novel pro-inflammatory role for distinct CAF subtypes, which act as a promoter of T cell activation and cytotoxicity. In light of this immunostimulatory function, we have termed the HNCAF-0/3 phenotype as T cell-stimulating CAF (tsCAF). Based on previous studies identifying TGFβ1 signaling through SMAD3 as a regulator of PD-1 expression, we hypothesize that tsCAF may repress SMAD3 to transcriptionally inhibit PD-1/TIM-3 expression [115]. Although our data strongly suggests that HNCAF-0/3 are immunostimulatory, the inability to sort these cells to obtain pure populations is a major limitation of this study as the CAF cells used in our functional studies do not encompass a pure population at this point. Proteomic based approach to select and

sort $CD29^{+}FAP^{+}$ CAF into tsCAF and HNCAF-1 is an active area of investigation. Recently, tumor restrictive $CD105^{-}$ CAF have been demonstrated in murine models of PDAC which are mediated by the adaptive immune system [116]. $CD105^{-}$ CAF highly overlap with the myCAF gene signature, which has also previously been demonstrated to be tumor constraining [117]. Since the tsCAF we describe here are distinct from the myCAF population both molecularly and by surface marker expression, we believe they represent a distinct CAF population from $CD105^{-}$ CAF.

Interestingly, we found that co-culture of HNCAF-0/3 with CD8 T cells induced a tissue-resident memory (Trm) phenotype that co-expressed NKG2A. NKG2A is an inhibitory receptor that we and others found to be highly enriched in tumor-infiltrating Trm^{+} CD8 T cells in HNSCC [118] [119]. NKG2A ligation with its ligand HLA-E reduces cytotoxicity and effector function and is therefore a novel immunotherapy target [120]. Clinical trials combining NKG2A blockade (monalizumab) and other checkpoint inhibitors have shown promising results in HNSCC [118]. While it is not clear why NKG2A is upregulated on tumor-infiltrating CD8 T cells, our findings suggest HNCAF characterization may also inform clinical development of NKG2A inhibition along with other immune checkpoint inhibitors. We found that upregulation of NKG2A required contact between the HNCAF-0/3 and activated CD8 T cells, which suggests that induction is mediated by either a surface ligand on HNCAF-0/3 or a component in the extracellular matrix produced by the CAF. To our knowledge, NKG2A expression on CD8 T cells has never been associated with fibroblasts and our finding here provides a clear link between intra-tumoral NKG2A expression on CD8 T cells and the tumor stroma. Future studies will need to be

performed on HNSCC specimens to determine if NKG2A expression on CD8 T cells is localized to stromal regions or associated with increased HNCAF-0/3 cells.

Plasticity of CAF subtypes have also been well demonstrated [102] [121], and our CAF atlas provides an excellent framework to develop strategies to force CAF differentiation towards the pro-inflammatory tsCAF phenotype rather than the immunosuppressive HNCAF-1 phenotype, to be combined with immunotherapy. Our characterization of each CAF subpopulation characterization through their master regulatory network lends itself towards this strategy. Further investigation of the signals that induce tsCAF formation and activation, possibly by targeting master regulators of the two subtypes, either genetically [66] [65] or pharmacologically, via the OncoTreat algorithm [24] is currently underway. Critically, this study highlights a much greater molecular heterogeneity of CAF subtypes than previously appreciated and demonstrates the critical need to functionally characterize their pleiotropic effects in terms of cancer progression, outcome, and response to immunotherapy and other cancer treatments.

4.5 Methods

Clinical Design and Tissue Collection

Biospecimens were harvested from a window of opportunity trial of locally advanced HNSCC patients (oral cavity, oropharynx, larynx, hypopharynx) who were candidates for primary surgical intervention with curative intent (NCT03238365). All enrolled patients were treated with 1 month of 240mg nivolumab every 2 weeks for 2 doses prior to definitive surgery (N=50). Half of the patients received tadalafil, and no differences were noted in response rates between the two cohorts [122]. Consented patients were required to have fresh pre-nivolumab biopsy as

well pre and post imaging. Meta-clinical annotation using pathological criteria was used to delineate paired subject specimens as responders vs. non-responders. For both pre and post treatment timepoints, fresh specimens were collected for frozen fixation, paraffin embedded fixation, and processed for both bulk and single-cell transcriptomic sequencing. Due to insufficient RNA quality, only 36 of the 50 patients were used for bulk and single-cell RNA sequencing.

Clinical Design and Tissue Collection

Fresh head and neck squamous cell carcinoma tumor specimens were collected in DMEM supplemented with streptomycin (200 mg/ml), penicillin (200 U/ml), and amphotericin B (250 mg/mL). Tumor specimens were minced to 2-4 mm sized pieces in separate 6-cm dishes and digested to single cell suspension using the Miltenyi Biotec human tumor dissociation kit (Miltenyi Biotec #130-095-929) on the Miltenyi gentleMACS Octo dissociator (Miltenyi Biotec #130-096-427) following manufacturer's instructions. Dissociated cells were aliquoted for single-cell sequencing, flow cytometry analysis, or CAF culture.

Single-Cell RNA-Sequencing

Samples were processed to generate single-cell gene expression profiles (scRNA-Seq) using the 10X Chromium 3' Library and Gel Bead Kit (10X Genomics), following the manufacturer's user guide. After GelBead in-Emulsion reverse transcription (GEM-RT) reaction, 12-15 cycles of polymerase chain reaction (PCR) amplification were performed to obtain cDNAs used for RNAseq library generation. Libraries were prepared following the manufacturer's user guide and sequenced on the Illumina NovaSeq 6000 Sequencing System. Gene expression data were

processed with “cellranger count” on the pre-built human reference set of 30,727 genes to generate counts matrices. Cell Ranger performed default filtering for quality control, and produced for each sample a barcodes.tsv, genes.tsv, and matrix.mts file containing counts of transcripts for each sample, such that the expression of each gene is in terms of the number of unique molecular identifiers (UMIs) tagged to cDNA molecules corresponding to that gene. These data were loaded into the R version 3.6.1 programming environment, where the publicly available Seurat package v3.0 was used to further quality-control filter cells to those with fewer than 10% mitochondrial RNA content, more than 1,500 unique UMI counts, and fewer than 15,000 unique UMI counts.

Single-Cell Data Processing

Gene Expression UMI count matrices for each sample were normalized and scaled in R using the Seurat SCTransform command to perform a regularized negative binomial regression based on the 3000 most variable genes. Scaled data from each patient were batch-corrected by Seurat using the functions FindIntegrationAnchors and IntegrateData, with default parameters. The resulting dataset included 22906 high-quality cells (mean UMI count 4802) across four patients, including both pre-treatment and post-treatment time points for each patient (Patient1: 5857 pre-treatment, 7360 post-treatment, Patient2: 4938 pre-treatment, 1550 post-treatment, Patient3: 487 pre-treatment, 1741 post-treatment, Patient4: 401 pre-treatment, 572 post-treatment). The batch-corrected dataset was projected into its first 50 principal components using the RunPCA function in Seurat, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. The data were clustered by the Louvain algorithm with silhouette score resolution-optimization

selecting the resolution with maximum bootstrapped silhouette score in the range of resolution from 0.01 to 1.0 incremented by 0.01 [35]. This resulted in an initial coarse clustering on gene expression (Fig. 29A). Within each cluster metaCells were computed for downstream regulatory network inference by summing SCTransform-corrected template counts for the 10 nearest neighbors of each cell by Pearson correlation distance.

For each single cell, inference of cell type was performed using the SingleR package and the preloaded Blueprint-ENCODE reference, which includes normalized gene expression values for 259 bulk RNASeq samples generated by Blueprint and ENCODE from 43 distinct cell types representing pure populations of stroma and immune cells [49] [50]. The SingleR algorithm computes correlation between each individual cell and each of the 259 reference samples, and then assigns both a label of the cell type with highest average correlation to the individual cell and a p-value computed by wilcox test of correlation to that cell type compared to all other cell types. Highest-Frequency SingleR labels within each cluster among labels with $p<0.05$ are projected into the Gene Expression UMAP space in Fig. 29B, such that localization of SingleR labels is highly concordant with the unsupervised clustering. Unsupervised Clusters determined by the resolution-optimized Louvain algorithm are therefore labelled as a particular cell type based on the most-represented SingleR cell type label within that cluster.

Differential gene expression analysis between single cell clusters throughout the manuscript is computed by the MAST hurdle model, as implemented in the Seurat FindAllMarkers command, with a log-fold change threshold of 0.5 and minimum fractional expression threshold of 0.25,

indicating that the resulting gene markers for each cluster are restricted to those with log fold change greater than 0 and non-zero expression in at least 25% of the cells in the cluster.

Regulatory Network Inference

From the integrated dataset, metaCells were computed within each gene expression-inferred cluster by summing SCTransform-corrected template counts for the 10 nearest neighbors of each cell by Pearson correlation distance. 200 metaCells per cluster were sampled to compute a regulatory network from each cluster. All regulatory networks were reverse engineered by the ARACNe algorithm. ARACNe was run with 100 bootstrap iterations using 1785 transcription factors (genes annotated in gene ontology molecular function database as GO:0003700, "transcription factor activity", or as GO:0003677, "DNA binding" and GO:0030528, "transcription regulator activity", or as GO:0003677 and GO:0045449, "regulation of transcription"), 668 transcriptional cofactors (a manually curated list, not overlapping with the transcription factor list, built upon genes annotated as GO:0003712, "transcription cofactor activity", or GO:0030528 or GO:0045449), 3455 signaling pathway related genes (annotated in GO biological process database as GO:0007165, "signal transduction" and in GO cellular component database as GO:0005622, "intracellular" or GO:0005886, "plasma membrane"), and 3620 surface markers (annotated as GO:0005886 or as GO:0009986, "cell surface"). ARACNe is only run on these gene sets so as to limit protein activity inference to proteins with biologically meaningful downstream regulatory targets, and we do not apply ARACNe to infer regulatory networks for proteins with no known signaling or transcriptional activity for which protein activity may be difficult to biologically interpret. Parameters were set to zero DPI (Data Processing Inequality) tolerance and MI (Mutual Information) p-value threshold of 10^{-8},

computed by permuting the original dataset as a null model. Each gene list used to run ARACNe is available on github [35].

Protein Activity Inference

Protein activity was inferred for all cells by running the metaVIPER algorithm, using all cluster-specific ARACNe networks, on the SCTransform-scaled and Anchor-Integrated gene expression signature of single cells from each patient. Because the SCTransform Anchor-Integrated scaled gene expression signature is already normalized as an internal signature comparing all cells to the mean expression in the dataset, VIPER normalization parameter was set to "none." The resulting VIPER matrix included 1239 proteins with activity successfully inferred from ARACNe gene regulatory networks. VIPER-Inferred Protein Activity matrix was loaded into a Seurat Object with CreateSeuratObject, then projected into its first 50 principal components using the RunPCA function in Seurat, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Clustering was performed by resolution-optimized Louvain algorithm, as for gene expression (Fig.30A, Fig. 29A), and SingleR-inferred cell type labels were carried over to identify cluster-by-cluster cell type labels (Fig. 30D). Differential Protein Activity between clusters identified by resolution-optimized Louvain was computed using bootstrapped t-test, run with 100 bootstraps, and top proteins for each cluster were ranked by p-value (Fig. 30B). The entire pipeline is implemented as in (15). Cluster cell counts were normalized to a fraction of the total sample separately for each patient and separately for pre-treatment and post-treatment samples, with differences in pre-treatment vs post-treatment frequency distribution plotted in Fig. 30C.

Association of HNCAF Signatures with Response to Immunotherapy

Fibroblast clusters including 5,414 cells from overall VIPER clustering of all cells were further isolated and sub-clustered (Fig. 31A), with differential protein activity and frequency pre-treatment vs post-treatment compared as in the analysis of initial clustering for all cells. A proteomic gene set for each head and neck cancer-associated fibroblast (HNCAF) cluster was defined based on proteins differentially upregulated in each cluster. In the dataset of clinical trial patients profiled by bulkRNASeq that had been annotated with subsequent response (n=9) or non-response (n=19) to αPD-1 immunotherapy (39), we applied VIPER transformation using the single-cell ARACNe networks on z-score scaled log10(TPM) counts from pretreatment bulk-RNA-Seq data, and computed a differential protein activity signature ranking proteins by most upregulated in responders to most downregulated in responders. Enrichment of each HNCAF cluster marker set in the VIPER-transformed signature of responders vs nonresponders from bulkRNASeq was computed by Gene Set Enrichment Analysis (GSEA), with normalized enrichment score and p-value determined by 1000 random permutations of gene labels (Fig. 31D). Insufficient number of HPV+ samples prevented CAF enrichment analysis in HPV+ vs HPV- groups.

Clinical association of HNCAF cluster 0 and cluster 1 signatures with outcome was further tested in TCGA head and neck cancer cohort processed by ARACNe and VIPER as above. Sample-by-Sample Normalized Enrichment Scores were computed ranking VIPER-inferred protein activity in each patient sample from highest to lowest activity and then applying GSEA. Normalized Enrichment scores for HNCAF cluster signatures were binarized to less than zero (low) or

greater than zero (high), and Kaplan-Meier curve showing association with Overall Survival time was plotted along with the log-rank p-value (Fig. 36, Fig. 37), such that HNCAF-0 enrichment is associated with improved overall survival (p=0.014, median survival time = 602 days vs 1671 days) and HNCAF-1 enrichment is associated with worse overall survival (p=0.011, median survival time = 1718 days vs 773 days). We further plotted the sample-by-sample enrichment of these HNCAF populations among different TCGA tumor types with high stromal involvement (HNSC, PAAD, SARC, UCS, BRCA, CHOL, LIHC) and plotted the distribution of these enrichment scores by tumor type to assess relative tumor-type specificity of the identified HNCAF signatures (Fig. 41).

Digital Spatial Profiling

NanoString GeoMX Digital Spatial profiling was further applied, profiling IO360 immune gene panel expression among three regions of interest (ROIs) from one patient and four ROIs from another. Anti-CD8, anti-αSMA, anti-PanCK, and DAPI stains were used for morphology identification and ROIs were selected based on high abundance of tumor (PanCK), cytolytic T cells (CD8), and fibroblasts (αSMA). ROIs were split into PanCK-positive and PanCK-negative components, with gene expression evaluated separately in each. In order to further assess spatial co-localization of HNCAF subtypes with functionally exhausted T-cells, we applied segment-by-segment gene set enrichment of HNCAF-0 and HNCAF-1 markers as well as enrichment of a published T-cell exhaustion signature [123], and correlate normalized enrichment scores for these populations between spatial segments (Fig. 38G-H).

CAF Phenotyping

In order to assess phenotypic concordance between prior fibroblast categorizations, including CAF-S1/S2/S3/S4 subtypes described in the setting of breast cancer [101] and iCAF/myCAF subtypes described in the setting of pancreatic cancer [33], we have performed pairwise gene set enrichment of fibroblast phenotype marker gene sets among our HNCAF clusters identified by scRNA-Seq. Published iCAF/myCAF VIPER-inferred marker gene sets [33] were directly tested by GSEA for enrichment in each single-cell, with resulting enrichment scores plotted by HNCAF cell cluster in Fig. 34C-D, such that cells in HNCAF-1 are enriched for iCAF signature and cells in HNCAF-2 are enriched for myCAF signature. For CAF-S1/S2/S3/S4 phenotype-matching, we sorted S1/S2/S3/S4 cells by FACS using the gating strategy described by [101] (Fig. 32), and subsequently performed bulk-RNA Sequencing of each sorted population, applied VIPER using single-cell derived ARACNe networks, and computed differential protein activity of each population against the mean to define population-specific signatures, with genes ranked from most differentially-upregulated protein activity to most differentially-downregulated protein activity in CAF-S1/S2/S3/S4. We then performed pairwise Gene Set Enrichment Analysis of HNCAF cluster marker gene sets (by VIPER protein activity) among CAF-S1/S2/S3/S4 gene signatures (Fig. 34A). We highlight the findings that CAF-S1 gene signature was significantly enriched for the gene sets of HNCAF-0 (NES=7.43, $p=1.1 \times 10^{-13}$), HNCAF-1 (NES=6.54, $p=6 \times 10^{-11}$), and HNCAF-3 (NES=6.24, $p=4.4 \times 10^{-10}$), CAF-S2 gene signature was not significantly enriched for any HNCAF population, CAF-S3 signature was significantly enriched for HNCAF-4 gene set (NES=3.09, $p=2 \times 10^{-3}$), and CAF-S4 signature was significantly enriched for HNCAF-2 gene set (NES=6.7, $p=2.2 \times 10^{-11}$). This phenotypic classification scheme is shown on Fig. 33D and highlights the distinction between our HNCAF categorization observed from scRNA-Seq and prior CAF classification paradigms.

Receptor-Ligand Interactions

Receptor-Ligand Interactions were inferred between coarse-grained cell types using 2,557 high-quality receptor-ligand interactions reported in the RIKEN FANTOM5 database [124]. This list of receptor-ligand pairs was filtered to identify pairs where the ligand was significantly upregulated, at the gene expression level, in at least one subpopulation, across patients, and the corresponding receptor was significantly activated in another subtype, based on VIPER protein activity analysis, as proposed in [35]. We further filtered these interactions to those detected in cancer-associated fibroblasts and plotted the number of unique receptor-ligand interaction pairs inferred between fibroblasts and each other detected subpopulation (Fig. 29E).

CAF isolation and culture

Fresh head and neck squamous cell carcinoma tumor specimens were processed to single cell suspension as described above. For HNCAF-0/3, tumor single cell suspension was cultured in DMEM supplemented with 10% FBS, streptomycin (100 µg/ml), and penicillin (100 U/ml) for 2-3 weeks at 37°C until fibroblasts grew out. For HNCAF-1, tumor single cell suspension was cultured in pericyte medium (ScienCell #1201) supplemented with 2% FBS, streptomycin (100 µg/ml), and penicillin (100 U/ml) for 2-3 weeks at 37°C until fibroblasts grew out. To verify CAF identity, RNA was isolated from CAF lysates using TRIzol (Invitrogen #10296010) and sent for bulk RNA sequencing. Gene set enrichment analyses for the HNCAF subtype protein activity signatures were then performed on the bulk sequencing data, along with inference of cell type proportions by CIBERSORTx. Fibroblasts were passaged when cultures reached ~80% confluence and all experiments were performed with CAF under 10 passages.

T cell isolation

CD3+ T lymphocytes were isolated from the peripheral blood of healthy human donors. Peripheral blood mononuclear cells (PBMCs) were isolated using Ficoll-Paque Plus, following manufacturer's instructions. CD3+ T cells were isolated from PBMCs using magnetic bead sort with the MojoSort Human CD3 T Cell Isolation Kit (Biolegend #480022) according to manufacturer's instructions. For isolation of CD3+ tumor-infiltrating lymphocytes (TILs), fresh head and neck squamous cell carcinoma tumor specimens were processed to single cell suspension as described above. CD3+ tumor-infiltrating lymphocytes were isolated from the tumor single cell suspension using magnetic bead sort with the MojoSort Human CD3 T Cell Isolation Kit.

T cell and CAF Coculture Assays

25,000 primary CAF were plated in DMEM supplemented with 10% FBS in 96 well plates. After CAF had attached to the plate, 50,000 CD3+ T cells were added to the CAF with or without CD3/CD28 activation beads (Gibco # 11131D) and cocultured at 37°C for 5-7 days with or without 20 ng/mL TGFβ. Media was renewed on days 3 and 5. Cocultures with tumor-infiltrating lymphocytes were only cultured for 3 days to preserve TIL viability. Following incubation, T cells were harvested and stained with Live/Dead Aqua (1:1600) for 15 minutes in PBS. Cells were then washed and stained for 15 minutes with an antibody cocktail containing anti-CD4-APC/Fire 810 (1:1000, SK3), anti-CD8-BB515 (1:200, RPA-T8), anti-PD-1-BV421 (1:100, EH12.2H7), anti-TIM-3-BV786 (1:100, F38-2E2), anti-NKG2A-PE (1:200, S19004C), and anti-CD103 (1:400, Ber-ACT8). Cells were then washed, fixed, permeabilized and stained

with an intracellular antibody cocktail containing anti-Perforin-PerCP/Cy5.5 (1:100, B-D48) and anti-Granzyme B-Alex Fluor 700 (1:100, QA16A02). Cells were subsequently analyzed by flow cytometry using the Cytek Aurora.

Transwell T cell and CAF Coculture Assays

100,000 primary CAF were plated in DMEM supplemented with 10% FBS in the lower chamber of the transwell (0.4 μm pore size, Corning Polycarbonate Membrane Transwells #3401). 200,000 CD3+ T cells were plated in DMEM supplemented with 10% FBS in the upper chamber of the transwell. Cells were incubated at 37°C for 7 days. Media was renewed on days 3 and 5. Following incubation, T cells were stained and analyzed by flow cytometry using the Cytek Aurora as described above.

ELISA

The level of IFNɣ in cell culture supernatants was measured using an ELISA MAX Deluxe kit (Biolegend #430104) following manufacturer's instructions. Supernatants were collected from CAF-T cell cocultures as described above.

Statistical Analysis

All quantitative and statistical analyses were performed using the R computational environment and packages described above with the exception of CAF composition and co-culture experiments. Statistical analyses of these assays were performed using Prism 9 software (GraphPad). Differential gene expression was assessed at the single-cell level by the MAST single-cell statistical framework as implemented in Seurat v3 [32], and differential VIPER

activity was assessed by t-test, each with Benjamini-Hochberg multiple-testing correction. Comparisons of cell frequencies were performed by non-parametric Wilcox rank-sum test, and survival analyses were performed by log-rank test. In all cases, statistical significance was defined as an adjusted p-value less than 0.05. Details of all statistical tests used can be found in the corresponding figure legends.

Chapter 5: Single-Cell-Derived Proteomic Signatures Strongly Predict anti-PD1 Response in CTLA-4 Naive and CTLA-4 Pre-Treated Melanoma

5.1 Summary

Anti-PD1 Checkpoint immunotherapy has been transformative in the treatment of melanoma, and is now considered first-line therapy, with response rates of 40-45% in patients with melanoma given pembrolizumab or nivolumab in the first line [2]. To date, PD-L1 expression in tumor tissue has been the most-utilized biomarker for anti-PD1 response, yet this has been poorly predictive, with negative predictive value as low as 58% for anti-PD1 (nivolumab) and 45% for combination with anti-CTLA4 immunotherapy (ipilimumab) [2]. Here, we leverage published single-cell RNA-Seq datasets profiling melanoma patients who did or did not respond to anti-PD1 checkpoint blockage [17] [18] in combination with a protein activity inference pipeline [35] and random forest feature selection, in order to identify a single-cell signature of inferred protein activity which is strongly predictive of anti-PD1 immunotherapy response across bulk-RNA-Seq datasets [7] [125]. Crucially, distinct predictive signatures were identified for treatment-naïve patients receiving anti-PD1 (nivolumab or pembrolizumab) as first-line therapy and for patients previously treated with anti-CTLA4 (ipilimumab). Optimal predictive power was achieved using immune cell population marker proteins, with out-of-bag training AUC in Riaz et al. for the treatment-naïve signature of 0.972 (95% CI 0.908 – 1.0), and out-of-bag training AUC in Riaz et al. for the CTLA4-experienced signature of 0.909 (95% CI 0.731 – 1.0). We independently validated these predictive models in the Liu et al. cohort, achieving an test AUC for treatment-naïve

signature of 0.839 (95% CI 0.733 -0.944), and a test AUC for CTLA4-experienced signature of 0.893 (95% CI 0.771 – 1.0).

5.2 Results

We have identified two significant single-cell RNA-Sequencing datasets which profile melanoma patients who did or did not respond to anti-PD1 checkpoint immunotherapy. The first dataset includes 16,292 cells across 48 nivolumab-treated patients, including 31 non-responders and 17 responders, defined by RECIST criteria [17]. The manuscript primarily focuses on characterization of changes induced in the melanoma micro-environment by treatment in responders versus non-responders, but it also includes pre-treatment profiles from 8 subsequent responders and 10 subsequent non-responders. Each sample includes only sorted CD45-positive immune cells. We have analyzed this subset of pre-treatment cells from subsequent responders vs non-responders by both gene expression clustering using the standard Seurat v3 pipeline, as well as by protein activity inference, as described previously in [35]. For clusters identifed by each method, we report their

frequency in responders vs non-responders and predictive value by random forest model with out-of-bag AUC (Figure 43).

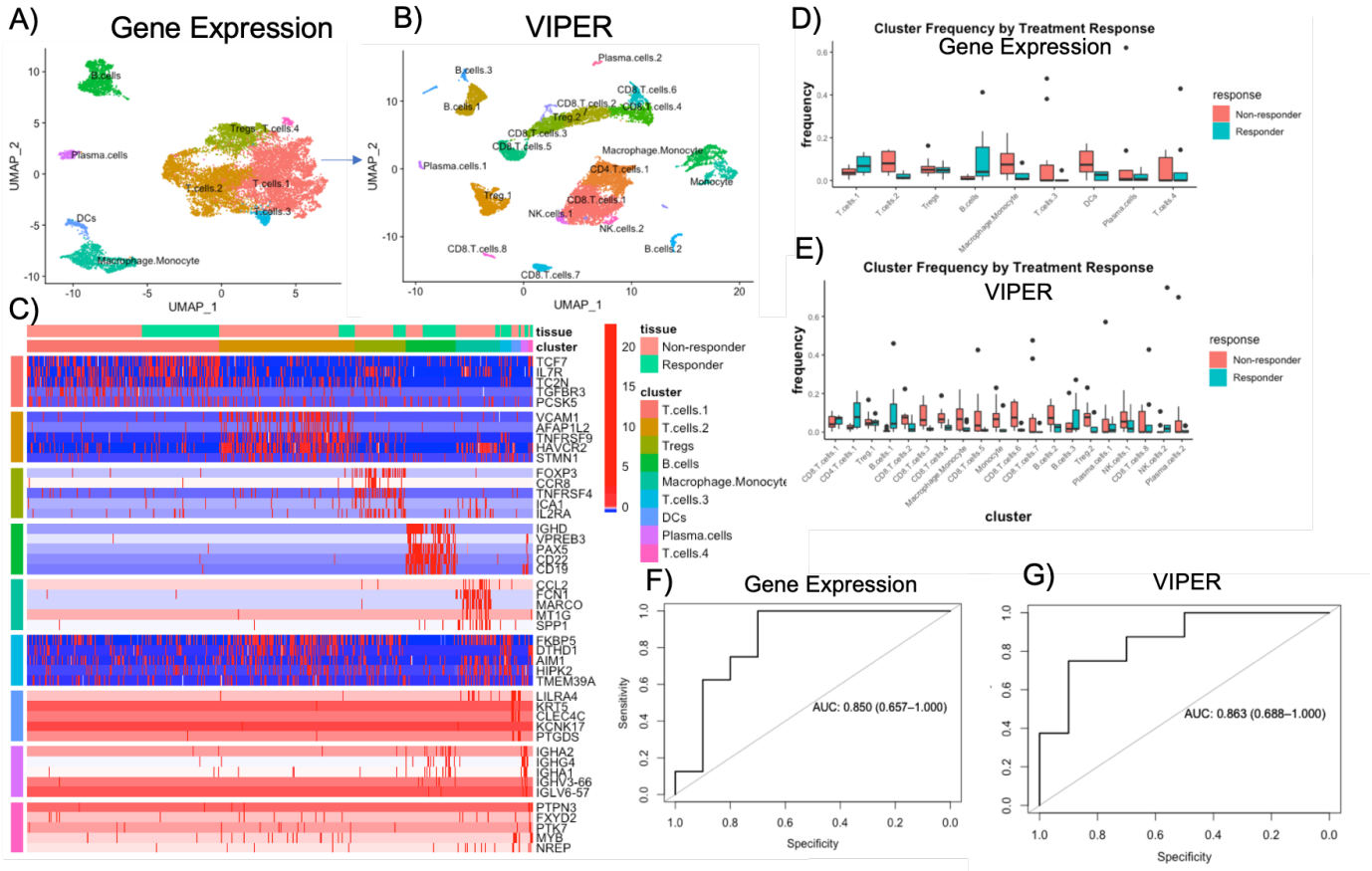


Figure 43: Single-Cell Profiling of Melanoma Immune Infiltrate and Predictive Power with Respect to Anti-PD1 Response.

A) UMAP Plot of Gene Expression based clusters from Pre-Treatment samples in Sade-Feldman et al, B) UMAP plot of VIPER Protein-Activity based re-clustering of data from A. C) Gene Expression heatmap of top5 differentially upregulated genes from clusters shown in A. D) Boxplot of frequencies for each immune population from A at baseline in subsequent responders vs non-responders to immunotherapy. E) Boxplot of frequencies for each immune population from B at baseline in subsequent responders vs non-responders to immunotherapy. F) Out-of-bag AUC from random forest classifier of response vs non-response using frequencies of each cluster in A as features. 95% Confidence Interval is shown. G) Out-of-bag AUC from random forest classifier of

response vs non-response using frequencies of each cluster in B as features. 95% Confidence Interval is shown.

In addition to these data, we have also identified a separate single-cell RNA-sequencing cohort which profiles 16 melanoma patients, including 15 non-responders and 1 responder [18]. This dataset did not pre-sort for CD45-positive immune cells and therefore included primarily tumor cells, which were identified in the original study for each patient, with significant batch effect across patients. The dataset includes 1,193 pre-treatment tumor cells. Therefore, we supplement the data from Sade-Feldman et al. with the gene expression and inferred protein-activity profiles of these tumor cells. Notably, although gene expression based clustering of tumor cells showed strong batch effect

between patients, VIPER protein activity based clustering resolved batch effect into a smaller number of tumor cell phenotypes shared across patients (Figure 44).

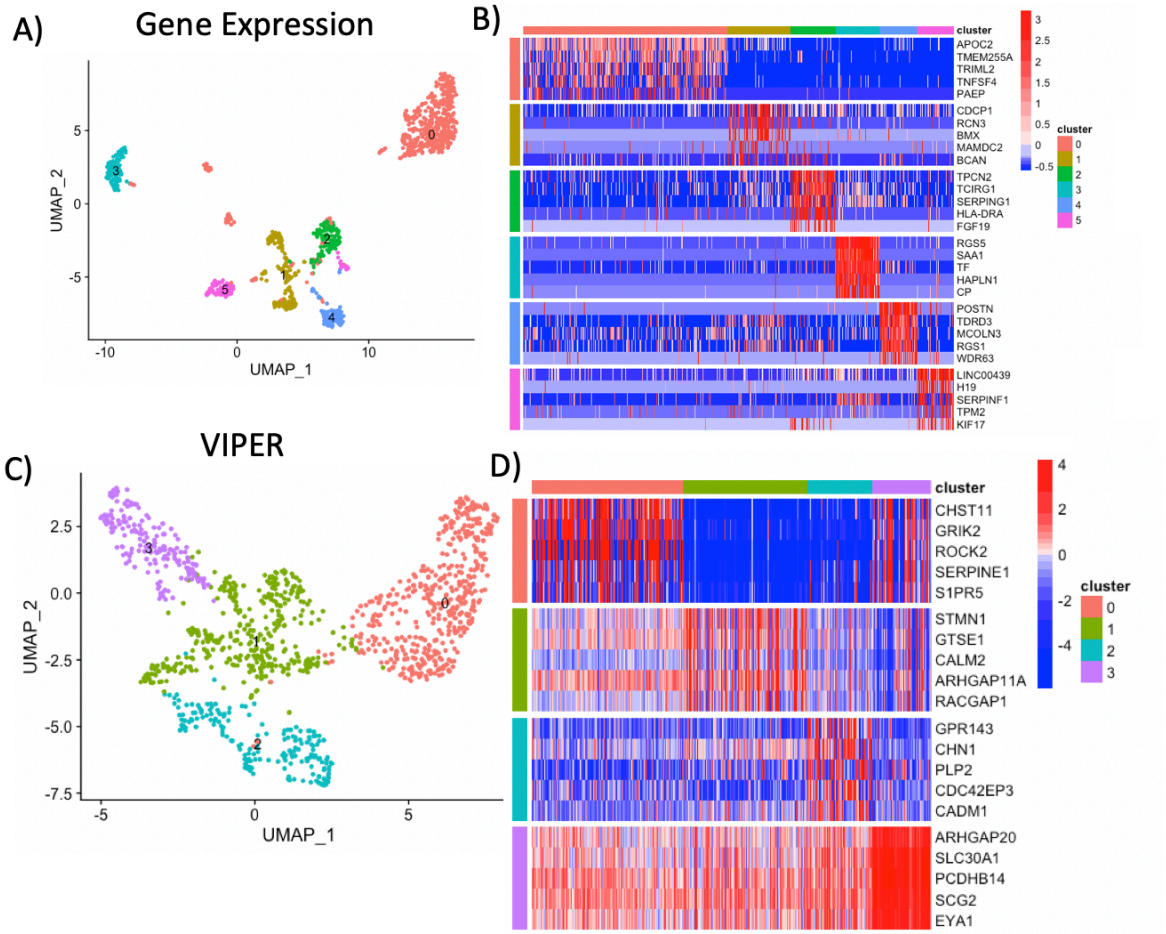


Figure 44: Single-Cell Profiling of Melanoma Tumor Cells Prior to Anti-PD1 Checkpoint Immunotherapy.

A) UMAP Plot of Gene Expression based clusters from Pre-Treatment tumor cells in Jerby-Arnon et al. B) Gene Expression heatmap of top5 differentially upregulated genes from clusters shown in A. C) UMAP Plot of VIPER Protein Activity based clusters from Pre-Treatment tumor cells in Jerby-

Arnon et al. D) Gene Expression heatmap of top5 differentially upregulated genes from clusters shown in C.

From single-cell RNA-Sequencing data, we first perform coarse clustering on gene expression, and generate clsuter-specific gene regulatory networks by ARACNe. With these, we run VIPER to infer cell-by-cell regulatory protein activity, and re-cluster [35]. We then generate for each cluster a signature of all regulatory proteins significantly up-regulated compared to the centroid of all clusters. We define this as the candidate Master Regulator feature set, generated separately for immune cells [17] and tumor cells [18].

We transfer this single-cell information into the context of two larger bulk-RNA Sequencing cohorts by normalizing all bulk cohort gene expression to TCGA as an external reference, and inferring protein activity by VIPER using the single-cell-derived gene regulatory networks. To date, we have identified two major public datasets profiling both bulkRNASeq and Tumor Mutational Burden in melanoma patients treated with anti-PD1 immunotherapy, which allows direct comparison of our novel predictor to Tumor Mutational Burden as an independent predictor of immunotherapy response. These datasets are [7], and [125]. Both datasets contain treatment-naïve patients (N=23 and N=64, respectively) and patients previously treated with anti-CTLA4 checkpoint immunotherapy (N=26 and N=39, respectively), and report response to anti-PD1 checkpoint immunotherapy by RECIST criteria. We construct and report discovery of distinct predictive signatures for pre-treated and treatment-naïve patients, separately.

Taking the full regulatory protein activity feature set identified from single-cell population markers, we first perform Boruta Random Forest Feature Selection to construct a machine-learning predictor

of treatment response in the Riaz et al. cohort, converging to a minimum set of proteins most strongly predictive of response to anti-PD1 immunotherapy. For treatment-naïve patients, this protein set consists of the following: (immune cell cluster markers: ASF1B, CD1E, CENPK, SPAG5, GMCL1, MED23,SLC19A1, TBC1D31,NEK4,ZGPAT, APAF1, PLAGL2, RBBP8, IFNAR1, STOML1, PLAA, MCMBP, MCM3, AKAP8, PEAK1, CDCA7L, RAB14, GOT2; tumor cell cluster markers: UHRF1, ASF1B, CENPK, IQGAP3, SPAG5, GMNN, DEPDC1, CDCA7L, TYMS, SLC19A1, FANCI, OXTR, ARHGAP25). For anti-CTLA4 pre-treated patients, the protein set consists of (immune cell cluster markers: KLRD1,CBLB, ZNF324B, TSHZ3, SLC1A5, PLCD1, RBBP8, SLC38A5, MCM6, PRDX4, CDK2, RAB18, NSFL1C, RRP8, ZNF747, EDARADD, CNIH1, CENPI, ZC3H12D, NCOA3, TAF4B, CERKL, CD84, ZNF292; tumor cell cluster markers: CD84, CXCL10, CD27, IPCEF1, LBH, SLC38A1, CCDC88C, FGFRL1, SHANK2, SHOX2, MCF2L, DAPK2, ZNF396, ICOSLG, MAP3K5). We evaluate the predictive power of each protein set in the

Riaz et al. cohort by out-of-bag AUC and validate independently in the Liu et al. cohort. The entire analysis pipeline is described in Figure 45.

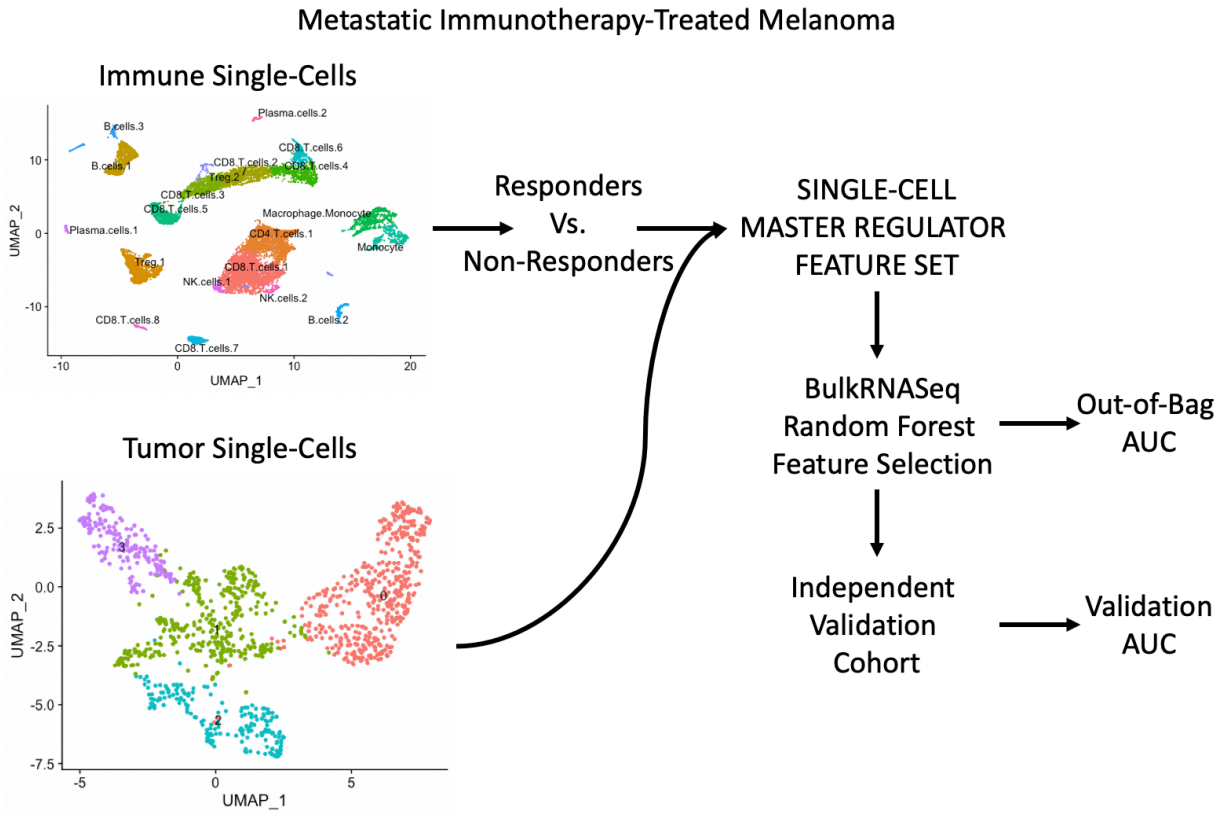


Figure 45: Analysis Pipeline For Selection of Treatment Response Predictive Signature in Bulk-RNA-Seq Data From Single-Cell RNA-Seq Immune and Non-Immune Master Regulator Feature Sets.

The protein activity of all immune and non-immune single-cell markers in responders vs non-responders to anti-PD1 immunotherapy are visualized among the Riaz et al. cohort in Figure 46.

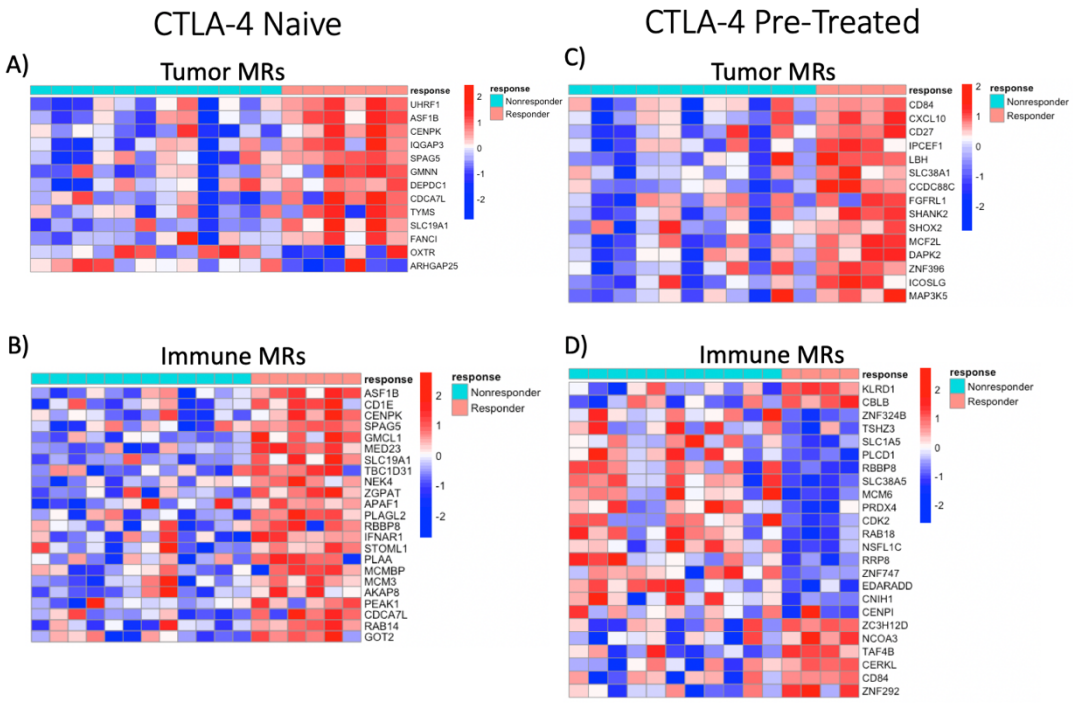


Figure 46: Protein Activity of Top Predictive Immune and Non-Immune Feature Sets in Subsequent Responders vs Non-Responders to anti-PD1 Immunotherapy

A) VIPER protein activity heatmap of tumor cell derived features most predictive of anti-PD1 treatment response in CTLA4-naïve patients. B) VIPER protein activity heatmap of immune cell derived features most predictive of anti-PD1 treatment response in CTLA4-naïve patients. C) VIPER protein activity heatmap of tumor cell derived features most predictive of anti-PD1 treatment response in CTLA4 pre-treated patients. D) VIPER protein activity heatmap of immune cell derived features most predictive of anti-PD1 treatment response in CTLA4 pre-treated patients.

Optimal predictive power is achieved with immune cell population marker proteins, with out-of-bag training AUC for treatment-naïve signature of 0.972 (95% CI 0.908 – 1.0), and out-of-bag training

AUC for CTLA4-experienced signature of 0.909 (95% CI 0.731 – 1.0). We independently validate these predictive models in the Liu et al. cohort, achieving an test AUC for treatment-naïve signature of 0.839 (95% CI 0.733 -0.944), and a test AUC for CTLA4-experienced signature of 0.893 (95% CI 0.771 – 1.0). Predictive power for all feature sets alone and in combination is shown in Figure 47. In all cases this drastically improved over tumor mutational burden as a predictor of response to

immunotherapy, which never achieved statistically significant predictive power (maximum test AUC = 0.505, with 95% CI 0.390 – 0.619).

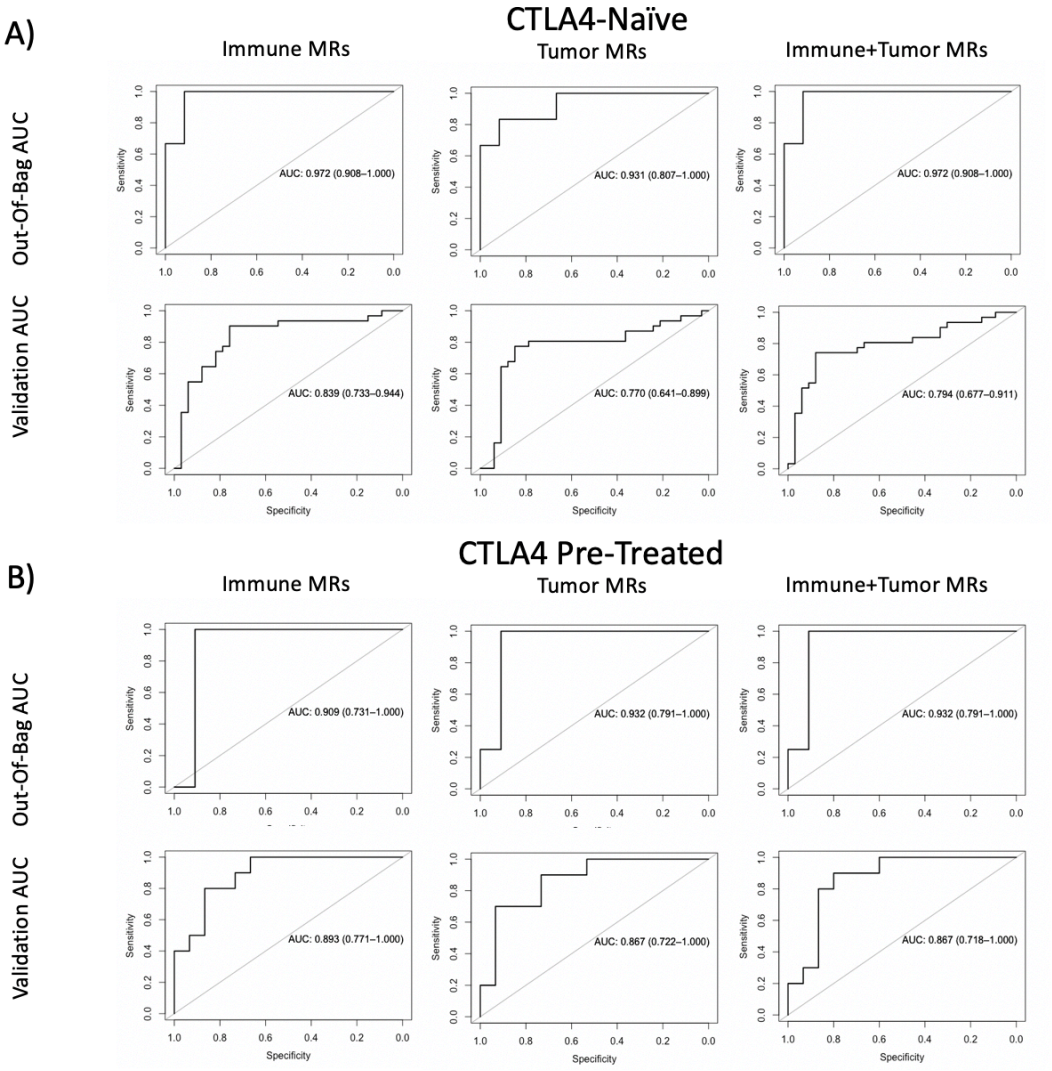


Figure 47: Predictive Power of Single-Cell-Derived Protein Activity Feature Sets For Classifying Response vs Non-Response to anti-PD1 Immunotherapy.

A) For CTLA4-Naïve patients, ROC curves for (left-to-right) immune cell derived features, tumor cell derived features, and combination of both, showing AUC with 95% confidence intervals for out-of-bag training data (top) and validation cohort (bottom). B) For CTLA4 Pre-Treated patients, ROC curves for (left-to-right) immune cell derived features, tumor cell derived features, and combination

of both, showing AUC with 95% confidence intervals for out-of-bag training data (top) and validation cohort (bottom).

5.3 Discussion

Solid tumors consist of a heterogeneous micro-environment of tumor cells and immune cells, including various cytotoxic and immunosuppressive populations. Mechanisms of immune tolerance in tumors are common, most notably the activity of immune checkpoints that inhibit cytotoxic T-cell activity. In particular, the interaction between PD-1 and PD-L1 has been successfully targeted in the clinic by the anti-PD1 antibodies pembrolizumab and nivolumab. However, many patients evade response by engaging additional immunosuppressive factors, including recruitment of immunosuppressive cell populations. These mechanisms are incompletely understood, and there is currently a lack of clinically relevant predictors for treatment response.

Existing predictors based on PD-L1 staining have shown poor predictive value of only and predictors based on mutational load and T-cell infiltration have been similarly limited [2], such that none to date are part of routine clinical care in prioritization of treatment options. The increasing use of immunotherapy in the first and second line across a range of tumor types means that identifying patients likely to respond to it could lead to dramatic improvement in clinical outcomes. More importantly, identifying patients unlikely to respond to immunotherapy could lead to prioritization of alternative therapies and rapid enrollment in combination therapy trials to overcome immunotherapy resistance.

Our approach represents a novel pipeline for discovery of immunotherapy response predictors which also serves to identify potentially mechanistic regulators of response vs non-response, which may

directly suggest combination-therapy candidates to overcome resistance among patients predicted not to respond. Our approach presents specific advantages relative to previously reported biomarker approaches. Specifically, immuno-histochemical profiling is very limited in the number of markers which can be concurrently profiled by fluorescent-tagged antibodies, such that biomarker discovery is limited to cell markers already suspected to associate with clinical outcome a priori. This severely limits potential for novel biological insights. On the other hand, traditional bulk RNA-Sequencing is limited by its high background noise and inability to distinguish transcriptional contribution of different cell types, and single-cell RNA-Sequencing without our VIPER analytic approach suffers from substantial data dropout, with loss of up to 95% of genes in any given cell. Our approach enables cell-by-cell master regulatory protein discovery across all cells without dropout, profiling up to 6000 signaling and transcriptional regulatory proteins simultaneously, then selecting the most strongly predictive proteins by a random forest machine learning algorithm. This approach in itself is novel, and furthermore none of the proteins in the identified predictive signatures have been previously reported, highlighting their novelty and need for follow-up mechanistic studies.

The chief advantage of the invention over existing approaches for predictive biomarker discovery is that we have shown a higher predictive area-under-the-curve than reported in any prior approach (genomic, transcriptomic, or immuno-histochemical). Furthermore, the fact that all proteins discovered to predict response to therapy with this approach are regulatory proteins active in modulating cell state opens up opportunity for a mechanistic rather than simply predictive understanding of response to therapy. In addition to providing an improved biomarker of response to anti-PD1 immunotherapy, the mechanistic nature of the discovered master regulators also represents a potential avenue for therapeutic intervention, both by direct targeting of discovered proteins

differentially activated among immunotherapy-resistant patients, and through drug screening for off-target transcriptional effects predicted to invert the observed anti-PD1 resistance signature.

Furthermore, although our current discovery has been made in the context of melanoma, as additional single-cell and bulk RNA-Sequencing data accrue across tumor types in the setting of immunotherapy-treatment, the predictive discovery approach we have developed may readily be extended to other tumor types and to broader development of treatment response biomarkers for other therapeutic regimens from master regulators of cell populations in the tumor immune micro-environment. We are therefore working to continue these analyses in novel clinical datasets and planning to extend to datasets profiling response to immunotherapy in additional tumor types (e.g. renal cell carcinoma). We are also planning functional experiments knocking down the identified Master Regulators in the predictive response signatures to assess their therapeutic potential in an immune-competent mouse model. Ultimately, resistance-associated proteins in this predictive signature may be inhibited by a variety of pharmaceutical or gene therapy approaches, and we anticipate future utility of combining such treatments with anti-PD1 immunotherapy, with or without contemporaneous targeted therapy, radiation therapy, chemotherapy and/or surgery.

Chapter 6: Prostate Cancer Hormonal Therapy Induces a Complex Immune Infiltrate Counteracted by Influx of Tumor-Infiltrating Regulatory T-cells

The following is adapted from:

Obradovic, A. Z.*, Dallos, M. C.*, Zahurak, M. L., Partin, A. W., Schaeffer, E. M., Ross, A. E., Allaf, M. E., Nirschl, T. R., Liu, D., Chapman, C. G., O'Neal, T., Cao, H., Durham, J. N., Guner, G., Baena-Del Valle, J. A., Ertunc, O., De Marzo, A. M., Antonarakis, E. S., & Drake, C. G. (2020). T-cell infiltration and adaptive treg resistance in response to androgen deprivation with or without vaccination in localized prostate cancer. *Clinical Cancer Research*, *26*(13), 3182–3192. https://doi.org/10.1158/1078-0432.ccr-19-3372
*These authors contributed equally

6.1 Summary

Statement of Translational Relevance: In preclinical models of prostate cancer, androgen deprivation therapy (ADT) promotes immunogenic cell death, transiently mitigates T cell tolerance to tumors and augments vaccine-induced antigen-specific CD8+ T cell responses. However, there are limited data on the immunologic effects of ADT on the tumor microenvironment (TME) in patients. In a neoadjuvant trial, we treated men with high-risk localized prostate cancer with either ADT or ADT plus low-dose cyclophosphamide and a cell-based vaccine (Cy/GVAX), prior to radical prostatectomy. ADT induced a complex immune cell infiltrate and increased intratumoral cytolytic CD8+ T cells. However, this CD8+ T cell increase was accompanied by a proportional increase in FoxP3+ regulatory T cells (Tregs), proving strong evidence for adaptive Treg resistance. When given prior to surgery, Cy/GVAX modestly augmented the immunologic effects of ADT and decreased disease recurrence compared to ADT alone. These data support the observation that ADT has pro-inflammatory effects. However, these antitumor effects appear to be counterbalanced by a proportional increase in local

immunosuppression.

Purpose: Previous studies suggest that androgen deprivation therapy (ADT) promotes antitumor immunity in prostate cancer. Whether a vaccine-based approach can augment this effect remains unknown.

Experimental Design: Therefore, we conducted a neoadjuvant, randomized study to quantify the immunologic effects of a granulocyte-macrophage colony-stimulating factor (GM-CSF)-secreting allogeneic cellular vaccine in combination with low-dose cyclophosphamide (Cy/GVAX) followed by degarelix versus degarelix alone in patients with high-risk localized prostate adenocarcinoma who were planned for radical prostatectomy.

Results: Both Cy/GVAX plus degarelix and degarelix alone led to significant increases in intratumoral CD8+ T cell infiltration and PD-L1 expression as compared to a cohort of untreated, matched controls. However, the CD8+ T cell infiltrate was accompanied by a proportional increase in regulatory T cells (Treg), suggesting that adaptive Treg resistance may dampen the immunogenicity of ADT. Although Cy/GVAX followed by degarelix was associated with a modest improvement in time-to-PSA progression and time-to-next treatment as well as an increase in PD-L1, there was no difference in the CD8 T-cell infiltrate as compared to degarelix alone. Gene expression profiling demonstrated that CHIT1, a macrophage marker, was differentially upregulated with Cy/GVAX plus degarelix compared to degarelix alone.

Conclusions: Our results highlight that ADT with or without Cy/GVAX induces a complex

immune response within the prostate tumor microenvironment. These data have important implications for combining ADT with immunotherapy. In particular, our finding that ADT increases both CD8+ T cells and Tregs, supports the development of regimens combining ADT with Treg-depleting agents in the treatment of prostate cancer.

6.2 Introduction

Prostate cancer remains the second most common cause of cancer-related mortality in men and definitive local therapy represents the only treatment modality with the potential for cure [126]. Despite advances in surgical approaches, patients with high-risk localized prostate cancer continue to have a high likelihood of disease recurrence following definitive local therapy [127] [128]. To date, no neoadjuvant therapy preceding prostatectomy has demonstrated sufficient efficacy to warrant FDA approval.

In contrast to traditional therapies which decrease tumor bulk prior to surgery, immunotherapy has the potential to re-engage systemic anti-tumor immune responses, thereby eradicating distant micro-metastases. Although the development of sipuleucel-T for castration-resistant prostate cancer (CRPC) demonstrated the potential for immunotherapy in prostate cancer, immune checkpoint inhibitors have not yielded significant responses, except perhaps when used in combination [129] [130] [131] [132] [133] [134] [135]. One significant challenge to inducing anti-tumor immunity in prostate cancer is the non-inflamed tumor microenvironment (TME) [136]. Prostate tumors also generally have a low mutational burden and low PD-L1 expression; these factors predict response to immunotherapy in other tumor types [37] [137]. In addition, prostate tumors demonstrate multiple mechanisms of immune escape including defective antigen

processing, decreased MHC class I expression, and infiltration with regulatory T cells (Tregs), myeloid-derived suppressor cells, and M2 macrophages [138] [139] [140] [141].

Prostate GVAX is an allogeneic cell-based prostate cancer vaccine composed of two irradiated cell lines (PC3 and LNCaP) that have been genetically modified to secrete granulocyte-macrophage colony-stimulating factor (GM-CSF) [138]. The release of GM-CSF by these modified tumor cells promotes the recruitment of dendritic cells and subsequent presentation of tumor antigens to T-cells with associated activation of anti-tumor immune responses. Prior randomized controlled trials of GVAX as monotherapy or in combination with docetaxel in mCRPC failed to show a survival benefit over chemotherapy, suggesting that allogeneic cell-based immunotherapy may be insufficient on its own to generate a robust T cell response against prostate cancer [138]. This may be particularly relevant in advanced metastatic CRPC, wherein a more immunosuppressive TME predominates [139]. However, preclinical studies demonstrate that administering low-dose cyclophosphamide prior to a cell-based GM-CSF-secreting vaccine can increase CD8+ T cell infiltration in the prostate, and transiently deplete regulatory T cells (Tregs) [140] [141]. These preclinical data are supported by clinical trials combining GVAX with low-dose cyclophosphamide in breast cancer, colorectal cancer and pancreatic cancer [4] [142].

In addition, prior studies in murine models show that castration results in *de novo* presentation of prostate-restricted antigens in tumor-draining lymph nodes, with transient mitigation of T cell tolerance [4]. ADT can also induce a pro-inflammatory immune cell infiltrate, supporting the hypothesis that androgen ablation may augment vaccine-induced effector T cell responses,

particularly during the peri-castration period [4]. Whether similar immune modulation occurs in patients remains poorly understood.

To address these questions, we conducted a randomized neoadjuvant study to test the hypothesis that the combination of low-dose cyclophosphamide plus GVAX (Cy/GVAX) could augment the ADT-induced immune response in men with localized high-risk prostate cancer. The LHRH antagonist degarelix acetate was selected as ADT for this study based on its rapid onset-of-action allowing shorter time-to-surgery, lack of transient increase in testosterone reducing risk of tumor flare, and the observation that degarelix leads to a robust immune cell infiltrate in pre-clinical models, peaking around 2 weeks after administration [4]. A secondary endpoint of the study was to test whether ADT plus Cy/GVAX prolongs time to PSA recurrence as compared to ADT alone. We also sought to more deeply profile the immunological changes in the prostate TME mediated by ADT with or without Cy/GVAX.

6.3 Results

Thiry-two patients were recruited to the study with 16 randomized to each arm. 1 patient randomized to degarelix alone and 2 patients randomized to degarelix plus Cy/GVAX withdrew consent before study drug initiation. Therefore, 29 patients received study treatment (Figure 48).

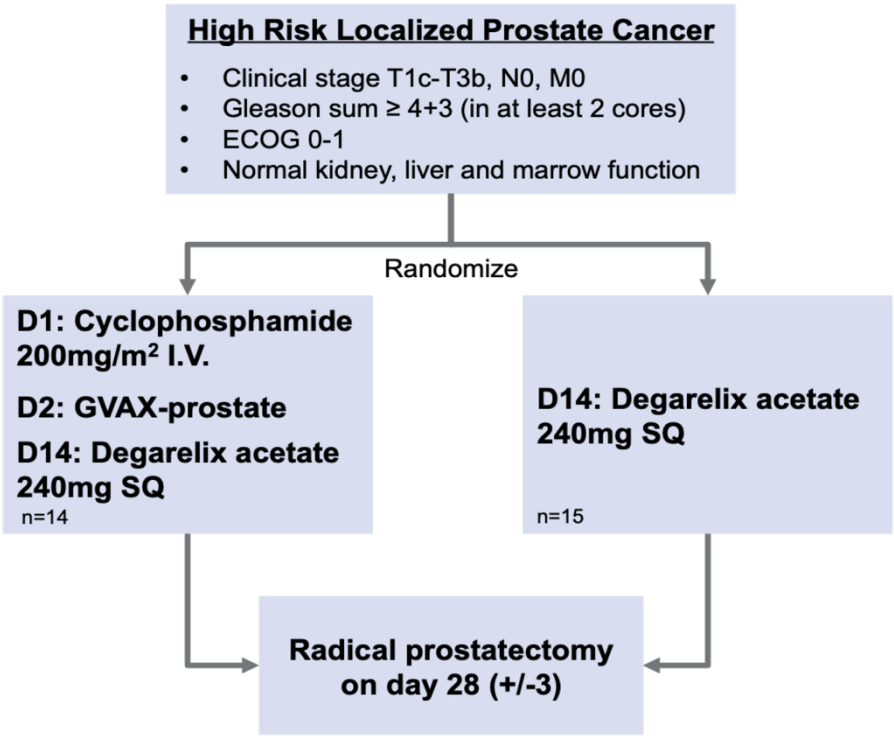


Figure 48: Clinical trial design and patient disposition diagram.

Patients with high-risk localized prostate cancer (T1c–3b N0 M0, Gleason 7–10) were randomized 1:1 to degarelix (240 mg SQ) vs. Cyclophosphamide (200 mg/m2 IV) / GVAX (2.5×108 PC3 cells, 1.6×108 LNCaP cells) given 2 weeks before degarelix. All patients then underwent radical prostatectomy 2 weeks after degarelix. Abbreviations: Eastern Cooperative Oncology Group (ECOG); subcutaneously (SQ); intravenously (IV).

15 patients received degarelix alone and 14 received degarelix plus Cy/GVAX (one patient in this group withdrew following cerebrovascular ischemia and was subsequently lost to follow-up). Clinical characteristics of the two treatment groups were similar with respect to age, risk status, Gleason sum, tumor stage, regional nodal involvement, and surgical margins (Figure 49). 64% of patients had Gleason ≥8 disease, 56% had pathological stage T3b, and 18% were found to have N1 disease at the time of surgery.

Clinical Variable	Degarelix (N=15)	Degarelix + Cy/GVAX (N=13)
Median age (interquartile range), years	58 (55-64.5)	61 (54-63)
Very high risk (%)	6 (40%)	8 (61%)
Gleason sum, *n* (%)		
7	6 (40%)	4 (31%)
8	1 (7%)	2 (15%)
9	8 (53%)	6 (46%)
10	0 (0%)	1 (8%)
T Stage, *n* (%)		
pT2	4 (27%)	3 (23%)
pT3a	5 (33%)	6 (46%)
pT3b	6 (40%)	4 (31%)
ECOG Status, *n* (%)		
0	15 (100%)	13 (100%)
Regional Lymph Node Involvement, *n* (%)	2 (13%)	3 (23%)
Positive margin, *n* (%)	7 (47%)	5 (38%)
Recurred, *n* (%)	9 (60%)	4 (31%)
Developed Metastasis, *n* (%)	2 (13%)	3 (23%)

Figure 49: Patient baseline demographics and disease characteristics.

Clinical variables for patients treated with degarelix alone versus degarelix plus Cy/GVAX. *Gleason sums for the histologic pattern of carcinoma range from 7-10 with higher scores indicating a higher-grade tumor.

Safety

Both degarelix alone and degarelix plus Cy/GVAX were well-tolerated. A single grade 3 ALT elevation was reported in the degarelix plus Cy/GVAX group, with no other treatment-related grade 3 or 4 adverse events reported (Figure 50). All enrolled patients successfully underwent radical prostatectomy, with no significant unexpected surgical complications or toxicities reported. Significant surgical complications were defined as blood loss in excess of 2500mL, operative time in excess of 3.5 hours, hospital stay in excess of 4 days or systemic symptoms including fever, rash or myelosuppression.

	Degarelix (N=15)			Degarelix + Cy/GVAX (N=14)		
Adverse Events	Grade 1	Grade 2	Grade 3	Grade 1	Grade 2	Grade3
General Disorders						
Injection-site reaction	10 (66%)	0	0	11 (79%)	2 (14%)	0
Fatigue	4 (27%)	0	0	3 (21%)	0	0
Chills	2 (13%)			1 (7%)		
Fever	0	0	0	2 (14%)	0	0
Flu like symptoms	0	0	0	0	1 (7%)	0
Malaise	0	0	0	1 (7%)	0	0
Edema, limbs	1 (7%)	0	0	0	0	0
Localized Edema	0	0	0	1 (7%)	0	0
Gastrointestinal Disorders						
Abdominal Pain	2 (13%)	0	0	1 (7%)	0	0
Nausea	0	0	0	1 (7%)	0	0
Vascular Disorders						
Hot flashes	6 (40%)	0	0	8 (57%)	0	0
Reproductive System Disorders						
Erectile dysfunction	1 (7%)	0	0	2 (14%)	0	0
Testicular disorder	0	0	0	1 (7%)	0	0
Urinary disorders	0	0	0	0	0	0
Urinary incontinence	1 (7%)	0	0	1 (7%)	0	0
Laboratory Abnormalities						
Elevated ALT	1 (7%)	0	0	0	0	1 (7%)
Elevated AST	2 (13%)	0	0	0	1 (7%)	0
Skin Disorders						
Rash (systemic)	0	0	0	3 (21%)	2 (14%)	0
Dizziness	0	0	0	1 (7%)	0	0
Musculoskeletal Disorders						
Arthralgias	1 (7%)	0	0	2 (14%)	0	0
Myalgias	1 (7%)	0	0	1 (7%)	0	0
Nervous System Disorders						
Lethargy	0	0	0	1 (7%)	0	0
Dizziness	0	0	0	1 (7%)	0	0
Headache	0	0	0	2 (14%)	0	0
Ischemia, cerebrovascular	0	0	0	1 (7%)	0	0
Surgical Complications						
Post-op hematoma (pelvic)	0	0	0	0	1 (7%)	0

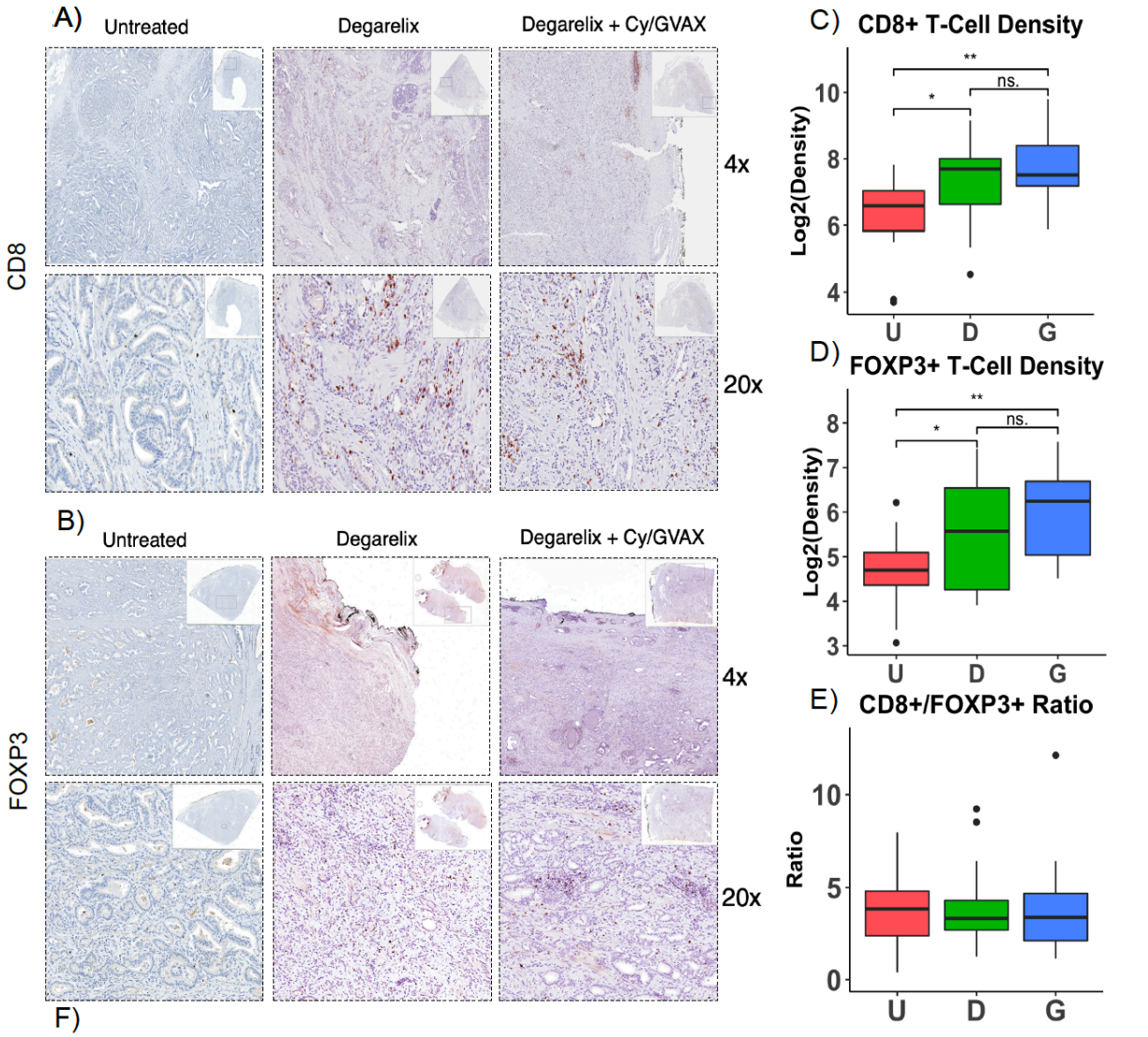
Figure 50: Adverse events reported by treatment group.

Adverse events for patients treated with degarelix alone versus degarelix plus Cy/GVAX were reported for all patients in the study, including one patient in the degarelix plus Cy/GVAX group that subsequently went off-study following cerebrovascular ischemia.

Degarelix (ADT) Induces CD8 T Cell Infiltration with a Proportional Increase in Tregs

Prostatectomy samples from both treatment arms, degarelix and degarelix + Cy/GVAX, showed significantly increased intratumoral CD8+ T cell density by IHC as compared to untreated

matched controls (Figure 51C). However, this CD8 infiltration was balanced by a proportionally increased infiltration with Tregs, such that the CD8/Treg ratio remained consistent across all treatment groups (Figure 51D, 51E). While there was a significant increase in both CD8+ T cell and Treg infiltrate with degarelix versus controls and degarelix + Cy/GVAX versus controls, there was no statistically significant difference between the degarelix and degarelix + Cy/GVAX treatment groups (Figure 51), suggesting that the GVAX vaccine did not induce additional CD8 infiltration in this setting as compared to degarelix alone.



F)

	Untreated (U) N = 20	Degarelix (D) N = 15	Degarelix + Cy/GVAX (G) N = 13	P-value
CD8+ T cell density (mean, 95%CI)	96 (72–120)	205 (121–289)	263 (129–397)	<0.01 (G vs U) 0.03 (D vs U) 0.35 (G vs D)
Treg cell density (mean, 95%CI)	29 (21–36)	59 (34–85)	78 (48–107)	<0.01 (G vs U) 0.02 (D vs U) 0.18 (G vs D)
CD8+ / Treg ratio (mean, 95%CI)	3.7 (2.9–4.6)	4.0 (2.7–5.3)	3.9 (2.2–5.7)	0.87 (G vs U) 0.68 (D vs U) 0.66 (G vs D)

Figure 51: Degarelix and degarelix + GVAX increase CD8+ and FOXP3+ T cell infitration in prostate tumors.

A) Representatitive H&E and immunohistochemistry (IHC) for CD8+ T cells, visualized at 4x and 20x magnification B) Representative H&E and IHC for FOXP3+ T cells, visualized at 4x and 20x magnification C) Boxplots of Log2(CD8+ T cell density), quantified from IHC as

represented in A. D) Boxplots of Log2(FOXP3+ T cell density), quantified from IHC as represented in B. E) Boxplots of the CD8+/FOXP3+ T cell ratio, quantified from IHC as represented in A and B. F) Table of mean CD8+ T cell density (cells/mm^2), mean Treg density (cells/mm^2), and CD8/Treg ratio for each treatment group and untreated controls, with 95% confidence intervals and p-values by Gleason-stratification-adjusted ANOVA reported for each comparison of groups; * = $p<0.05$, ** = $p<0.01$, and *** = $p<0.005$

Since FOXP3 can potentially be expressed in other T-cell populations, we also analyzed our transcriptomic data to identify whether treatment led to increased expression of other Treg markers including GITR (TNFRSF18), CTLA-4 and CD25 (IL2RA). We observed increased expression of GITR, CTLA-4 and CD25 with both degarelix alone and degarelix plus Cy/GVAX compared to untreated controls (Figure 52). However, there was no difference in expression of these markers between degarelix and degarelix plus Cy/GVAX.

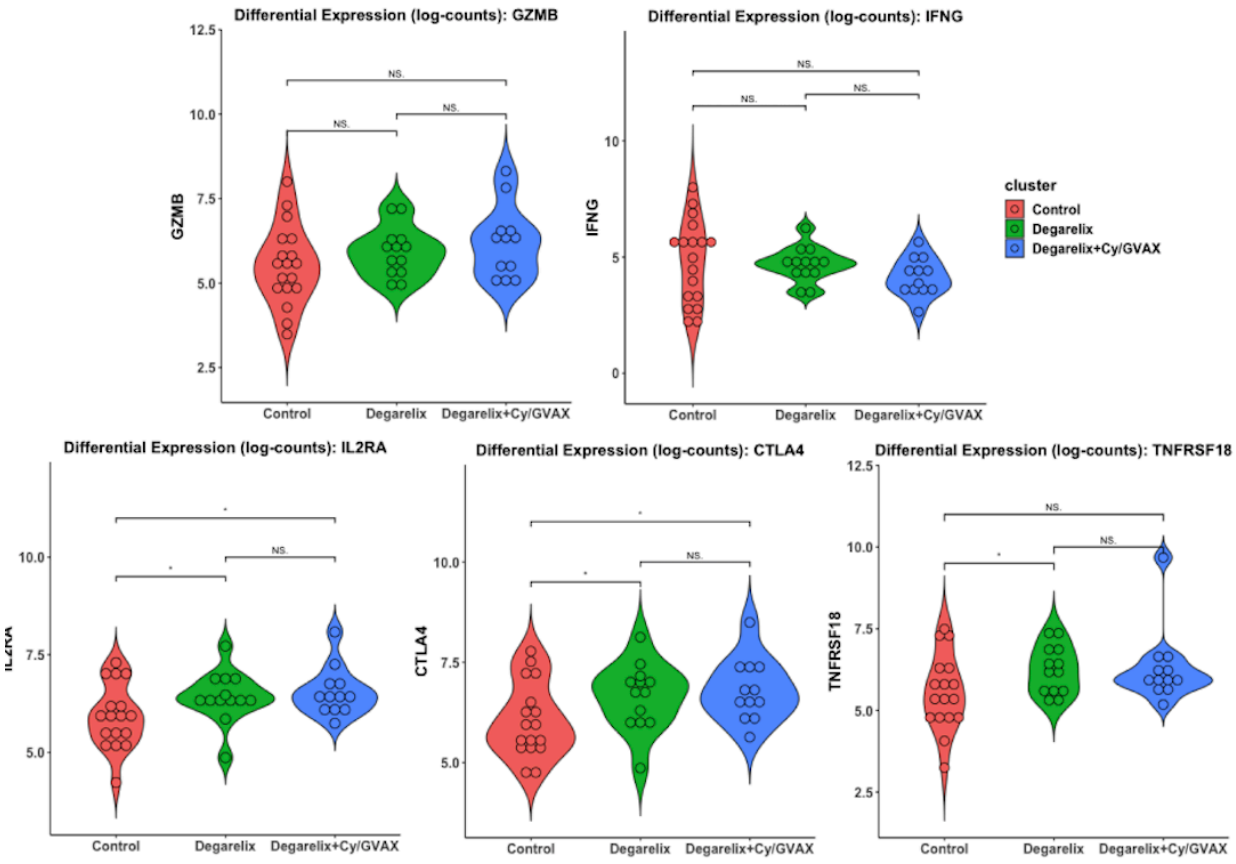


Figure 52: Expression of Cytotoxicity and Treg related genes by treatment group.

Violin-plots of log-scaled post-normalization Nanostring gene counts for specific genes of interest in each treatment group and untreated controls. GZMB (granzyme) and IFNG (interferon gamma) correspond to cytotoxicity-related genes, and show no statistically significant expression difference across groups. IL2RA (CD25), TNFRSF18 (GITR), and CTLA4, correspond to Treg-related genes, such that expression of each gene is significantly up-regulated in both treatment groups relative to untreated control, but significantly different between treatment arms. Raw Wilcox test p-values are shown as * = $p<0.05$, ** = $p<0.01$, and *** = $p<0.005$.

Increased PD-L1 Expression after GVAX Vaccination

Consistent with prior reports, tumor cell PD-L1 expression was minimal in untreated patients (Figure 53). Degarelix alone appeared to modestly increase PD-L1 expression, consistent with the notion that cytokine secretion from infiltrating CD8+ T cells may drive up-regulation of immune

checkpoints. Tumor samples from patients treated with degarelix + Cy/GVAX were found to have increased PD-L1 staining compared to patients treated with degarelix alone, with a higher proportion of samples exceeding 5% PD-L1 positivity (Figure 53); this trend was not statistically significant. Although there appeared to be some areas of PD-L1 staining in inflammatory cells in the stroma, the majority of cells staining positive for PD-L1 were tumor cells. Taken together, these data suggest that while the GVAX vaccine does not significantly increase CD8+ T cell density, the infiltrating immune cells induced by GVAX may be capable of promoting PD-L1 up-regulation.

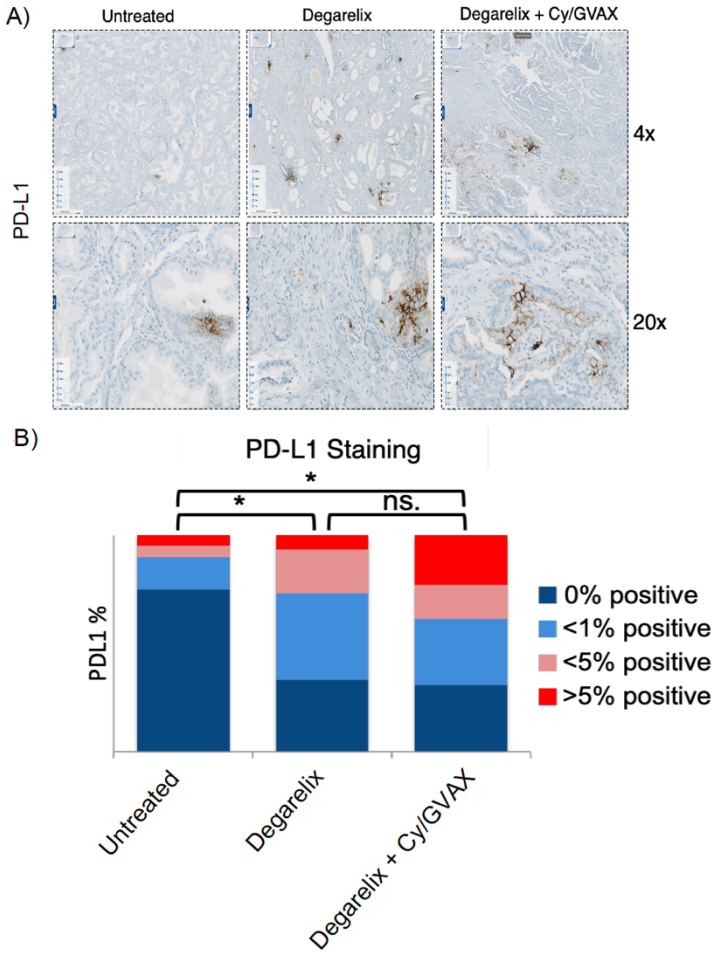


Figure 53: Degarelix and degarelix + Cy/GVAX increase PD-L1 expression in prostate tumors.

A) Representative IHC for PD-L1, visualized at 4x and 20x magnification B) Stacked barplot of %PD-L1 positive cells, showing relative proportion of samples with 0% PD-L1 staining, <1% PD-L1 staining, <5% PD-L1 staining, and >5% PD-L1 staining in tumor cells in each treatment group and a cohort of untreated matched controls. Distributions of %PD-L1 categories may be visually compared between groups, such that the degarelix + Cy/GVAX group has the highest proportion of samples with PD-L1 > 5%. Proportions of samples with %PD-L1 > 0 were also compared between groups by Fisher's exact test, with p-values shown above the plot for each comparison, where * = $p<0.05$, ** = $p<0.01$, and *** = $p<0.005$.

Degarelix and Degarelix plus Cy/GVAX Induce Complex Changes in Immune Gene Expression

Pairwise differential gene expression was performed on normalized Nanostring data from prostatectomy samples, comparing untreated control patients, degarelix-treated patients, and degarelix + Cy/GVAX treated patients. This analysis identified 98 genes up-regulated in both degarelix and degarelix + Cy/GVAX vs control (Figure 54A). CHIT1, a macrophage activation marker, was the only gene significantly up-regulated in degarelix + Cy/GVAX vs degarelix (Figure 54B) [143]. The CIBERSORT algorithm was used to de-convolute and infer the abundance of immune cell subtypes in each sample from Nanostring gene expression. Fractional contributions of immune cell populations were then compared between treatment groups (Figure 54C). These data show that a complex immune infiltrate was present in these prostatectomy samples at time of surgery, with significant populations of B cells, CD4 T cells, M1 macrophages, M2 macrophages, and mast cells. Summing the inferred abundance of each cell type yielded a total immune infiltrate estimate from gene expression data. Those data showed that that total immune infiltrate was significantly increased in both degarelix and degarelix + Cy/GVAX compared to control, but not in degarelix + Cy/GVAX as compared to degarelix alone (Figure 54D). CIBERSORT analysis also revealed an increased infiltrate of CD8+ T cells, M2 macrophages, and gamma-delta T-cells

in both treatment groups as compared to untreated controls, with a raw p-value < 0.05. Although the CD8+ T cell increase is consistent with the IHC data (Figure 48), these differences based on gene-expression analysis were not statistically significant after adjustment for multiple testing (Figure 54E). To further assess whether treatment could increase T-cell activation, we evaluated interferon-γ and granzyme B expression levels and demonstrated no significant difference in expression levels between the treatment groups (Figure 52).

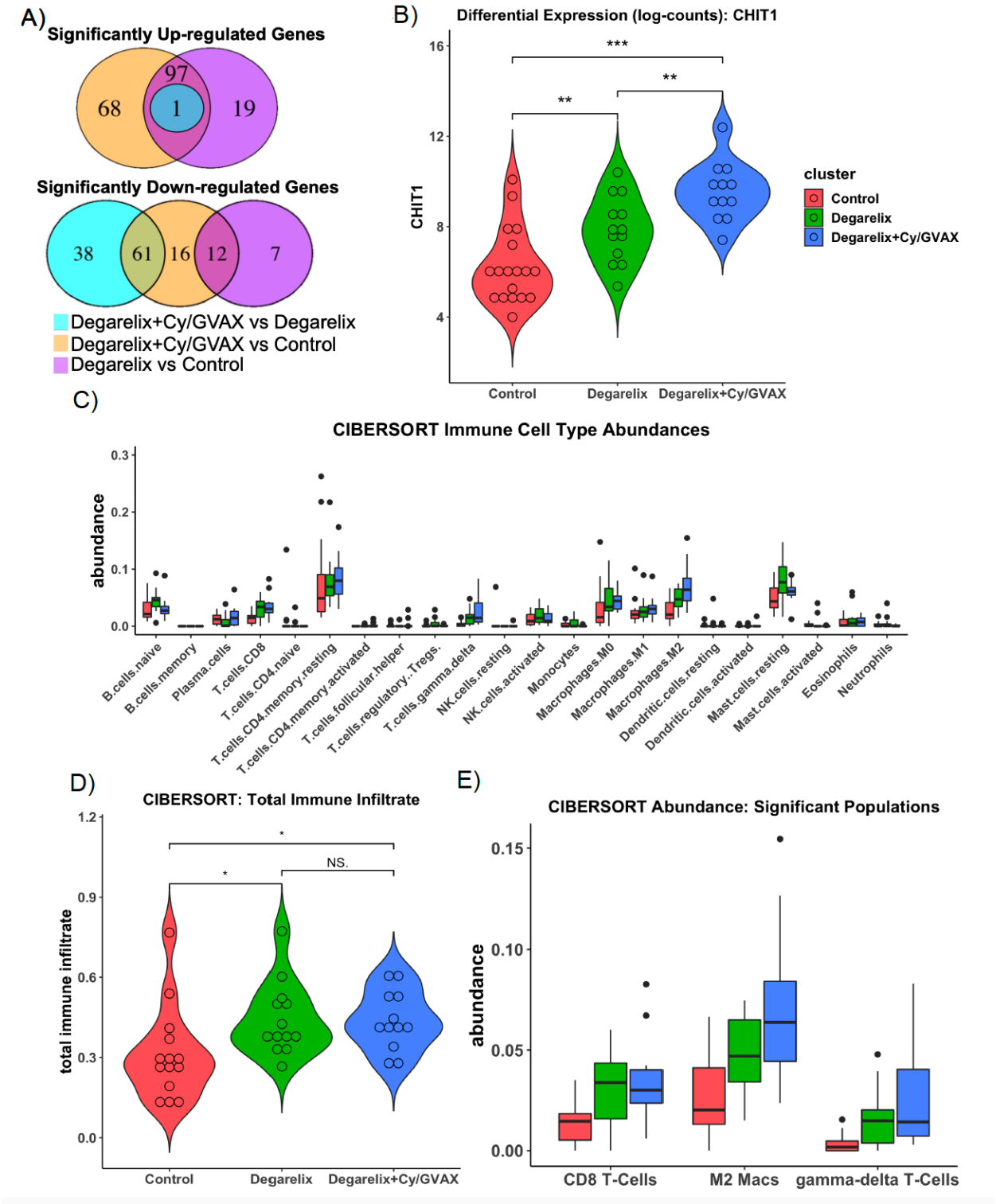


Figure 54: Degarelix and degarelix + Cy/GVAX induce complex changes in immune gene signatures in primary prostate tumors.

A) Differential expression of immune related genes by Nanostring Immune Profiling Panel in

primary prostate tumors after degarelix, degarelix + Cy/GVAX, and untreated matched controls. Euler plots showing number of genes with Benjamini-Hochberg corrected t-test p-value < 0.01 for each pairwise comparison of groups, such that "up-regulated genes" refers to genes that have higher mean frequency in the degarelix + Cy/GVAX group than in the degarelix group (cyan), higher mean frequency in the degarelix group than the untreated control group (purple), and higher mean frequency in the degarelix + Cy/GVAX group than the untreated control group (orange), and "down-regulated genes" refers to genes that have lower mean frequency in degarelix + Cy/GVAX vs degarelix (cyan), degarelix vs controls (purple), and degarelix + Cy/GVAX vs controls (orange), respectively. B) Violin-plot of log-scaled post-normalization Nanostring gene counts for CHIT1 in each treatment group and untreated controls. In A, CHIT1 is the sole gene significantly upregulated in each comparison. C) Boxplot of immune cell type absolute abundances as inferred by CIBERSORT, colored by treatment group and reported for all samples with CIBERSORT p-value<0.05. D) Violin-plot of total immune cell infiltrate for each sample by treatment group, such that total immune cell infiltrate represents the sum of CIBERSORT immune cell abundances as shown in C. E) Boxplot of immune cell populations for which t-test comparing abundance between groups showed an uncorrected p-value<0.05. P-values were obtained by unpaired t-test with Benjamini-Hochberg multiple-testing correction and shown on B and D with * = p<0.05, ** = p<0.01, and *** = p<0.005.

Degarelix plus Cy/GVAX is Associated with Clinical Outcome

At 24 months post-prostatectomy, 69% of patients were free of PSA recurrence in the Cy/GVAX plus degarelix treatment group as compared to 40% in the degarelix-only group (Figure 49). Initial univariate cox regression of treatment group against time-to-PSA recurrence stratified by Gleason sum 7 versus Gleason sum greater than 7 yielded a hazard ratio of 0.44 (95% Confidence Interval 0.13-1.43, p = 0.17), with time-to-next-treatment yielding a hazard ratio of 0.41 (95% Confidence Interval 0.13-1.36, p = 0.15). After determining informative clinical variables for prediction of time-to-PSA recurrence using backwards feature selection by the Akaike Information Criterion, multiple Cox regression was performed accounting for interactions between patient age, tumor

stage, Gleason sum, and treatment group. Using this regression analysis, treatment with Cy/GVAX plus degarelix showed an increased time to PSA recurrence as compared to that observed in patients treated with degarelix alone, with a hazard ratio of 0.29 (95% Confidence Interval 0.08-1.00, p = 0.05) (Fig. 55A, 55B). Backward feature selection converged to the same set of clinical variables for prediction of time-to-next treatment, where there was a statistically significant treatment effect for degarelix plus Cy/GVAX compared to degarelix alone, with a hazard ratio of 0.26 (95% Confidence Interval 0.071-0.97, p = 0.046) (Fig. 55C, 55D).

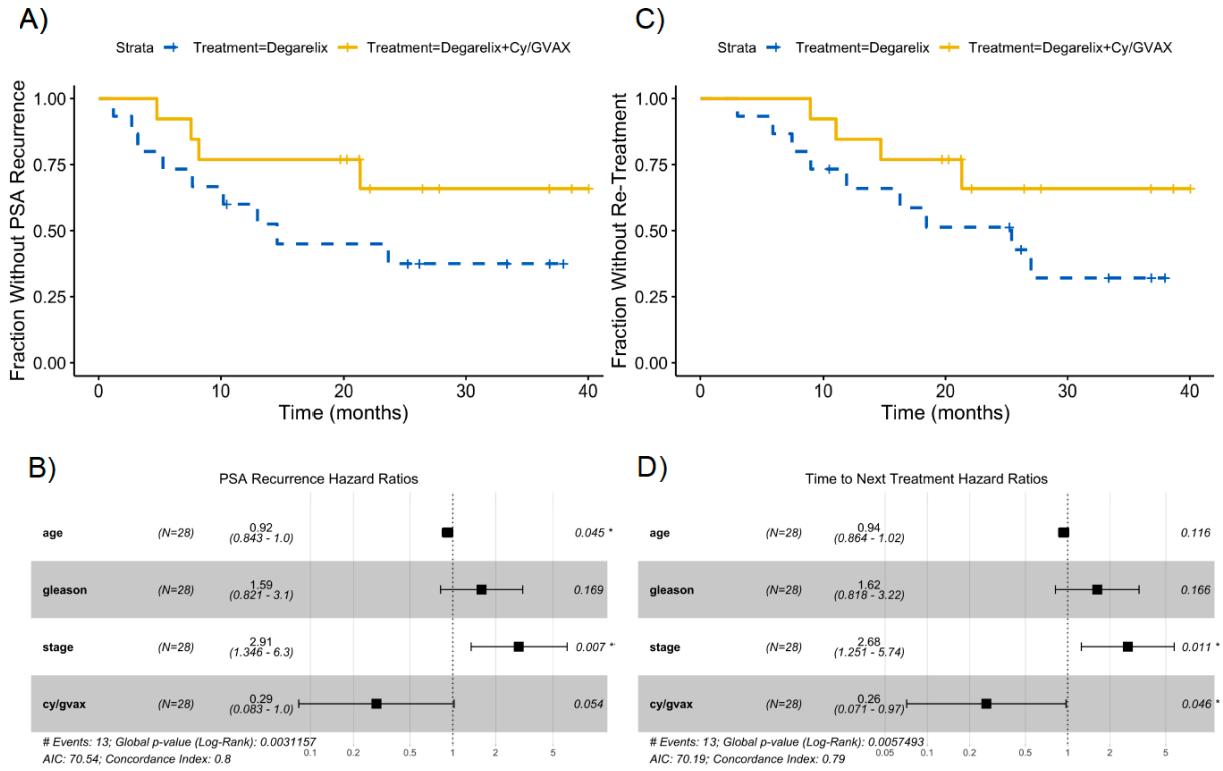


Figure 55: Combination of Cy/GVAX with degarelix improves time-to-PSA recurrence and increases time-to-next treatment.

A) Kaplan-Meier curves comparing time-to-PSA recurrence of patients treated with degarelix + Cy/GVAX vs degarelix alone. Informative clinical variables for multivariate analysis were selected by backward feature selection using the Akaike Information Criterion. B) Forest plot showing time-to-PSA recurrence hazard ratios with 95% confidence interval for multiple cox regression of progression-free-survival against Cy/GVAX status, patient age, tumor stage, and

Gleason score. P-values for each variable are reported, as is the overall log-rank p-value, Akaike Information Criterion value, and concordance index for the regression C) Kaplan-Meier curves comparing time to next treatment for patients treated with degarelix + Cy/GVAX vs degarelix alone, with log-rank p-value reported from multiple cox regression of time-to-next-treatment against Cy/GVAX status, patient age, tumor stage, and Gleason score. Informative clinical variables were selected as in A D) Forest plot showing time-to-next-treatment hazard ratios with 95% confidence interval for multiple cox regression of time-to-next-treatment against Cy/GVAX status, patient age, tumor stage, and gleason score. P-values for each variable are reported, as is the overall log-rank p-value, Akaike Information Criterion value, and concordance index for the regression.

There was no significant difference observed between the two treatment groups in prediction of time-to-metastasis, where backward feature selection converged to a null model, and univariate cox regression with treatment group yielded a p-value of 0.46 (Figure 56). This may be due to the overall low rate of metastases in this patient population, with only 5 cases of metastasis observed across the two treatment groups (Figure 49). There was also no significant difference in time-to-testosterone recovery between the two treatment groups (Figure 56), suggesting that the improved time to PSA recurrence cannot be accounted for by differences in the duration of a castrate level of testosterone. Correlation with recurrence is shown in Figure 57 for each variable considered in the first step of the backward feature selection model, such that CD8+ and FOXP3+ density as well as PD-L1 level were each negatively correlated with recurrence, but were not individually predictive of time-to-recurrence and were not additionally informative after accounting for treatment group, age, stage, and Gleason sum.

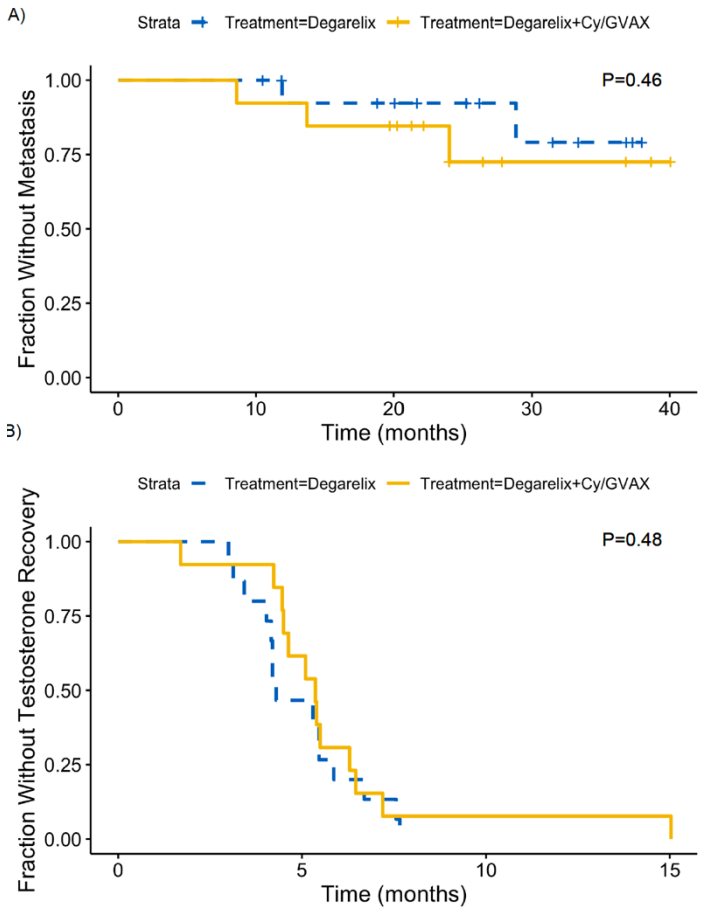


Figure 56: No evidence for effect of Cy/GVAX on metastasis or time-to-testosterone recovery.

A) Kaplan-Meier curves comparing time to metastasis for patients treated with degarelix + Cy/GVAX vs degarelix alone. When selecting for informative clinical variables by backwards feature selection as in Figure 55, Akaike Information Criterion converged on a null model with no informative variables, indicating that no set of clinical variables provided predictive value. Therefore, a univariate cox regression was performed of time to metastasis against treatment group, and the log-rank p-value for that regression is reported here. B) Kaplan-Meier curves comparing time-to-testosterone recovery for patients treated with degarelix + Cy/GVAX vs degarelix alone. As in A, backwards feature selection of clinical variables converged to a null model with no informative features. Therefore, a univariate cox regression was performed of time-to-metastasis against treatment group, and the log-rank p-value for that regression is reported.

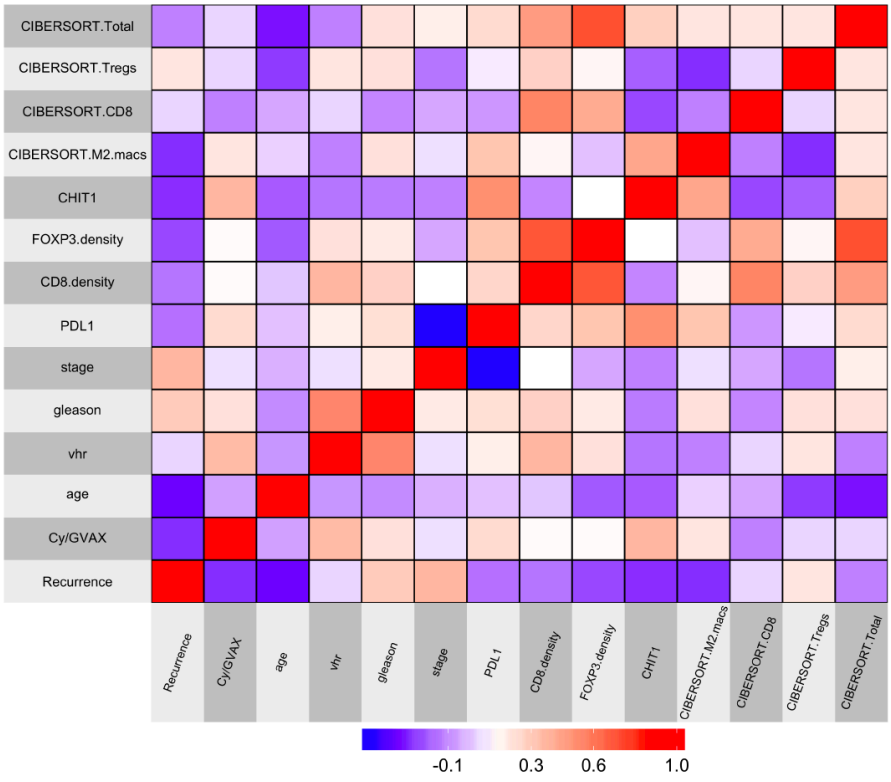


Figure 57: Correlation plot of clinical, genetic, and immunohistochemical features.

Pairwise Pearson correlations between disease recurrence, clinical variables (treatment group, patient age, stage, gleason sum, very-high-risk status), CD8+ T cell and FOXP3 density as determined by IHC in Figure 51, CHIT1 gene expression (normalized Nanostring counts), total CIBERSORT immune infiltrate, and CIBERSORT M2 macrophage abundance, CD8 abundance, and Treg abundance, as determined from Nanostring profiling in Figure 53.

6.4 Discussion

This study demonstrates that neoadjuvant ADT (degarelix acetate) with or without the addition of GVAX immunotherapy and low-dose cyclophosphamide promotes a complex immune response within the prostate TME. Treatment was well-tolerated and did not lead to unexpected surgical complications, providing proof-of-concept for an immunotherapy-based neoadjuvant approach to prostate cancer treatment. Importantly, we found that ADT significantly increases the intra-

tumoral CD8+ T cell infiltrate in prostate cancer. However, our comprehensive analyses of the immune TME showed that ADT induces other important immunologic changes, with both pro-inflammatory and immunosuppressive effects. Perhaps most strikingly, we observed that the CD8+ T cell infiltrate was accompanied by a proportional increase in Tregs, a key immunosuppressive cell population that mediates immune resistance in multiple tumor types [144]. The addition of cyclophosphamide, which has previously been shown to transiently deplete Tregs, did not appear to significantly deplete Tregs in this setting. The addition of Cy/GVAX to ADT did lead to a modest increase in PD-L1 expression as well as a statistically significant increase in the macrophage marker CHIT1, perhaps suggesting increased immunologic activity for the combination therapy. When accounting for patient age, tumor stage and Gleason sum in a multiple regression model selected by unbiased AIC backward feature selection [145] [146] [147], there were significant improvements in time-to-PSA recurrence and time-to-next therapy in patients treated with Cy/GVAX plus degarelix compared to degarelix alone, suggesting the possibility that the combination regimen has some clinical activity.

Prior pre-clinical and clinical studies showed that androgen deprivation can re-model the immune TME in prostate tumors towards a pro-inflammatory state. Our group previously demonstrated in the MycCaP murine model that ADT initially leads to a pro-inflammatory immune cell infiltrate in prostate tumors with increases in CD8+ T cells, Tregs, macrophages and NK cells [4]. However, this infiltrate is transient and appears to dissipate with the emergence of castration-resistance. Other groups have also shown that androgen ablation can increase B-cell infiltration, which may promote progression to castration-resistance through B-cell-derived lymphotoxin production [148]. In patients, the androgen-receptor blocker flutamide was shown to induce T cell infiltration

and increase expression of pro-inflammatory immune-related genes (interferon-ɣ, TNF-α, Granzyme A) in prostate cancers when given prior to prostatectomy [143] [149]. Several prior studies have also investigated the use of neo-adjuvant vaccine-based immunotherapy approaches to enhance anti-tumor immune responses. For example, the autologous cellular vaccine, sipuleucel-T, was shown to promote lymphocyte recruitment and enhance T_H1 responses when given in the neoadjuvant setting [150] [151].

The findings reported here are largely consistent with these prior observations and suggest that ADT may prime prostate-specific T cell responses. We observed that ADT led to a robust increase in CD8+ T cells, which was not further enhanced by Cy/GVAX. One possible reason for the lack of further CD8+ infiltration with Cy/GVAX could be the allogeneic nature of the GVAX vaccine relative to the patients' tumors. The vaccine cell line PC3 was originally derived from a skull metastasis, and LNCAP is originally derived from a lymph node metastasis, and it is possible that neither consistently shared tissue-specific antigens with the primary prostate tumors in the treated patients. It may also be the case that GM-CSF was insufficiently able to activate dendritic cells, as it has been found that modified versions of GVAX expressing dendritic cell activating molecules such as STING were far more effective in preclinical models [152]. It should also be noted that the prostate cancer microenvironment is particularly immunosuppressive, such that CD8 T-cells isolated from the prostate remain refractory to stimulation even in ex vivo experiments [153], indicating that improved depletion of Tregs may also improve response to GVAX. Of note, there was also increased PD-L1 expression with ADT, which did appear to be augmented by the addition of Cy/GVAX. The significance of this upregulation of PD-L1 is unclear but could reflect an adaptive response to interferon-ɣ produced by activated T-lymphocytes. Future mechanistic work

is required to better understand this observation. Furthermore, and consistent with the hypothesis that counter-regulatory mechanisms can function to maintain immune evasion, we observed an increase in Treg infiltration with ADT. This process of adaptive Treg resistance has not previously been described in the setting of neo-adjuvant ADT, although increases in Treg density have been observed in response to a range of therapies across a number of tumor types, highlighting the notion that adaptive Treg resistance may be a broad-based mechanism that can attenuate maximal responses to immunotherapy in patients with diverse malignancies.

Interestingly, in both treatment groups, differential gene expression analysis showed that degarelix treatment upregulated CHIT1, a marker of macrophage activation shown to regulate many inflammatory processes through stimulation of inflammatory mediators such as IL8, MMP9, CCL2, CCL5, and CCL11, and correlated with levels of IL-1β and TNFα [154]. Given that macrophages are key antigen-presenting cells, this finding corroborates the notion that ADT enhances prostate-antigen presentation and thereby promotes prostate-specific T-cell responses. CHIT1 expression appeared to be further upregulated by the addition of Cy/GVAX to ADT.

Limitations of this study include the relatively small number of patients in each treatment arm and our inability to capture serial immunologic changes within the prostate TME over time. We hypothesized that 2 weeks of ADT would be optimal to elicit robust immunologic responses, since pre-clinical data suggest that the immunologic effects of ADT are transient, with the initial immune infiltrate evolving over time into a more suppressive one, dominated by Tregs [4]. The optimal duration of ADT prior to radical prostatectomy remains unknown and it is possible that the single dose of degarelix acetate used in this study was insufficient to sustain a clinically significant

immune response. Our study used cyclophosphamide in combination with GVAX based on the hypothesis that low-dose cyclophosphamide would be capable of depleting Tregs and therefore augmenting an anti-tumor immune response. This approach was supported by preclinical studies which showed significant augmentation of anti-tumor immunity upon administration of cyclophosphamide approximately 24 hours prior to vaccination with GVAX [155] [156]. The dosage of cyclophosphamide used here reflects the dosage in a breast cancer study that also showed augmentation of anti-tumor immunity with administration of cyclophosphamide prior to a GM-CSF secreting vaccine [157]. However, we observed no difference in Treg density with the addition of Cy/GVAX to degarelix. One possibility is that the dosing regimen of cyclophosphamide used in this study was not optimal for Treg depletion. Since the completion of our study, emerging data showed that oral cyclophosphamide may be more effective for Treg depletion [158] [159]. Given these limitations, future studies may be required to fully characterize the evolution of the immune TME over time and to optimize neoadjuvant immunotherapy in patients with prostate cancer.

However, these results do provide important insights into the immunologic effects of ADT, either alone or in combination with an allogeneic cell-based vaccine. Importantly, the complexity of the immune response to ADT suggests that selectively targeting immunosuppressive cell populations may be essential for maximizing the immunogenicity of neoadjuvant ADT. The observation that ADT can induce adaptive Treg resistance provides a strong rationale for novel strategies aimed at depleting Tregs within the prostate TME. Finally, future mechanistic studies aimed at comprehensively understanding how androgen deprivation regulates anti-tumor immunity in prostate cancer are warranted.

6.5 Methods

Patients

Men with intermediate to high-risk localized prostate adenocarcinoma, defined as clinical stage T1c-T3b, N0, M0 and a Gleason sum ≥ 4+3 (grade group ≥3) in at least two cores were considered eligible if they were planning to undergo prostatectomy. All patients were required to have an Eastern Cooperative Oncology Group performance status of 0 or 1; and normal kidney, liver, and marrow function. Patients with nodal (N1) or distant (M1) metastases were excluded. Key additional exclusion criteria included prior immunotherapy or vaccine therapy for prostate cancer, prior radiation, hormonal, or chemotherapy, autoimmune disease requiring corticosteroids, or known allergy to cyclophosphamide or G-CSF/GM-CSF. Written informed consent was obtained from all patients, and studies were conducted in accordance with the U.S. Revised Common Rule and approved by Institutional Review Board.

Study Design and Treatment

Patients were randomized 1:1 to degarelix alone (240 mg subcutaneously) versus cyclophosphamide (200 mg/m2 intravenously) and GVAX (2.5×10^8 PC3 cells, 1.6×10^8 LNCaP cells) given 2 weeks before degarelix. Randomization was stratified by Gleason sum: ≤7 vs 8-10. All patients underwent radical prostatectomy 2 weeks after degarelix (Figure 48). Prostatectomy specimens were assessed for Gleason grade, nodal involvement, and pathological stage using standard methods. Following pathological review of prostatectomy specimens, a tumor block was selected from the highest-grade tumor located in the prostate and microtome sections were prepared for biological analysis of the TME, including immunohistochemical staining for CD8, FOXP3, and PD-L1, with additional sections for expression profiling (Nanostring). In addition, a

contemporaneous cohort of matched-controls (Cohort C) who did not receive any neoadjuvant therapy provided untreated radical prostatectomy tumor samples, which were compared to post-treatment prostatectomy samples from study Cohort A (degarelix alone) and Cohort B (Cy/GVAX plus degarelix) in genetic and immunohistochemical analysis. Patients were subsequently followed for biochemical (PSA) and metastatic disease progression.

Outcomes:

The co-primary endpoints of the trial were safety and CD8+ T cell density (CD8+ cells/mm^2) in the prostate tumor tissue following neoadjuvant therapy. Safety was assessed using NCI Common Toxicity Criteria version 4.03. Secondary endpoints included feasibility, Treg density (FoxP3+ cells/mm^2) in the prostate gland, CD8 to Treg ratio, time-to-PSA recurrence, time-to-next anti-cancer therapy, and time-to-metastatic progression. Time-to-PSA recurrence was defined as the interval from time of prostatectomy to the time when the PSA was ≥0.2ng/mL for the first of at least two serial rises in PSA (≥2 weeks apart).

Immunohistochemistry

CD8 staining was performed by steaming slides for 45 minutes in Dako Target Retrieval Solution (Agilent Technologies, Inc, Wilmington, DE), followed by incubation with a mouse anti-human monoclonal anti-CD8 antibody for 45 minutes at room temperature (Agilent Technologies, Inc, Wilmington, DE). For FoxP3 staining, slides were steamed for 45 minutes in Dako Target Retrieval Solution (Agilent Technologies, Inc, Wilmington, DE) and then incubated with a mouse monoclonal anti-FoxP3 antibody overnight at 4C (eBioscience, San Diego, CA, 1:250 dilution). For CD8, the secondary antibody used was the UltraVision Quanto Detection System HRP DAB

(ThermoFisher, Waltham, MA). For Foxp3, the secondary antibody was the PowerVision+ kit (Leica Biosystems, Buffalo Grove, IL). Staining was visualized using 3,3'-Diaminobenzidine (DAB) (Sigma, Saint Louis, MO, FAST 3,3'-Diaminobenzidine Tablets) and slides were counterstained with hematoxylin. For CD8 and Foxp3, IHC stained slides were scanned using an Aperio ScanScope CS. Sections for tumor for image analysis were performed using ImageScope by selecting regions of invasive carcinoma and carefully excluding regions in which inflammatory infiltrates involved benign glands. CD8 and Foxp3 cell data were obtained using positive IHC cell counting algorithms implemented in Aperio Spectrum software by applying Hue, Saturation and Brightness (HSB) color space. Cell numbers were normalized to the overall areas/region of interest and annotated a trained pathologist to provide cell density, which was assessed for each patient and compared across study arms. PD-L1 IHC staining and scoring was performed as previously described [137]. Although some PD-L1 expression has previously been reported on immune cells in prostate cancer, such cells are morphologically identified as primarily macrophages; here we analyzed and report tumor-cell PD-L1 expression.

Expression Profiling

Immune gene expression in the prostate TME was profiled using the Nanostring IO360 Immune Panel [142]. Sufficient tissue for analysis was available from 13 patients from arm A (degarelix) and 12 patients from arm B (degarelix + Cy/GVAX) as well as 18 untreated matched-control patients. Nanostring count data were normalized by first thresholding to exceed mean + 1 standard deviation of negative controls, then scaling each sample by a positive control normalization factor to correct for total counts, and additionally, scaling with a set of pre-defined housekeeping genes, as described in the Nanostring documentation [144]. Three housekeeping genes (FCF1, POLR2A,

and TUBB) were excluded from the normalization process due to high cross-sample variance, and two additional genes (CC2D1B and GUSB) were excluded due to poor correlation with other housekeeping genes. This scaling corrected for background noise and differences in total gene count across samples, allowing for differential gene expression between groups to be calculated by unpaired t-test. For each pairwise comparison, we performed Benjamini-Hochberg multiple-testing correction and reported the number of differentially expressed genes with p-value<0.05.

Nanostring data were used to computationally infer an absolute abundance of immune cell types in each sample in order to compare the two study arms with each other and with the untreated group. These analyses were performed using the CIBERSORT algorithm, which de-convolutes gene expression matrices to a mixture of known immune cell types by fitting to a validated reference matrix of 22 immune cell subtypes, where each cell subtype has a defined set of differentially expressed genes [148]. This approach was limited by the fact that Nanostring profiles a limited set of targeted genes rather than the whole-transcriptome, so not all differentially expressed genes in the CIBERSORT reference matrix were captured. However, Nanostring specifically targets immune-related genes, and there are a significant number of differentially expressed genes captured for each immune cell subtype by the Nanostring panel. CIBERSORT was able to de-convolute immune cell composition from these genes with a p-value of <0.05 for 13 treatment arm A samples, 10 treatment arm B samples, and 12 untreated control samples.

Statistical Analysis

Our primary hypothesis was that men receiving Cy/GVAX followed by ADT would have a 2-fold (100%) increase in CD8+ T cell infiltration as compared to men receiving ADT alone. With 16

patients per arm, and assuming an 86% coefficient of variation for the average CD8+ T cell density, a one-sided 0.05 α-level t-test of the logarithms of these ratios would provide 82% power to detect a 2-fold (100%) increase in CD8+ T cell density between treatment groups. Thus, the trial was powered to recruit 32 patients, with a total of 29 patients ultimately recruited. The primary statistical endpoint of this study was CD8+ T cell density quantified by the number of nuclei of staining positive for CD8 per mm2. Following a log transformation, the mean CD8+ T cell densities were compared between treatment arms using a two-way analysis of variance (ANOVA) with the stratification variable, Gleason score, treated as a block factor. Event time distributions for PSA recurrence, time to metastasis, and time to next cancer treatment were estimated with the method of Kaplan and Meier and compared using a stratified Cox proportional hazards model. For all comparisons of differential gene expression, t-tests were applied to the normalized Nanostring counts matrix, and p-values corrected for multiple testing by the Benjamini-Hochberg method. Similarly, t-tests with Benjamini-Hochberg correction were applied to the inferred CIBERSORT immune cell abundance matrices, and to the IHC density values for CD8 and FOXP3. In a secondary analysis, hypothesis testing for unbiased association of clinical variables with time-to-PSA recurrence and time-to-next treatment was performed using multiple Cox regression with backward feature selection using the Akaike Information Criterion [145] [146] [147], and visualized using hazard ratio forest plots and Kaplan-Meier survival curves. The same multiple Cox regression with backward feature selection was performed to test for association of clinical variables with metastasis and time to testosterone recovery (Figure 56). Pearson correlation was also calculated between all clinical, gene expression, and IHC variables as well as correlation of each variable with disease recurrence, visualized in Figure 57. Statistical analyses were performed using R version 3.5.3 and SAS version 9.2.

Chapter 7: Combination of Hormonal Therapy with Immunotherapy Induces T-cell Expansion in Metastatic and Primary Prostate Cancer

7.1 Summary

We and others have shown that the tumor microenvironment (TME) in primary prostate and castration-resistant prostate cancer (CRPC) are relatively void of immune cells. Treatment with Androgen Deprivation Therapy (ADT) is known to induce a complex immune infiltrate in localized prostate cancer, in both animal models and humans. The baseline features of the TME and tumor cells in metastatic tumor sites and the effect of ADT alone or in combination with checkpoint blockade in metastatic, hormone-sensitive prostate cancer (mHSPC) have not been well-described. Here, we present single-cell RNA-sequencing data of longitudinal metastatic tumor biopsies at baseline and on-treatment from patients enrolled on our phase 2 clinical trial (NCT03951831), which aims to test the hypothesis that ADT-induced immune infiltrate can be further augmented with anti-PD-1 inhibition in men with mHSPC. Using protein activity inference, we comprehensively describe the baseline TME and tumor sub-clusters, highlight significant changes induced with treatment, and features of both the immune micro-environment and tumor cells themselves associated with differences in clinical response. We describe a treatment-resistant tumor sub-cluster phenotype that increases in frequency at time of metastatic progression, markers for which are enriched across several bulk-RNA-Seq cohorts in patients with worse clinical outcomes. Our study outlines several potential druggable targets in this tumor

population and in resistance-associated immune populations which may advance treatment and improve outcomes for men with mHSPC.

7.2 Introduction

Localized prostate cancer (PC) is an immunologically 'cold' tumor microenvironment (TME) with a relative dearth of immune cells compared to other tumor types [160] [161]. Preclinical studies and analyses of human primary PC samples show that the immune cell infiltration that does exist is likely immune tolerant and suppressive given the presence of terminally differentiated cytotoxic T cells and T regulatory cells [153] [162] [163]. Several studies show that androgen deprivation therapy (ADT), the backbone of therapy for advanced prostate cancer, induces immunogenic changes in the TME of hormone-sensitive prostate cancer. This is due to several mechanisms including: thymic regeneration and increased production of naïve T cells; decreased tolerance and clonal expansion of effector T cells; stimulation of an antigen-specific adaptive immune response; and by driving a robust and functional immune infiltrate into primary prostate tumors with chemokine and cytokine secretion [164] [12] [149] [165] [166] [4] [167] [168] [169] [170] [171]. However, these favorable effects on the immune system are not durable, and they are often counter balanced by a concomitant increase in immunosuppressive cell compartments or interference with T cell priming [4] [172] [173] [174]. Therefore, combination therapy with ADT and immunotherapies that leverage the positive immune effects and mitigate the immunosuppressive compartments induced by ADT makes rational sense. Investigating the optimal timing and sequencing of specific combination therapies for men with prostate cancer is a promising and active area of clinical and translational research.

Most preclinical and translational studies evaluating the immunogenic effects of ADT to date have focused on primary prostate cancer. The TME of metastatic hormone-sensitive prostate tumors has not been well characterized, nor have the effects of ADT pressure on the TME of metastatic tumors. This is likely due to the challenges associated with tissue acquisition from metastatic tumor biopsies. It is not known whether metastatic hormone-sensitive tumors have immunologically 'cold' TMEs like primary prostate cancer. Studies using digital spatial imaging and single-cell sequencing show that there is a paucity of immune cells in the more advanced castrate-resistant setting, like primary PC [175] [176]. Data comparing PD-L1 protein expression between primary and metastatic castrate-resistant prostate cancer (mCRPC) show that there were notable differences between the two disease stages (7.7% of cases had detectable PD-L1 expression in primary PC and 31.6% of mCRPC cases had detectable expression) [137]. Whether the TME differs by metastatic niches, or the changes induced by ADT pressure are similar across all metastatic niches is also not well described.

In this study, we comprehensively characterized the TME and tumor cells of metastatic hormone-sensitive prostate cancer across several metastatic niches (bone, lymph node, liver, and lung) using high-throughput droplet-based single-cell RNA sequencing (scRNASeq) and our previously developed pipeline for Virtual Inference of Protein Activity by Enriched Regulons (VIPER) [26] [21] [35]. This has enabled deep sub-clustering of immune cell subpopulations and tumor cells as well as amplification of biological signal-to-noise to eliminate data dropout for key regulatory and signaling proteins. We apply this method to a series of paired metastatic tumor biopsies (baseline and on-treatment) from eight patients enrolled in a phase 2 clinical trial that is testing the activity of ADT and an anti-PD-1 antibody (cemiplimab-rwlc) with docetaxel

in men diagnosed with metastatic hormone-sensitive prostate cancer (NCT03951831). In addition to characterizing the differences in TME composition across baseline samples by metastatic site, we also investigated TME and tumor cell transcriptional perturbation and phenotypic changes with treatment, either with ADT alone or ADT plus anti-PD-1 therapy. We next highlight associations between early PSA responses and characteristics of the baseline TME composition and tumor cells. Finally, we apply the CLIA-certified OncoTarget algorithm for inference of druggable proteins active in tumor cell sub-clusters at the single-cell level to reveal candidate drugs for upfront treatment of tumor cells associated with treatment resistance, and we performed gene set enrichment analysis (GSEA) testing association of distinct tumor cell sub-cluster gene sets with recurrence-free survival outcomes in external, publicly available datasets.

7.3 Results

Gene expression and protein activity clustering reveal a robust immune infiltrate in metastatic hormone-sensitive prostate cancer. Given that primary prostate cancer (PC) is a relative immune desert with low proportions of immune cell subpopulations [160] [161], we sought to determine if the tumor microenvironment (TME) in patients with metastatic hormone-sensitive prostate cancer was similarly immunologically 'cold'. We collected baseline, pre-treatment metastatic needle-core biopsies from 10 patients (**Figure 58**) across 4 different metastatic niches (bone, lymph nodes, liver, and lung), isolated and analyzed all live cells using single-cell RNA sequencing (scRNASeq).

Patient	Biopsy Location	Baseline (no. cells)	ADT only (no. cells)	ADT + aPD-1 (no. cells)	Recurrence (no. cells)
1	Bone	2360	0	686	0
3	Bone	1027	995	0	764
5	Lymph Node	4090	5153	0	0
6	Liver	0	0	2475	0
7	Lung	699	0	1956	0
8	Bone	606	1784	0	0
10	Bone	521	0	97	0
12	Bone	0	913	0	0
13	Lymph Node	1212	0	0	0
14	Liver	2161	0	0	0

Figure 58: Tissue site and cellular yield per biopsy sample.

We performed both gene expression-based clustering and protein activity-based clustering, using VIPER [35]. We correlated these scRNASeq data to a preexisting dataset of lineage-sorted bulk RNA sequencing with SingleR, commonly used to phenotype single-cell subpopulations [48]. Gene expression-based clustering revealed 15 overall clusters across all metastatic sites, including 12 distinct immune cell clusters, as well as fibroblasts, endothelial, and epithelial clusters (Figure 59).

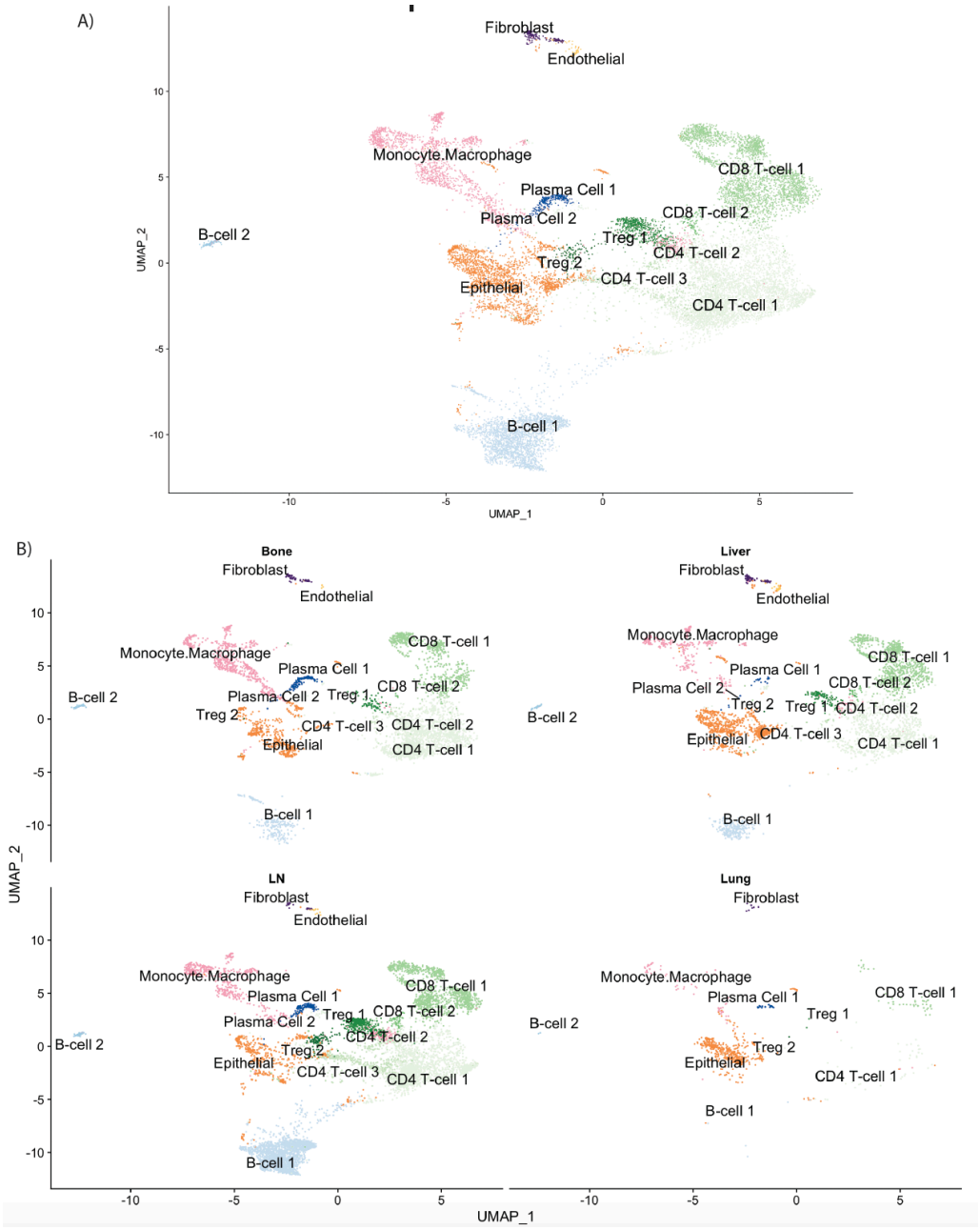


Figure 59: Gene Expression Clustering

A) UMAP plot showing clustering of all cells in tumor micro-environment across all patients, clustered on gene expression instead of VIPER-inferred protein activity. Cell types are inferred by SingleR. B) UMAP plot from A, split by metastatic tissue site.

Clustering was performed using the Seurat Louvain algorithm with resolution optimized in the range of 0 to 1 at increments of 0.01 by sub-sampled silhouette score [35]. Clusters were visualized using 2D UMAP dimensionality reduction. Inspection of the top five most differentially upregulated transcripts by cluster (Figure 60) further confirmed the ascribed cellular identity of clusters assigned by SingleR.

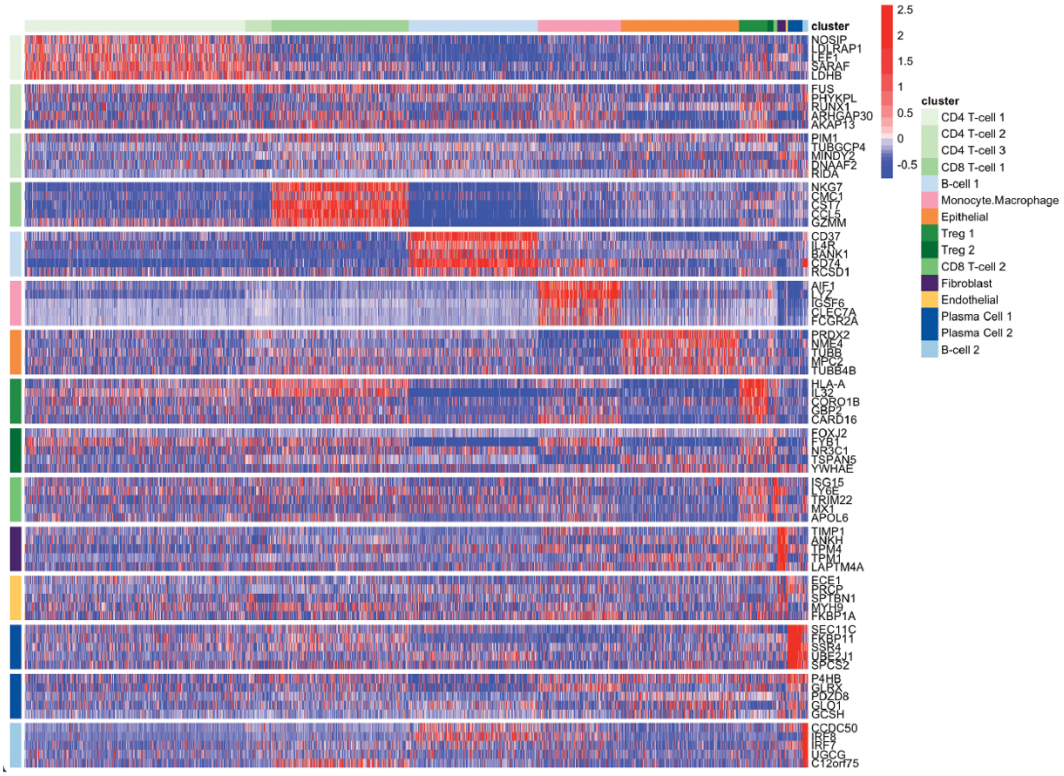


Figure 60: Top Gene Expression Cluster Markers

Heatmap of top 5 most differentially upregulated genes for each cell type cluster from aggregate single-cell RNA-Sequencing data across all patient samples, with clusters corresponding to Figure 59. Each row represents a protein, grouped by cluster in which they are the most active, with cluster labels on the x and y-axes.

For example, granzyme M (GZMM) and natural killer granule 7 (NKG7) were differentially upregulated in CD8 T cells, and CD37 in B cells. The sorted bulk RNA seq reference used by SingleR does not contain tumor cells, therefore tumor cells with epithelial origin such as prostate cancer cells are labelled by SingleR as 'epithelial cells.' These can then be confirmed as tumor cells by expression of tumor marker genes such as KLK3 and presence of inferred Copy Number Variations (Figure 61).

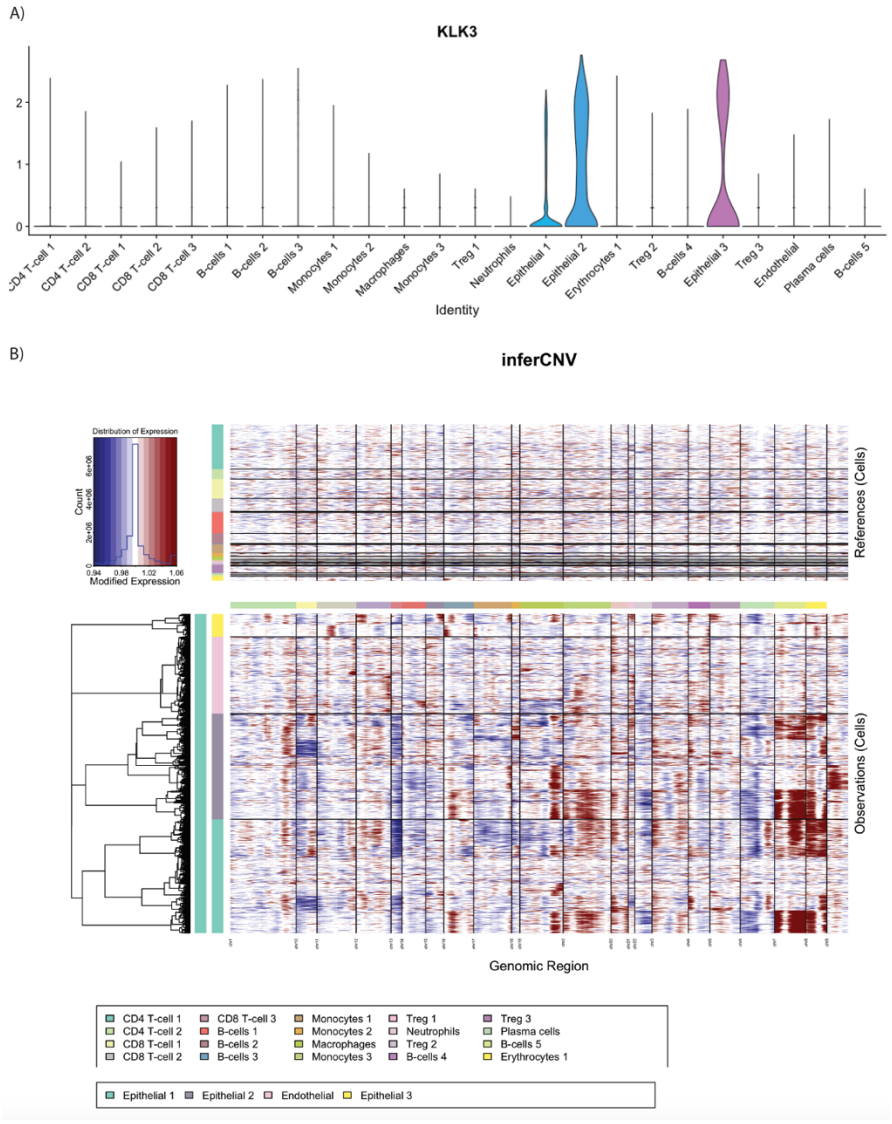


Figure 61: Identification of Tumor Cells by Marker Expression and Copy Number Variation

A) log10 normalized expression of prostate cancer tumor marker protein KLK3 in each cell cluster identified by VIPER, such that expression is non-zero only in Epithelial cell clusters. B) InferCNV plot of cell-by-cell copy number variations, where all immune-lineage cells are taken as a copy-number-normal reference for inference of variations in copy number in Epithelial cell clusters and Endothelial cell cluster as a control. Each epithelial cell cluster is notable copy number aberrant across multiple chromosomes, while endothelial cells are grossly copy number normal.

Due to high levels of data dropout, single-cell RNA-Seq is inherently noisy, and we therefore employed a previously published analysis pipeline which uses the VIPER algorithm for inference of protein activity from single-cell gene expression data to mitigate dropout and amplify detection of transcriptional regulatory proteins and signaling molecules through their effect on downstream transcriptional state [35]. We re-clustered on inferred protein activity to increase resolution of the immune and stromal cell subpopulations. This analysis revealed 24 distinct protein activity-based clusters, primarily immune cells but also including erythrocytes, endothelial and three ‘epithelial’ or tumor cell clusters (Figure 62).

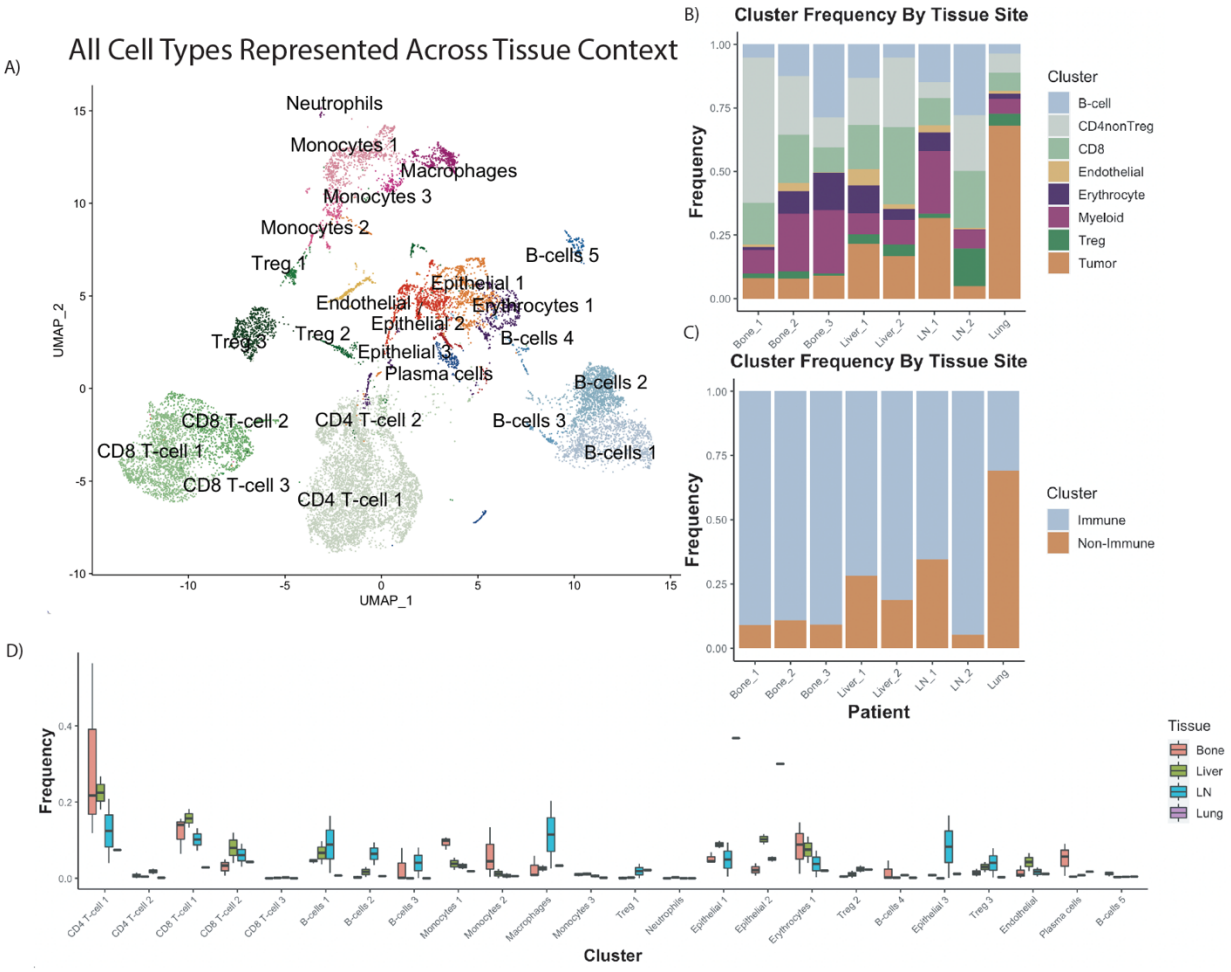


Figure 62: Baseline Composition of Micro-Environment by Tissue Site.

A) Uniform Manifold Projection (UMAP) plot constructed from VIPER-inferred protein activity of all cells in aggregate across baseline pre-treatment patient samples. Cells are clustered by resolution-optimized Louvain algorithm with cell type inferred by SingleR B) Stacked barplot of the frequency of each major cell lineage within each baseline patient sample, with each column representing a unique patient and patients grouped by metastatic site. Cell clusters from A are aggregated by shared cell type. C) Stacked barplot of immune vs non-immune cell frequencies, from B. D) Boxplot showing distribution of frequencies for each cell cluster in A at baseline, comparing tissue sites.

VIPER increased granularity with respect to the number of myeloid, lymphoid, and epithelial clusters compared to gene expression clustering. While gene expression clustering yielded only

one large homogeneous monocyte / macrophage combined cluster, VIPER-based protein activity inference showed three distinct monocyte cellular subpopulations, one each of macrophage and neutrophil clusters. Additionally, there was further refinement of the T cell clusters specifically with additional clusters of T-regulatory (Treg) and CD8 T cell clusters using protein activity-based analysis compared to gene expression. Furthermore, five B cell clusters and one plasma cell cluster were identified with VIPER as compared with one and two clusters, respectively, with gene expression. Overall, the mean proportion of immune cells across all metastatic sites was 87% (range: 30.9% [lung] – 94.6% [lymph node]) and exceeds the sparse immune infiltration typically seen in primary prostate cancer [163] [177].

Protein activity reveals distinct differences in immune cell subpopulations in different metastatic sites. The proportional composition of immune cell subpopulations in the TME can vary depending on the tissue type [178] [179]. To compare cellular composition across different metastatic sites, we collapsed the initial 24 VIPER clusters into eight lineage-specific meta-clusters (e.g., B cells, CD4 non-Treg, CD8, endothelial, erythrocyte, myeloid, Treg, and tumor) and visualized frequencies across the four different metastatic sites prior to treatment with ADT or anti-PD-1 therapy (Figure 62B, 62C). Overall, bone, liver, and lymph node were more immune infiltrated, with mean proportion of immune cells of 90.5% and 77.9%, and 94.6% respectively, compared to 30.9% in lung. We next compared the frequency of the 24 different cellular subpopulations identified across the four different metastatic sites (bone, lymph node, liver, lung) (Figure 62D), with top protein activity markers of each population shown in Figure 63.

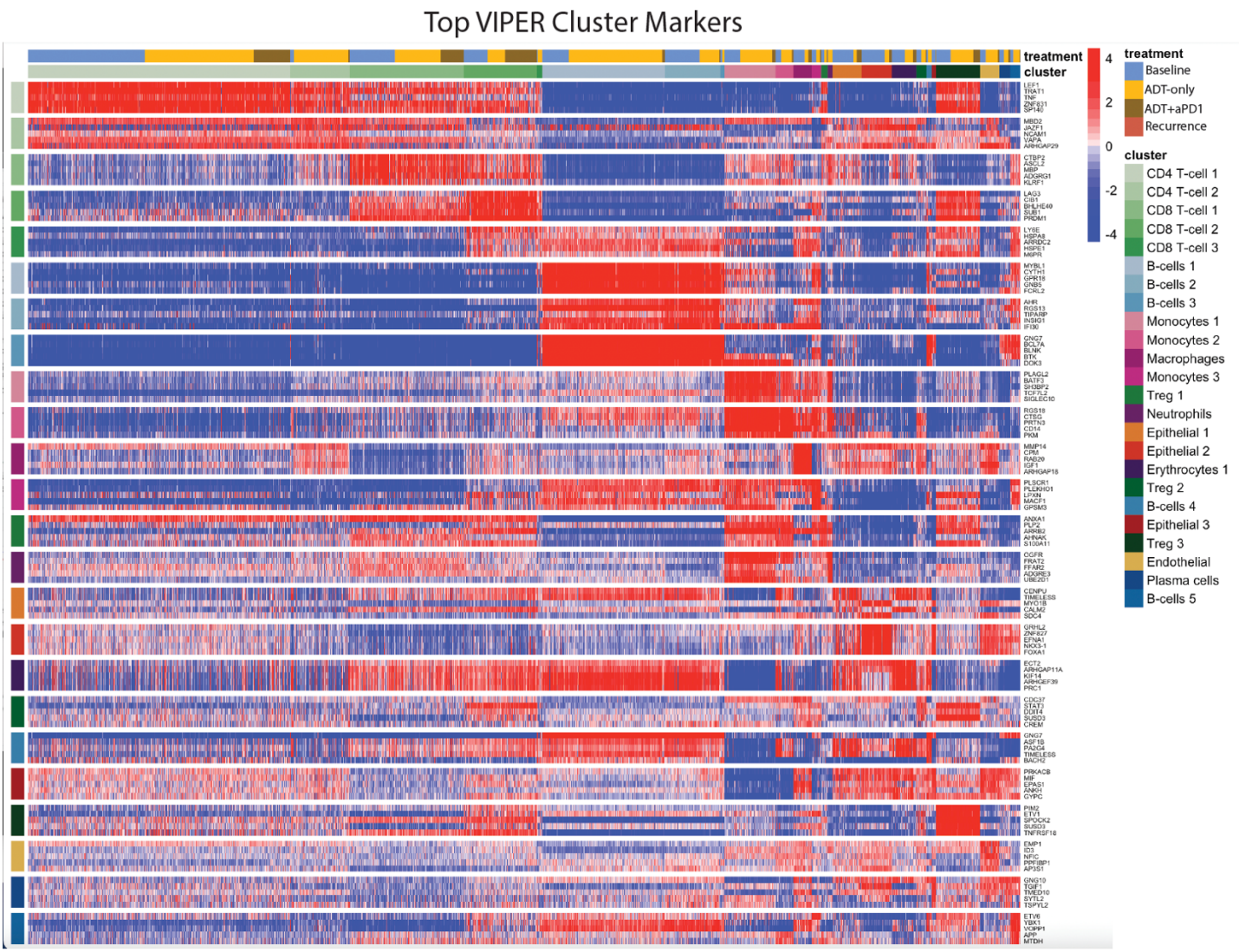


Figure 63: Top Protein Activity Cell Cluster Markers

Heatmap of top 5 most differentially activated proteins for each cell type cluster from aggregate single-cell RNA-Sequencing data across all patient samples. Each row represents a protein, grouped by cluster in which they are the most active, with cluster labels on the x and y-axes. Each column represents a single cell. Above the x-axis cluster label there is also a treatment label indicating timepoint at which a given cell was profiled.

In the **bone** samples, as expected, more plasma cells were seen relative to other sites.

Additionally, there was an increased frequency of monocytes (monocytes 1 and 2) relative to

other metastatic sites, with increased protein activity of transcriptional repressors (BATF3), transcriptional activators (SH3BP2), protein members in G-protein signaling (RGS18), and serine proteases (PRTN3). We also saw a large population of erythrocytes in the bone metastases with high protein activity of epithelial cell transforming 2 (ECT2), Rho GTPase Activating Protein 11A (ARHGAP11A), and Kinesis Family Member 14 (KIF14), proteins that are involved in mitosis, cell-cycle arrest, and microtubule motor proteins respectively. These likely represent dividing erythroid progenitor cells in the bone marrow captured incidentally by the process of bone marrow biopsy. In **lymph node** samples, a robust B cell population was seen (B cell 2). There was also a higher proportion of T regulatory cells (Treg 3), with elevated activity of TNFSRF18 (GITR), in the lymph nodes compared to other metastatic sites. Of interest, this specific T regulatory population had high activity levels of ETS Variant Transcription Factor 1 (ETV1) which is a gene known to be overexpressed in prostate cancer [180] [181]. To our knowledge, this has not been previously described in T regulatory cells of prostate cancer tumor metastases and supports recent findings that immune cells can express tumor marker genes [182]. The **liver** metastases had immune infiltrations similar in both overall proportion and in subpopulation frequencies to bone metastases. Notably across all tissues there was a large proportion of CD8 T cells (CD8 T cell 1 and 2) and CD4 non-Treg T cells (CD4 T cell 1). The CD8 T cell 2 cluster was chiefly defined by increased protein activity of lymphocyte activation gene 3 protein (LAG3), an inhibitory immune receptor [183]. Finally, the single lung metastasis profiled was notably the most immune depleted at baseline, with only 30.9% immune cells overall.

Treatment with combination ADT and anti-PD-1 results in a dramatic expansion of CD8 T

cells across several metastatic sites. To compare the immunologic effects of ADT alone and ADT plus anti-PD-1 (cemiplimab-rwlc) in the TME of the four different metastatic sites, we compared cluster frequencies and visualized changes in the microenvironment by UMAP plots and stacked bar graphs (Figure 64).

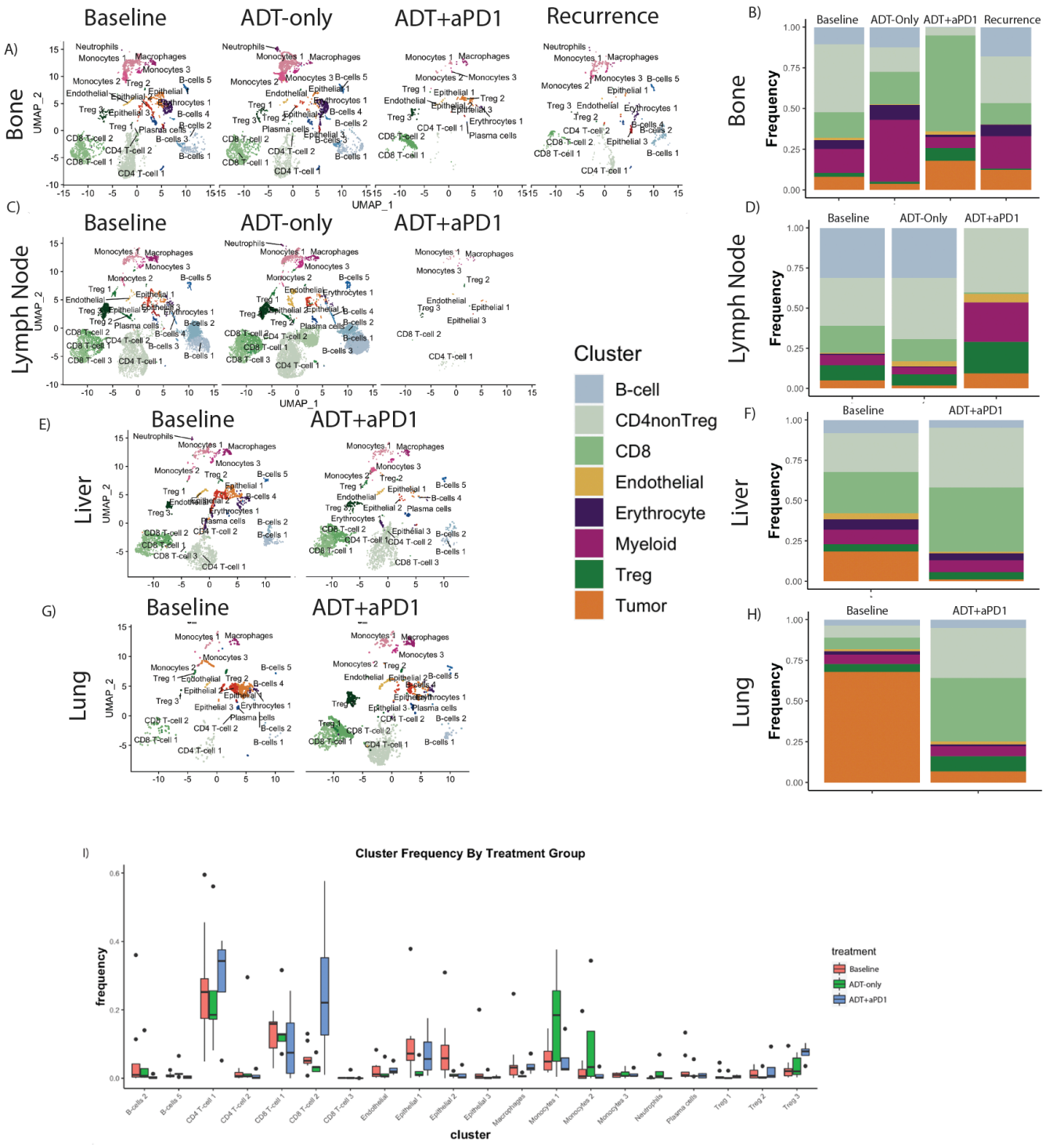


Figure 64: Treatment with ADT+aPD1 Induces Dramatic Changes in the Tumor Micro-Environment

A) UMAP plot of all cells from patients with metastatic Bone lesions, split by treatment time-point (Baseline, ADT-only, ADT+aPD1, and post-treatment Recurrence) and labelled by cell

cluster. B) Stacked barplot showing the relative frequency of each major cell lineage by treatment time-point for patients with metastatic Bone lesions, with each column representing aggregate of all samples profiled at a specific treatment time-point. C) UMAP plot, as in A, for patients with metastatic Lymph Node lesions. D) Stacked barplot, as in B, for patients with metastatic Lymph Node lesions. E) UMAP plot, as in A, for patients with metastatic Liver lesions. F) Stacked barplot, as in B, for patients with metastatic Liver lesions. G) UMAP plot, as in A, for patients with metastatic Lung lesions. H) Stacked barplot, as in B, for patients with metastatic Lung Lesions. I) Boxplot showing distribution of frequencies for each cell cluster, comparing frequencies across treatment time-points including Baseline, ADT-only, and ADT+aPD1.

All patients on the clinical trial were required to have a baseline metastatic biopsy as well as an on-treatment biopsy. Patients were randomized to one of two time points for the on-treatment biopsy, either four weeks after beginning ADT (degarelix) initiation or after ADT plus two cycles of anti-PD-1 antibody. Tumor biopsies at the time of disease progression were optional. In each patient, the same site of disease was sampled for both the baseline and on-treatment biopsies to compare changes in the TME within the same tissue type. Overall, enough patients with bone and lymph node metastases were enrolled to enable collection of biopsy samples at baseline and both on-treatment time points. Critically, we were also able to obtain a tumor progression biopsy from a patient with subsequent tumor recurrence in the bone after 11 months on treatment. Liver and lung biopsy samples were collected at baseline and after ADT with two cycles of anti-PD-1. No patients with lung and liver metastases were randomized to have their on-treatment biopsy after ADT alone.

Treatment pressure can induce complex changes in the TME. We thus used our single-cell data to interrogate the dynamic shifts in cellular composition with ADT and ADT plus anti-PD-1 as

compared to baseline samples. Collapsing sub-clusters into B-cells, CD4 non-Tregs, Tregs, CD8 T-cells, Myeloid cells, Endothelial cells, Erythrocytes, and Tumor cells, we have characterized changes induced in the tumor micro-environment by treatment in each metastatic tissue site. In the **bone**, treatment with ADT resulted in an increased proportion of myeloid cells with a relative decrease in CD4 non-T reg cells and tumor cells (Figure 64A-B). However, myeloid cell abundance decreased with a concomitant dramatic increase in CD8 T cells after treatment with the combination of ADT and anti-PD-1. In the single bone tumor progression biopsy, the relative frequencies of each cell type returned to a distribution resembling the baseline samples, albeit with a greater proportion of tumor cells. In the **lymph node** samples, treatment with ADT resulted in a mild expansion of CD4 non-Treg cells (Figure 64C-D). Unlike the bone samples, myeloid cells were not expanded in the lymph node with ADT alone. These findings confirmed our suspicion that treatment-induced immunologic changes vary based on the metastatic niche and that different strategies to alter tumor immunology may be warranted depending on the site of disease. Few cells were recovered in lymph nodes after the combination of ADT and anti-PD-1. However, of the cells recovered, there was a greater proportion of Treg cells and myeloid cells, with virtually no CD8 T-cells or B cells present. Surprisingly, in both the bone and lymph nodes, there was a relative increase in tumor cells after the combination of ADT and anti-PD-1 compared to baseline and ADT only samples. This is in contrast to observations from the **viscera (liver and lungs),** where combination ADT and anti-PD-1 demonstrated a substantial reduction in the overall proportion of tumor cells (Figure 64E-H). Additionally, the expansion of the myeloid compartment seen in bone samples after ADT alone and in the lymph nodes with combination therapy was not observed in the viscera **(liver and lungs)**. However, similar to the bone samples, there was a dramatic increase in the CD8 T cells with combination of ADT and

anti-PD-1 therapies in the liver and lung samples, which was not seen in either case with ADT alone.

In fact, the only tissue site in which dramatic expansion of CD8 T-cells with combination therapy was not observed was lymph node, which was also severely limited in sample size in terms of the number of cells profiled from the treatment time-point following combination ADT and anti-PD1. Taken together, these data demonstrate that anti-PD-1 immunotherapy increased CD8 T cell infiltration into metastatic sites in combination with ADT to an extent that was not observed with ADT alone. Further interrogating the relative expansion or depletion of each cell sub-cluster in response to treatment reveals a few specific phenotypes most responsive to therapy (Figure 64I). Overall, after ADT alone, the median level of CD8 T cells cluster 2 decreased relative to baseline ($p=0.034$), while there was a concordant increase in monocytes ($p=0.036$). Notably after treatment with ADT and anti-PD-1 we observed a dramatic expansion in CD4 T cells cluster 1 (characterized by high protein activity of TNF; Figure 63), CD8 T cells cluster 2 (characterized by high protein activity of LAG3; Figure 63), and Treg cells cluster 3 (characterized by high protein activity of TNFRSF18; Figure 63), with p-value = 0.033, 0.026, and 0.008, respectively. These three populations represent the bulk of tumor-infiltrating immune cells induced by anti-PD1 therapy.

Association of immune subpopulations at baseline with subsequent PSA response.

We sought to compare baseline cellular subpopulation frequencies by treatment response to determine if the presence of any subpopulations at baseline was associated with PSA response to therapy as a potential biomarker or mechanistic target for depletion in therapy-resistant patients.

We first categorized patients into three treatment response groups (early PSA response, stable disease, or late progressors) based on PSA log10 fold-change (Figure 65A), such that early responders represent patients with excellent rapid decrease in PSA to below 1% of initial value, indicating excellent response to therapy, and late progressors represent patients who initially respond to therapy with a decrease in PSA level but begin a rapid increase in PSA by week 28. We then compared the frequencies of each cell cluster between patients categorized as 'early responders' versus 'late progressors'

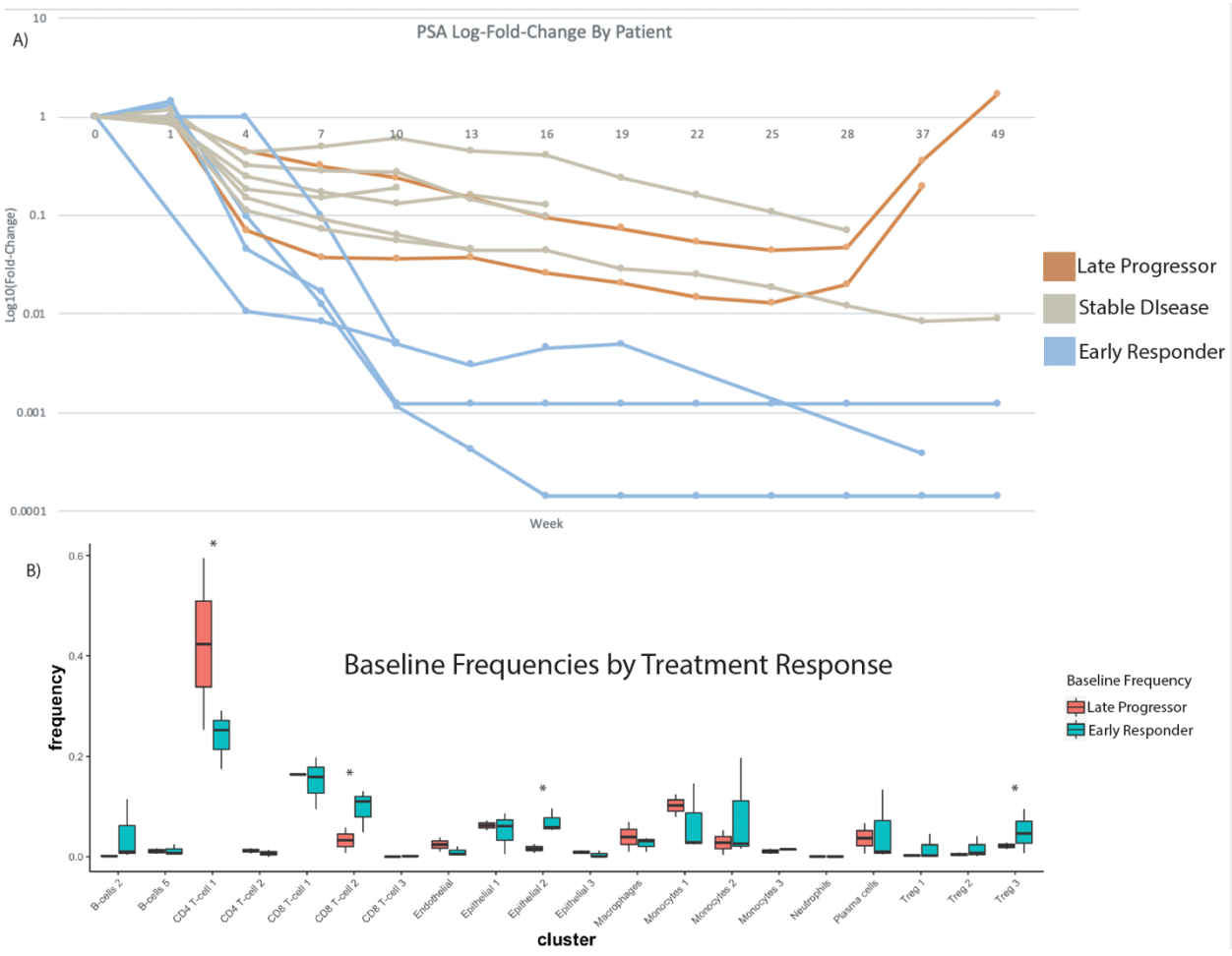


Figure 65: Differences in Baseline Immune Composition Associate with Differences in Treatment Response

A) Spider-plot of log10(Fold-Change) from Baseline in Prostate-Specific Antigen (PSA) over time with treatment, for each patient, such that three patients, labelled in blue, exhibited rapid and dramatic decrease to below 1% of initial PSA and were identified as Early Responders to treatment, and two patients, labelled in orange, initially responded to treatment with a rapid increase in PSA observed after on-treatment week 28. These were considered Late Progressors on-treatment. The remaining patients, in grey, generally trended toward a decreasing PSA, though not as rapidly as the Early Responders. B) Boxplot showing distribution of frequencies at Baseline for each cell cluster, comparing frequencies in Early Responders vs Late Progressors, such that clusters with significant difference at baseline ($p<0.05$ by Student's T-test) included CD4 T-cell 1, CD8 T-cell 2, Treg 3, and Epithelial 2.

Overall, higher levels of CD8 T cells cluster 2 at baseline were significantly associated with an early PSA response ($p = 0.007$). Interestingly, this is the same population of LAG3+ CD8 T-cells expanded by anti-PD1 immunotherapy. Additionally, we found that the specific subpopulation of TNFRSF18+ T regulatory cells (Treg cluster 3) was also associated with an early PSA response, though not with statistical significance ($p = 0.14$). Conversely, CD4 T cells cluster 1 was significantly associated with late PSA progression ($p = 0.026$). One of the most differentially active proteins by VIPER in CD4 T cell 1 cluster was tumor necrosis family (TNF), a multifunction proinflammatory cytokine implicated in tumor progression [184] [185] [186]. These data may suggest modulation of TNF as a potentiating adjunct to the administered combination ADT plus anti-PD1 therapy.

Tumor cell clustering revealed a phenotypic shift in tumor cells across metastatic sites and with treatment. Initial analysis of tumor cells by protein activity-based clustering resulted in three 'epithelial' clusters (epithelial 1, epithelial 2, and epithelial 3). As discussed, we performed copy number alteration (CNA) inference and profiling of KLK3 prostate tumor marker

expression to confirm that the three epithelial clusters indeed represented tumor cells (Figure 61). Tumor cells were identified in all baseline samples across metastatic sites and all three epithelial / tumor clusters were observed in bone, lymph node, and lung samples. In order to provide more granularity of the transcriptomic program in all tumor cells across different metastatic sites, we pursued additional sub-clustering of the tumor cells after excluding all other cell types. This analysis yielded eight tumor sub-clusters (Figure 66A).

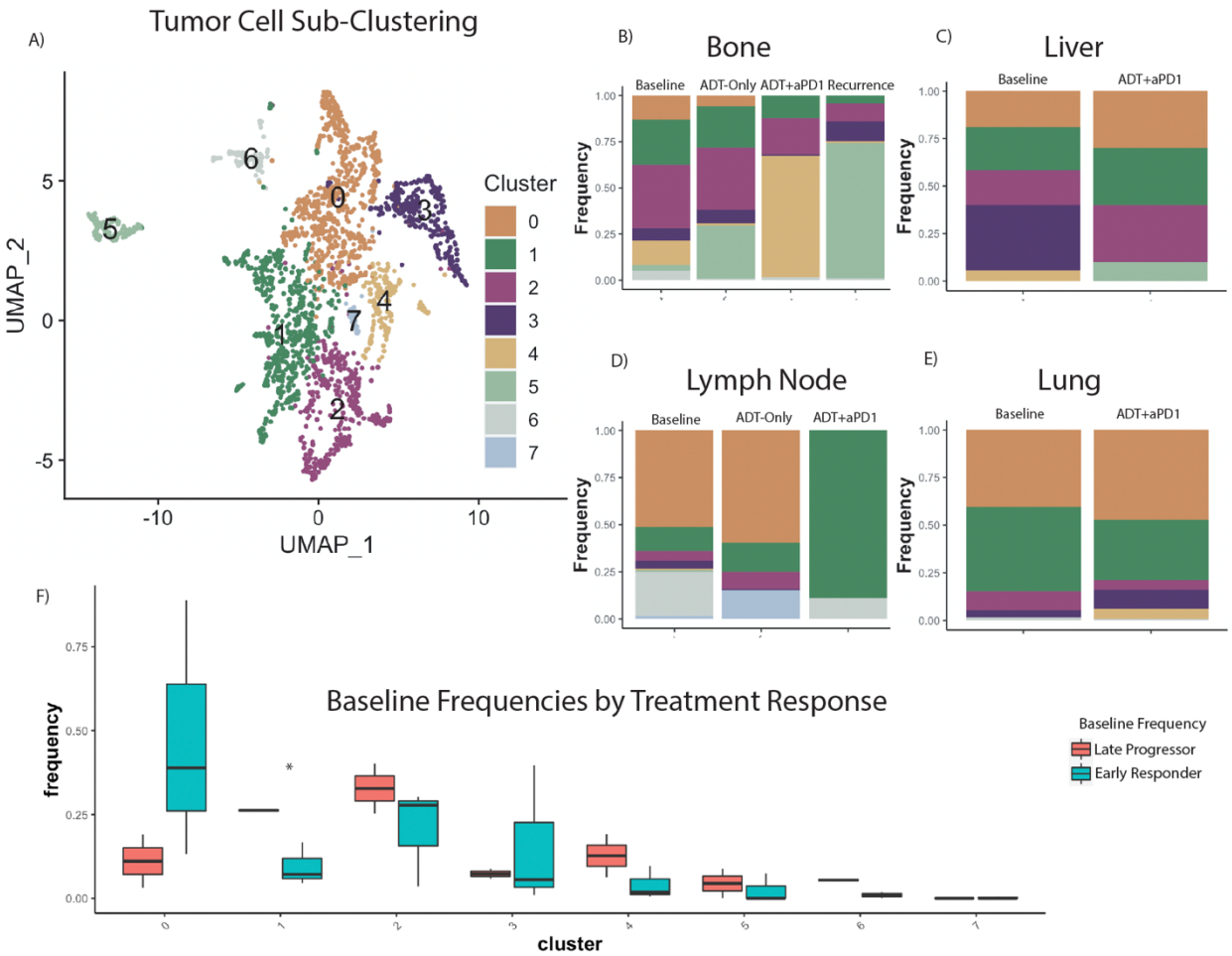


Figure 66: Sub-Clustering Reveals Heterogeneity of Tumor Cells by Tissue Site

A) UMAP plot showing sub-clustering by resolution-optimized Louvain algorithm of only tumor cells (Epithelial 2 and Epithelial 3 from Figure 62A). Plot shows aggregate of all 2,550 tumor cells across all patients at all time-points. B) Stacked barplot of tumor cluster frequency by treatment time-point in patients with metastatic Bone tumors. C) Stacked barplot of tumor cluster

frequency by treatment time-point in patients with metastatic Lymph Node tumors. D) Stacked barplot of tumor cluster frequency by treatment time-point in patients with metastatic Liver tumors. E) Stacked barplot of tumor cluster frequency by treatment time-point in patients with metastatic Lung tumors. F) Boxplot showing distribution of frequencies at Baseline for each tumor sub-cluster, comparing frequencies in Early Responders vs Late Progressors, such that the only cluster with significant difference at baseline ($p<0.05$ by Student's T-test) was cluster 1, with higher baseline frequency in Late Progressors.

These were assessed for enrichment in hallmarks of cancer pathways among differentially activated proteins (Figure 67), such that cluster 0 is most defined by enrichment of androgen response, clusters 1 and 2 are defined by upregulation of E2F targets, Myc targets, and G2M checkpoint, clusters 3 and 4 are defined by upregulation of TNFa signalling and interferon response, cluster 5 is defined by heme metabolism, cluster 6 by unfolded protein response and androgen response, and cluster 7 by the reactive oxygen species pathway.

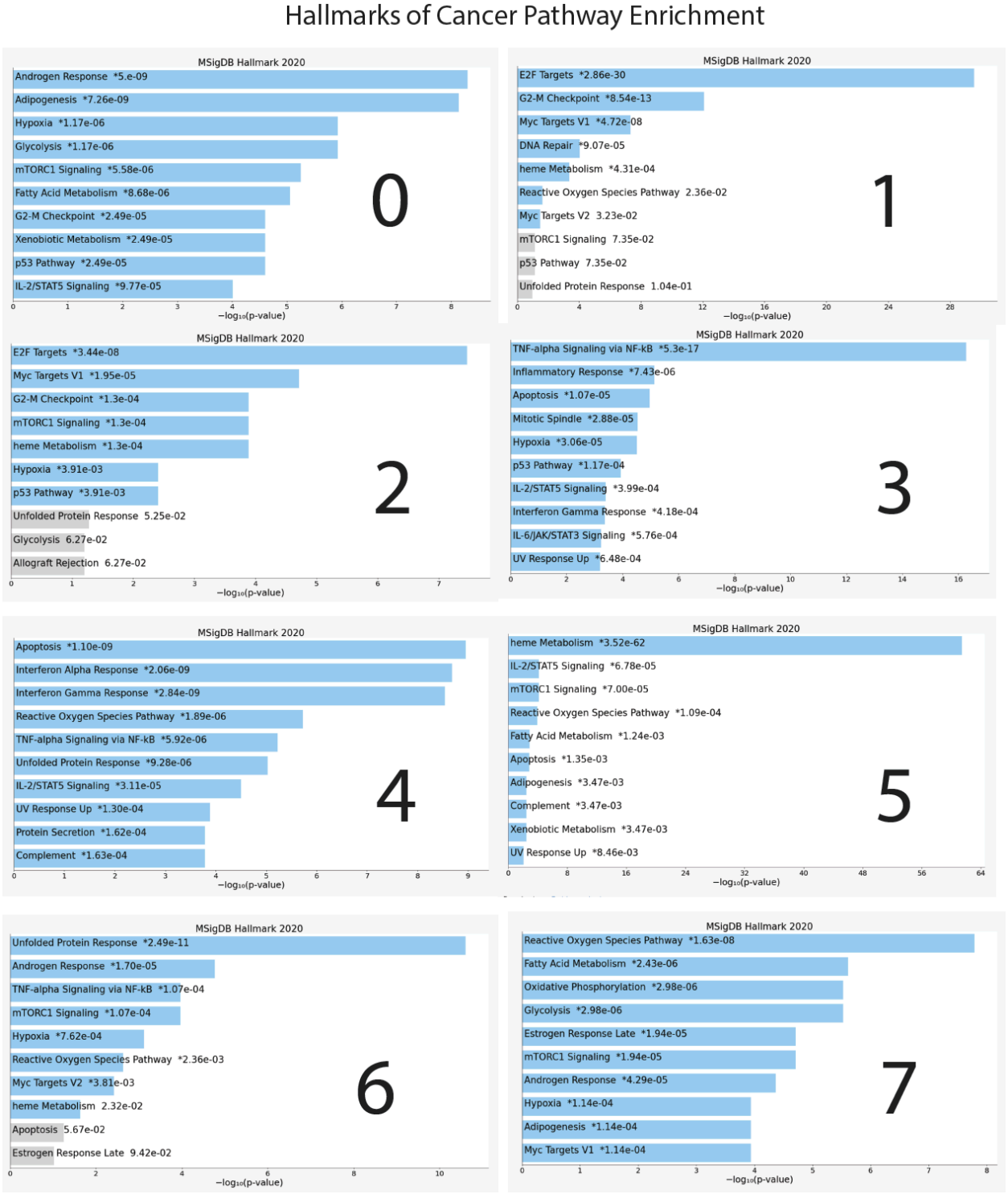


Figure 67: Hallmarks of Cancer Enriched Pathways in Tumor Cell Sub-Clusters

For each tumor cell sub-cluster identified in Figure 66, plots of the top10 enriched pathways from Hallmarks of Cancer. Pathway enrichment is computed on genes differentially expressed in

each tumor sub-cluster relative to other tumor sub-clusters. -Log10(p-values) are plotted on the x-axes, such that statistically significant enriched pathways are shaded in blue.

Top differentially upregulated proteins in each cluster are reported in Figure 68.

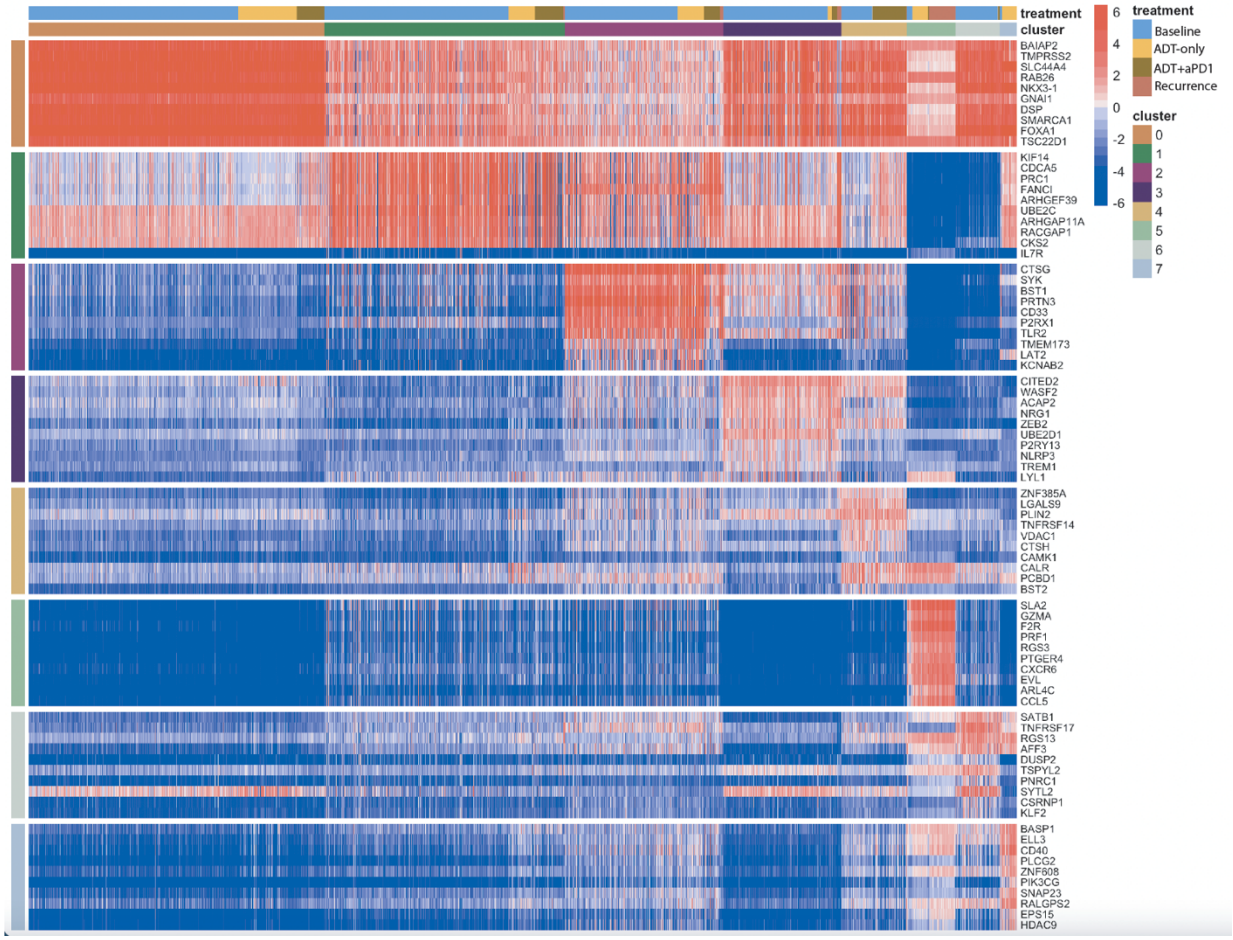


Figure 68: Top Protein Activity Tumor Sub-Cluster Markers

Heatmap of top 10 most differentially activated proteins for each tumor sub-cluster in Figure 66 from aggregate single-cell RNA-Sequencing data across all patient samples. Each row represents a protein, grouped by cluster in which they are the most active, with cluster labels on the x and y-axes. Each column represents a single cell. Above the x-axis cluster label there is also a treatment label indicating timepoint at which a given cell was profiled.

Stacked frequency bar plots (Figure 66B-E) show changes in the relative proportions of each

tumor subcluster across different metastatic sites at baseline, after ADT only, after combination ADT and anti-PD-1, and at recurrence in available biopsy samples. At baseline, there was wide variability in the composition of tumor subclusters across the different metastatic sites. The bone and lymph node samples were more heterogenous, with nearly all tumor sub-clusters identified within these samples. In contrast, the liver and lung samples had fewer tumor cell sub-clusters represented. Notably, lung was comprised almost entirely of tumor sub-cluster 0 and 1. After treatment, the proportions of tumor subclusters differentially changed between different metastatic sites. In bone samples, there was a relative increase in cluster 5 after administration of ADT. With the addition of anti-PD-1 to ADT cluster 5 nearly disappeared and subcluster 4 comprised nearly 65% of the tumor cells. However, in the recurrent bone sample, subcluster 5 returned and comprised nearly 75% of all tumor cells, while sub-cluster 4 represented only 1% of all tumor cells at recurrence (Figure 66B). In the lymph node baseline samples, sub-cluster 0 accounted for nearly 50% of tumor cells. After ADT, the relative proportion of subcluster 0 increased to ~65% of all tumor cells present. Similar to bone samples, after ADT only the proportions of tumor sub-clusters in the lymph nodes changed slightly but overall remained heterogeneous. However, with the addition of anti-PD-1 therapy, a predominant sub-cluster (sub-cluster 1) emerged in the lymph node samples. The liver and lung samples each maintained heterogeneity without emergence of a predominant subcluster after treatment with combination ADT and anti-PD-1. There were no patients with liver and lung metastases randomized to the ADT only timepoint.

Tumor subclusters present at baseline differ by subsequent treatment response and druggable protein profile. We next and most critically compared the frequency of tumor

subclusters at baseline by subsequent PSA treatment response in order to assess if there were significant differences in tumor phenotypes associated with ‘early responders’ versus ‘late progressors’. Sub-cluster 0, which had the highest upregulation of androgen response, was significantly higher at baseline in ‘early responders’ ($p = 0.05$) compared to ‘late progressors.’ Conversely, sub-clusters 1 and 2 were higher at baseline in ‘late progressors’, who were ultimately non-responsive to treatment ($p = 0.0008$, $p = 0.08$, respectively). Interestingly, two proteins (TMPRSS2 and NKX3-1) among the top differentially active proteins in subcluster 0 are mediated by the androgen receptor (AR). There were no other AR-mediated proteins observed in the top differential protein list for the other tumor subclusters. This may indicate that a higher proportion of AR-mediated proteins within tumor cells is predictive of treatment response and may be a novel predictive biomarker in future studies. More interestingly, clusters 1 and 2 share high differential protein activity of KIF14 as the most-upregulated protein, which has previously been described as a candidate oncogene correlating with poor prognosis in prostate cancer [187]. This suggests that targeting these populations may represent an attractive prospect to overcome any treatment resistance mediated by these tumor cells.

To this end, we have utilized the OncoTarget algorithm [24] to identify potential druggable proteins active in each tumor cell subcluster according to compounds listed in DrugBank [188] (Figure 69).

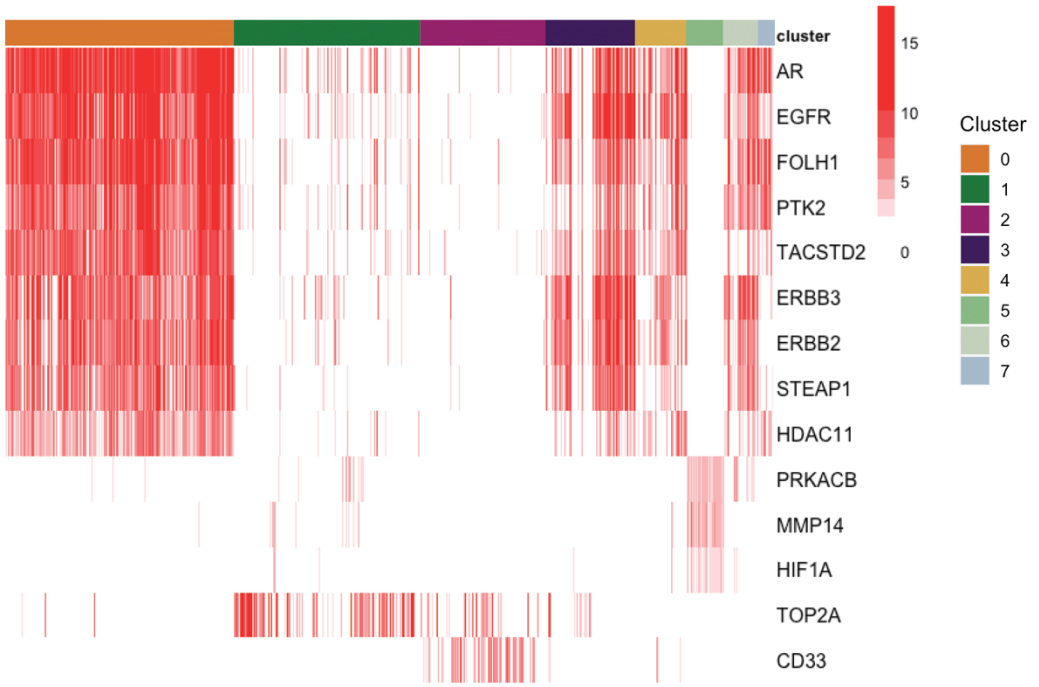


Figure 69: OncoTarget Druggable Proteins in Each Tumor Cell Sub-Cluster

Single-Cell heatmap of all druggable proteins from DrugBank, active with median -log10(p-value) > 5 in any tumor cell sub-cluster, as inferred by OncoTarget. Clusters 1 and 2 are the most phenotypically distinct with respect to druggable protein activity, as they do not have high activity of AR and are characterized instead by high activity of TOP2A. Cluster 5, which was specifically enriched in post-treatment recurrence for the Late Progressor sample with post-recurrence single-cell RNA-Seq, is also phenotypically distinct, with activity of PRKACB, MMP14, and HIF1A.

Amazingly, clusters 1 and 2, which associated at baseline with subsequent tumor progression on treatment, lacked activity of druggable proteins shared by other tumor cell clusters, most notably activity of the androgen receptor protein (AR). AR activity was highest in cluster 0, which was most enriched in subsequent early responders. Other druggable proteins with high activity in this cluster as well as clusters 3, 4, 6, and 7 included EGFR, FOLH1, PTK2, TACSTD2, ERBB3, ERBB3, STEAP1, and HDAC11 (Figure 69). Clusters 1 and 2, while they lack these druggable

targets, share elevated activity of druggable protein topoisomerase 2-alpha (TOP2A), for which 36 FDA-approved and investigational compounds exist as drugs annotated in DrugBank, including Doxorubicin and Etoposide. Cluster 2 also has uniquely elevated activity of CD33, druggable by investigational compounds Gemtuzumab ozogamicin and AVE9633. Finally, cluster 5, which was dramatically expanded in the bone metastasis patient profiled at time of recurrence (5B), also had a unique druggable protein profile lacking activity of AR and the other druggable proteins identified in clusters 0, 3, 4, 6, and 7. Instead, cluster 5 exhibited high activity of PRKACB, MMP14, and HIF1A. This represents a unique profile of druggable proteins for which subsequent patients may be assessed and which may represent a rare and more aggressive prostate tumor cell phenotype, but which did not associate with differences in subsequent response to treatment at baseline across patients (Figure 66F), instead being represented primarily in a single outlier patient.

Association of Tumor Cell Cluster Signatures with Clinical Outcome is Validated in External Bulk-RNA-Seq Cohort In order to assess the generalizability of our findings with respect to tumor cell sub-cluster association with treatment response, we have defined each set of differentially active proteins in each sub-cluster as a unique protein activity signature for that cluster (Figure 68). With these signatures and our protein activity inference algorithm, we can test for the enrichment of each cluster within larger cohorts of bulk-RNA-Sequencing data. While there are no previously published cohorts of metastatic prostate cancer patients treated with combination ADT plus anti-PD1 immunotherapy, we can assess the general prognostic significance of each tumor cell population across treatments in the TCGA. By Cox regression on patient-by-patient normalized enrichment scores (Figure 70A), enrichment of tumor cell sub-

cluster 1 is significantly associated with shorter recurrence-free-survival (hazard ratio 1.37, p = 0.002). The leading-edge genes in the cluster 1 protein activity signature most enriched in patients with recurrence vs non-recurrence are reported in Figure 70B and include KIF14 as well as TOP2A, both previously noted as biologically significant markers of clusters 1 and 2.

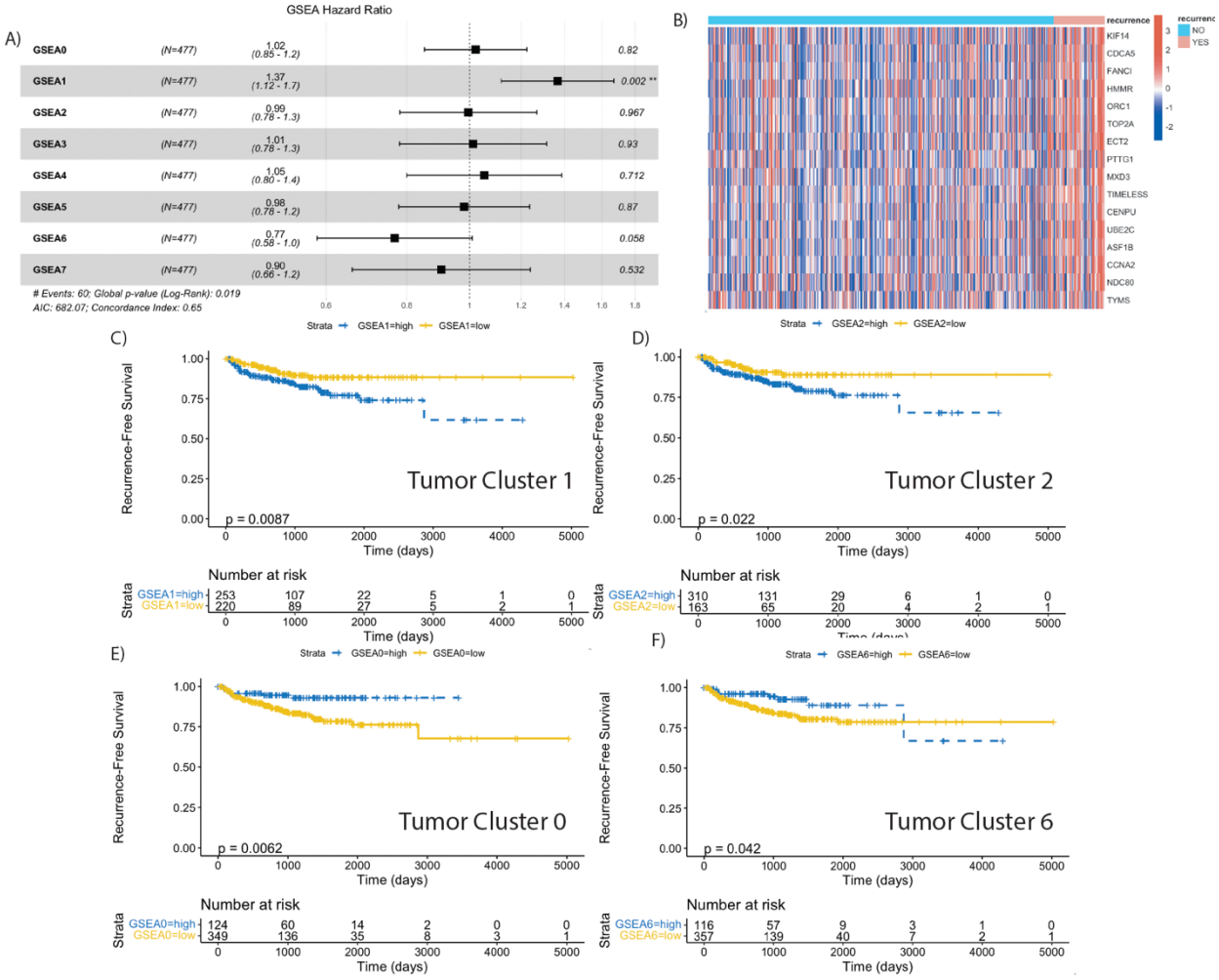


Figure 70: Tumor Single-Cell Sub-Cluster Signatures Associate with Differential Outcomes in TCGA.

A) Forest plot of Cox regression hazard ratios testing association in TCGA of patient-by-patient Normalized Enrichment Score for each tumor sub-cluster gene set with Recurrence-Free survival. Cluster 1 gene set enrichment is significantly associated with worse survival outcomes (p = 0.002). B) Heatmap of Leading-Edge Gene Set from Cluster 1 comparing all Recurrent vs Non-Recurrent patients in TCGA. C) Kaplan-Meier curve testing association of binarized cluster

1 gene set enrichment (greater than 0 = high, less than 0 = low) with recurrence-free survival in TCGA, such that cluster 1 enrichment significantly associates with worse recurrence-free survival. D) Kaplan-Meier curve testing association of binarized cluster 2 gene set enrichment (greater than 0 = high, less than 0 = low) with recurrence-free survival in TCGA, such that cluster 2 enrichment significantly associates with worse recurrence-free survival. E) Kaplan-Meier curve testing association of binarized cluster 0 gene set enrichment (greater than 0 = high, less than 0 = low) with recurrence-free survival in TCGA, such that cluster 0 enrichment significantly associates with improved recurrence-free survival. F) Kaplan-Meier curve testing association of binarized cluster 6 gene set enrichment (greater than 0 = high, less than 0 = low) with recurrence-free survival in TCGA, such that cluster 6 enrichment significantly associates with improved recurrence-free survival, up to 2800 days. Kaplan-Meier curves are not shown for the remaining clusters as log-rank p-values for these were not statistically significant ($p>0.05$).

Furthermore, log-rank tests on enrichment scores binarized to “high” vs “low” show significant association of both cluster 1 ($p = 0.0087$) and cluster 2 ($p = 0.022$) with shorter recurrence-free survival (Figure 70C-D), and significant association of cluster 0 ($p = 0.0062$) and cluster 6 ($p = 0.042$) with improved recurrence-free survival (Figure 70E-F). No other cluster signature was associated with survival at statistical significance of $p<0.05$. In two smaller datasets specifically profiling metastatic prostate tumors (East Coast Stand Up to Cancer, West Coast Stand Up to Cancer) [189] [190], trends were observed toward association of clusters 1 or 2 with worse overall survival (Figure 71, Figure 72), such that cluster 2 was significantly associated with worse overall survival in East Coast SU2C (Figure 72B), though cluster 1 did not reach statistical significance. However, neither of these datasets included recurrence-free survival or PSA response as clinical metadata.

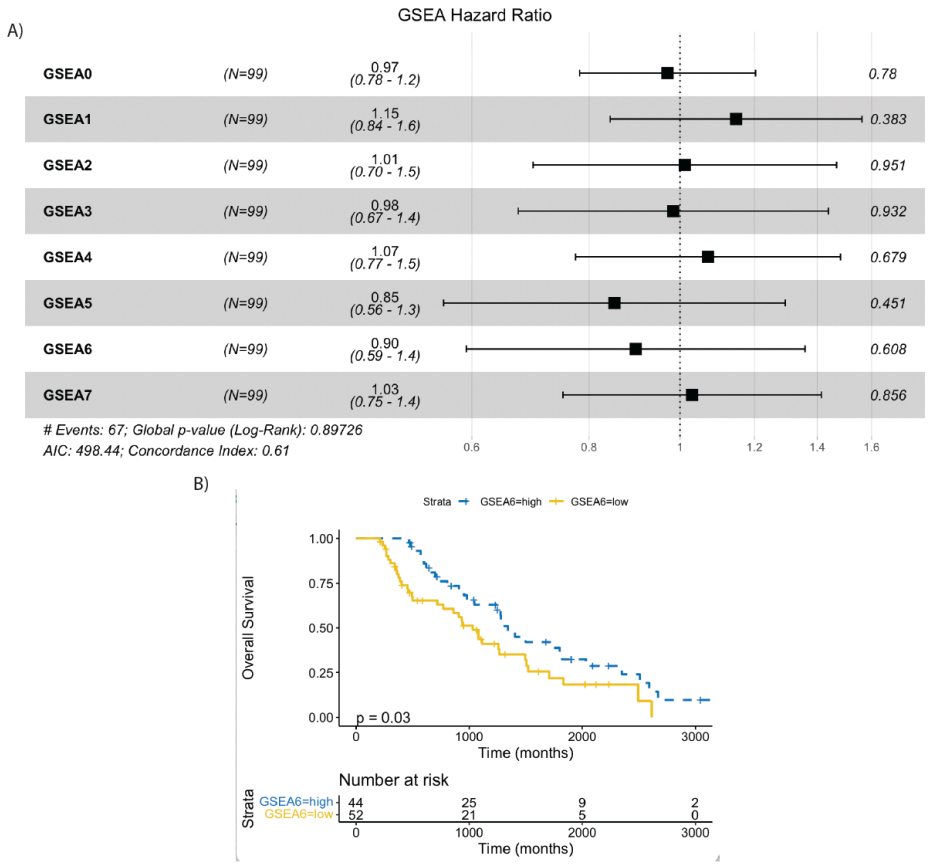


Figure 71: Tumor Single-Cell Sub-Cluster Signatures and Outcome in West Coast SU2C

A) Forest plot of Cox regression hazard ratios testing association in West Coast Stand Up to Cancer (SU2C) dataset of patient-by-patient Normalized Enrichment Score for each tumor sub-cluster gene set with overall survival. B) Kaplan-Meier curve testing association of binarized cluster 6 gene set enrichment (greater than 0 = high, less than 0 = low) with survival, such that cluster 6 enrichment significantly associates with improved survival. Kaplan-Meier curves are not shown for the remaining clusters as log-rank p-values for these were not statistically significant ($p>0.05$).

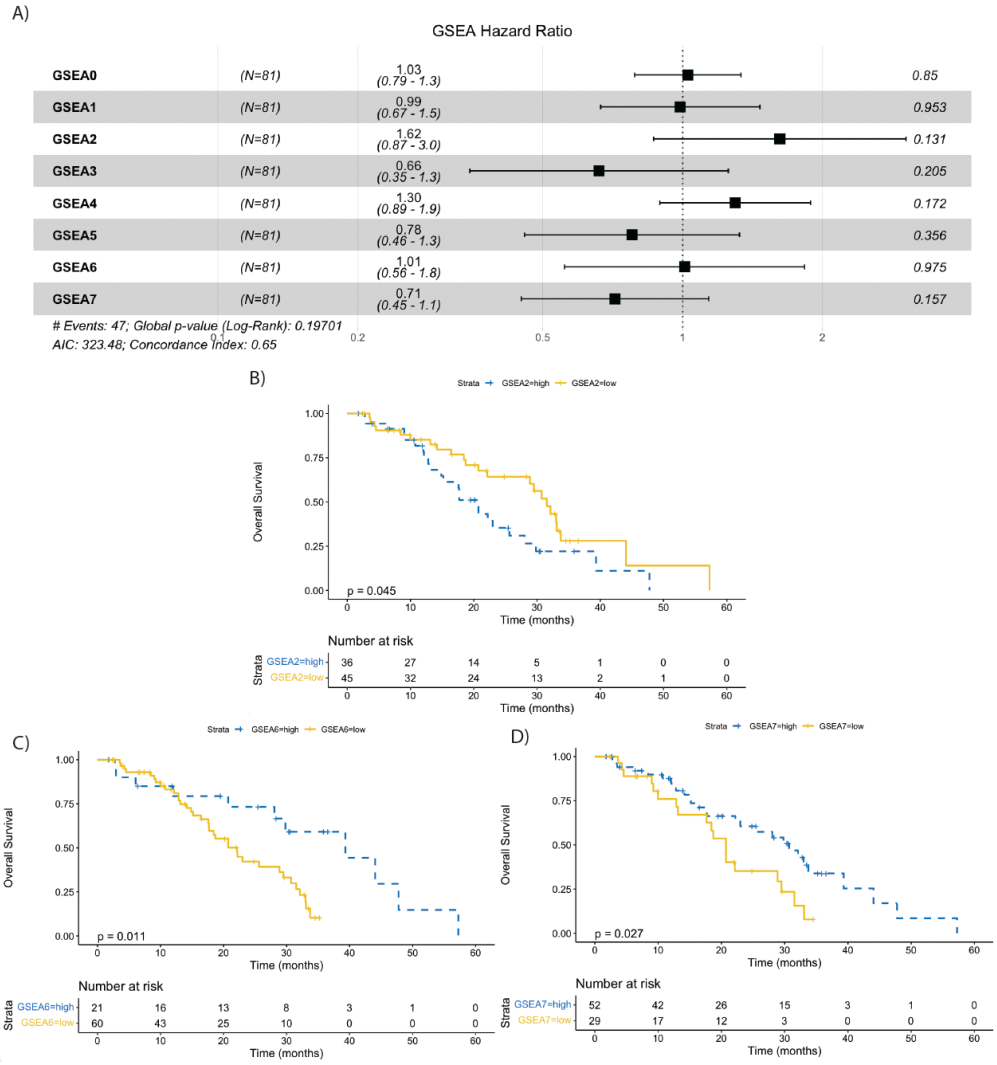


Figure 72: Tumor Single-Cell Sub-Cluster Signatures and Outcome in West Coast SU2C

A) Forest plot of Cox regression hazard ratios testing association in East Coast Stand Up to Cancer (SU2C) dataset of patient-by-patient Normalized Enrichment Score for each tumor sub-cluster gene set with overall survival. B) Kaplan-Meier curve testing association of binarized cluster 2 gene set enrichment (greater than 0 = high, less than 0 = low) with survival, such that cluster 2 enrichment significantly associates with worse survival. C) Kaplan-Meier curve testing association of binarized cluster 6 gene set enrichment (greater than 0 = high, less than 0 = low) with survival, such that cluster 6 enrichment significantly associates with improved survival. D) Kaplan-Meier curve testing association of binarized cluster 7 gene set enrichment (greater than 0 = high, less than 0 = low) with survival, such that cluster 6 enrichment significantly associates

with improved survival. Kaplan-Meier curves are not shown for the remaining clusters as log-rank p-values for these were not statistically significant ($p>0.05$).

Taken together, these results are highly concordant with the finding in our own data that cluster 0 enrichment at baseline associates with better treatment response while clusters 1 and 2 associate with worse treatment response. This suggests that while we are able to dramatically increase intra-tumoral CD8 T-cell infiltration by combination of anti-PD1 with ADT in metastatic prostate cancer, patients may further benefit from targeting an independent treatment resistance mechanism mediated by the TOP2A+ tumor cell sub-clusters we have identified.

7.4 Discussion

Evaluation of primary prostate cancer and metastatic, castration-resistant disease using high-throughput, transcriptomic sequencing [176] [182] [191] [192] [193], has shown that the TME is relatively immune-depleted. We used scRNAseq to comprehensively characterize the TME of metastatic, hormone-sensitive prostate cancer (mHSPC), across a variety of tissue types. Using longitudinal samples from 10 patients over a treatment course with ADT and anti-PD-1 antibody, we describe the baseline TME and tumor cells, the specific changes induced with treatment, and associated baseline features with PSA response. We leveraged our expertise in inferred protein-activity computational methods to increase resolution of the immune and tumor cell subpopulations as compared to conventional gene-expression and transcriptomic methods.

Profiling transcriptomes from a cumulative 40,270 single-cells, our study uncovered a previously unknown rich immune infiltrate in untreated metastatic, hormone-sensitive tumors. In our analyses the baseline bone, lymph node, and liver samples were similarly immune infiltrated

while the lung metastasis was relatively immune-depleted (Figure 62B-C) and distinctly different, this with the caveat of a small sample size. These data add to the notion that pulmonary-tropic and non-pulmonary metastatic mHSPC may be biologically and/or immunologically fundamentally different [194]. Regarding prostate tumor cells, our data adds to the body of literature demonstrating intra- and inter-patient tumor cell heterogeneity [195] and expands this knowledge by highlighting the different frequencies of the eight tumor-cell sub clusters (Figure 66A), using protein-based activity, between tissue types (Figure 66B-E). Furthermore, we show phenotypic changes in tumor cell types induced by treatment, associate baseline tumor cell phenotypes with clinical response, and define pathways enriched longitudinally and upon recurrent, progressive disease.

Our study is limited regarding the total number of patient samples per tissue type and analysis of a single metastatic site per patient. However, the longitudinal analysis of prostate cancer tumor metastases using single-cell sequencing over a course of treatment, to our knowledge, has never been reported. Our analyses are potentially biased towards more aggressive biology given that only patients with evaluable disease at the time of on-treatment biopsy were able to safely undergo metastatic biopsy. As such, we are unable to comment regarding the on-treatment changes in the TME and tumor cell profiles of participants who were rapidly responding to therapy. The time points for on-treatment biopsies are fixed due to the nature of a clinical trial and are not based on tumor kinetics although may be an option in future studies. Here we comment on cell types that were represented across all tissue types to avoid analyzing subpopulations that may be less relevant to tumor-immune crosstalk given their expected presence in a specific metastatic niche, *i.e.* common progenitor cells in the bone marrow and B

cells in the lymph nodes. This approach highlights broad changes in transcriptional program across tissue types in lieu of a deep dive on tissue specific idiosyncrasies.

The observation that treatment induces changes in both the TME cellular composition and transcriptional program of tumor cells, so called lineage plasticity, is consistent with other studies in prostate cancer [196] [197] [198]. However, the TME changes we observed with administration of ADT in mHSPC are opposite of those described in primary prostate cancer. In primary prostate cancer, an immune infiltrate rich in T cells invades the TME after ADT administration [199] [4]. In mHSPC biopsy samples we observed a decrease in CD4 and CD8 T cells after ADT, whereas the combination of ADT and anti-PD-1 immunotherapy was effective at recruiting CD8 effector T cells. It is possible that these observed differences between primary and metastatic HSPC are due to the baseline TME composition, i.e. an 'immune desert' vs. 'immune replete' respectively, and immunomodulatory factors already present in the milieu.

Importantly, in our study, we observed significant increases in TNF+ CD4 non-T reg T cells (CD4 1), LAG3+ CD8 T cells (CD8 2), and GITR+ T regs (Treg 3) after combination therapy. This highlights the notion that combination therapy with ADT and anti-PD-1 therapy in men with mHSPC is an immunologically active combination, even in bone metastasis. Lymphocyte activation gene-3 (LAG3, CD223), a CD4 homologue that binds to MHC class II [183], is upregulated on CD8 T cells after antigen experience and represents an 'exhausted' state, which negatively regulates their activation and homeostasis [200] [201]. Dual-inhibition of PD-1 and LAG-3 was recently shown to improve outcomes in patients with melanoma in a large phase 3 clinical trial [202]. In prostate cancer models that are resistant to single-agent PD-1, dual

blockade of PD-1 and LAG3 improved vaccine efficacy providing evidence that combination immune checkpoint therapy may be key to improving clinical response outcomes [203]. Taken together, these data suggest that LAG3 may be a potential adjunct to combination immune checkpoint therapy in prostate cancer. Glucocorticoid-induced TNFR-Related (GITR) protein, an immune checkpoint receptor, belongs to Tumor Necrosis Factor Receptor Superfamily (TNFRSF). GITR is preferentially expressed on CD8 and T reg cells and agonistic antibodies are shown to potentiate the former and reduce functionality of the latter [204] [205] [206]. Although several preclinical and early phase studies have shown that anti-GITR agonist antibodies are safe, clinical results have been modest thus far [207] [208]. Trials of dual immune checkpoint blockade targeting GITR and PD-1 have shown slight advantage over single-agent anti-GITR agonists antibodies [209] [210], although data in mHSPC is limited. Tumor Necrosis Factor (TNF), a major inflammatory cytokine with signaling potential both as a membrane-bound protein and as a soluble ligand, was initially implicated as an anti-tumor cytokine but has since been implicated, in complete contraindication to its name, in tumor progression [184] [185] [211]. Given the ever-increasing number of patients treated with combination immune checkpoint therapy, more patients are developing immune-related adverse events (irAEs) that frequently require treatment with immunosuppressive therapies like anti-TNFa inhibitors. As such, much has been learned about the effects of anti-TNFa inhibitors in patients with cancer [185]. In several studies aimed to abrogate irAEs upfront using concomitant TNFa inhibitors with combination immune checkpoint blockade, also showed improvements in anti-tumour efficacy [186] suggesting a role for TNFa blockade as an anti-tumor agent. Our group has shown that elevated TNFa levels are indeed associated with PSA progression in men with biochemically recurrent prostate cancer [212]. Indeed, in Myc-CaP androgen-sensitive murine models, TNFa

signaling signatures were observed to be upregulated following ADT. Additionally, intratumoral TNFa was upregulated in microdissected human mCRPC lesions [89], and furthermore associated with failure to respond to enzalutamide. Taken together, these data in the context of increased protein activity of TNFa in CD4 T cells after ADT and anti-PD-1 therapy shown here (Figure 63, 67), suggest that inhibiting TNFa, concurrently or sequentially, with ADT and immune checkpoint therapy in hormone-sensitive prostate cancer may be an effective treatment combination.

Due to the tropism of metastatic prostate cancer to bone, tissue-specific changes occurring in the TME of bone samples were of particular interest. Notably, following ADT alone, we observed a relative increase of myeloid cells (Figure 64A-B) and a noticeable decrease in CD4 non-T reg T cells as well as tumor cells. As discussed earlier, the observed expansion of CD8 T cells was most pronounced in the bone TME after combination therapy compared to other tissue types (Figure 64B). This highlights the notion that moving away from a 'one-therapy-treats-all' treatment paradigm and towards more precision-based, targeted, and tissue-based algorithms is likely on the horizon.

Given that our transcriptomic data comes from a prospective clinical trial with close monitoring of clinical response, associations between baseline biopsy features and PSA response was possible. Thus far, the primary endpoint of our phase 2 trial, the rate of undetectable PSA (< 0.02 ng/dL) at 37 weeks after combination therapy, is 42%. This compared to 32% in the ADT plus docetaxel arm of the phase 3 CHAARTED trial (E3805) [213]. Although the proportional difference in the rates of undetectable PSA between our small phase 2 study and CHAARTED

are not statistically significant, due to current statistical power and enrollment, we remain encouraged. It is, however, somewhat perplexing that the rate of undetectable PSA is not overwhelmingly lower given the robust influx of CD8 effector T cells we observed with combination therapy.

New observations from our data sheds light on issues using classically defined immune markers as predictive biomarkers. As previously noted, the association of baseline immune sub-clusters with PSA response showed that TNFa+ CD4 T cells (CD4 1) were statistically significantly associated with late PSA progression, and also that LAG3+ CD8 T cells (CD8 2) and GITR+ T reg cells (Treg 3) were both statistically significantly associated with early PSA response. In the case of Tregs this is somewhat paradoxical. T regulatory cells are classically considered to be an immune suppressive subpopulation of T cells. However, a recent analysis of Tregs using high-dimensional flow cytometry from an NPK-C1 transplantable prostate tumor model revealed significant phenotypic diversity within Tregs, including a Treg subpopulation enriched in regressing tumors [214]. Thus, our data supports this preclinical observation that 'favorable' Tregs may be present in the TME at various states of treatment pressure. Further, taken together, these data demonstrate the capacity of single-cell and high-dimensional data to provide more granularity on immune subpopulations and may even challenge the definitions of classical 'immunosuppressive' or 'immune effector cells.'

We observed both expected and unexpected findings upon association of baseline tumor subclusters with PSA response. Subcluster 0, which has high protein activity levels of two androgen receptor (AR) regulated proteins (TMPRSS2 and NKX3-2), was, not surprisingly,

associated with early PSA response (Figure 66F). Using the 'hallmarks of cancer' pathway enrichment analysis[37], we confirmed that subcluster 0 was indeed enriched with androgen response genes (Figure 67). We confirmed this positive association between this gene set and clinical outcomes in the TCGA dataset using GSEA (Figure 70E). In contrast, subclusters 1 and 2, defined by upregulation of E2F targets, Myc targets, and G2M checkpoint (Figure 67), were associated with late PSA progression (Figure 66F). These associations were validated in a similar manner (Figures 70C-D, 72). Interestingly, we did not observe tumor subcluster 3, the subcluster enriched with TNFa signaling genes (Figure 67), to be associated with worse outcomes (Figure 66F). This association is opposite to what we observed in the TNFa+ CD4 T cells (CD4 1), which was associated with decreased clinical outcomes. Taken together, these observations highlight the notion that the specific cell of protein activity and/or expression (CD4 T cells vs. tumor cells) indeed matters and is a reflection of the underlying immunologic processes at play.

Because certain tumor subclusters were associated with PSA progression or prominent on disease progression, we used OncoTarget (see methods) to examine potential druggable proteins active in the tumor subclusters. Subclusters 1 and 2 showed high activity of TOP2A and low AR activity suggesting these subclusters may be upfront resistant to AR-targeted therapies (Figure 69). Although subclusters 1 and 2 were associated with PSA progression in our dataset (Figure 66F), and validated as such in external datasets, they were not the dominant tumor subclusters seen in the tumor progression biopsy (Figure 66B). More metastatic biopsies at the time of tumor recurrence would be helpful to delineate if there is a role for targeting tumor subclusters 1 and 2 with topoisomerase inhibition in mHSPC. Targeting these tumor cell populations upfront to eradicate them prior to combination ADT, anti-PD-1 therapy may be one approach for future

clinical trials. Of interest particular interest was tumor subcluster 5, the dominant cluster seen in the bone progression biopsy. Figure 69 shows that potential druggable proteins in subcluster 5 include PRKACB (cAMP-dependent protein kinase catalytic subunit beta), MMP14 (matrix metallopeptidase 14), and HIF1a (Hypoxia Inducible Factor 1 Subunit Alpha). Taken together, our combined dataset and analyses highlight several potential targets for worthy of further investigation in mHSPC.

Building on this rich longitudinal transcriptomic dataset, we propose that the "holy grail" treatment for men with mHSPC will require a multi-pronged and adaptive combination regimen to elicit complete and durable responses. We propose a term 'Highly Active Anti-Tumor Therapy' (or HAATT), that includes a treatment regimen with the following properties: 1) strong upfront tumor cell killing perhaps directed at known resistant tumor cell clones, 2) activation of CD8 effector T cell function via combination immune checkpoint therapy (anti-PD-1 with anti-LAG3, anti-GITR agonist antibodies, or novel agents targeting costimulatory agonists like 4-1BB or B7-H3), 3) depletion or blocking of T regulatory cells, and 4) targeting immunosuppressive or tumor-permissive molecules in the TME (i.e. cytokines, chemokine receptors, metabolomic pathways, or transcription factors). As demonstrated, it is imperative to review the underlying tumor immunology and biology continuously and critically with advanced techniques when conceiving of the next biologically plausible clinical trials.

7.5 Methods

Study design and participants

The PRIME-CUT (modulating the Prostate cancer Immune Microenvironment with Chemoimmunotherapy for metastatic prostate cancer) study is an open-label, single-arm, phase 2 study (NCT03951831) conducted at Columbia University Irving Medical Center (New York, NY) that tests the clinical activity of phased administration of ADT, anti-PD-1 therapy, and docetaxel in men with newly diagnosed metastatic, hormone-sensitive prostate cancer. The study was approved by the institutional review board and all participants provided written consent. Key eligibility criteria included diagnosis of metastatic, hormone-sensitive prostate cancer with a robust testosterone level (>150ng/dL). Prior ADT was allowed but not if within preceding six months. Recruitment was restricted to patients with metastatic lesions amenable to biopsy. Patients with bone metastases were allowed. Treatment consisted of ADT (degarelix 240mg SC one dose, followed by leuprolide 22.5mg SC every 12 weeks) followed by anti-PD-1 antibody (cemipliamb-rwlc 350mg IV every 3 weeks) beginning four weeks after ADT initiation. To prime the immune system, a two-cycle lead-in of anti-PD-1 therapy is administered prior to docetaxel (75mg/m^2 every 3 weeks for six cycles). Participants receive ADT and anti-PD-1 antibody until study completion (52 weeks) or until lack of clinical benefit or intolerable side effects. The primary endpoint is the rate of undetectable prostate specific antigen (PSA) at 6 months after chemotherapy initiation (37 weeks on-study) and will be compared to the historical control (i.e., CHAARTED) [213]. Secondary endpoints include time to progression to CRPC and rate of radiographic response upon study completion. To ensure patient safety within the limitations of a small phase 2 study, toxicity is monitored using a Bayesian method which provides continuous monitoring boundaries for termination of the trial if the toxicity rate is unacceptable [215].

Sample collection

Twelve patients were enrolled on the phase 2 trial from May 2019 through December 2020. At enrollment all patients were scheduled for an interventional radiology-guided baseline biopsy of the most accessible metastasis. Subjects were randomized to one of two time points, either week 4 or week 10 on-study, for an on-treatment biopsy as well. Subsequent biopsies at the time of disease progression were optional. In all cases, patients had their on-treatment biopsy at the same location as their baseline biopsy. Patients randomized to the week 4 time point have been treated with four weeks of ADT (degarelix) alone. Patients randomized to the week 10 time point have been treated with 10 weeks of ADT (4 weeks of degarelix and 6 weeks of leuprolide) as well as two cycles (6 weeks) of anti-PD-1-therapy. We subsequently refer to these time points as 'ADT only' and 'ADT+aPD1,' respectively. Given the phased administration of ADT and anti-PD-1 therapy and the differing time points for on-treatment biopsies, these data present a unique opportunity to comprehensively define and compare the specific treatment effects of ADT only and ADT+aPD1 on the transcriptional program of immune cell subpopulations and tumor cells in the TME across a variety of tissue types.

Of the twelve patients enrolled, two patients' samples baseline and on-treatment samples were excluded from these analyses due to either an insufficient number of viable cells for loading onto the 10X Genomics instrument or a lack of tumor cells identified in the biopsy sample using copy number inference (see methods below). Here, we report on 10 patients' samples from bone, lymph node, liver, and lung metastases. We recovered an adequate number of cells in both baseline and on-treatment biopsy samples from six patients (four patients with bone metastases, one patient with lymph node metastases, and one patient with lung metastases). In four patients,

only one of the two samples yielded adequate cells for sequencing and analysis (one baseline lymph node sample, one baseline liver sample, and one on-treatment bone sample [ADT only], and one on-treatment [ADT+aPD1] liver sample (Figure 59**)**.

Tissue Dissociation

Fresh tumor was minced to 2-4 mm sized pieces with micro-scissors. For bone metastases, minced tissue was resuspended and examined under microscope. If already found to be dissociated to a single-cell suspension, the entire sample was passed through a 70-um filter for downstream processing. For bone metastases not found to be dissociated and for all other tissue sites, tissue was digested with 2.5mL of in-house digestion medium (collagenase IV 800 U/ml, DNase 0.1 mg/ml in L-15 medium with 1g/L glucose, 5% FBS, 15 mM HEPES) in glass bottle with stir bar in 37C water bath for 30min. Digested tissue was passed through a 70-um filter. If pink, single-cell suspensions were incubated with ACK red cell lysis buffer and washed. All spins were performed at 300g, for 5min, at 4 degrees Celsius. Dissociated cells were aliquoted for single-cell sequencing, loading 10,000-20,000 cells per sample.

Single-cell RNA sequencing and data processing

Samples were processed for single-cell gene expression capture (scRNASeq) using the 10X Chromium 3' Library and Gel Bead Kit (10x Genomics), following the manufacturer's user guide at the Columbia University Human Immune Monitoring Core (HIMC). After GelBead in-Emulsion reverse transcription (GEM-RT) reaction, 12-15 cycles of polymerase chain reaction (PCR) amplification were performed to obtain cDNAs used for RNAseq library generation. Libraries were prepared following the manufacturer's user guide and sequenced on Illumina

NovaSeq 6000 Sequencing System. Single-cell RNASeq data were processed with Cell Ranger software at the Columbia University Single Cell Analysis Core. Illumina base call files were converted to FASTQ files with the command “cellranger mkfastq.” Expression data were processed with “cellranger count” on the pre-built human reference set of 30,727 genes. These data were loaded into the R version 3.6.1 programming environment, where the publicly available Seurat package was used to further quality-control filter cells to those with fewer than 10% mitochondrial RNA content, more than 1,500 unique UMI counts, and fewer than 15,000 unique UMI counts. Gene Expression count matrices for each sample were processed in R using the Seurat SCTransform command to perform a regularized negative binomial regression based on the 3000 most variable genes. Samples were then combined by the Seurat Anchor Integration algorithm. The resulting matrix was clustered by the Louvain Algorithm, with resolution selected automatically to maximize clustering silhouette score, as previously described [35]. Gene Expression data were projected into their first 50 principal components using the RunPCA function in Seurat, and further reduced into a 2-dimensional visualization space using the RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Differential Gene Expression between clusters was computed by the MAST hurdle model for single-cell gene expression modeling, as implemented in the Seurat FindAllMarkers command, with log fold change threshold of 0.5 and minimum fractional expression threshold of 0.25, indicating that the resulting gene markers for each cluster are restricted to those with log fold change greater than 0 and non-zero expression in at least 25% of the cells in the cluster.

Semi-Supervised Cell Type Calling

For each single cell gene expression sample, cell-by-cell identification of cell types was performed using the SingleR package [48] and the preloaded Blueprint-ENCODE reference, which includes normalized expression values for 259 bulk RNASeq samples generated by Blueprint and ENCODE from 43 distinct cell types representing pure populations of stroma and immune cells [49] [50]. The SingleR algorithm computes correlation between each individual cell and each of the 259 reference samples, and then assigns both a label of the cell type with highest average correlation to the individual cell and a p-value computed by Wilcox test of correlation to that cell type compared to all other cell types. Cell-by-cell SingleR labels were restricted to those with $p<0.05$, and unsupervised clusters are labelled as a particular cell type based on the most-represented SingleR cell type label within that cluster. Since tumor cells are not represented within the Blueprint-ENCODE reference, tumor cells are typically assigned as epithelial, since prostate cancer is epithelial in origin. The tumor cell identity of these cells was further confirmed by expression of KLK3, a prostate cancer marker gene, as well as by inference of copy number variations using the InferCNV algorithm with all lymphoid and myeloid cell clusters specified as a copy-number-normal reference.

Regulatory network and protein activity inference

Protein activity was inferred from gene expression at the single-cell level according to the pipeline described [216] and previously used for analysis of single cell ccRCC samples [35]. From the combined dataset, metaCells were computed within each gene expression cluster by summing SCTransform-corrected template counts for the 10 nearest neighbors of each cell by Pearson correlation distance. 200 metaCells per cluster were sampled to compute a regulatory network from each cluster. All regulatory networks were reverse engineered by the ARACNe

algorithm. ARACNe was run with 100 bootstrap iterations using 1785 transcription factors (genes annotated in gene ontology molecular function database as GO:0003700, “transcription factor activity”, or as GO:0003677, “DNA binding” and GO:0030528, “transcription regulator activity”, or as GO:0003677 and GO:0045449, “regulation of transcription”), 668 transcriptional cofactors (a manually curated list, not overlapping with the transcription factor list, built upon genes annotated as GO:0003712, “transcription cofactor activity”, or GO:0030528 or GO:0045449), 3455 signaling pathway related genes (annotated in GO biological process database as GO:0007165, “signal transduction” and in GO cellular component database as GO:0005622, “intracellular” or GO:0005886, “plasma membrane”), and 3620 surface markers (annotated as GO:0005886 or as GO:0009986, “cell surface”). ARACNe is only run on these gene sets so as to limit protein activity inference to proteins with biologically meaningful downstream regulatory targets, and we do not apply ARACNe to infer regulatory networks for proteins with no known signaling or transcriptional activity for which protein activity may be difficult to biologically interpret. Parameters were set to zero DPI (Data Processing Inequality) tolerance and MI (Mutual Information) p-value threshold of 10-8, computed by permuting the original dataset as a null model. Protein activity was inferred by running the VIPER algorithm with all ARACNe networks on the combined SCTransform-scaled and Anchor-Integrated gene expression signature of all single cells from each patient. The resulting protein activity matrix was loaded into a Seurat Object with CreateSeuratObject, then projected into its first 50 principal components using the RunPCA function, and further reduced into a 2-dimensional visualization space using RunUMAP function with method umap-learn and Pearson correlation as the distance metric between cells. Differential protein activity between clusters identified by resolution-

optimized Louvain was computed using Student's t-test, and top proteins for each cluster were ranked by p-value.

Association of cell clusters with tissue site and PSA response

Cell counts per cluster were normalized within each individual sample to cluster frequencies, and subsequent comparisons were made between cluster frequencies in different tissue sites, treatment time-points, and at baseline between patients who responded or did not respond to treatment, as assessed by change in PSA over time (Figure 65). This was done for clusters identified by the Louvain algorithm in combined dataset of all cells, representing the entire tumor micro-environment. Separately, tumor cell clusters were isolated as a new Seurat object on which principal components and UMAP projection were re-computed from the VIPER-inferred protein activity matrix. These were subsequently sub-clustered by resolution-optimized Louvain algorithm [35]. Differential protein activity was computed for tumor cell sub-clusters by Student's t-test, with results shown in Figure 68, and pathway enrichment within each cluster was assessed by the Enrichr browser tool [217] (Figure 67). Tumor cell counts within each sub-cluster were normalized to the total count of all tumor cells to compare relative frequencies of each tumor cell population at baseline in patients with early response to treatment (defined by reduction to less than 1% of initial PSA within 10 weeks of treatment) vs patients with initial response followed by progression of PSA on-treatment (late progressors). The same was done to compare frequencies of each population by tissue site and by treatment time-point.

Tumor cell sub-cluster OncoTarget analysis and association with outcome in external datasets

Druggable protein activity within tumor-cell sub-clusters was evaluated by the OncoTarget algorithm, in which the log(UMI/million) gene expression of each tumor cell was scaled by z-score against the log(TPM) gene expression of the entire TCGA database as a reference, and VIPER was applied using the set of single-cell ARACNe networks. From the resulting protein activity matrix, a subset of proteins was selected for which a known drug targeting that protein exists in the DrugBank online database [188]. Cell-by-cell protein activity scores for these proteins were transformed into p-values by fitting to the analytical normal distribution and corrected for multiple hypothesis testing within each cell by Bonferonni's method. Corrected p-values were converted to -log10(p-value) for ease of visualization, and the matrix was subsequently subset to proteins with mean -log10(p-value) of at least 5 in any of the identified tumor cell sub-clusters. The resulting druggable protein activity matrix is shown in Figure 69. Further, a protein signature for each tumor cell sub-cluster was defined based on the set of all proteins differentially upregulated in that cluster. Then for each of three independent external prostate cancer bulk-RNA-Seq datasets, (TCGA, East Coast SU2C, West Coast SU2C) [218] [189] [190], enrichment of these protein signatures was assessed as follows. First, the bulk-RNA-Seq dataset was internally scaled by z-score, then VIPER protein activity inference was performed using the single-cell ARACNe networks, and finally enrichment of each tumor sub-cluster signature was determined in each bulk-RNA-Seq sample by Gene Set Enrichment Analysis (GSEA) [104], where genes were ranked by highest to lowest protein activity. The resulting normalized enrichment scores were tested against recurrence-free-survival time in TCGA or overall survival time in SU2C by Cox regression (Figure 69A). Since enrichment of tumor cell cluster 1 was found to be significantly associated with shorter recurrence-free-survival in TCGA, the leading-edge genes within the protein signature responsible for this association

were further identified by GSEA on the set of all proteins ranked by differential activity in TCGA samples with vs without recurrence. Activity of these proteins for all TCGA samples is shown in Figure 69B. Finally, patient-by-patient enrichment scores were binarized to less than zero = “low” and greater than zero = “high” and assessed for effect on survival by log-rank test and Kaplan-Meier curve, with all statistically significant results shown in Figure 69C-F.

Quantification and Statistical Analysis:

All quantitative and statistical analyses were performed using the R computational environment and packages described above. Differential gene expression was assessed at the single-cell level by the MAST single-cell statistical framework as implemented in Seurat v3, and differential VIPER activity was assessed by t-test, each with Benjamini-Hochberg multiple-testing correction. Comparisons of cell frequencies were performed by non-parametric Wilcox rank-sum test, and survival analyses were performed by log-rank test and cox regression. In all cases, statistical significance was defined as an adjusted p-value less than 0.05. Details of all statistical tests used can be found in the corresponding figure legends.

7.6 Preliminary Data from Primary Prostate Cancer

In addition to the above clinical trial profiling response to combination androgen deprivation therapy (ADT) plus checkpoint immunotherapy in metastatic prostate cancer, we are also collecting single-cell RNA-Sequencing data from a separate clinical trial of combination ADT plus anti-CTLA4 checkpoint immunotherapy in high-risk primary prostate cancer. This trial (NEO-RED-P) consists of two separate arms, each aiming to enroll 10 patients, such that

treatment arm A will consist of ADT alone, and treatment arm B will combine ADT with an anti-CTLA4 checkpoint inhibitor immunotherapy (BMS-986218). Patients in both arms will undergo radical prostatectomy for subsequent single-cell RNA-Seq analysis, along with a comparator set of 10 treatment-naïve prostatectomy specimens. Clinical trial design is outline in Figure 73.

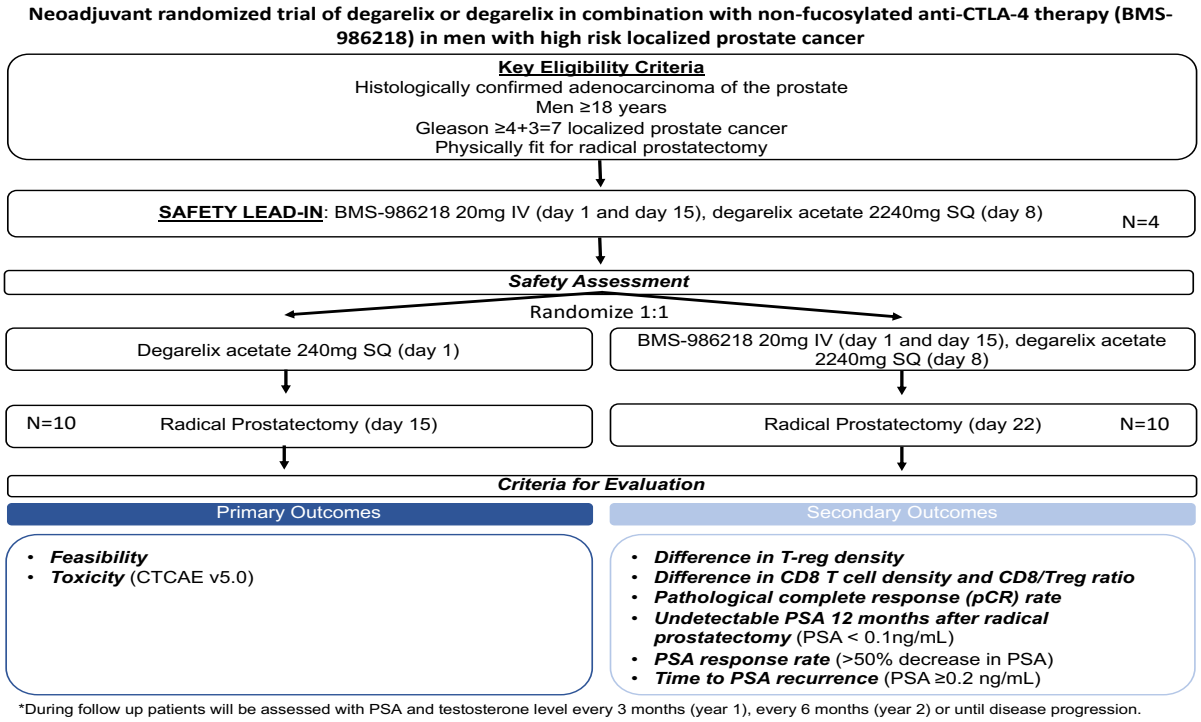


Figure 73: NEO-RED-P Primary Prostate Cancer Clinical Trial Design

The trial opened in 2021 and has currently enrolled and sequenced 4 untreated controls, 6 patients on combination ADT plus aCTLA4, and 2 ADT-only patients, with an aggregate Quality Control filtered cell count of 5,357 cells from untreated controls, 1,397 cells from ADT-only, and 14,684 cells from combination therapy. Results from preliminary comparison of these patients by single-cell protein activity inference analysis are presented below, and the trial is rapidly continuing enrollment, aiming to reach full enrollment over the next year. Ultimately, patients will have prostate-specific antigen (PSA) checked every 3 months for the first year of

follow up or until relapse of disease. For year 2 of follow-up patients will have PSA measured every six months until disease relapse, and tumor micro-environment between responders vs non-responders to therapy will be compared separately in each treatment arm.

Clustering on VIPER-inferred protein activity of cells aggregated across patients reveals a significant immune heterogeneity, with thirty-two distinct clusters of cells identified, and their cell lineage and respective frequencies in untreated vs ADT alone vs combination therapy characterized in Figure 74.

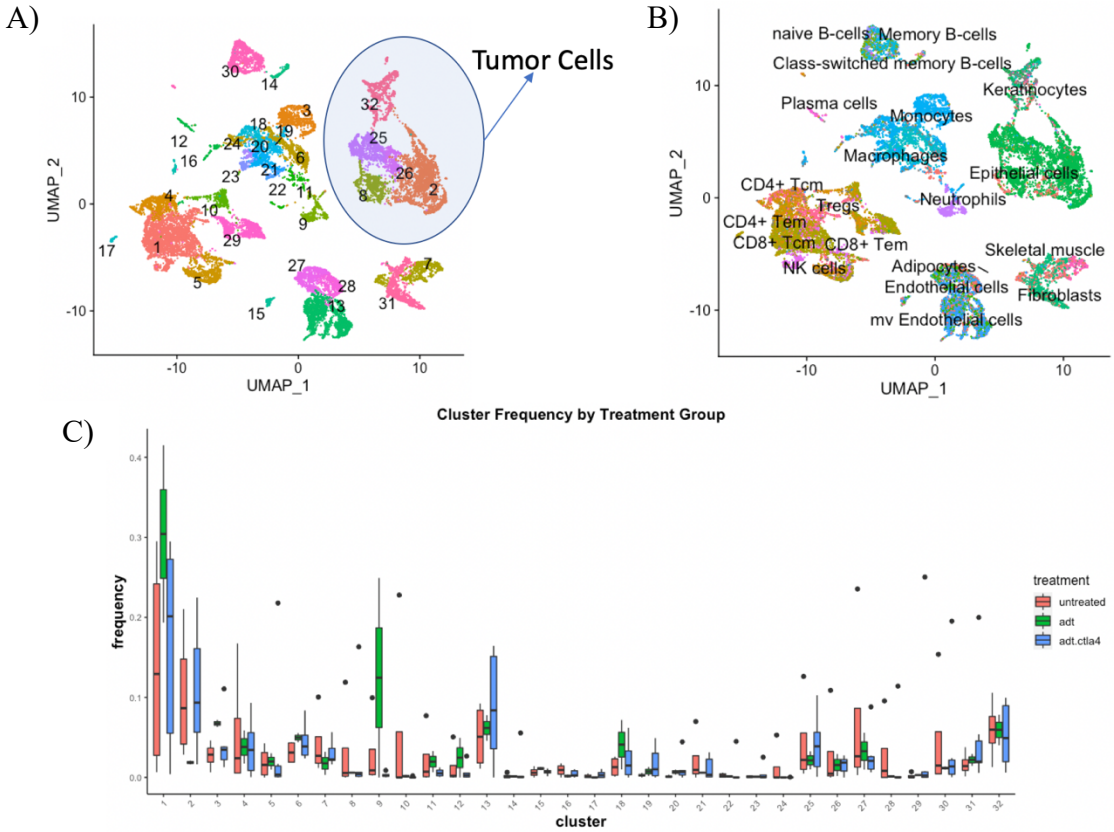


Figure 74: Clustering of all Cells Across Patients on VIPER-Inferred Protein Activity

A) UMAP Projection of unsupervised clustering performed on all cells aggregated across patients. Tumor cells (as defined by Copy Number Variation Inference and expression of tumor marker gene KLK3) are circled. B) SingleR cell type labels overlaid on UMAP plot from A, such

that major cell types represented include Epithelial (Tumor) cells, Myeloid cells (Macrophage/Monocyte), T-cells (CD4, CD8, and Treg), Endothelial cells, Fibroblasts, and B-cells. C) Boxplot of cluster frequencies from A, split by treatment group, comparing untreated patients, patients treated with ADT alone, and patients treated with combination ADT + aCTLA4.

Consistent with our prior work [199], we see few immune cells in primary prostate cancer at baseline, particularly few cytotoxic CD8 T-cells, with diverse subsequent influx of immune cells after ADT. Indeed, we see influx of CD8 T-cells into the prostate tumor micro-environment counterbalanced by influx of immunosuppressive regulatory T-cells (Treg), such that the Treg to CD8 ratio actually increases with ADT versus treatment-naïve baseline (Figure 75). With the addition of aCTLA4 combination therapy, we actually see less overall immune infiltration than with ADT alone. However, we crucially see an inversion of the Treg to CD8 ratio, with increase in CD8 infiltrate relative to baseline and decre B) ı Treg infiltrate (Figure 75). It remains to be seen on clinical follow-up whether this corresponds to generation of a more active anti-tumor immune response, but the changes observed in the micro-environment from this initial patient cohort are encouraging.

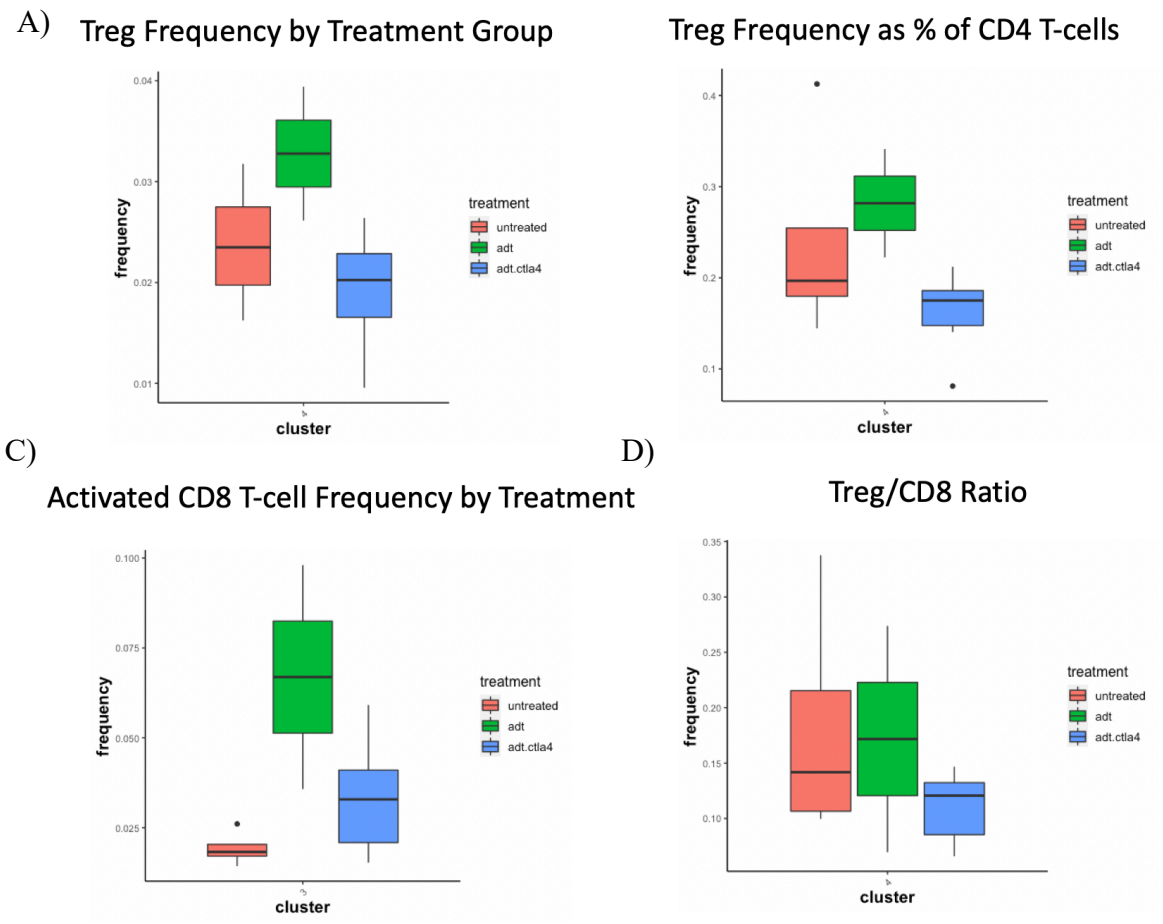


Figure 75: Prostate Tumor Infiltration by Activated CD8 T-cells and Immunosuppressive Regulatory T-cells with ADT versus Combination ADT + aCTLA4.

A) Boxplot showing distribution of Treg frequencies as a percentage of all cells in tumor micro-environment among treatment-naïve patients (red), patients treated with ADT only (green), and patients treated with combination ADT + aCTLA4 (blue). B) Boxplot showing distribution of Treg frequencies as a percentage of total CD4+ T-cells by treatment group. C) Boxplot showing distribution of cytotoxic CD8 T-cell frequencies by treatment group. D) Boxplot showing distribution of Treg/CD8 ratio by treatment group.

Moreover, five distinct sub-clusters of tumor cell phenotypes are observed across patients, with differing representation in different patients (Figure 76A). Although it is too early to correlate prevalence of these tumor sub-phenotypes with differences in clinical outcome among patients in

each treatment arm, we are able to infer significant diversity in activity of druggable proteins among these clusters by OncoTarget (Figure 76B), as well as differences in relative enrichment of oncogenic pathways among proteins active in each cluster (Figure 76C).

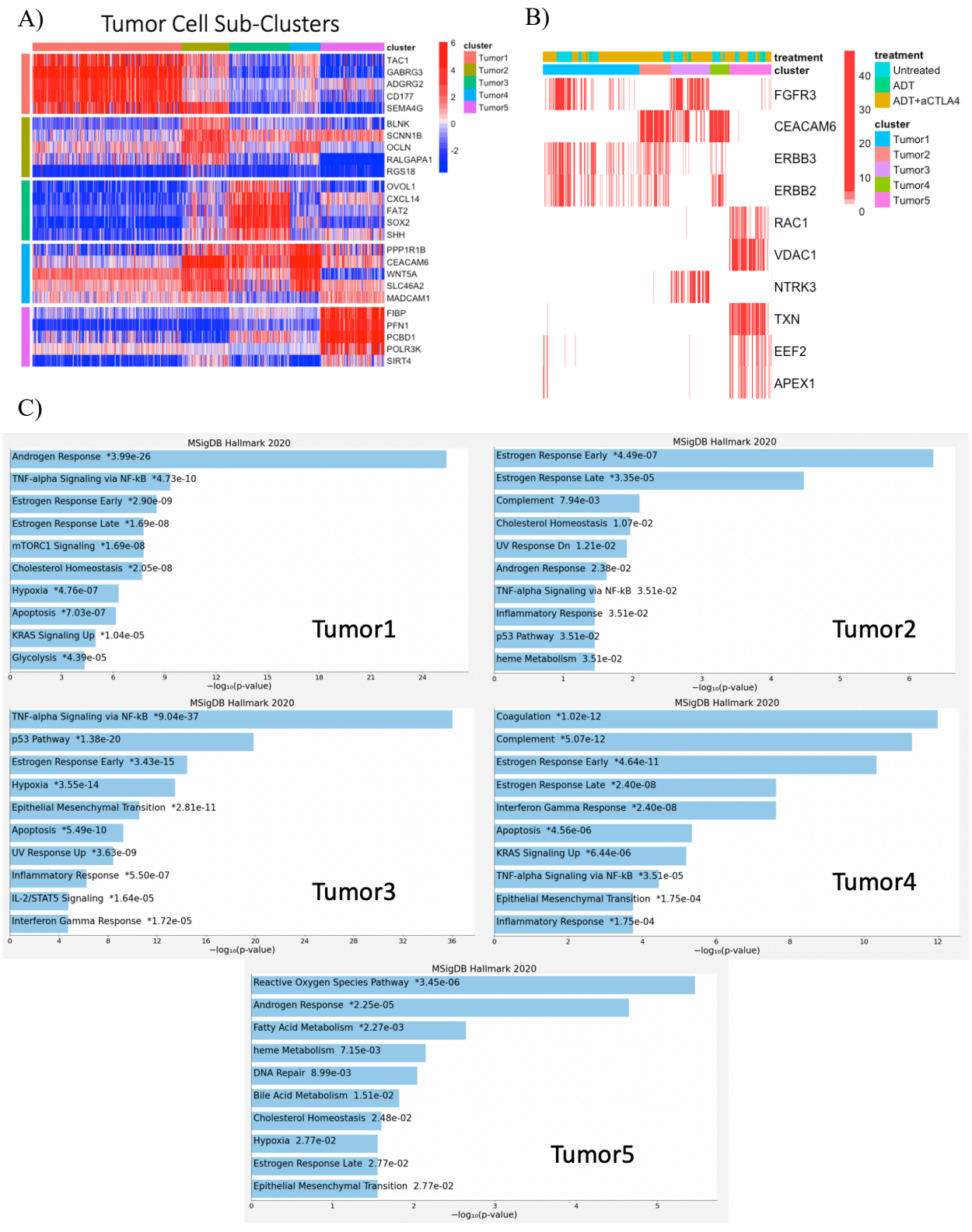


Figure 76: Sub-Clustering of Tumor Cells Reveals Phenotypic Diversity, with Distinct Upregulated Pathways and Druggable Protein Targets.

A) Heatmap of VIPER-inferred protein activity for the top 5 most differentially up-regulated proteins in the five tumor cell sub-clusters observed across patients. B) Heatmap of -log10(Bonferroni-corrected p-values) cell by cell for all druggable protein targets inferred by OncoTarget in any tumor cell cluster with mean -log10(p-value) exceeding 2. C) For each tumor cell cluster in A, plots of the top10 most-enriched pathways among differentially active proteins, from MSigDB Hallmarks of Cancer.

Nevertheless, if we take the single-cell-derived marker proteins of each tumor cell cluster and test for their enrichment patient-by-patient in TCGA by Gene Set Enrichment Analysis (GSEA), we find by Cox regression and by Kaplan-Meier analysis that Tumor cell cluster 2 is enriched among patients with worse clinical outcome, and Tumor cell cluster 4 is enriched among patients with better clinical outcome (Figure 77). The relative favorability of unfavorability of these tumor cell phenotypes with respect to immunotherapy response will be borne out by further study as the clinical trial continues to accrue patients and extend length of follow-up.

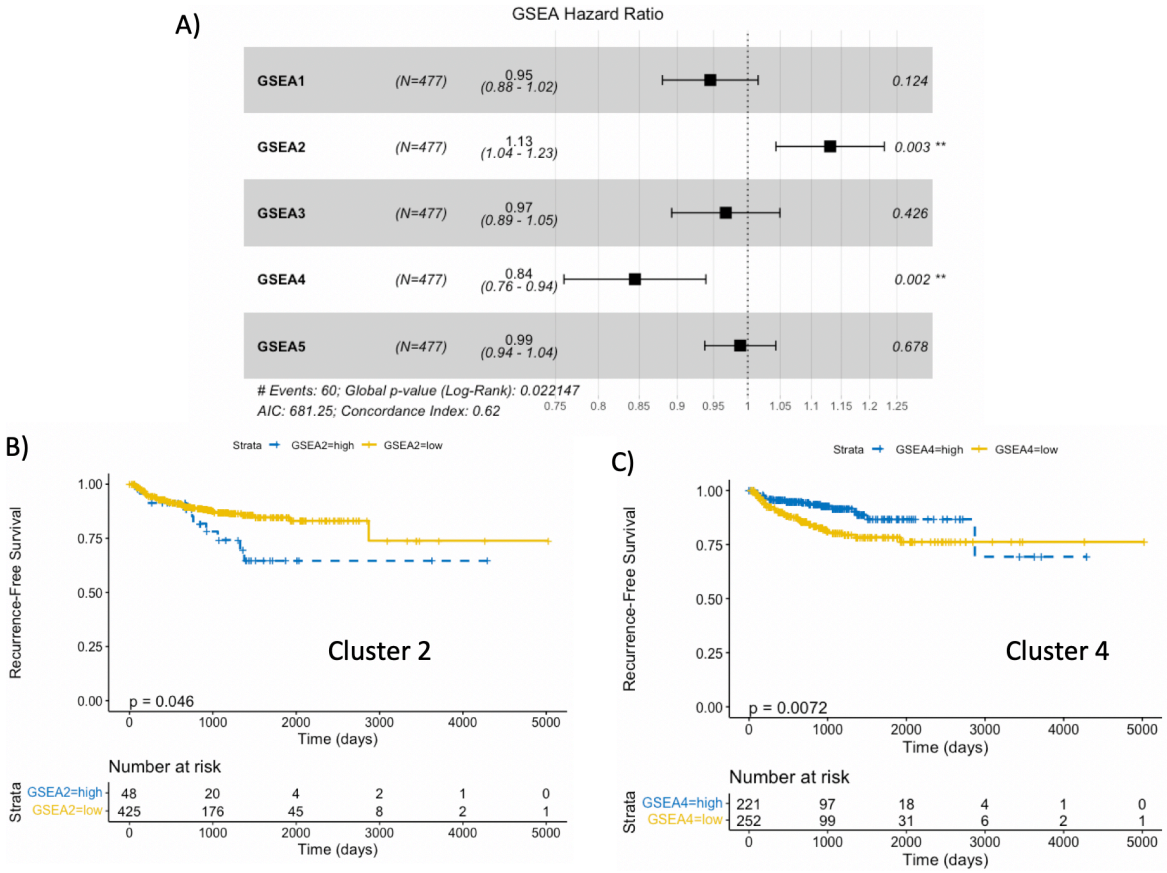


Figure 77: Single-Cell Tumor Sub-Cluster Marker Enrichment Associates with Clinical Outcome in TCGA

A) Cox regression forest plot showing proportional hazards with 95% confidence intervals and p-values for association of each tumor cluster gene set with Recurrence-Free Survival Time in TCGA. Enrichment of Tumor Cell Cluster 2 markers (GSEA2) was found to associate with worse clinical outcome (p=0.003), and enrichment of Tumor Cell Cluster 4 markers (GSEA4) was found to associate with better clinical outcome (p=0.002). B) Kaplan-Meier curve showing association of binarized Tumor Cell Cluster 2 marker enrichment (>0 = "high", <0 = "low") with Recurrence-Free Survival Time in TCGA. Enrichment for Cluster 2 associates with shorter Recurrence-Free Survival Time with log-rank p-value = 0.046. C) Kaplan-Meier curve showing association of binarized Tumor Cell Cluster 4 marker enrichment (>0 = "high", <0 = "low") with Recurrence-Free Survival Time in TCGA. Enrichment for Cluster 4 associates with longer Recurrence-Free Survival Time with log-rank p-value = 0.0072.

Chapter 8: Elucidating and Targeting Proteomic Master Regulators of Tumor-Infiltrating Regulatory T-cells as a Novel Immunotherapeutic Approach

8.1 Summary

Tumor infiltrating regulatory T cells (TI-Tregs) suppress natural and therapeutically stimulated antitumor immunity. While TI-Tregs are highly attractive therapeutic targets across cancers, the main challenge in designing TI-Treg inhibitors is the need to preserve peripheral Treg (P-Treg) function, which is critically required to suppress autoimmunity. To address this problem, we have identified and validated novel regulators of the human TI-Treg phenotype, leveraging a large collection of transcriptional profiles of multiple T cell populations sorted from tumor and blood of patients with several distinct tumor types. We performed VIPER (Virtual Inference of Protein Activity) analysis on this dataset to systematically identify 17 Master Regulator (MR) proteins predicted to determine the TI-Treg phenotype. To identify actionable therapeutics that modulate the TI-Treg MR program, we first performed a systematic *ex vivo* drug screen against human P-Tregs and TI-Tregs, generating perturbational RNA-Seq profiles for an unbiased set of drugs from a panel of 1,554 total compounds which exhibited greatest inhibition of Treg growth ex vivo. We then applied the OncoTreat algorithm to identify three candidate drugs that invert the TI-Treg MR activity signature and exhibit preferential cytotoxic activity on TI-Tregs relative to P-Tregs. We validated the top target, Gemcitabine (Gem), *in vivo* using the MC38 model and found that at low doses Gem inhibits tumor growth in immune-competent but not immune-incompetent mice. Further, Gem synergizes with anti-PD-1 checkpoint blockade and preferentially inhibits a subset of TI-Tregs with high activity of TI-Treg MRs, assessed by

single-cell RNA-Seq. Additionally, we validated the functional relevance of TI-Treg MRs in a pooled *in vivo* CRISPR screen using the CHIME CRISPRko hematopoietic stem cell transplant model[1]. We validated 7 candidate MRs for which knockout inhibits TI-Treg recruitment and/or retention in the TME without perturbing P-Tregs. The top candidate, TRPS1 (Transcriptional Repressor GATA Binding 1), was also confirmed to inhibit TI-Treg tumor infiltration with respect to infiltration by CD4nonTregs. To further validate the biological effect of TRPS1, we generated single gene knockout CHIME chimeras and found that tumor growth was strongly inhibited in mice bearing knockout of this gene in hematopoietic cells. Together, these studies identified a collection of novel TI-Treg MRs with potentially transformative clinical utility as therapeutic targets for the specific inhibition of TI-Tregs, and additionally discovered related small molecule compounds with preferential activity against TI-Tregs. These compounds, particularly low-dose Gemcitabine, are attractive candidates to use in combinatorial strategies to sensitize tumors to checkpoint inhibitors, with potential to improve treatment outcomes across a broad range of tumor types.

8.2 Introduction

To manifest as clinically relevant disease, cancer must evade a complex host-protective immune response, the outcome of which is largely determined by the balance of inflammatory (anti-tumor) and tolerogenic (pro-tumor) immune cell function in the tumor microenvironment (TME). The regulatory T cell (Treg) lineage, characterized by expression of the transcription factor FoxP3, promotes tumor growth and immunotherapy resistance by establishing a tolerogenic TME (Figure 78A). As such, increased Treg infiltration in the tumor is generally correlated with poor prognosis across most human malignancies and increased resistance to checkpoint targeting

immunotherapy [135] [199]. While this makes Tregs attractive therapeutic targets, several factors have prevented clinical translation. First, any Treg-directed therapy must target only tumor infiltrating Tregs (TI-Tregs), while sparing peripheral Tregs (P-Tregs), to avoid catastrophic autoimmunity [219]. Second, from a molecular perspective, Tregs express genes that broadly recapitulate those in activated T cells, thus complicating design of targeting strategies that would spare anti-tumor cytotoxic $CD8^+$ and inflammatory $CD4^+$ T cell function [220]. Unfortunately, the Treg-targeting agents that satisfy these criteria, although effective in murine models, have not been effective in human patients [214] [221] [222], thus highlighting the need for more specific actionable vulnerabilities of TI-Tregs.

Our work improves upon several recent studies that have assembled transcriptional profiles of Tregs from independent cancer patients. De Simone *et al.* profiled Tregs from up to ten non-small cell lung cancer (NSCLC) or colorectal cancer (CRC) [223] patients; Plitas *et al.* sequenced Tregs from tumor, blood, and normal ductal epithelial tissue from a breast cancer cohort [224]; Zheng *et al*. performed single cell RNA sequencing (scRNA-Seq) on T cells from 6 patients with hepatocellular carcinoma (HCC) [225]; finally, Magnuson *et al.* profiled cryopreserved Tregs from 12 CRC patients [226]. Each team sequenced Tregs flow sorted from tumor versus peripheral blood and most groups included FoxP3- CD4+ conventional T cells (Tconv), as additional controls.

One limitation of all of these studies was a failure to include canonically activated Tconv cells as an additional control, as Tregs commonly overexpress markers of T cell activation [220]. We overcame this limitation by sequencing αCD3/αCD28-stimulated CD4 T cells from 8 patients,

thus more rigorously filtering shared markers. The samples profiled here were also patient-matched, thus reducing inter-patient variability. Patients were selected across four tumor types that were not represented in the existing datasets, including: prostate adenocarcinoma, bladder cancer, clear cell renal carcinoma, and glioblastoma. Thus, Treg transcriptional profiles were acquired from both immunotherapy-sensitive and immunotherapy-resistant tumors and across highly disparate tissue contexts (Figure 78B).

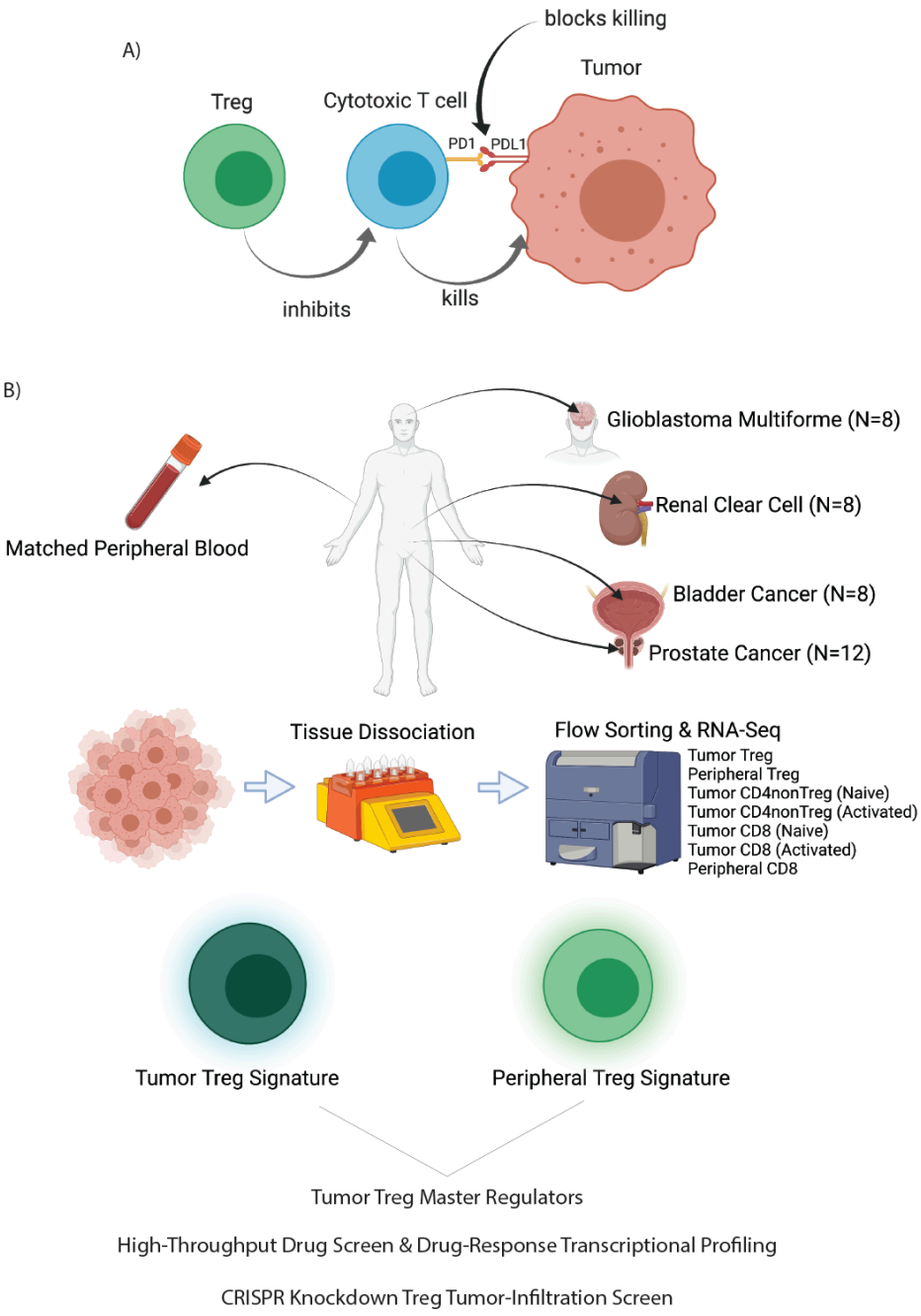


Figure 78: Graphical Abstract and Experimental Design

A) Illustrates immunosuppressive function of Tregs. B) Clinical data collected profiling Tumor vs Peripheral Tregs with additional control populations and outline of downstream analyses and validation experiments.

Given the challenge of elucidating complex mechanisms from scRNA-Seq analysis, due to massive gene dropout effects which result in less than 4,000 genes being sequenced, we first de-noised the scRNA-Seq data and then transformed it into highly reproducible and fully quantitative protein activity profiles, using VIPER (Virtual Inference of Protein-activity by Enriched Regulon analysis) [214]. Since VIPER infers protein activity based on the differential expression of a protein's regulatory targets, akin to a highly multiplexed gene reporter assay, the most differentially active proteins represent candidate MR proteins that are most likely to mechanistically control the transcriptional state of these cells. This approach has never been previously applied to address this question.

To validate candidate MR proteins, we further leveraged an optimized CHIME (chimeric immune editing) platform designed to support *in vivo* CRISPR-Cas9 perturbations in immune competent animals, without requiring artificial adoptive transfer [227]. Finally, we identified and experimentally validated drugs capable of rescuing the signature of P-Tregs in TI-Tregs, thus abrogating the tumor infiltrating phenotype, using the CLIA-certified OncoTreat algorithm.

8.3 Results

Discovery of Treg Tumor-infiltration Master Regulators from Human RNA-Sequencing Data using VIPER: Tumor and matched peripheral blood were collected from 32 patients, including 8 glioblastoma, 8 bladder adenocarcinoma, 8 clear cell renal carcinoma, and 12 prostate adenocarcinoma patients. The following T cell lineages were freshly sorted from each

patient by FACS: TI-Tregs, P-Tregs, peripheral blood CD4 T cells, tumor-infiltrating CD8 T cells, and peripheral blood CD8 T cells. As additional controls for T cell activation, aliquots of peripheral blood CD4 and CD8 T cells were stimulated for 24-hours with anti-CD3/anti-CD28 beads. RNA from each population was isolated and sequenced; in total we have generated a 236-sample transcriptional dataset. Gene expression clustering of these data was extremely noisy (Figure 79A), with only weak separation based on T-cell subtype. By contrast, clustering based on VIPER-inferred protein activity showed clear separation of naïve and activated cells in a reduced principal component space, with tumor-infiltrating cells in-between (Figure 79B). Neither gene expression nor protein activity-based clustering stratified the data according to tumor type, suggesting a rather homogeneous molecular state for Tregs across different tumors. To define MRs with differential activity in TI-Tregs as compared to P-Tregs, naïve CD4, and activated CD4 T cells, we applied a Random Forest algorithm for selection of top discriminative features. This was done by splitting the dataset into 75% training and 25% testing, with test-AUC for a given number of top differentially active proteins compared against test-AUC for a randomly sampled set of control proteins with differential activity p-value = 1.0, converging on a number of features which maximizes AUC while maintaining statistically significant difference compared to AUC from an equivalent number of randomly sampled control proteins (Figure 79C, 79D). This analysis yielded 15 proteins significantly up-regulated in TI- vs. P-Tregs with VIPER activity computed on gene signature normalized against CD4non-Tregs (Figure 79F), with predictive test AUC = 0.982 (Figure 79D). A further 7 proteins were found to be differentially active in in TI-Tregs as compared to all controls, with VIPER activity computed on gene signature normalized against Tumor CD8+ T-cells (Figure 79E), with predictive test AUC = 0.988 (Figure 79C). These gene lists significantly overlap, and the union of the gene sets

comprises 17 putative MRs of Treg Tumor-Infiltration in total, including: EGR1, NR3C1, PBX4, MAFB, ID2, STAT4, NR4A3, NR4A1, TRPS1, EGR3, BANP, ZEB2, KLF4, GLI1, CSRNP2, KDM2B, and FOSL2.

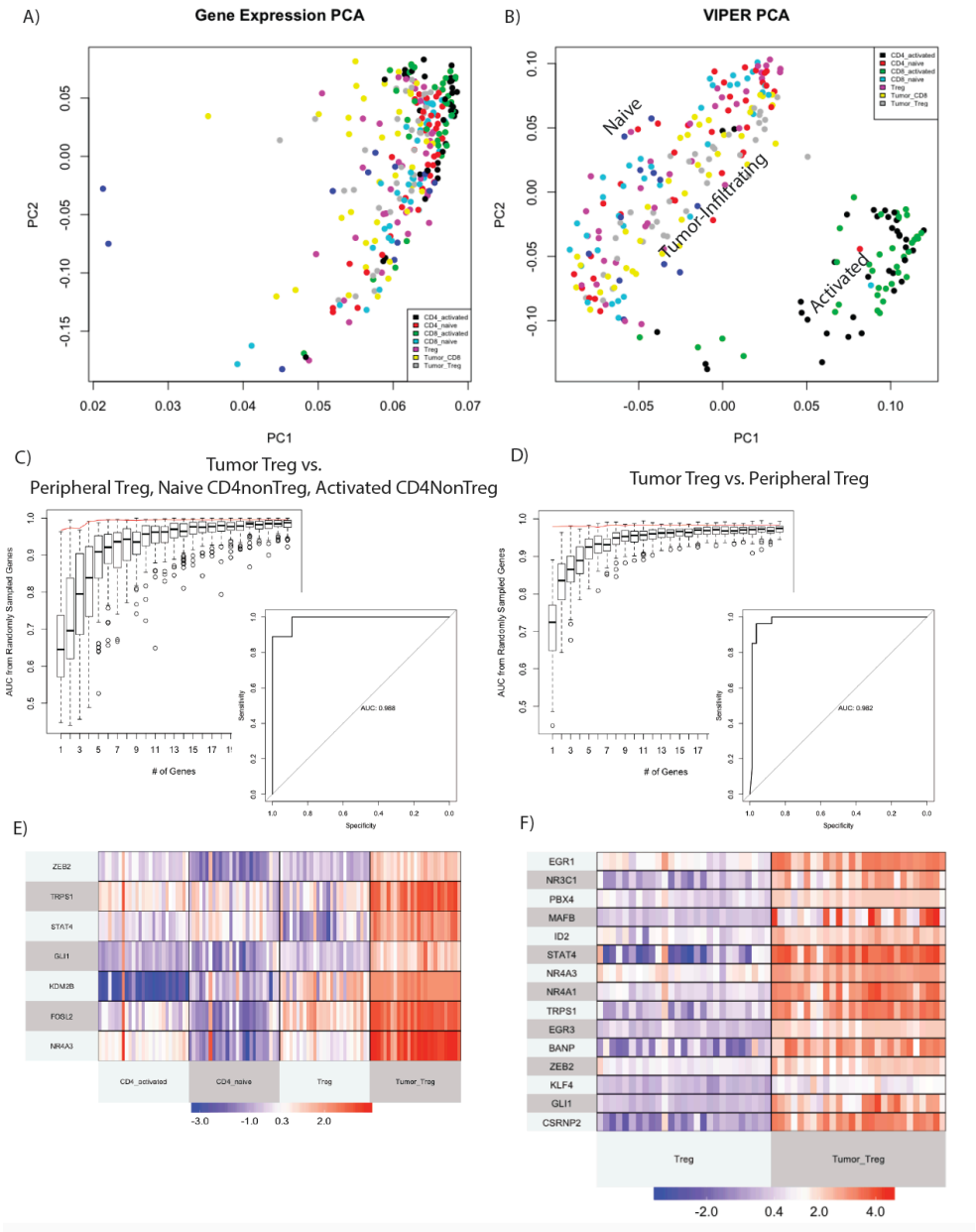


Figure 79: VIPER Enables Definition of Tumor vs Peripheral Treg Master Regulator Signature

A) Principal Component Analysis (PCA) plot of Gene Expression colored by T-cell subtype where black indicates activated CD4nonTregs, green indicates activated CD8 T-cells, red indicates naïve CD4nonTregs, cyan indicated naïve CD8 T-cells, yellow indicates Tumor CD8 T-cells, purple indicates Peripheral Tregs, and grey indicates Tumor Tregs. B) PCA plot of VIPER-inferred protein activity, colored as in A, showing spatial separation of T-cell sub-types not seen in A. C) Random Forest Feature Selection of VIPER Master Regulators Up-Regulated in Tumor Tregs vs Peripheral Tregs, Naïve CD4nonTregs, and Activated CD4nonTregs, normalized against Naïve CD8 T-cells. Boxplot shows distribution of test-AUCs for randomly sampled number of genes corresponding to x-axis, with red line indicating actual AUC of selected Master Regulator gene set. AUC of master regulator gene set for selected number of AUCs is shown in inset to the right. D) Random Forest Feature Selection of VIPER Master Regulators Upregulated in Tumor Tregs vs Peripheral Tregs, normalized against Naïve CD4nonTregs. Boxplot shows distribution of test-AUCs for randomly sampled number of genes corresponding to x-axis, with red line indicating actual AUC of selected Master Regulator gene set. AUC of master regulator gene set for selected number of AUCs is shown in inset to the right. E) Heatmap of VIPER Protein Activity for Master Regulators Selected in C. F) Heatmap of VIPER Protein Activity for Master Regulators Selected in D.

Drug Screening in Human Tregs Identifies Tumor-Treg Directed Therapeutic Candidates: We next performed drug screening on human tumor Tregs to compile a resource of the transcriptional effect of FDA-approved and investigational oncology compounds on TI-Tregs, in an unbiased manner. For these studies, we first performed a single-dose viability screen on flow sorted and *ex vivo* expanded peripheral human Tregs using a library of 1,554 FDA-approved and investigational compounds plated in 96-well format (Figure 80A). From these, we selected 195 compounds that inhibited Treg viability by 60% or more, then performed a secondary dose-response titration with these compounds in order to identify IC20 doses at which 80% of plated peripheral Tregs remain viable at 24 hours of incubation. We next performed a high-throughput PLATE-Seq screen using the top 86 growth-inhibiting compounds at their IC20 concentrations

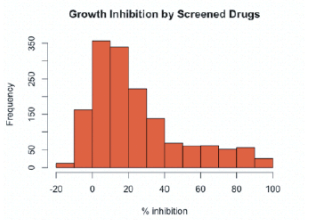
on Tregs flow sorted from a fresh clear cell renal carcinoma tumor and expanded *ex vivo*. The screen was run in duplicate and allowed for measurement of TI-Treg viability in addition to generating transcriptional profiles of Tregs exposed to each compound. (Figure 80A). This screen identified seven compounds with significantly greater growth inhibition on tumor derived Tregs than peripheral Tregs. Notably, each drug was found to significantly inhibit the protein activity of the 17-gene master regulator signature of TI-Tregs (Figure 80B).

A)

Single-Dose (5uM) Growth Inhibition Screen – Human Peripheral Tregs

1554 drugs (Selleckchem FDA-approved + Investigational Oncology)

195 compounds with >= 60% growth inhibition -> Dose-Response Curves to Identify IC20 Doses



E.g., Gemcitabine

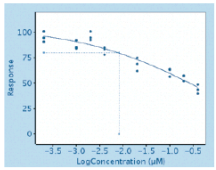


PLATE-SEQ on Tumor Tregs Sorted from ccRCC patients

+ % growth inhibition vs peripheral Treg (at IC20 dose)

Compound	Concentration (uM)	Viability	Viability StdDev
Pralatrexate(Folotyn)	0.2	38.589	0.75882
Triapine	1.5	49.584	5.2708
Malotilate	1.5	51.017	2.0257
Floxuridine	0.5	57.195	4.6382
PKC412	0.07	61.155	4.9233
Gemcitabine (Gemzar)	0.01	64.505	7.1326
BI2536	1	69.218	3.5761

B)

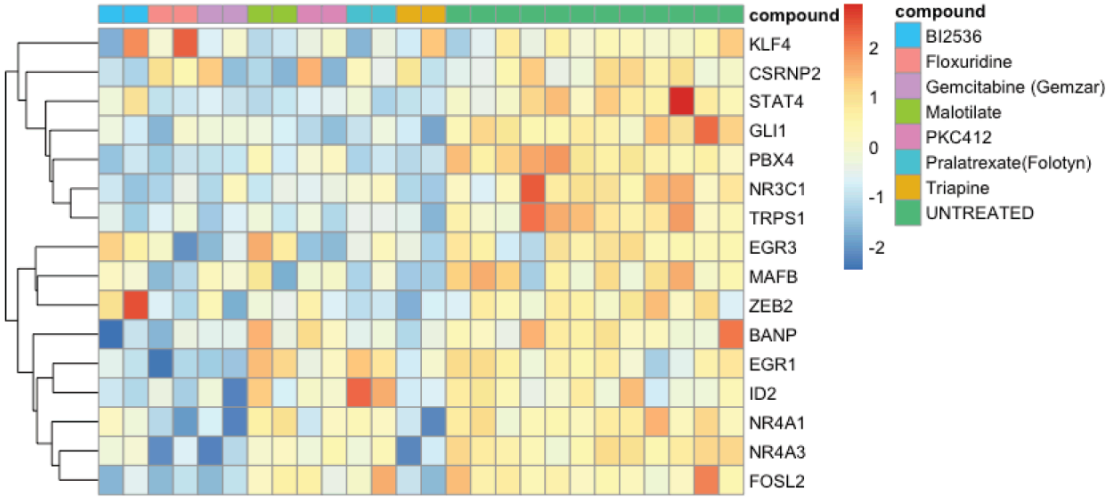


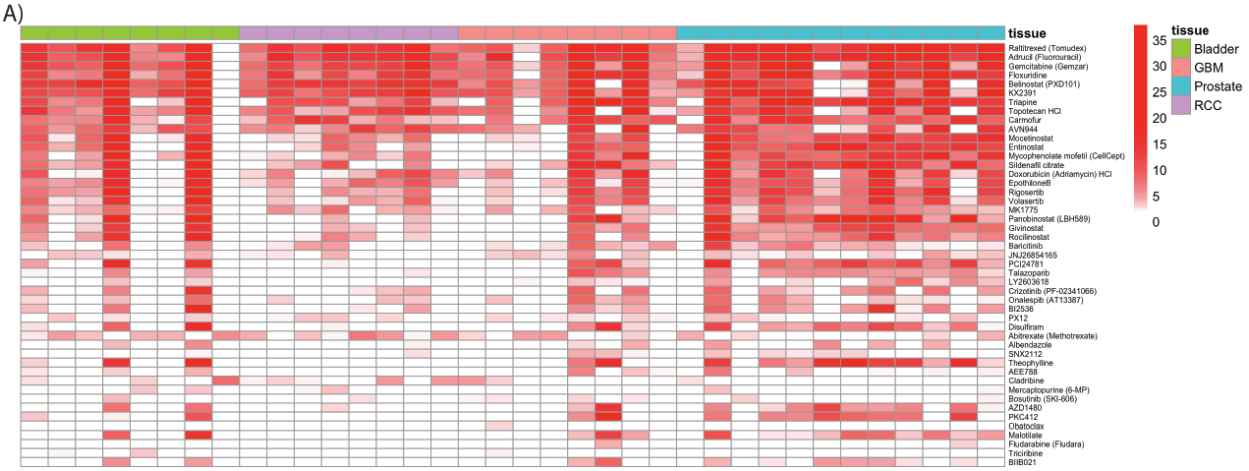
Figure 80: High-Throughput Drug Screening Platform Identifies Potential Drug Candidates with Tumor-Treg Directed Toxicity

A) Experimental design of High-Throughput Treg-Directed Drug Toxicity Screen, such that an initial set of 1,554 FDA-approved and investigational Oncology compounds are screened at single-dose for peripheral Treg growth inhibition, then 195 compounds with >60% inhibition at 5uM are dose-titrated to define IC20 dose, and subsequently assessed for growth inhibition on sorted Tumor Tregs at peripheral-Treg IC20 dose, and effect on transcriptional state by PLATE-

Seq. Table of all seven drugs found to have significantly higher toxicity in tumor Tregs relative to peripheral Tregs, and their respective in vitro doses. B) Heatmap of VIPER protein activity for Tumor vs Peripheral Treg MRs defined in 79E, 79F comparing transcriptional effect of drugs in A vs untreated control, with downregulation of nearly all identified Master Regulators by these drugs.

In a parallel analysis, we utilized the CLIA-certified OncoTreat algorithm to rank each drug for its ability to invert the TI-Treg signature in our original human Treg transcriptional dataset, in a patient-by-patient manner (Figure 81A).

A)



B)

Drug	-log10(Bonferroni-Corrected p-value)	Concentration
Raltitrexed (Tomudex)	16.0120428	10uM
Gemcitabine (Gemzar)	14.0907063	0.01uM
Adrucil (Fluorouracil)	13.8004313	9uM
Floxuridine	13.0250749	0.5uM
Triapine	12.5682276	1.5uM
KX2391	12.4838562	0.5uM
Topotecan HCl	12.3860404	0.065uM
Volasertib	11.1958988	0.24uM
EpothiloneB	9.67898425	8.5uM
Doxorubicin (Adriamycin) HCl	9.47823228	0.16uM
Mycophenolate mofetil	8.5938642	10uM
Belinostat (PXD101)	7.07684924	0.14uM
Rigosertib	7.05275929	0.08uM
Panobinostat (LBH589)	6.93053883	0.07uM
Sildenafil citrate	6.49302924	2uM
Carmofur	5.82534374	1uM
Theophylline	5.52492604	10uM
Onalespib (AT13387)	5.23334434	0.035uM
Talazoparib	5.05827958	0.5uM
Entinostat	4.62376284	1uM
PCI24781	4.61786217	0.2uM
PKC412	4.51702539	0.07uM
AVN944	4.23345055	0.215uM
Mocetinostat	3.62168835	0.5uM
Vorinostat (SAHA; MK0683)	3.4288806	2.5uM
BI2536	3.24077657	1uM
Rocilinostat	2.89525071	1uM
Pralatrexate(Folotyn)	2.79171615	0.2uM
Givinostat	2.437055	0.16uM
17-AAG (Tanespimycin)	2.39042227	0.085uM
MK1775	2.27975304	0.13uM
Malotilate	1.75540515	1.5uM

Figure 81: Tumor-Treg OncoTreat Drug Predictions, Expanded List of All Statistically Significant Compounds.

A) Patient-by-Patient Drug predictions according to inversion of patient-specific Tumor Treg vs Peripheral Treg protein activity signature by drug-treatment protein activity signature. Each drug predicted to invert Tumor Treg signature with -log10(Bonferroni-Corrected p-value) < 0.01 in a particular patient is colored red. Patients are grouped by tumor type. B) Table of all drugs significantly down-regulating Tumor-Treg MRs identified in Figure 79E, 79F, ordered by p-value. Drugs also identified by growth screen to have differentially higher toxicity in Tumor Tregs vs Peripheral Tregs are highlighted in yellow. All seven of these are identified as statistically significant hits down-regulating Tumor-Treg MRs.

By this approach we identified three of the seven above drugs that consistently inverted the TI-Treg signature across nearly all patients (representing four tested tumor types): gemcitabine, triapine, and floxuridine (Figure 82A). As additional analytical validation, we ranked compounds by their overall ability to invert the 17-gene TI-Treg MR signature across all patients in aggregate. We found all seven drugs with differentially greater inhibition of TI-Tregs ex vivo were identified as statistically significant, with gemcitabine, floxuridine, and triapine in the top 5 drugs most strongly inhibiting tumor Treg MR activity (Figure 81B).

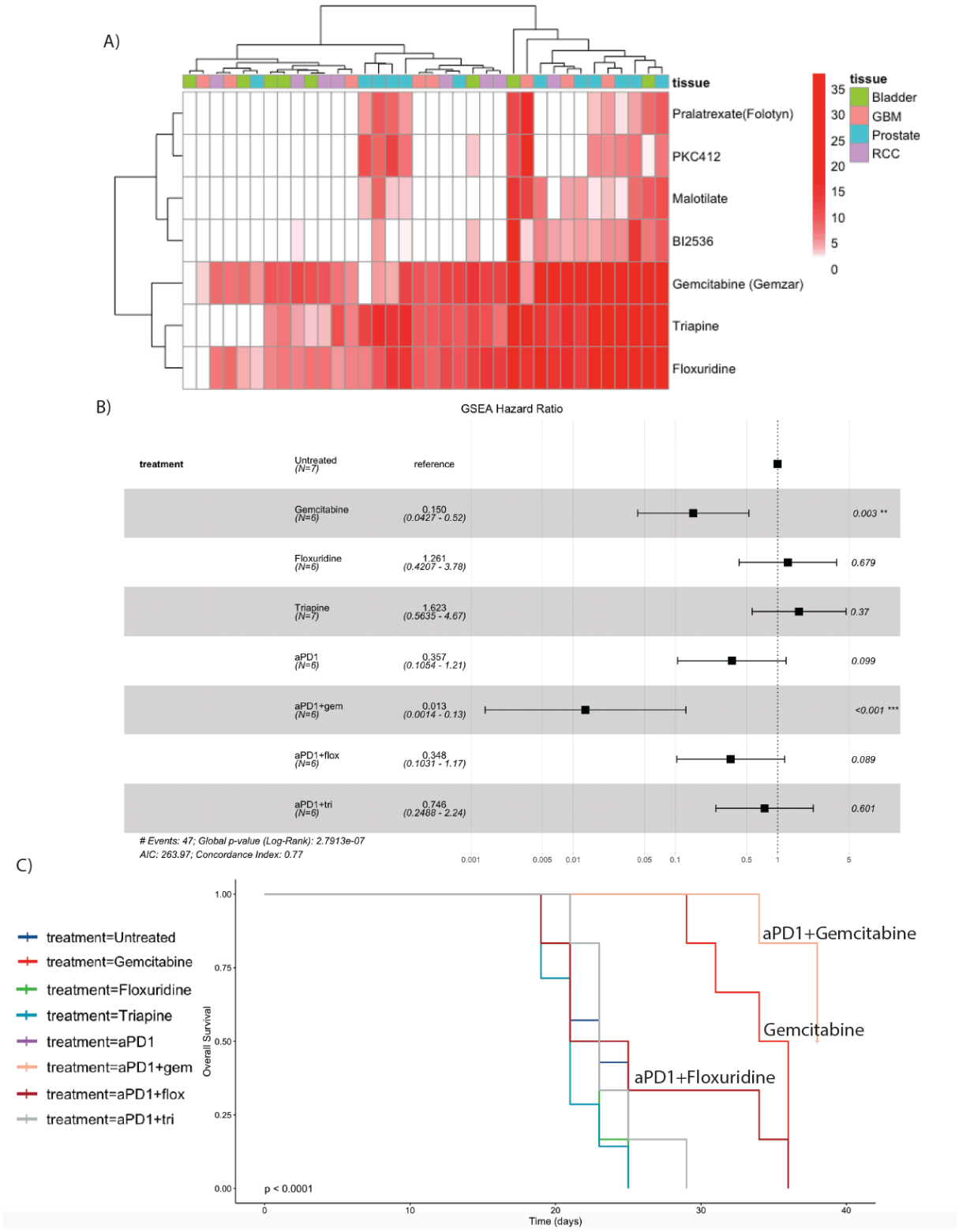


Figure 82: Transcriptional Profiling Identifies Drugs Acting on Functionally Validated Tumor Treg Master Regulator Proteins.

A) Patient-by-Patient Drug predictions according to inversion of patient Tumor Treg vs Peripheral Treg protein activity signature by drug-treatment protein activity signature. Each drug predicted to invert Tumor Treg signature with -log10(Bonferroni-Corrected p-value) < 0.01 in a particular patient is colored red. Patients are grouped by tumor type. Subset to drugs identified by tumor Treg growth screen in 80A, 80B, with columns colored by tumor type and clustered by unsupervised hierarchical clustering. B) Forest-plot showing the result of multiple cox regression assessing treatment effect on time-to-death for each of the treatments Gemcitabine, Floxuridine, Triapine, anti-PD1, anti-PD1+Gemcitabine, anti-PD1+Floxuridine and anti-PD1+Triapine, versus untreated control. Hazard ratios are shown with 95% confidence interval and p-value, such that anti-PD1+Gemcitabine most improves survival, followed by Gemcitabine monotherapy. C) Kaplan-Meier survival curves for treatments shown in B.

Low-Dose Gemcitabine Alone and in Combination with anti-PD1 Inhibits tumor growth in immune-competent mice by differential depletion of Tumor-Specific Tregs.

We first tested whether gemcitabine, floxuridine, or triapine are efficacious *in vivo* at doses approximating the IC20 values of the above PLATE-Seq assay, where differential activity against TI-Tregs is observed. We implanted C57BL/6J mice with subcutaneous MC38 carcinomas, and initiated therapy 12 days later when MC38 tumors are largely resistant to anti-PD-1 immunotherapy. Gemcitabine was administered IP on days 12, 15, and 18 at 12 mg/kg, or $1/10^{th}$ of the lowest conventional clinical-equivalent dose in mice (120 mg/kg). Floxuridine and triapine were IP daily from day 12-18, also at $1/10^{th}$ the standard murine dose. Additional cohorts of mice received combination therapy with anti-PD-1 administered IP on days 12, 15, and 18. We found only low-dose Gem exhibited single-agent efficacy against established MC38 (Figure 82B-C), and the combination of anti-PD-1 and low-dose Gem were synergistic, curing 50% of mice versus 0% individually (Fisher's Exact Test p-value = 0.07). To test whether low-dose Gem functions via immune-dependent versus immune-independent mechanisms, we performed

parallel dose titrations of Gem in immune-competent C57BL/6J mice and immune-deficient NSG (NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ) mice (Figure 83A-C). The standard clinical dose of 120mg/kg inhibits tumor growth in both C57BL/6J and NSG mice, with cox regression p-values <0.001 (Figure 83D), but with no difference in efficacy between C57BL/6J and NSG ($p = 0.19$, Figure 83E). However, at 12mg/kg, efficacy was rapidly lost in NSG mice but not in C57BL/6J ($p = 0.012$, Figure 83E), such that 12mg/kg Gem dose achieved cure in 40% of C57BL/6J as monotherapy (equivalent to the observed cure rate at 120mg/kg in C57BL/6J mice) and 60% cure in combination with anti-PD1 checkpoint immunotherapy. At 1.2mg/kg there remains a trend toward improved outcomes in C57BL/6J vs NSG mice, but it is no longer statistically significant ($p = 0.09$, Figure 83E). These results suggest that low dose Gemcitabine has immunogenic activity.

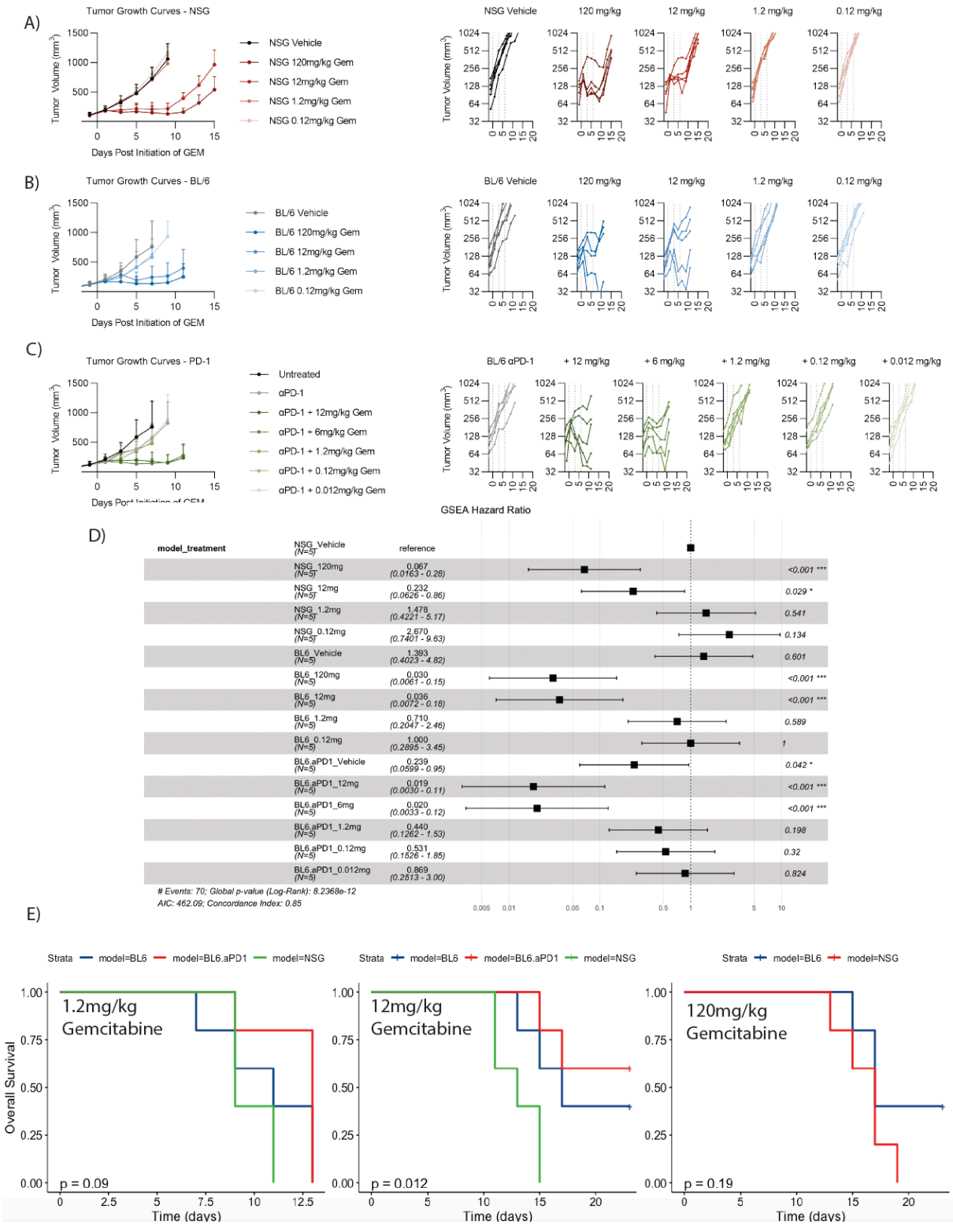


Figure 83: Low-Dose Gemcitabine Inhibits Tumor Growth Only in Immune-Competent Mice

A) Tumor growth curves with varying doses of gemcitabine (0.12mg/kg, 1.2mg/k, 12mg/kg, 120mg/kg) in NSG immune-incompetent mouse model, such that 120mg/kg represents the lowr end of clinically administered gemcitabine dosing. B) Tumor growth curves with varying doses of gemcitabine (0.12mg/kg, 1.2mg/k, 12mg/kg, 120mg/kg) in BL6 immune-competent mouse model. C) Tumor growth curves with varying doses of gemcitabine (0.12mg/kg, 1.2mg/k, 6mg/kg, 12mg/kg, 120mg/kg) in combination with anti-PD1 checkpoint immunotherapy in BL6 immune-competent mouse model. D) Forest-plot showing the result of multiple cox regression assessing treatment effect on time-to-death for each of the treatments in A, B, and C as well as untreated controls. Hazard ratios are shown with 95% confidence interval and p-value. E) Kaplan-Meier survival curves for 1.2mg/kg, 12mg/kg, and 120mg/kg gemcitabine dosing (left-to-right) comparing treatment efficacy in each mouse model, such that 1.2mg/kg shows non-significant trend toward improved survival in BL6 mice ($p = 0.09$) but no cures, 120mg/kg shows no significant difference in efficacy between BL6 vs NSG mouse ($p = 0.19$), and 12mg/kg shows significant improvement in outcome for immune-competent BL6 mice, with 40% of mice cured in BL6 by gemcitabine monotherapy and 60% of mice cured by gemcitabine + anti-PD1.

To test the hypothesis that low-dose Gem modulates TI-Tregs, we performed single cell RNA sequencing of MC38 tumor- and spleen-derived Tregs 24 hours after exposure to a single dose of 12 mg/kg Gem as well as 24 hours after vehicle control. For this study, we implanted *FoxP3*$^{Yfp\text{-}Cre}$ mice with MC38 to facilitate flow-sorting of TCR-β^+ $CD4^+$ $FoxP3^+$ Tregs from tumor and spleen specifically by the YFP marker. We identified five unique clusters of Tregs by VIPER analysis of the single cell data (Fig 84A), of which cluster 3 was highly enriched for human TI-Treg MRs (Fig 84B-C). In Vehicle-treated control, cluster 3 represented 7.8% of splenic Tregs, compared to 30.1% of TI-Tregs ($p = 1.78e\text{-}84$), while Gemcitabine-treatment reduced this cluster by roughly 50% to only 14.9% of TI-Tregs, while no change was observed in the spleen (Fig 84D). Furthermore, this resulted in a proportional increase in Tregs from cluster 1, which exhibit signs of interferon exposure (high IFI16 activity). These data suggest low-dose Gem has

antagonistic effects on tumor Tregs expressing the TI-Treg signature *in vivo*, and that pro-inflammatory effects of low-dose Gem can be uncoupled from direct tumor cytolysis in the context of anti-PD-1 checkpoint blockade.

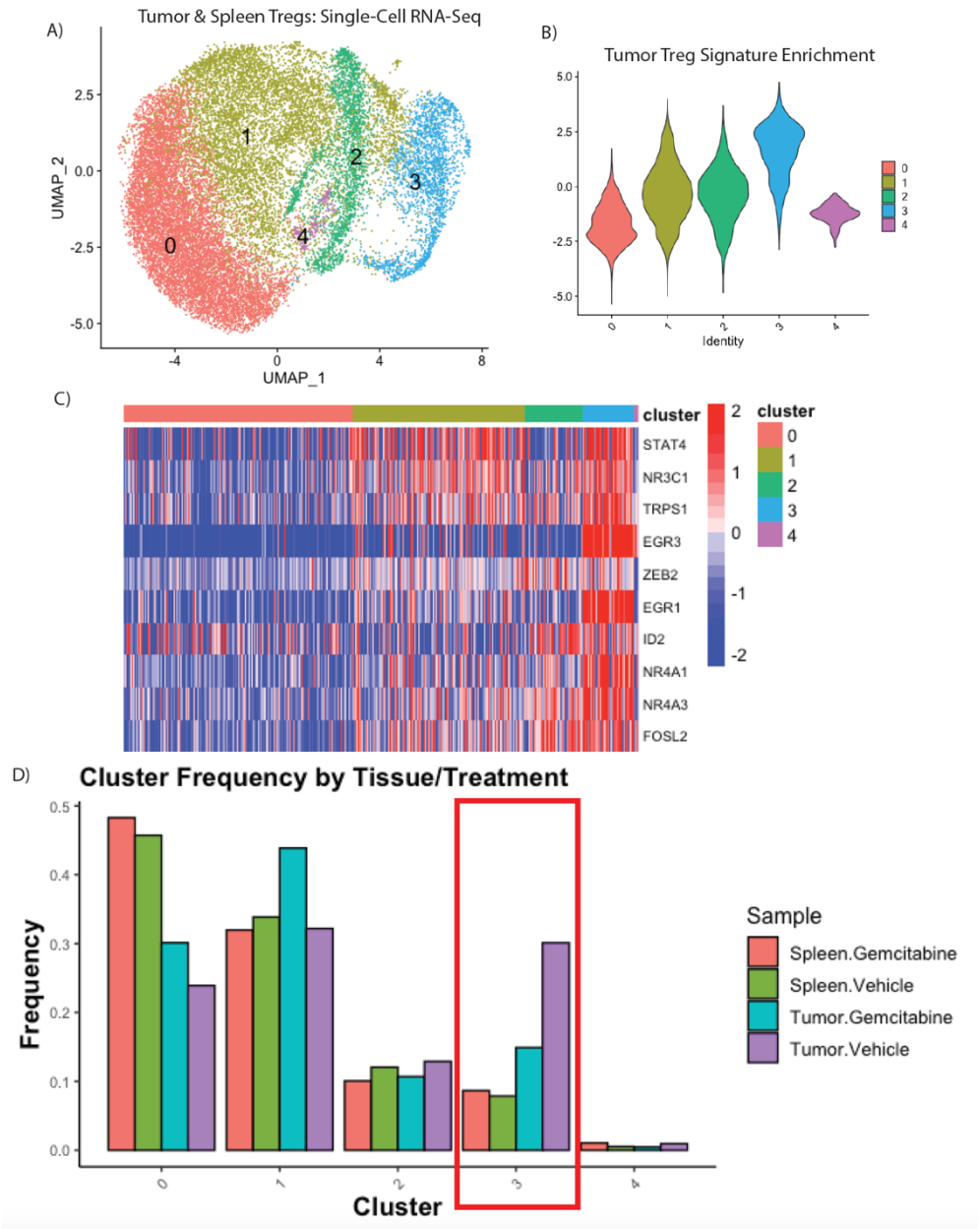
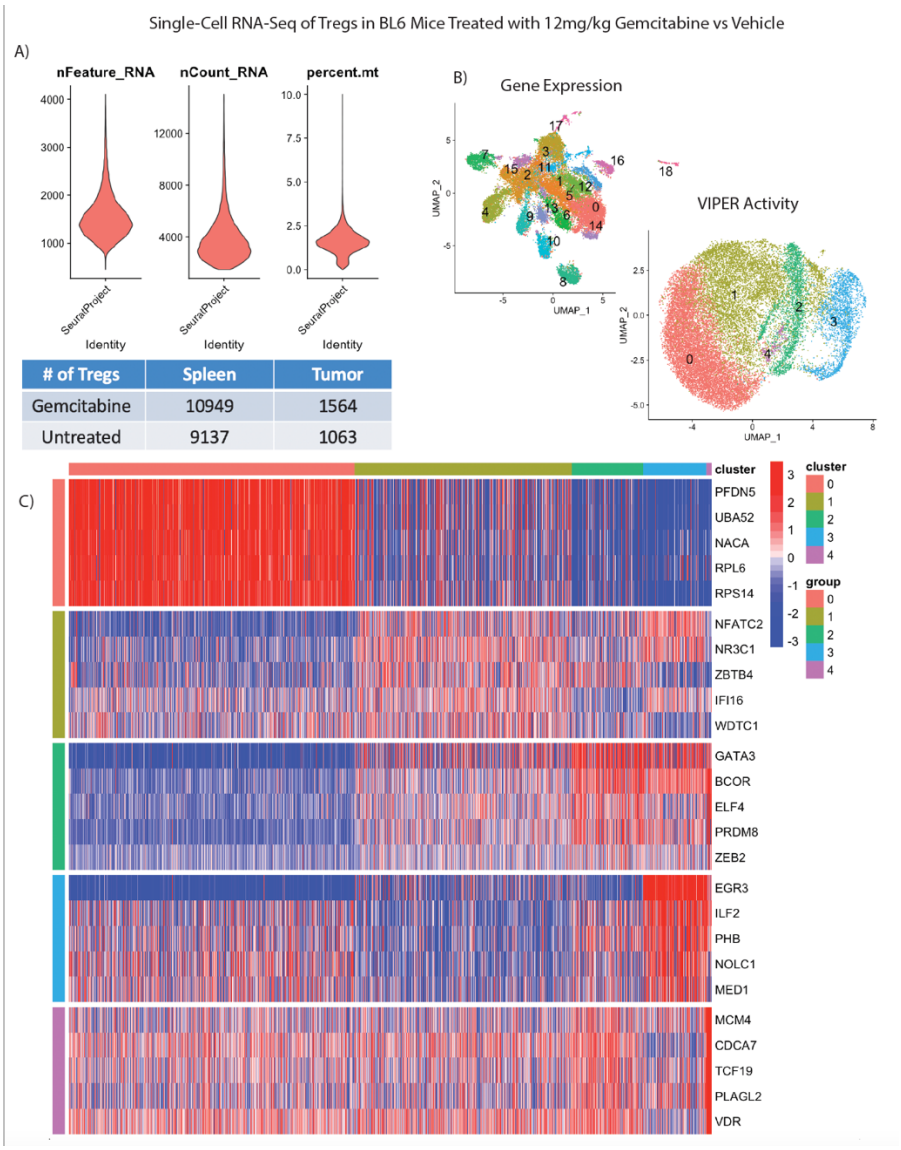


Figure 84: Single-Cell RNA-Sequencing Confirms Depletion by Low-Dose Gemcitabine of Tumor-Specific Tregs

A) UMAP plot and unsupervised clustering by VIPER-inferred protein activity of Tregs from Untreated and Gemcitabine-treated Tumor and Spleen B) Distribution of Tumor-Treg MR signature (Figure 79E, 79F) normalized enrichment score by Gene Set Enrichment Analysis (GSEA), grouped by cluster, such that cluster3 is most enriched for the Tumor Treg signature. C) Heatmap of cell-by-cell protein activity for each Tumor-Treg MR identified by single-cell RNASeq, grouped by cluster. D) Barplot of Cluster Frequency in each sample, such that cluster3 has a baseline frequency of 7.8% in spleen of vehicle-control sample and 30.1% in tumor ($p = 1.78e-84$, OR = 0.198 [95% CI: 0.169-0.231]), with Frequency of only 14.9% in tumor of gemcitabine-treated sample ($p=1.51e-20$, OR = 0.407 [95% CI: 0.334-0.494)



# of Tregs	Spleen	Tumor
Gemcitabine	10949	1564
Untreated	9137	1063

Figure 85: Single-Cell RNA-Seq Characterization of Tumor-Infiltrating and Peripheral Tregs With or Without Low-Dose Gemcitabine Treatment

A) Number of Tregs per sample post-quality control filtering. Violinplot of data quality showing distribution of nFeature_RNA (number of unique genes profiled), nCount_RNA (number of unique molecular identifiers profiled), and percent.mt (percentage of mitochondrial transcripts) per cell. B) Clustering of Tregs by Gene Expression (top-left) and VIPER protein activity inference (bottom-right), showing noisiness of clustering by gene expression due to cross-sample batch effects. C) Top5 most differentially upregulated proteins per Treg cluster.

Pooled CRISPRko-screen in Mouse Hematopoietic Stem Cell Model Validates Functional Role of Tumor Treg Master Regulators: To functionally validate individual candidate MRs for their ability to regulate TI-Treg recruitment and/or retention with the tumor microenvironment, we performed a pooled, *in vivo* CRISPR knockdown screen using the established CHIME (CHimeric IMmune Editing) system[1]. In brief, we sorted Lin$^-$Sca-1$^+$c-Kit$^-$ hematopoietic stem cells from mice constitutively expressing Cas9, and lentivirally transduced them with a sgRNA library targeting 34 genes—including 17 MRs, 13 randomly selected negative control genes with differential protein activity p-value = 1.0 between TI-Tregs and P-Tregs, and 4 positive controls known to be toxic to Tregs (Fig 86A). In this model, we transduced hematopoietic stem cells sorted from Cas9+ mice using a library of 102 sgRNAs (34 genes, 3 guides / gene) cloned in pXPR_053-vector, which also encodes a Vex fluorophore, then implanted them into recipients irradiated at two doses of 600 rads, allowing the immune system to reconstitute over 10 weeks. During this time, all Vex+ immune lineage cells, including Tregs, harbored CRISPR mediated knockouts. This system allowed us to implant syngeneic MC38 tumors, allow two weeks for tumor growth, then flow-sort Vex+ Tregs as well as CD4nonTregs from tumors and spleen to compare guide frequencies in Tumor Tregs vs peripheral Tregs. Upon engraftment and reconstitution of the hematopoietic system, roughly 25-40% of immune cells harbored CRISPR gene knockdown. We implanted two cohorts of CHIME chimeras with syngeneic MC38 tumors (the second cohort implanted with Vex+ Lin$^-$Sca-1$^+$c-Kit$^-$ hematopoietic stem cells harvested from the bone marrow of the first), allowed 18 days for tumor growth, then flow sorted CD4$^+$CD25$^+$ Tregs and CD4$^+$CD25$^-$ Tconv cells from tumors and spleen (Figure 86B) and performed gDNA sequencing to compare guide frequencies in TI-Tregs vs P-Tregs and Tconv.

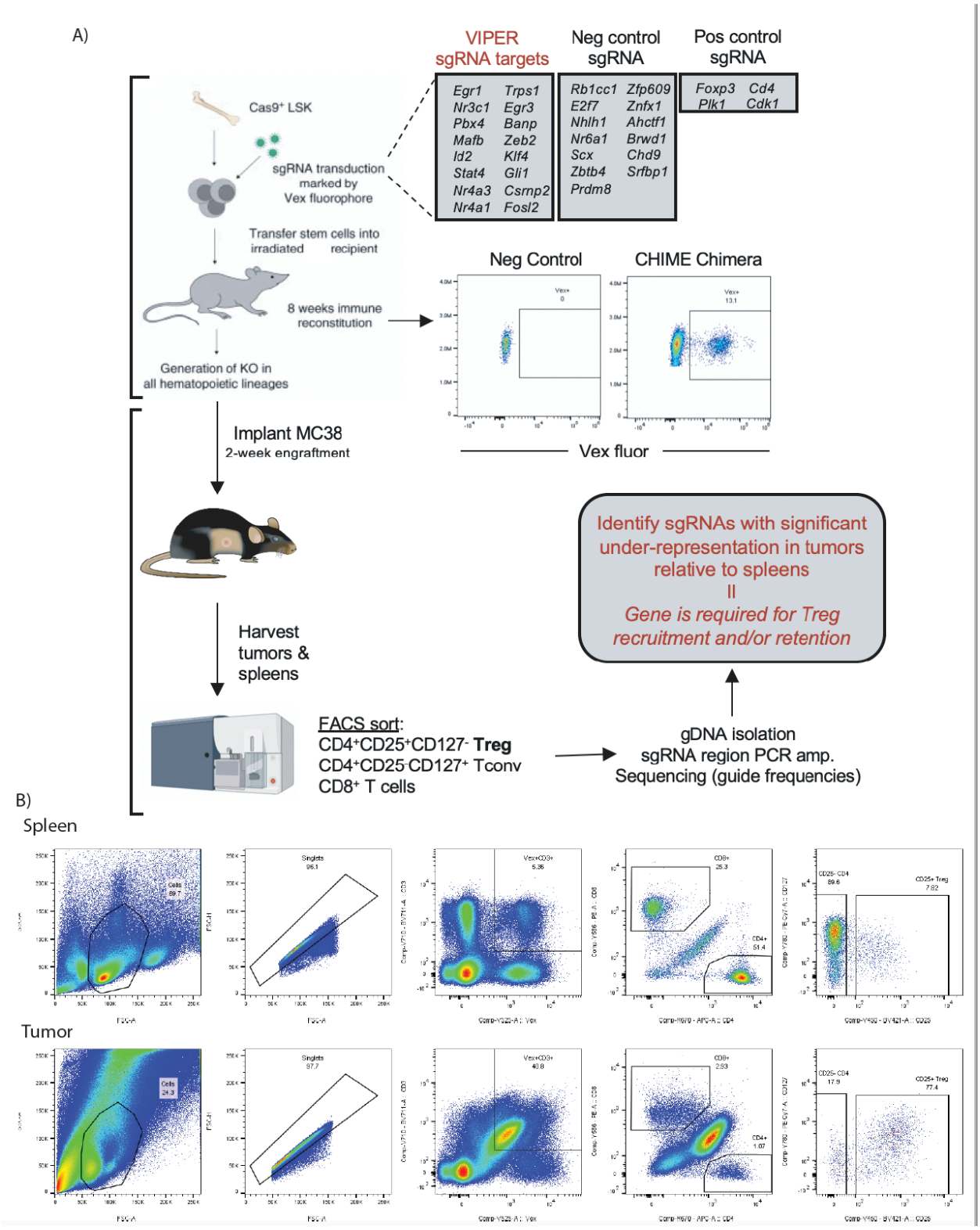
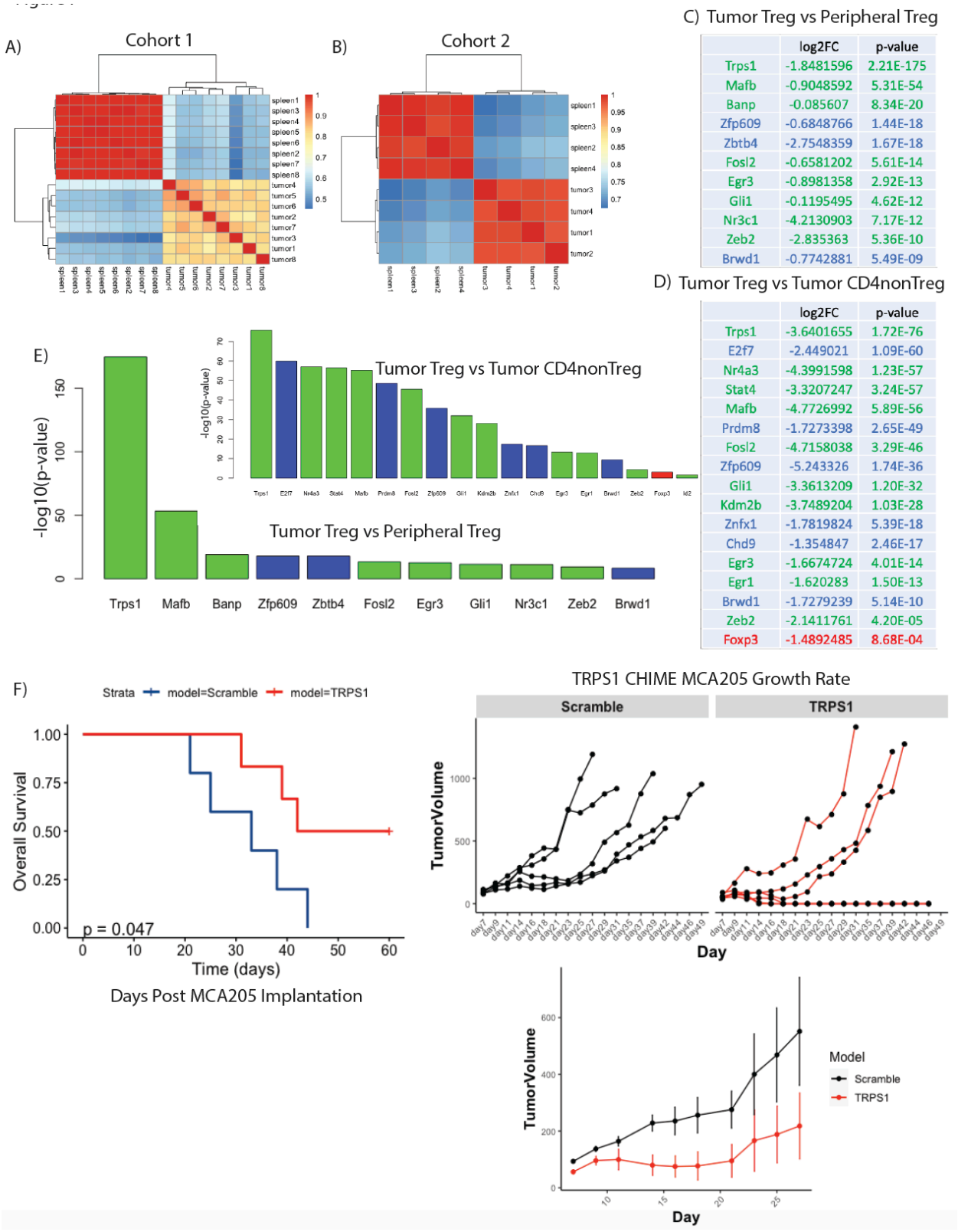


Figure 86: Chimeric Immune Editing Mouse Model Enables Validation of Treg Tumor-Infiltration Master Regulators

A) Experimental design for CRISPRko Validation of Tumor vs Peripheral Treg Master Regulators MR targets, randomly sampled negative control genes, and Treg-toxic positive control genes are listed in the Figure. B) Representative flow cytometry gating for Vex+ CRISPR-transduced Tregs, CD4nonTregs, and CD8 T-cells in Spleen (top) and Tumor (bottom).

Across both mouse cohorts, we noted strong correlation between replicates of sorted spleen and tumor Tregs (Figure 87A,87B), such that the predicted Master Regulators *Trps1, Mafb, Banp, Fosl2, Egr3, Gli1, Nr3c1,* and *Zeb2* were significantly depleted in TI-Tregs relative to P-Tregs (Figure 87C, 87E), indicating a significant degree of dependence on activity of these proteins for successful Treg tumor infiltration and/or retention. Importantly, *Trps1, Mafb, Fosl2, Egr3, Gli1,* and *Zeb2* were also significantly depleted in TI-Tregs relative to tumor CD4 Tconv, (Figure 87D, 87E) supporting Treg-specificity in the function of these genes, such that knocking them down depleted Tregs to a greater extent than CD4 Tconv, a clinically desirable outcome.



	log2FC	p-value
Trps1	-1.8481596	2.21E-175
Mafb	-0.9048592	5.31E-54
Banp	-0.085607	8.34E-20
Zfp609	-0.6848766	1.44E-18
Zbtb4	-2.7548359	1.67E-18
Fosl2	-0.6581202	5.61E-14
Egr3	-0.8981358	2.92E-13
Gli1	-0.1195495	4.62E-12
Nr3c1	-4.2130903	7.17E-12
Zeb2	-2.835363	5.36E-10
Brwd1	-0.7742881	5.49E-09

	log2FC	p-value
Trps1	-3.6401655	1.72E-76
E2f7	-2.449021	1.09E-60
Nr4a3	-4.3991598	1.23E-57
Stat4	-3.3207247	3.24E-57
Mafb	-4.7726992	5.89E-56
Prdm8	-1.7273398	2.65E-49
Fosl2	-4.7158038	3.29E-46
Zfp609	-5.243326	1.74E-36
Gli1	-3.3613209	1.20E-32
Kdm2b	-3.7489204	1.03E-28
Znfx1	-1.7819824	5.39E-18
Chd9	-1.354847	2.46E-17
Egr3	-1.6674724	4.01E-14
Egr1	-1.620283	1.50E-13
Brwd1	-1.7279239	5.14E-10
Zeb2	-2.1411761	4.20E-05
Foxp3	-1.4892485	8.68E-04

Figure 87: CRISPRko of Master Regulators Reproducibly Inhibits Treg Tumor-Infiltration and Tumor Growth

A) Correlation of sgDNA frequency distribution between replicates of Spleen and Tumor Tregs in experimental cohort 1. B) Correlation of sgDNA frequency distribution between replicates of Spleen and Tumor Tregs in experimental cohort 2. C) Table of log2(Fold Change) and Bonferroni-corrected p-values Stouffer-integrated across experimental cohorts for genes with consistent and statistically significant depletion of targeting gDNAs in Tumor Tregs vs Spleen Tregs. D) Table of log2(Fold Change) and Bonferroni-corrected p-values Stouffer-integrated across experimental cohorts for genes with consistent and statistically significant depletion of targeting sgDNAs in Tumor Tregs vs Tumor CD4-nonTregs. E) Barplot of -log10(P-values) from C and D (inset). F) Tumor growth curves and Kaplan-Meier plot for CRISPRko validation using sgRNAs targeting TRPS1 vs or non-targeting scramble-control, showing significant difference in tumor growth ($p<0.05$).

These findings were consistent across both CRISPR screen cohorts (Figure 88). Notably, knockdown of the positive control Foxp3 is expected to be differentially more lethal to Tregs than CD4nonTregs and is also differentially depleted in this comparison. The most statistically significant hit both in terms of relative depletion in TI-Tregs vs P-Tregs ($p=2.21*10^{-175}$) and relative depletion in TI-Tregs vs Tumor CD4 Tconv ($p=1.72e^{-76}$) was Trps1, a gene with unknown function in T-cells.

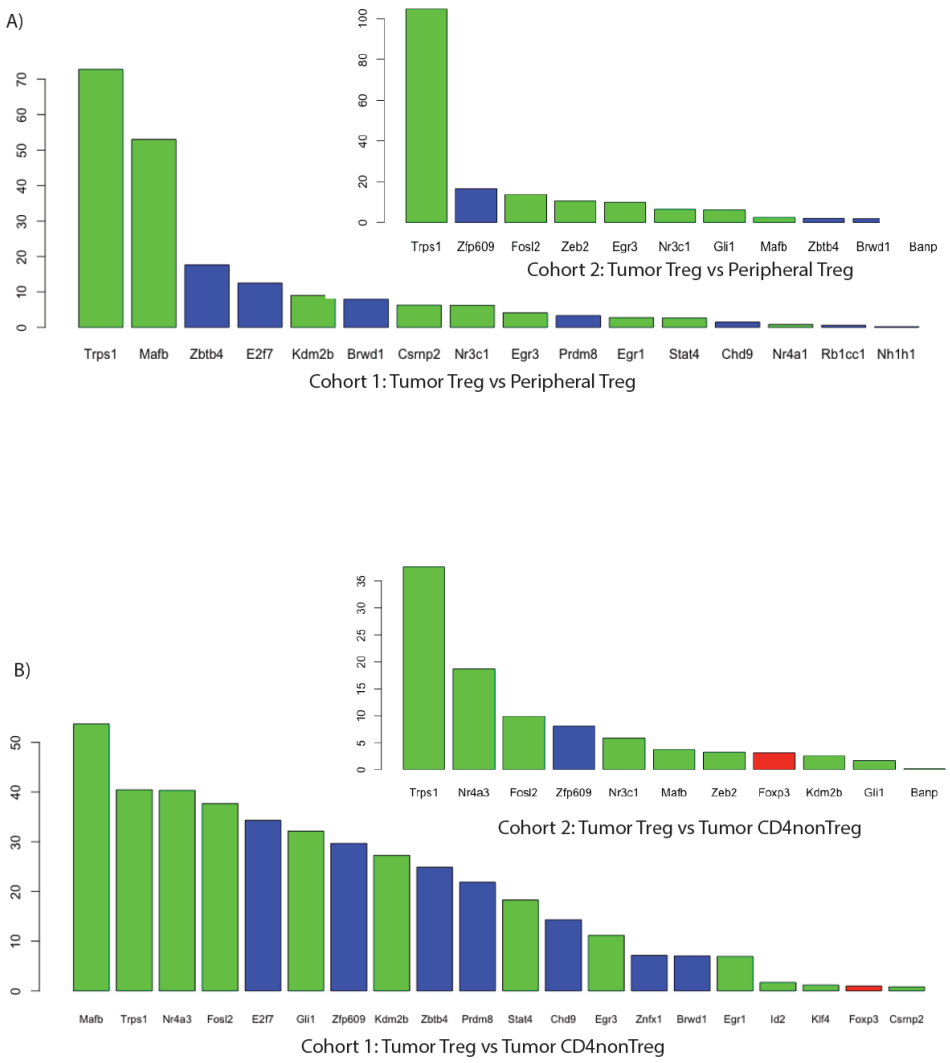


Figure 88: Cohort-Specific CRISPR Validation Results

A) Barplot of -log10(Bonferroni-Corrected P-values) for genes with statistically significant depletion of targeting gDNAs in Tumor Tregs vs Peripheral Tregs, separately for experimental cohort 1 (left) and experimental cohort 2 (right). B) Barplot of -log10(Bonferroni-Corrected P-values) for genes with statistically significant depletion of targeting gDNAs in Tumor Tregs vs Tumor CD4nonTregs, separately for experimental cohort 1 (left) and experimental cohort 2 (right).

Loss of Trps11 in Hematopoietic Cells Significantly Inhibits Tumor Growth

Due to the discovery from high-throughput CRISPRko screen that Trps1 knockout has the strongest effect on inhibiting Treg tumor-infiltration and the greatest differential effect on Tumor Tregs vs Tumor CD4-non-Tregs, we performed a follow-up single-gene knockdown of Trps1 in the hematopoietic lineage in a cohort of 6 bone marrow transplant chimeras. For comparison, we also performed transduction of hematopoietic stem cells with a non-targeting scramble guide as a control and transplanted into a comparator cohort of 5 mice (Figure 87F). Upon simultaneous implantation of both cohorts with MCA205, an immune-resistant fibrosarcoma tumor model, there was a significantly slower tumor growth among mice where Trps1 was targeted by CRISPRko ($p<0.05$), with dramatic biological effect, such that three of the six mice in this cohort spontaneously rejected the tumor and were cured, and three of the remaining four had dramatically slowed rate of tumor growth relative to mice with scramble control guide (Figure 87F). This indicates that inhibition of TRPS1 represents a potent immune-stimulatory pathway, which has never previously been described.

8.4 Discussion

Treg immunosuppression in the TME is a major barrier to antitumor immunity and undermines efficacy of checkpoint blockade immunotherapy, which remains effective only in a minority of cancer patients [219] [228]. Limitations in the prior research has led to prioritization of tumor Treg targeting approaches that have failed to affect this clinical reality. We have tailored our approach to overcome these limitations, as solving the problem of Tumor-Specific Treg depletion would have extremely broad clinical impact, leading to both conceptual advancements in the understanding of Treg biology and a selection of actionable targets with potential for rapid translation into clinical trials. Furthermore, by generating a resource of high-throughput drug

perturbations on transcriptional profile of human Tumor Tregs, we can apply the New York and California CLIA-certified OncoTreat algorithm [21] to rank current FDA-approved and investigational drug compounds according to their ability to invert the identified Tumor vs Peripheral Treg Master Regulator activity, which may further accelerate drug development and clinical trials directed at the identified Master Regulators.

By applying VIPER followed by Random Forest feature selection to a unique bulk-RNA-Sequencing database of purified Tregs, CD4nonTregs, and CD8 T-cells sorted from tumor and peripheral blood of 36 patients across a variety of tumor types, we have discovered a set of 17 master regulators (EGR1, NR3C1, PBX4, MAFB, ID2, STAT4, NR4A3, NR4A1, TRPS1, EGR3, BANP, ZEB2, KLF4, GLI1, CSRNP2, KDM2B, FOSL2) specifically enriched in Tumor Tregs relative to Tregs in the peripheral circulation and non-Treg subsets. We have found that the activity of these proteins is differentially upregulated in tumor-infiltrating regulatory T cells (Tregs) relative to peripheral blood Tregs, peripheral blood conventional CD4 T cells, and *ex vivo* activated T cell controls in samples freshly isolated from human patients, demonstrating relative specificity to Tumor Tregs, along with a generalized effect across tumor types (Figure 79).

Downregulating or inhibiting these proteins was found to decrease the ability of regulatory T cells (Treg) to infiltrate tumors and locally exert their immunosuppressive function, thus augmenting the immune response to cancer. Since this predominantly targets Tumor-Infiltrating Tregs rather than peripheral Tregs, it therefore limits adverse effects resulting from systemic Treg inactivation, which may otherwise lead to autoimmune diseases and chronic inflammation.

We have functionally validated the MRs by CRISPRko in a Chimeric Immune Editing (CHIME) MC38 mouse tumor model (Figure 86), such that the top-most-depleted targeted MRs in Tumor Tregs relative to Peripheral Tregs were found to be depleted with highly significant p-values both in Tumor Tregs relative to Peripheral Tregs and in Tumor Tregs relative to Tumor CD4nonTregs (Figure 87), and included TRPS1 as the most-strongly-depleted MR.

These data show that the loss of TRPS1 dramatically impairs the ability of Tregs to infiltrate the tumor mass, relative to control tissue (spleen). Further, the retained presence of TRPS1 and other master regulator knockdown cells among circulating Tregs shows that TRPS1, MAFB, BANP, FOSL2, EGR3, GLI1, NR3C1, and ZEB2 are not required for survival or homeostatic regulation of extra-tumoral Tregs. The ability of CRISPRko perturbed HSCs to successfully engraft and reconstitute all subsets of T-cells further indicates that inhibition of the Master Regulators is not broadly toxic to immune cells. Instead, these data show that Master Regulator activity mainly controls Treg targeting to tumors. Finally, we observe a dramatic inhibition of tumor growth in mice with single-gene CRISPRko of TRPS1 in the hematopoeitic lineage, highlighting inhibition of TRPS1 as a novel target for Treg-directed immunotherapy, particularly exciting since the role of TRPS1 in T-cells has never previously been described.

TRPS1 is a transcription factor that represses GATA-regulated genes and binds to a dynein light chain protein. As such, TRPS1 may be therapeutically inhibited by the action of drugs which are found to invert the expression of its downstream regulon targets. Alternatively, TRPS1 may be inhibited by small molecule inhibitors, antisense oligonucleotides, epigenetic modulators, protein mimetics, intracellular targeted antibodies, specific degraders such as proteolysis targeting

chimera (PROTAC) molecules, and gene therapy. Analogous approaches may be taken to target the other discovered Master Regulators, alone or in combination, anticipating further synergy in combination with conventional immunotherapy, antibody-drug conjugates, targeted therapy, radiation therapy, chemotherapy, and/or surgery.

Furthermore, we have generated a unique drug perturbation PLATE-SEQ resource from human Tumor-Infiltrating Tregs, which we have used to effectively identify drugs with a transcriptional effect on Tumor-Treg Master Regulators and differential toxicity to Tumor vs Peripheral Tregs (Figure 80), specifically highlighting low-dose gemcitabine as having great potential for rapid translation to clinical trials. Critically, the dosages at which gemcitabine was found to be toxic to Tumor Tregs in vitro were orders of magnitude lower than doses which are clinically administered as cytotoxic tumor-cell-directed therapy. Therefore, low-dose gemcitabine combination with checkpoint immunotherapy is likely to exert differential depletion of Tregs relative to non-Treg populations and Tumor Tregs relative to non-Tumor Tregs, representing potential for novel drug synergy to overcome Treg-mediated treatment resistance with minimal drug toxicity. We have found that each of the drugs observed to have differential toxicity in TI-Tregs vs P-Tregs also reduced activity of the CRISPR-validated Master Regulator gene set we have identified (Figure 80B), suggesting that a more intentional development of drugs targeting these master regulators as an intended rather than secondary effect may represent an even better synergy with immunotherapy.

Critically, gemcitabine administered in vivo at a dose of 12mg/kg alone and in combination with anti-PD1 checkpoint immunotherapy was found to significantly inhibit tumor growth in immune-

competent, but not NSG mouse model (40% cure rate in immune-competent mice as monotherapy, 60% cure rate in combination with anti-PD1, and 0% cure rate in NSG mice, $p = 0.012$). This is especially significant because typical cytotoxic doses of gemcitabine administered as anti-tumor therapy exceed 120mg/kg, at which there is no statistically significant difference in survival between immune-competent and immune-incompetent mice ($p = 0.19$). Therefore, the immunogenic effects of gemcitabine are more pronounced at low doses, where we observe differential toxicity in TI-Tregs relative to P-Tregs, confirmed in vivo by single-cell RNA-Sequencing (Figure 84). The Tregs depleted by gemcitabine are found to be tumor-specific and characterized by high activity of the Master Regulator protein signature we have discovered, including TRPS1, suggesting inhibition by gemcitabine of these MRs as the putative mechanism of its immunogenicity in the low-dose regimen.

The paradigm we have presented of large-scale drug screening ex vivo followed by CRISPR validation of putative regulatory proteins in chimeric mouse model represents a novel approach to discovery of Treg-directed immunotherapy targets, and can be extended in the future to other immunosuppressive cell types, as well as serving as a resource for further investigation of potential Treg-depleting therapies. This may discover previously unknown treatment synergies and represents a paradigm-shift by informing development of rational pan-cancer treatment approaches directed at Treg-mediated treatment resistance. The utility of this approach is exemplified by our discoveries of low-dose gemcitabine's differential toxicity to tumor-infiltrating Tregs with potential for rapid clinical translation, and of TRPS1 as a mechanistic regulator of Treg tumor-infiltration with significant effect on tumor growth and attractive prospects as a novel drug target.

8.5 Methods

Clinical Sample Collection, Sorting, and RNA-Sequencing: Tissue was collected from treatment-naïve resected tumors across patients with four tumor types, including 8 patients with glioblastoma multiforme, 8 patients with clear cell renal carcinoma, 8 patients with bladder cancer, and 12 patients with prostate cancer (Figure 79). For each patient, 50ml of peripheral blood was drawn at the same time that tumor was resected. Tumors were dissociated with the GentleMACS OctoDissociator following manufacturer's instruction, and subsequently Tregs and CD8 T-cells were flow-sorted from tumor along with Tregs, naïve CD4nonTregs, and naïve CD8 T-cells from peripheral blood. An aliquot of naïve CD8 and CD4nonTreg were stimulated ex vivo with IL2 and anti-CD3/anti-CD28 beads for 24 hours to induce T-cell activation. Flow-sorted and ex-vivo-stimulated populations were processed to prepare cDNA libraries following Illumina user guide and were sequenced on Illumina NovaSeq 6000 Sequencing System.

Gene Expression and VIPER Analysis: Gene Expression was combined across all samples and scaled to log10(Transcripts Per Million + 1). Gene Expression was subsequently scaled across rows by z-score transformation and used as input for Principal Component Analysis (Figure 80A) and differential gene expression. Log10(TPM+1) matrix was separately used to infer gene regulatory network structure by the ARACNe algorithm. ARACNe was run with 100 bootstrap iterations using 1785 transcription factors (genes annotated in gene ontology molecular function database as GO:0003700, "transcription factor activity", or as GO:0003677, "DNA binding" and GO:0030528, "transcription regulator activity", or as GO:0003677 and GO:0045449, "regulation of transcription"), 668 transcriptional cofactors (a manually curated list, not overlapping with the

transcription factor list, built upon genes annotated as GO:0003712, “transcription cofactor activity”, or GO:0030528 or GO:0045449), 3455 signaling pathway related genes (annotated in GO biological process database as GO:0007165, “signal transduction” and in GO cellular component database as GO:0005622, “intracellular” or GO:0005886, “plasma membrane”), and 3620 surface markers (annotated as GO:0005886 or as GO:0009986, “cell surface”). ARACNe is only run on these gene sets so as to limit protein activity inference to proteins with biologically meaningful downstream regulatory targets, and we do not apply ARACNe to infer regulatory networks for proteins with no known signaling or transcriptional activity for which protein activity may be difficult to biologically interpret. Parameters were set to zero DPI (Data Processing Inequality) tolerance and MI (Mutual Information) p-value threshold of 10^{-8}, computed by permuting the original dataset as a null model.

Using the ARACNe gene regulatory network structure, VIPER protein activity inference was performed on gene expression signature. First directly on z-score-scaled gene expression signature for all T-cell subtypes, used for Principal Component Analysis and clustering (Figure 80A). Then separately scaling Tumor and Peripheral Tregs against naïve CD4nonTregs by viperSignature command in Rstudio for comparison of Tumor Treg vs Peripheral Treg (Figure 80D, 80F), and scaling all Tregs and CD4nonTregs against naïve CD8nonTregs by viperSignature for comparison of Tumor Treg vs all Treg and CD4nonTreg controls (Figure 80C, 80E).

Random Forest Feature Selection: The full dataset was randomly split into 75% training data and 25% testing data. On training data, a Random Forest Model was built with VIPER-inferred

protein activity to classify Tumor Treg vs Peripheral Treg (Figure 80D) or Tumor Treg vs all Controls (Figure 80C), taking the list of all differentially active proteins (t-test p-value < 0.01) as an initial feature set. Features were ranked by mean decrease in model accuracy and included one-by-one to construct random forest models with feature selection. Predictive power was assessed by Area-Under-ROC-Curve (AUC) in the held-out testing data, and a null model of AUC was constructed from random sampling of the same number of genes (from the set of genes with differential activity p-value =1.0) 1000 times. For each comparison, the maximum number of discriminative genes was selected for which AUC vs null model remained statistically significant (Figure 80C, 80D). These genes are shown in Figure 80E and 80F and aggregated into a combined list of 17 putative Tumor Treg vs Peripheral Treg Master Regulators with Activity specifically upregulated in Tumor Tregs.

CRISPR Validation in Chimeric Immune Editing Model: Confirmatory evidence that the predicted proteins regulate tumor Treg infiltration was generated in murine models in a pooled CRISPR screen (Figure 82); by comparing the differential representation of gene-knockout Tregs in tumor versus non-tumor tissue (spleen, as a control), for each candidate Master Regulator gene. For these studies, Hematopoetic Stem Cells (HSCs) were extracted from Cas9+ mice and transduced with sgRNA library targeting 34 genes with 3 guides/gene. The transduced stem cells were reimplanted into irradiated recipient mice, allowing reconstitution of the entire immune system, including Tregs, with a unique pool of CRISPR knockout genes in place. Subsequent implantation of a subcutaneous MC38 murine colon adenocarcinoma tumor model allowed direct observation of differential infiltration of tumors by Tregs receiving selected CRISPR guides, in a single, high-throughput experimental screen. Critically, the experiment would not have been

possible on a genome-wide level without initial narrowing of candidate master regulators by VIPER protein activity analysis, due to fundamental limitations in achieving a sufficient number of tumor-infiltrating Tregs harboring guide DNAs for the full set of human genes. This is because we typically find fewer than 10,000 tumor-infiltrating Tregs in MC38 tumor model.

We designed the gDNA library with three guides per gene targeting the 17 predicted Tumor Treg MRs and 13 randomly sampled negative control genes (genes with p=1.0 comparing Tumor Treg to Peripheral Treg). We also included Treg context-specific positive controls such as Foxp3 and Cd4 and core-essential genes Cdk1 and Plk1 (these were not detected in any cells post-transduction, indicating successful gene-editing). For guide design, we used the Broad Institute Genetic perturbation platform (GPP) sgRNA designer-tool. Sorted Cas9+ hematopoietic stem cells were successfully transduced and implanted into irradiated recipient mice, A cohort of six replicate mice (cohort 1) and three replicate mice (cohort 2) were separately implanted and harvested. Vex+ gDNA-bearing Tregs and CD4nonTregs were flow-sorted from Tumor and spleen, separately.

Pelleted Tregs/CD4s were first pooled together, with entire tumor samples pooled and spleen samples pooled in proportion to ratio of sorted cell counts from Tumor. gDNA was extracted first by adding 400ul of RIPA buffer (with added RNAseA) on top of the pelleted Tregs/CD4s, followed by 1h incubation at 65C. This was followed by Phenol/Chloroform/Isoamyl alcohol-extractions and Isopropanol-precipitations. Extracted gDNA was divided into 8 replicates with equal volumes for cohort 1 and four replicates with equal volumes for cohort 2, each then

amplified by 2-step PCR and then sequenced. Correlation between replicates was assessed (Figure 83A, 83B).

After sorting and gDNA sequencing, differential frequency of guides in Tumor Treg vs Peripheral Treg and Tumor Treg vs Tumor CD4nonTreg are assessed by DESeq with Bonferroni correction on the p-values, separately (Figure 88), and then p-values were integrated by Stouffer's Method (Figure 87A-E).

CRISPRko library design : For CRISPRko screening we designed the target gene list to include 34 genes (3 sgRNAs / gene)—including 17 MRs and 13 negative control genes (genes whose loss is not predicted to differentially affect Tumor Tregs compared to Peripheral Tregs i.e. p=1.0 comparing Tumor Treg to Peripheral Treg), and 4 positive controls (2 genes whose loss is known to be toxic to Tregs (FOXP3 and CD4) and 2 core-essential genes (PLK1 and CDK1)). Positive control sgRNAs were not detected in any cells post-transduction, indicating successful gene-editing. For guide design, we used the Broad Institute Genetic perturbation platform (GPP) sgRNA designer-tool. The pooled guide-library was ordered from Twist-bioscience.

CRISPRko oligo synthesis and library cloning: Oligo libraries (102 oligos) were ordered from Twist-biosciences in following format (200mers):

ACACGTCATATAGATGCCGTCCTAGCGAGCGTGGAGTGAGCCATTGTGAGCGCTCAC

AATTATATATCTTGTGGAAAGGACGAAACACCGNNNNNNNNNNNNNNNNNNNGT

TTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCATCGGCAGCAA

CCAGATGGGCACAGGAAAGATACTTAACGCTT

From the initial oligo pool, this TREG sub-library was amplified first with KAPA polymerase (KK2502) with the following PCR primers and settings:

TREG_1F=AGCGTGGAGTGAGCC,
TREG_1R=TCTGGTTGCTGCCGA

DNA(oligo pool 1ng/ul)	2ul	
5xHF-buffer	5	
dNTPs	0.75ul	
Oligo_F(10uM)		0.75ul
Oligo_R(10uM)		0.75ul
KAPA pol		0.5ul
H2O		to 25ul
95C	3min	
98C	20s	
55C	15s	
72C	15s	
72C	1min	
4C	---	

The PCR product from PCR1 was gel purified with GeneJet gel purification-kit. The final 2nd PCR prior to the Gibson cloning-step was done with the following primers and settings:

TREG_2F: AGCGCTCACAATTATATATCTTGTGGAAAGGACGAAACACCG
TREG_2R: CGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC

DNA(product from 1st PCR)	3ng	
5xHF-buffer	5	
dNTPs	0.75ul	
Oligo_F(10uM)		0.75ul
Oligo_R(10uM)		0.75ul
KAPA pol		0.5ul
H2O		to 25ul
95C	3min	
98C	20s	
64C	15s	
72C	15s	
72C	1min	
4C	---	

Both of these amplifications were done with qPCR and the program was stopped before the amplification started to plateau. After PCR the insert was gel purified (GeneJet) and Gibson

cloned into BsmBI-digested pXPR_053 (Addgene# 113591). Gibson cloned insert and vector was column purified (GeneJet) and large-scale electroporated into Lucigen Enduro competent cells. The bacterial colonies were scraped from 24,5cm x 24,5cm agar plates, so that the estimated library complexity was approximately 1000 colonies / sgRNA.

Lentiviral packaging of the sgRNA library: 13 million 293T cells were seeded for each 15cm dish the night before transfection. The following morning, viral transfections were conducted with the following components:

- 22.1ug sgRNA containing pXPR_053 (Addgene 113591).
- 16.6ug PsPAX2 (Addgene 12260)
- 5.5ug PMD2G (Addgene 8454).
- 1660ul of sterile H2O.

After mixing the plasmids and H2O, 110,6ul of Fugene HD (Promega) was added to the mix. The transfection mixture was vortexed, then incubated for 10 minutes before adding dropwise to 293T cells. The transfection mixture was removed the following day and virus was collected at 48h and 72h after initial transfections. To remove cellular debris, the virus-containing supernatant was centrifuged 500 x g for 5min and filtered with 0.45um PES filters (Millipore), followed by ultracentrifugation (25,000rpm for 2h), dissolving the viral pellet into PBS, aliquoting the virus and storing the aliquots at -80C. Viral titer was measured with 293T cells by using violet-excited GFP in the pXPR_053-plasmid.

Cell culture and sgRNA transductions into hematopoietic LSK cells:

HEK293T cells: HEK293T cells used in this study were obtained from the American Type Culture Collection (ATCC) and cultured at 37 °C in a humidified incubator (5% CO2) with the following media: DMEM + 10% FBS, 1% L-Glutamine and 1% Penicillin/Streptomycin. Cell

line was tested for mycoplasma status before viral production. LSKs: After sorting the LSKs from donor mice, cells were sorted into 96-well plate (100k LSKs/well) and incubated overnight in SFEM media supplemented with 100 ng/mL of the following cytokines: SCF, TPO, Flt3-Ligand, and IL-7. Pen/Strep was used in all in vitro cultures. The following day, LSK cells were transferred into Retronectin-coated 24-well plate and sgRNA-containing Lentiviruses were added to the wells with MOI 30 (based on viral tittering in 293T cells). The final volume was adjusted to 400ul / well by adding cytokine supplemented SFEM stem cell media. The cells were centrifuged at 650 x g for 1.5 hours at 37°C with an acceleration of 2 and a brake of 1. After centrifugation, the plate was placed into 37C incubator for 1h, before adding 500 microliters of prewarmed stem cell media on top of the LSKs and overnight incubation. Transduced LSKs were implanted into donor mice irradiated with two doses of 600rads, spaced four hours apart, by intravenous tail vein injection immediately following the second irradiation.

Genomic DNA extraction: Since the number of Vex+ tumor Tregs was very low in any individual mouse and because the mice all share the same genetic background, we decided to pool all tumor Tregs and tumor CD4s together across mice before the gDNA extraction step in order to reliably purify gDNA with sufficient yield. After the gDNA extractions, the extracted gDNA was split evenly into 8 (for cohort 1) or 4 (for cohort 2) separate technical replicates and library prep PCRs and NGS were done individually to all these technical replicates. In other words, genomic DNA was extracted by pooling all the FACS sorted Vex+ tumor Tregs (or tumor CD4 cells) from all the mice within each cohort and lysing the cells with 400ul of RIPA-buffer + RNAseA, followed by 1h incubation in 65C. After this, 400ul of Phenol/Chloroform/Isoamyl alcohol was added, followed by 6 min centrifugation at room temperature. Finally, the gDNA

was recovered by Isopropanol precipitation. For spleen Tregs and spleen CD4 all the gDNA extractions were done individually for each mouse-sample (not pooled together at the lysis-stage as with tumor Tregs and tumor CD4s), since the number of extracted Vex+ cells was much higher than with tumor Tregs / CD4s. Otherwise, the protocol was identical compared to gDNA extractions from tumor Tregs and CD4s.

Preparation of NGS libraries from the extracted gDNA: NGS libraries were prepared from extracted gDNAs following a 2-step PCR protocol with 2 x KAPA Mastermix (KK2612, KAPA Biosystems). For spleen Tregs and CD4s, individually purified gDNAs were pooled before the NGS library prep PCRs. This was done by pooling Spleen Tregs and CD4s in the same ratio as Tumor Tregs and CD4s previously pooled for gDNA extraction as measured by Vex+ FACS cell count. Before the 1st PCR, all pooled Treg and CD4 samples were split into 8 or 4 (first and second cohort) technical replicates, which were amplified separately and with different sample indexes. Correlation between replicates by gDNA frequency was assessed in each cohort and for each set of replicates following library sequencing (Figure 83A, 83B). Both 1st and the 2nd PCRs were stopped before amplification started to saturate in order to avoid biases in the library coverage. The following primers and PCR programs were used for the NGS library preps:

TREG_NGS_1F: GGACTATCATATGCTTACCGTAACTTGAAAGTAATTGT
TREG_NGS_1R: GAAGATCCGGGTGACGCTGCGAACGGACGT

1st PCR:

gDNA	12.5 - 25% of pooled material (depending on the cohort)
2 x KAPA mastermix	12.5ul
Oligo_F(10uM)	1ul
Oligo_R(10uM)	1ul
H2O	to 25ul

95C 3min
98C 20s
60C 15s
72C 20s

72C 1min

TREG_NGS_2F:
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC
T(0-8nt stagger)TTGTGGAAAGGACGAAACACCG
TREG_NGS_2R:
CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTGACTGGAGTTCAGACGTGTGCT
CTTCCGATCTTCTACTATTCTTTCCCCTGCACTGT
NNNNNNNN = sample-index

1:50 diluted DNA template from PCR 1	8ul
2 x KAPA mastermix	12.5ul
Oligo_F(10uM)	1ul
Oligo_R(10uM)	1ul
H2O	to 25ul

95C 3min
98C 30s
52.5C 15s
72C 20s
72C 1min

After the 2nd PCR, samples were gel purified (GenJet), pooled and sequenced with Illumina. Sequencing reads were aligned to a reference of sgRNA template sequences by kallisto to determine a counts matrix of reads per guide for each sample. Differential frequency of guides targeting the same gene in Tumor Treg vs Peripheral Treg and Tumor Treg vs Tumor CD4nonTreg was assessed by DESeq with Bonferroni correction on the p-values, separately (Figure 81), and then p-values across cohorts were integrated by Stouffer's Method (Figure 83C-F).

High-Throughput Treg-Directed Drug Screening: From an initial library of 1,554 FDA-approved or investigational oncology compounds (SelleckChem), single-dose viability screening was performed in vitro on human Tregs sorted from Buffy Coat peripheral blood mononuclear cells (PBMCs). 195 compounds were identified which reduced peripheral Treg growth by at least 60% relative to DMSO control at 5uM. For these, dose-response titrations were performed to

identify the IC20 dose at which peripheral Treg growth is inhibited by 20%, either by direct toxicity to Tregs or inhibition of Treg cell division. Subsequently, Tumor-Infiltrating Tregs were sorted from a large clear cell renal carcinoma tumor and plated with Treg-expansion beads in culture for 5 days, resulting in 5-million Tumor-Infiltrating Tregs. These were suspended at 160,000cells/mL and divided among 2 replicate plates for downstream RNA-Sequencing (PLATE-Seq) and 1 plate for viability testing in comparison to peripheral Tregs at the peripheral Treg IC20 dose. Seven drugs with significantly greater toxicity to tumor Tregs vs peripheral Tregs were identified (Figure 84A).

Wells of drug-treated Tregs were RNA-Sequenced and each normalized with viperSignature against the internal DMSO-control wells on the same PLATE. VIPER was run on the normalized gene expression using the T-cell ARACNe network inferred from sorted bulk-RNA-Sequencing clinical data. Drugs were ranked on their overall inversion across patients of the 17-gene Master Regulator signature previously identified and validated by CRISPR (Figure 81B), as well as on their patient-by-patient inversion of Tumor-Treg vs Peripheral-Treg protein activity signature by OncoTreat (Figure 86).

Tumor-Growth Screens: We assessed tumor growth first in response to treatment with floxuridine, triapine, and gemcitabine relative to untreated control, with or without anti-PD1 immunotherapy (Fig 82B-C). 10 C57BL/6J mice per treatment arm were implanted with subcutaneous MC38 tumor cells. Treatment was initiated after 12 days of initial tumor growth, at which point mice were monitored for tumor volume until exceeding 1000mm^3 or ulceration exceeding a diameter of 5mm. Gemcitabine was administered IP on days 12, 15, and 18 at 12

mg/kg, or 1/10th of the lowest conventional clinical-equivalent dose in mice (120 mg/kg). Floxuridine and triapine were IP daily from day 12-18 at 1mg/kg and 5mg/kg, respectively, also reflecting 1/10th the standard murine dose. Mice receiving anti-PD-1 were administered anti-PD-1 IP on days 12, 15, and 18. Treatment response outcomes were assessed by cox proportional hazards model (Figure 82B), Kaplan-Meier curve (Figure 82C), and computation of mean tumor growth slope over time. By all criteria, gemcitabine was the only treatment found to significantly inhibit tumor growth, alone and in combination with anti-PD1.

To further assess the doses at which gemcitabine inhibits tumor growth and the immune-mediated effects of gemcitabine, we performed parallel dose titrations of Gem in immune-competent C57BL/6J mice and immune-deficient NSG (NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ) mice, administering doses ranging from 0.12mg/kg up to 120mg/kg as shown in Figures 83A-83C. Six mice were treated per treatment arm. Doses were administered IP on days 12, 15, and 18, and treatment response was assessed by cox proportional hazards model (Figure 83D) and Kaplan-Meier test (Figure 83E).

Finally, tumor growth was assessed in single-gene TRPS1 CRISPRko generated by the CHIME protocol described above, compared to transduction by CHIME with a non-targeting scramble control guide. These cohorts included 6 TRPS-KO mice and 5 Scramble-control mice. For these mice, we pooled two guides targeting TRPS1 and two non-targeting guides with approx. MOI 50 based on 293T cell line tittering. These guide sequences were:

TRPS1_1: AGAGGGGCAGACATCCTACG
TRPS1_2: AGCATCGGATGTCAAACAGG
Non-targeting guide 1: GCGAGGTATTCGGCTCCGCG
Non-targeting guide 2: GCTTTCACGGAGGTTCGACG

Following immune reconstitution, mice were initially implanted with subcutaneous MC38 tumor, which spontaneously regressed in both arms following initial tumor growth for two weeks post-implantation. Subsequently, these mice were implanted with subcutaneous MCA205, a more aggressive and immune-resistant fibrosarcoma cell line. Tumor volume was assessed every 48 hours following day 7 post-implantation, such that tumor volumes in TRPS1 mice3 were determined to be significantly lower than scramble controls by day 13 ($p < 0.05$). Treatment response was assessed by cox proportional hazards model and Kaplan-Meier test.

Single-Cell RNA-Seq Profiling of Gemcitabine Effect on TI-Tregs: To test the hypothesis that low-dose Gem modulates TI-Tregs, we performed single cell RNA sequencing of MC38 tumor- and spleen-derived Tregs 24 hours after exposure to a single dose of 12 mg/kg Gem as well as 24 hours after vehicle control. For this study, we implanted *FoxP3*$^{Yfp\text{-}Cre}$ mice with MC38 to facilitate flow-sorting of TCR-β^+ $CD4^+$ $FoxP3^+$ Tregs from tumor and spleen specifically by the YFP marker. Tissue was harvested at day 14 post tumor-implantation, and fresh tissue was minced to 2-4 mm sized pieces in a 6-cm dish and subsequently digested to single cell suspension using Multi Tissue Mouse Tumor Dissociation Kit 1 (Miltenyi Biotec) and a gentleMACS OctoDissociator (Miltenyi Biotec) according to the manufacturer's instructions.

Dissociated cells were flow-sorted for YFP^+ Tregs and processed for single-cell gene expression capture (scRNASeq) using the 10X Chromium 3' Library and Gel Bead Kit (10x Genomics), following the manufacturer's user guide at the Columbia University Genome Center. After GelBead in-Emulsion reverse transcription (GEM-RT) reaction, 12-15 cycles of polymerase

chain reaction (PCR) amplification were performed to obtain cDNAs used for RNAseq library generation. Libraries were prepared following the manufacturer's user guide and sequenced on Illumina NovaSeq 6000 Sequencing System. Single-cell RNASeq data were processed with Cell Ranger software at the Columbia University Single Cell Analysis Core. Illumina base call files were converted to FASTQ files with the command "cellranger mkfastq." Expression data were processed with "cellranger count" on pre-built mouse reference. Cell Ranger performed default filtering for quality control, and produced a barcodes.tsv, genes.tsv, and matrix.mts file containing transcript counts for each cell, such that expression of each gene is in terms of the number of unique molecular identifiers (UMIs) tagged to cDNA molecules corresponding to that gene.

These data were loaded into the R version 3.6.1 programming environment, where the publicly available Seurat package was used to further quality-control filter cells to those with fewer than 25% mitochondrial RNA content, more than 1,000 unique UMI counts, and fewer than 15,000 unique UMI counts. Pooled distribution of UMI counts, unique gene counts, and percentage of mitochondrial DNA after QC-filtering is shown in Figure 85A, along with the number of sorted Tregs captured per sample. Gene Expression UMI count matrix was processed in R using the Seurat SCTransform command followed by Seurat Anchor-Integration. The sample was clustered on gene expression by a Resolution-Optimized Louvain Algorithm [35]. Protein activity was inferred for all cells by VIPER using the SCTransform gene expression signature and the T-cell ARACNe network derived from sorted T-cell bulk-RNA-Seq. The single-cell data were then re-clustered on VIPER protein activity (Figure 85B). Top 5 most differentially upregulated proteins per cluster were assessed by t-test (Figure 85C). Enrichment of the TI-Treg

MRs was assessed by Gene Set Enrichment Analysis (GSEA) on a cell-by-cell basis, with normalized enrichment scores shown in Figure 84B and protein activity of the individual MRs shown in Figure 84C. Cluster frequencies were plotted for each sample (Vehicle-Treated Tumor, Vehicle-Treated Spleen, Gem-Treated Tumor, Gem-Treated Spleen), with pairwise comparisons in frequency assessed by Fisher's Exact test (Figure 84D).

Conclusion

In this thesis I have developed and validated a suite of novel tools leveraging protein activity inference in single cells to profile the micro-environment of distinct tumor types and presented a platform to accelerate both drug repurposing in the context of immunotherapy resistance and rapid validation of novel therapeutic targets. The algorithms developed and successfully applied to elucidate aspects of tumor immune biology include the analytical pipeline for single-cell protein activity inference described, comprehensively validated, and benchmarked in Chapter 1, as well as the novel resolution-optimized Louvain clustering approach and receptor-ligand interaction inference described in Chapter 2, and the single-cell OncoTreat/OncoTarget drug prediction algorithms first described in Chapter 3. With these tools, the micro-environment of various tumors and their response to immunotherapy may be profiled at greater resolution than previously possible, with detection of key signaling and regulatory proteins which control downstream cell phenotype. In aggregate, we have discovered that different tumor types have different populations of cells driving therapy response and resistance, with opportunity to tailor therapies both to particular tumor types and to individual patients, targeting both tumor cells and cells in their immune microenvironment.

Critically, this work has demonstrated the successful application of cell sub-population markers derived from single-cell protein activity inference to perform Gene Set Enrichment Analysis in larger clinically-annotated bulk-RNA Sequencing databases such as TCGA, enabling independent validation of clinically significant immune and tumor sub-cluster signatures initially discovered in small single-cell RNA-Seq patient cohorts using much larger bulk RNA-Seq

cohorts. In Chapter 2, we use this approach to discover the association of our newly discovered tumor-specific C1Q+/TREM2+/APOE+ Macrophages in clear cell renal carcinoma with early post-surgical recurrence in two independent bulk-RNA-Seq cohorts, where prognostic significance of this population and its VIPER-discovered markers was additionally validated by quantitative immunofluorescence. In Chapter 4, the same approach enabled us to identify which sub-populations of head-and-neck cancer fibroblasts were associated with improved or worsened response to anti-PD1 immunotherapy, which was confirmed by flow cytometry and functional co-culture experiments. In Chapter 7, this approach was extended to prostate tumor sub-clusters to confirm that populations with distinct inferred drug sensitivity to TOP2A which were observed to be enriched at baseline in non-responders to immunotherapy are also enriched in TCGA among patients with shorter recurrence-free survival times. Finally, in Chapter 5, we take the single-cell derived Master Regulator enrichment approach to the pure biomarker discovery task of classifying responders vs. non-responders to anti-PD1 checkpoint immunotherapy in melanoma, where we achieve excellent classification accuracy across two independent bulk-RNA sequencing cohorts, although the functional role of the biomarker proteins identified remains to be explored.

In a broad systems biology approach leveraging protein activity inference, we can discover that different cell clusters associate with outcome in different tumor types, such that we can use these populations as biomarkers for treatment prioritization, but also target their active Master Regulator proteins and inferred drug sensitivities in combination with immunotherapy to potentially sensitize non-responders. This analysis pipeline represents a substantial improvement on previous approaches of analyzing single cell data and has yielded multiple actionable therapy

targets across a broad range of tumor types and cell types. In particular, we have identified the following:

In Chapter 2, a population of recurrence-associated C1Q+/TREM2+/APOE+ Macrophages in clear cell renal carcinoma, but not in the other tumor types examined, indicating that intra-tumoral frequency of this population may effectively risk-stratify patients for prioritization of more aggressive up-front treatment options and targeting it for depletion may provide clinical benefit in ccRCC.

In Chapter 3, we observe a significant and diverse immune infiltrate in cholangiocarcinoma, such that traditional bulk RNA-Sequencing is dominated by non-tumor cells, but our approach allows us to isolate and infer the sensitivity of cholangiocarcinoma tumor cells themselves to plicamycin and dacinostat, validated in a pre-clinical Patient-Derived Xenograft model. Although limited to a case report, this precision medicine approach and the two drug candidates discovered merit clinical follow-up, since there are no effective treatment options for cholangiocarcinoma currently in the clinic.

In Chapter 4, we find distinct fibroblast sub-clusters in head and neck squamous cell carcinoma (HNSCC) associated with improved response to immunotherapy which decrease T-cell exhaustion in co-culture but have not been described in previous studies of fibroblasts in breast cancer or pancreatic adenocarcinoma, and appear to be HNSCC-specific. Stimulation of these fibroblasts ex vivo may represent a novel paradigm for combination with checkpoint inhibitor immunotherapy both in HNSCC and potentially if injected in other tumor contexts.

In Chapter 6, we discover that androgen deprivation therapy (ADT) in prostate cancer is actually immunogenic, stimulating a robust infiltration by cytotoxic CD8 T-cells which is then neutralized by an infiltration of immunosuppressive regulatory T-cells, or Tregs, although modestly improved outcomes were observed in combination of ADT with a GVAX immune-therapy. The discovery of immunogenic effect of ADT has already led to active clinical trials of combination ADT with immune checkpoint inhibitors, in both the primary and metastatic contexts, which we are profiling at the single-cell level and discuss further in Chapter 7, as well as motivating the inhibition of Tumor-Infiltrating Tregs, which are observed not only in prostate cancer but across a range of tumor types, which is further explored in Chapter 8.

In Chapter 7, we discover among other things that in combination of androgen deprivation plus immunotherapy-treated prostate cancer, infiltration at baseline by CD4 T-cells with high activity of TNF inferred by VIPER is associated with worse clinical outcomes, as well as that infiltrating CD8 T-cells have high activity of the immune checkpoint LAG3, suggesting a rationale for combination of anti-PD1 checkpoint immunotherapy in prostate cancer with both TNF and LAG3 inhibitors, which are currently in clinical investigation across a range of other tumor types. Furthermore, we discover recurrence-associated tumor cell subpopulations with inferred sensitivity to TOP2A inhibitors, the signature for which is also enriched among patients in TCGA with early post-treatment recurrence. This suggests TOP2A inhibitors as an additional candidate for combination therapy in metastatic prostate cancer.

Finally, in Chapter 8 we discover Master Regulators of tumor infiltration by immunosuppressive regulatory T-cells from a clinical database including prostate cancer, bladder cancer, renal cell carcinoma, and glioblastoma. We present here an entire platform for functional validation of immune Master Regulators in the tumor micro-environment by CRISPR as well as a high-throughput drug screening approach for rapid repurposing of FDA-approved and investigational oncology compounds with beneficial off-target effects on the Treg transcriptional profile resulting in reduced tumor-infiltration. By this approach, we make two clinically significant discoveries with potential pan-cancer implications as a novel approach for immunotherapeutic treatment, targeting tumor-infiltrating Tregs without significantly depleting circulating Tregs or tumor-infiltrating non-Treg T-cells. First, we identify very-low-dose gemcitabine as differentially toxic to tumor-infiltrating Tregs, with significant inhibition of tumor growth rate in mouse models alone and in combination with anti-PD1 checkpoint immunotherapy. In addition, we identify single-gene knockout of TRPS1 (the most strongly validated Tumor-Treg Master Regulator) in the hematopoietic lineage as having a significant effect on tumor growth, leading to spontaneous rejection in 50% of mice implanted with MCA205 tumor model.

This final discovery in effective inhibition of Tumor-Specific Tregs based on discoveries from protein activity inference represents a very useful paradigm for high-throughput ex vivo drug screening and functional validation of putative immune micro-environment treatment targets by hematopoietic stem cell CRISPR. Such an approach can readily be extended beyond Tregs to macrophages such as those discovered in clear cell renal carcinoma, or even to fibroblasts, in order to facilitate rapid drug repositioning and prioritize agents for clinical development and trials.

Future Directions

In addition to pursuing pre-clinical and clinical trials targeting the treatment resistance associated elements of immune micro-environment described above, we are actively pursuing additional work in progress clarifying the functional role of prognostic markers of melanoma anti-PD1 response discovered in Chapter 5, as well as extending the Treg drug screening platform described in Chapter 8 to other immunosuppressive cell types, including Macrophages and myeloid-derived suppressor cells. Work is ongoing in single-cell profiling of our clinical trial testing combination ADT with anti-CTLA4 immunotherapy in primary prostate cancer, awaiting follow-up to define treatment responders and non-responders. We are also collaborating with Abate-Shen lab to profile circulating tumor cells by single-cell RNA-Sequencing in a mouse model of prostate cancer metastasizing to the bone [221], from which we may characterize the transition from primary to metastatic prostate tumor phenotype with respect to both tumor cells and micro-environment. We are also accruing single-cell RNA-Sequencing data from a clinical trial of il1-beta inhibition in clear cell renal carcinoma, which has been shown in mice to remodel the tumor macrophage compartment and reduce tumor growth rate [229], with significant implications for differentiation or depletion of the C1Q/TREM2/APOE Macrophage phenotype we observe in treatment-naïve renal carcinoma.

On the technical side, we are working to further improve the analysis pipeline described in Chapter 1 and have implemented a protein-activity-based cell type inference to replace SingleR in the workflow, as well as working to incorporate updated versions of ARACNe. As the pipeline has been designed with intentional modularity, any future changes to the workflow will be benchmarked for improvement in technical and biological robustness against the benchmarks

described in Chapter 1. Furthermore, since we have identified that given high-quality context matched ARACNe networks protein activity inferences are robust to significant data dropout, we are actively compiling a resource of ARACNe networks for each cell type reported in the Human Single Cell Type Atlas [230], to be included in the final publication of our pipeline validation manuscript as a resource for the community. The Single Cell Type Atlas aggregates 4,012,680 single cells from 13 different healthy human tissues, reporting 192 individual cell type clusters corresponding to 12 major cell type lineages. Given the selection for high data quality in this resource, ARACNe networks compiled across the entire range of cell types and tissue sites represented in this atlas may be used by researchers with significantly lower-quality data (Mean UMI counts below 5000 UMIs/cell) to salvage protein activity inference using ARACNe networks in the same cell lineage as the cells being profiled.

Finally, the principles and discoveries demonstrated in this thesis extend beyond the study of tumor micro-environment response to immunotherapy to include other forms of cancer treatment, including an active study of response to radiation therapy by protein activity inference on single-cell RNA-Sequencing at varying levels of radiation over time, where we have discovered and are further studying a resistance to radiation among myeloid cells relative to T-cells. The analysis pipeline developed can be extended even to study of non-malignant pathologies, such as organ transplant rejection, autoimmune disease, etc., where the immunosuppressive populations we have identified may actually be of clinical benefit. In summary, this thesis presents a broad systems biology approach to studying drivers of single-cell state in association with disease and a set of discoveries in the field of cancer immunotherapy

identifying cell populations driving subsequent response or non-response, with significant clinical implications for combination therapy, and extensive potential for future study.

References

[1] L. Galluzzi et al., "Immunological Effects of Conventional Chemotherapy and Targeted Anticancer Agents," *Cancer Cell,* vol. 28, no. 6, pp. 690-714, 2015.

[2] G. T. Gibney et al., "Predictive Biomarkers for Checkpoint Inhibitor-Based Immunotherapy," *The Lancet Oncology,* vol. 17, no. 12, 2016.

[3] G. Schepisi et al., "Inflammatory Biomarkers as Predictors of Response to Immunotherapy in Urological Tumors," *Journal of Oncology,* pp. 1-11, 2019.

[4] Y.-C. Shen et al., "Combining Intratumoral Treg Depletion with Androgen Deprivation Therapy (ADT): Preclinical Activity in the Myc-CaP Model," *Prostate Cancer and Prostatic Diseases,* vol. 21, no. 1, p. 113–125, April 2018.

[5] W. Hugo et al., "Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma," *Cell,* vol. 168, no. 3, p. 542, 2017.

[6] J. Kalina et al., "Immune Modulation by Androgen Deprivation and Radiation Therapy: Implications for Prostate Cancer Immunotherapy," *Cancers,* vol. 9, no. 12, p. 13, 2017.

[7] N. Riaz et al., "Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab," *Cell,* vol. 171, no. 4, p. 934–949, November 2017.

[8] Z. Lopez-Bujanda and C. G. Drake, "Myeloid-Derived Cells in Prostate Cancer Progression: Phenotype and Prospective Therapies," *Journal of Leukocyte Biology,* vol. 102, no. 2, p. 393–406, 2017.

[9] R. Noy and J. W. Pollard, "Tumor-Associated Macrophages: From Mechanisms to Therapy," *Immunity,* vol. 41, no. 5, p. 866, 2014.

[10] S. Ostrand-Rosenberg et al., "Cross-Talk between Myeloid-Derived Suppressor Cells (MDSC), Macrophages, and Dendritic Cells Enhances Tumor-Induced Immune Suppression," *Seminars in Cancer Biology,* vol. 22, no. 4, p. 275–281, 2012.

[11] R. Saleh and E. Elkord, "Treg-Mediated Acquired Resistance to Immune Checkpoint Inhibitors," *Cancer Letters,* vol. 457, p. 168–179, 2019.

[12] C. G. Drake et al., "Androgen Ablation Mitigates Tolerance to a Prostate/Prostate Cancer-Restricted Antigen," *Cancer Cell,* vol. 7, no. 3, p. 239–249, 2005.

[13] J. A. Wargo et al., "Immune Effects of Chemotherapy, Radiation, and Targeted Therapy and Opportunities for Combination With Immunotherapy," *Seminars in Oncology,* vol. 42, no. 4, p. 601–616, 2015.

[14] L. A. Emens and G. Middleton, "The Interplay of Immunotherapy and Chemotherapy: Harnessing Potential Synergies," *Cancer Immunology Research,* vol. 3, no. 5, p. 436–443, 2015.

[15] C. W. Wanderley et al., "Paclitaxel Reduces Tumor Growth by Reprogramming Tumor-Associated Macrophages to an M1- Profile in a TLR4-Dependent Manner," *Cancer Research,* 2018.

[16] G. Zheng et al., "Massively Parallel Digital Transcriptional Profiling of Single Cells," *Nature Communications,* 16 January 2017.

[17] M. Sade-Feldman et al., "Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma," *Cell,* vol. 175, no. 4, p. 998–1013, November 2018.

[18] L. Jerby-Arnon et al., "A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade," *Cell,* vol. 175, no. 4, p. 984–997, November 2018.

[19] Q. Wang et al., "Single-Cell Profiling Guided Combinatorial Immunotherapy for Fast-Evolving CDK4/6 Inhibitor-Resistant HER2-Positive Breast Cancer," *Nature Communications,* vol. 10, no. 1, 2019.

[20] T. Stuart et al., "Comprehensive integration of single-cell data," *Cell,* vol. 177, no. 7, pp. 1888-902, 13 June 2019.

[21] M. J. Alvarez et al., "Functional Characterization of Somatic Mutations in Cancer Using Network-Based Inference of Protein Activity," *Nature Genetics,* vol. 48, no. 8, p. 838–847, 2016.

[22] A. Califano and M. J. Alvarez, "The Recurrent Architecture of Tumour Initiation, Progression and Drug Sensitivity," *Nature Reviews Cancer,* vol. 17, no. 2, p. 116–130, 2016.

[23] A. Lachmann et al., "ARACNe-AP: Gene Network Reverse Engineering through Adaptive Partitioning Inference of Mutual Information," *Bioinformatics,* vol. 32, no. 14, p. 2233–2235, 2016.

[24] M. J. Alvarez et al., "A Precision Oncology Approach to the Pharmacological Targeting of Mechanistic Dependencies in Neuroendocrine Tumors," *Nature Genetics,* vol. 50, no. 18, p. 979–989, June 2018.

[25] F. M. Giorgi et al., "Inferring Protein Modulation from Gene Expression Data Using Conditional Mutual Information," *PLoS ONE,* vol. 9, no. 10, 2014.

[26] H. Ding et al., "Quantitative Assessment of Protein Activity in Orphan Tissues and Single Cells Using the MetaVIPER Algorithm," *Nature Communications,* vol. 9, no. 1, 2018.

[27] He P et al., "The changing mouse embryo transcriptome at whole tissue and single-cell resolution," *Nature,* vol. 583, no. 7818, pp. 760-767, July 2020.

[28] Qian J et al., "A pan-cancer blueprint of the heterogenous tumor microenvironment revealed by single-cell profiling," *Cell Res,* vol. 30, no. 9, pp. 745-762, September 2020.

[29] J. Zhao et al., "Single-cell RNA sequencing reveals the heterogeneity of liver-resident immune cells in human," *Cell Discovery,* vol. 6, no. 22, 2020.

[30] G. Xu et al., "The differential immune responses to COVID-19 in peripheral and lung revealed by single-cell RNA sequencing," *Cell Discovery,* vol. 6, no. 73, 20 October 2020.

[31] E. Speranza et al., "Single-cell RNA sequencing reveals SARS-CoV-2 infection dynamics in lungs of African green monkeys," *Sci Transl Med,* vol. 13, no. 578, 27 Jan 2021.

[32] G. Finak et al., "MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data," *Genome Biology,* vol. 16, no. 278, 2015.

[33] E. Elyada et al., "Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts," *Cancer discovery,* vol. 9, pp. 1102-1123, 2019.

[34] W. Chung et al., "Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer," *Nature Communications,* vol. 8, no. 15081, 5 May 2017.

[35] A. Obradovic et al., "Single-Cell Protein Activity Analysis Identified Recurrence-Associated Renal Tumor Macrophages," *Cell,* vol. 184, no. 11, pp. 2988-3005, 27 May 2021.

[36] A. Butler et al., "Integrating single-cell transcriptomic data across different conditions, technologies, and species," *Nature Biotechnology,* vol. 36, pp. 411-420, 2018.

[37] V. Thorsson et al., "The Immune Landscape of Cancer," *Immunity,* vol. 48, no. 4, pp. 812-830, 17 April 2018.

[38] R. Kushwaha et al., "Interrogation of a context-specific transcription factor network identifies novel regulators of pluripotency," *Stem Cells,* vol. 33, no. 2, pp. 367-377, February 2015.

[39] P. Laise et al., "Pancreatic ductal adenocarcinoma comprises coexisting regulatory states with both common and distinct dependencies," *bioRxiv,* 2020.

[40] K. Basso et al., "Reverse engineering of regulatory networks in human B cells," *Nature Genetics,* vol. 37, no. 4, pp. 382-90, 2005.

[41] P. Rajbhandari et al., "Cross-Cohort Analysis Identifies a TEAD4-MYCN Positive Feedback Loop as the Core Regulatory Element of High-Risk Neuroblastoma," *Cancer Discovery,* vol. 8, no. 5, pp. 582-599, May 2018.

[42] M. Carro et al., "The transcriptional network for mesenchymal transformation of brain tumours," *Nature,* vol. 463, pp. 318-325, 2010.

[43] A. Aytes et al., "ETV4 promotes metastasis in response to activation of PI3-kinase and Ras signaling in a mouse model of advanced prostate cancer," *Proceedings of the National Academy of Sciences,* vol. 110, no. 37, pp. E3506-E3515, 2013.

[44] T. Zeleke et al., "Network-based assessment of HDAC6 activity is highly predictive of pre-clinical and clinical responses to the HDAC6 inhibitor ricolinostat," *medRxiv,* 2020.

[45] K. Mani et al., "Systems biology and personalized medicine in cancer," *Current Pharmacogenomics and Personalized Medicine,* vol. 8, no. 1, pp. 64-72, 2010.

[46] M. Stoeckius et al., "Simultaneous epitope and transcriptome measurement in single cells," *Nature Methods,* vol. 14, pp. 865-868, September 2017.

[47] A. Teschendorff et al., "Single-cell entropy for accurate estimation of differentiation potency from a cell's transcriptome," *Nature Communications,* vol. 8, no. 15599, 2017.

[48] D. Aran et al., "Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage," *Nature Immunology,* vol. 20, no. 2, pp. 163-172, February 2019.

[49] J. Martens et al., "BLUEPRINT: mapping human blood cell epigenomes," *Haematologica,* vol. 98, no. 10, pp. 1487-1489, 2013.

[50] The ENCODE Project Consortium et al., "An integrated encyclopedia of DNA elements in the human genome," *Nature,* vol. 489, p. 57–74, 2012.

[51] H. Koul et al., "Molecular aspects of renal cell carcinoma: a review," *American journal of cancer research,* vol. 1, no. 2, pp. 240-254, 2011.

[52] A. Sánchez-Gastaldo et al., "Systemic treatment of renal cell cancer: A comprehensive review," *Cancer Treatment Reviews,* vol. 60, pp. 77-89, 2017.

[53] T. Davoli et al., "Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy," *Science,* vol. 355, 2017.

[54] Y. Senbabaoglu et al., "Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures," *Genome Biology,* vol. 17, no. 231, 2016.

[55] S. Turajlic et al., "Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis," *The Lancet Oncology,* vol. 18, pp. 1009-1021, 2017.

[56] X. Li et al., "Prognostic value of CD44 expression in renal cell carcinoma: a systematic review and meta-analysis," *Scientific Reports,* vol. 5, no. 13157, 2015.

[57] S. Chevrier et al., "An Immune Atlas of Clear Cell Renal Cell Carcinoma," *Cell,* vol. 169, no. 4, pp. 736-749, 2017.

[58] J. Chao et al., "Unlocking the Complexities of Tumor-Associated Regulatory T Cells," *Journal of immunology,* vol. 200, pp. 415-421, 2018.

[59] K. Mani et al., "A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas," *Molecular Systems Biology,* vol. 4, no. 169, 2008.

[60] F. Vargas et al., "Fc-Optimized Anti-CD25 Depletes Tumor-Infiltrating Regulatory T Cells and Synergizes with PD-1 Blockade to Eradicate Established Tumors," *Immunity,* vol. 46, no. 4, pp. 577-586, 2017.

[61] E. Becht et al., "Prognostic and Theranostic Impact of Molecular Subtypes and Immune Classifications in Renal Cell Cancer (RCC) and Colorectal Cancer (CRC)," *OncoImmunology,* vol. 4, no. 12, July 2015.

[62] O. Khan et al., "TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion," *Nature,* vol. 571, pp. 211-218, 2019.

[63] A. Scott et al., "TOX is a critical regulator of tumour-specific T cell differentiation," *Nature,* vol. 571, p. 270–274, 2019.

[64] A. Dutta et al., "Identification of an NKX3.1-G9a-UTY transcriptional regulatory network that controls prostate differentiation," *Science,* vol. 352, pp. 1576-1580, 2016.

[65] F. Talos et al., "A computational systems approach identifies synergistic specification genes that facilitate lineage conversion to prostate tissue," *Nature communications,* vol. 8, no. 14662, 2017.

[66] K. Arumugam et al., "The Master Regulator Protein BAZ2B Can Reprogram Human Hematopoietic Lineage-Committed Progenitors into a Multipotent State," *Cell Reports,* vol. 33, no. 10, p. 108474, 8 December 2020.

[67] A. Barkal et al., "CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy," *Nature,* vol. 572, pp. 392-396, 2019.

[68] D. Arik et al., "Prognostic Significance of CD24 in Clear Cell Renal Cell Carcinoma," *Pathology & Oncology Research,* vol. 23, pp. 409-416, 2017.

[69] H. S. Tan et al., "KRT8 upregulation promotes tumor metastasis and is predictive of a poor prognosis in clear cell renal cell carcinoma," *Oncotarget,* vol. 8, no. 44, p. 76189, 29 September 2017.

[70] N. J. Farber et al., "Renal cell carcinoma: the search for a reliable biomarker," *Translational cancer research,* vol. 6, no. 3, pp. 620-632, 2017.

[71] J. J. Hsieh et al., "Chromosome 3p Loss–Orchestrated VHL, HIF, and Epigenetic Deregulation in Clear Cell Renal Cell Carcinoma," *Journal of Clinical Oncology,* vol. 36, no. 36, p. 3533–3539, 2018.

[72] M. Lizio et al., "Update of the FANTOM web resource: expansion to provide additional transcriptomic atlases," *Nucleic Acids Research,* vol. 47, pp. D752-D758, 2019.

[73] J. Jacobs et al., "CD70: An emerging target in cancer immunotherapy," *Pharmacology & therapeutics,* vol. 155, pp. 1-10, 2015.

[74] L. Jilaveanu et al., "CD70 expression patterns in renal cell carcinoma," *Human Pathology,* vol. 43, pp. 1394-1399, 2012.

[75] L. Roumenina et al., "Tumor Cells Hijack Macrophage-Produced Complement C1q to Promote Tumor Growth," *Cancer Immunology Research,* vol. 7, pp. 1091-1105, 2019.

[76] A. Ramirez et al., "Functional characterization of a novel TREM2 coding variant linked to familial Alzheimer's disease," *Alzheimer's & Dementia,* vol. 11, p. P500, 2015.

[77] P. Roussos et al., "The triggering receptor expressed on myeloid cells 2 (TREM2) is associated with enhanced inflammation, neuropathological lesions and increased risk for Alzheimer's dementia," *Alzheimers & Dementia,* vol. 11, pp. 1163-1170, 2015.

[78] X. Wang et al., "Overexpression of TREM2 enhances glioma cell proliferation and invasion: a therapeutic target in human glioma," *Oncotarget,* vol. 7, pp. 2354-2366, 2016.

[79] Y. Yao et al., "TREM-2 serves as a negative immune regulator through Syk pathway in an IL-10 dependent manner in lung cancer," *Oncotarget,* vol. 7, pp. 29620-29634, 2016.

[80] Y. Katzenelenbogen et al., "Coupled scRNA-Seq and Intracellular Protein Activity Reveal an Immunosuppressive Role of TREM2 in Cancer," *Cell,* vol. 182, no. 4, pp. 872-885.e19, 20 August 2020.

[81] E. Brunt et al., "cHCC-CCA: Consensus terminology for primary liver carcinomas with both hepatocytic and cholangiocytic differentation," *Hepatology,* vol. 68, no. 1, pp. 113-126, 2018.

[82] R. Dhanasekaran et al., "Treatment outcomes and prognostic factors of intrahepatic cholangiocarcinoma," *Oncology reports,* vol. 29, no. 4, pp. 1259-1267, 2013.

[83] A. Society et al., "Survival Rates for Bile Duct Cancer," 2021. [Online]. Available: https://www.cancer.org/cancer/bile-duct-cancer/detection-diagnosis-staging/survival-by-stage.html.

[84] L. Fabris et al., "The tumour microenvironment and immune milieu of cholangiocarcinoma," *Liver International,* vol. 39, pp. 63-78, 2019.

[85] E. Loeuillard et al., "Targeting tumor-associated macrophages and granulocytic myeloid-derived suppressor cells augments PD-1 blockade in cholangiocarcinoma," *The Journal of clinical investigation,* vol. 130, no. 10, 2020.

[86] S. Rizvi and G. Gores, "Pathogenesis, diagnosis, and management of cholangiocarcinoma," *Gastroenterology,* vol. 145, no. 6, pp. 1215-1229, 2013.

[87] K. Chang et al., "The Cancer Genome Atlas Pan-Cancer analysis project," *Nature genetics,* vol. 45, no. 10, pp. 1113-1120, 2013.

[88] R. Tobin et al., "IL-6 and IL-8 are linked with myeloid-derived suppressor cell accumulation and correlate with poor clinical outcomes in melanoma patients," *Frontiers in oncology,* vol. 9, no. 1223, 2019.

[89] Z. Lopez-Bujanda et al., "TGM4: an immunogenic prostate-restricted antigen," *Journal for immunotherapy of cancer,* vol. 9, no. 6, 2021.

[90] M. Gonzalez-Aparicio and C. Alfaro, "Significance of the IL-8 pathway for immunotherapy," *Human vaccines & immunotherapeutics,* vol. 16, no. 10, pp. 2312-2317, 2020.

[91] L. Li et al., "Human primary liver cancer organoids reveal intratumor and interpatient drug response heterogeneity," *JCI insight,* vol. 4, no. 2, 2019.

[92] D. Jung et al., "CG200745, an HDAC inhibitor, induces anti-tumour effects in cholangiocarcinoma cell lines via miRNAs targeting the Hippo pathway," *Scientific reports,* vol. 7, no. 1, pp. 1-13, 2017.

[93] K. Pant et al., "Role of histone deacetylases in carcinogenesis: potential role in cholangiocarcinoma," *Cells,* vol. 9, no. 3, p. 780, 2020.

[94] R. Ferris et al., "Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck," *New England Journal of Medicine,* vol. 375, no. 19, pp. 1856-1867, 2016.

[95] T. Seiwert et al., "Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial," *Lancet Oncology,* vol. 17, no. 7, pp. 956-965, 2016.

[96] R. Mehra et al., "Efficacy and safety of pembrolizumab in recurrent/metastatic head and neck squamous cell carcinoma: pooled analyses after long-term follow-up in KEYNOTE-012," *British Journal of Cancer,* vol. 119, no. 2, pp. 153-159, 2018.

[97] B. Burtness et al., "Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study," *Lancet,* vol. 394, no. 10212, pp. 1915-1928, 2019.

[98] C. Feig et al., "Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer.," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 110, no. 50, p. 20212, 2013.

[99] C. Dominguez et al., "Single-Cell RNA Sequencing Reveals Stromal Evolution into LRRC15(+) Myofibroblasts as a Determinant of Patient Response to Cancer Immunotherapy," *Cancer Discovery,* vol. 10, no. 2, pp. 232-253, 2020.

[100] Y. Kieffer et al., "Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer," *Cancer Discovery,* vol. 10, no. 9, pp. 1330-1351, 2020.

[101] A. Costa et al., "Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer," *Cancer Cell,* vol. 33, no. 3, pp. 463-479, 2018.

[102] D. Ohlund et al., "Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer," *Journal of Experimental Medicine,* vol. 214, no. 3, pp. 579-596, 2017.

[103] S. Puram et al., "Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer," *Cell,* vol. 171, no. 7, pp. 1611-1624, 2017.

[104] A. Subramanian et al., "Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 102, no. 43, pp. 15545-15550, 2005.

[105] M. E. Spector et al., "Prognostic value of tumor-infiltrating lymphocytes in head and neck squamous cell carcinoma," *JAMA Otolaryngology–Head & Neck Surgery,* vol. 145, no. 11, pp. 1012-1019, 2019.

[106] R. Uppaluri et al., "Neoadjuvant and Adjuvant Pembrolizumab in Resectable Locally Advanced, Human Papillomavirus-Unrelated Head and Neck Cancer: A Multicenter, Phase II Trial," *Clinical Cancer Research,* vol. 26, no. 19, pp. 5140-5152, 2020.

[107] N. Direkze et al., "Bone marrow contribution to tumor-associated myofibroblasts and fibroblasts," *Cancer Research,* vol. 64, no. 23, pp. 8492-5, 2004.

[108] A. Karnoub et al., "Mesenchymal stem cells within tumour stroma promote breast cancer metastasis," *Nature,* vol. 449, no. 7162, pp. 557-63, 2007.

[109] P. Mishra et al., "Carcinoma-associated fibroblast-like differentiation of human mesenchymal stem cells," *Cancer Research,* vol. 68, no. 11, pp. 4331-9, 2008.

[110] Y. Kojima et al., "Autocrine TGF-beta and stromal cell-derived factor-1 (SDF-1) signaling drives the evolution of tumor-promoting mammary stromal myofibroblasts," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 107, no. 46, pp. 20009-14, 2010.

[111] H. Tan et al., "TGFbeta1 is essential for MSCs-CAFs differentiation and promotes HCT116 cells migration and invasion via JAK/STAT3 signaling," *OncoTargets and Therapy,* vol. 12, pp. 5323-5334, 2019.

[112] J. Larkin et al., "Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma," *New England Journal of Medicine,* vol. 373, no. 1, pp. 23-34, 2015.

[113] M. Lakins et al., "Cancer-associated fibroblasts induce antigen-specific deletion of CD8 (+) T Cells to protect tumour cells," *Nature Communications,* vol. 9, no. 1, p. 948, 2018.

[114] L. Gorchs et al., "Human Pancreatic Carcinoma-Associated Fibroblasts Promote Expression of Co-inhibitory Markers on CD4(+) and CD8(+) T-Cells," *Frontiers in Immunology,* vol. 10, p. 847, 2019.

[115] B. Park et al., "TGFbeta1-Mediated SMAD3 Enhances PD-1 Expression on Antigen-Specific T Cells in Cancer," *Cancer Discovery,* vol. 6, no. 12, pp. 1366-1381, 2016.

[116] C. Hutton et al., "Single-cell analysis defines a pancreatic fibroblast lineage that supports anti-tumor immunity," *Cancer Cell,* vol. 39, no. 9, pp. 1227-1244, 2021.

[117] B. Ozdemir et al., "Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival," *Cancer Cell,* vol. 25, no. 6, pp. 719-34, 2014.

[118] P. Andre et al., "Anti-NKG2A mAb Is a Checkpoint Inhibitor that Promotes Anti-tumor Immunity by Unleashing Both T and NK Cells," *Cell,* vol. 175, no. 7, pp. 1731-1743, 2018.

[119] N. van Montfoort et al., "NKG2A Blockade Potentiates CD8 T Cell Immunity Induced by Cancer Vaccines," *Cell,* vol. 175, no. 7, pp. 1744-1755, 2018.

[120] N. Lee et al., "HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 95, no. 9, pp. 5199-204, 1998.

[121] G. Biffi et al., "IL1-Induced JAK/STAT Signaling Is Antagonized by TGFbeta to Shape CAF Heterogeneity in Pancreatic Ductal Adenocarcinoma," *Cancer Discovery,* vol. 9, no. 2, pp. 282-301, 2019.

[122] A. Luginbuhl et al., "Discordant treatment response in primary tumors and lymph node metastases after four weeks of preoperative PD-1 blockade in head and neck squamous cell carcinoma (HNSCC)," *Journal of Clinical Oncology,* vol. 37, pp. 6016-601, 2019.

[123] L. Zhang et al., "Lineage tracking reveals dynamic relationships of T cells in colorectal cancer," *Nature,* vol. 564, no. 7735, pp. 268-272, 2018.

[124] M. Lizio et al., "Gateways to the FANTOM5 promoter level mammalian expression atlas," *Genome Biology,* vol. 16, p. 22, 2015.

[125] D. Liu et al., "Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma," *Nature Medicine,* vol. 25, no. 12, pp. 1916-1927, December 2019.

[126] R. Siegel et al., "Cancer statistics," *CA: A Cancer Journal for Clinicians,* vol. 69, pp. 7-34, 2019.

[127] C. Pound et al., "Natural history of progression after PSA elevation following radical prostatectomy," *JAMA,* vol. 281, pp. 1591-7, 1999.

[128] A. D'Amico et al., "Combination of the preoperative PSA level, biopsy gleason score, percentage of positive biopsies, and MRI T-stage to predict early PSA failure in men with clinically localized prostate cancer," *Urology,* vol. 55, pp. 572-7, 2000.

[129] S. Slovin et al., "Ipilimumab alone or in combination with radiotherapy in metastatic castration-resistant prostate cancer: results from an open-label, multicenter phase I/II study," *Annals of Oncology,* vol. 24, pp. 1813-21, 2013.

[130] E. Kwon et al., "Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial," *Lancet Oncology,* vol. 15, pp. 700-12, 2014.

[131] J. De Bono et al., "KEYNOTE-199: Pembrolizumab (pembro) for docetaxel-refractory metastatic castration-resistant prostate cancer (mCRPC)," *J Clin Oncol 2018;36,* vol. 36, 2018.

[132] T. Beer et al., "Randomized, Double-Blind, Phase III Trial of Ipilimumab Versus Placebo in Asymptomatic or Minimally Symptomatic Patients With Metastatic Chemotherapy-Naive Castration-Resistant Prostate Cancer," *Journal of Clinical Oncology,* vol. 35, pp. 40-7, 2017.

[133] P. Kantoff et al., "Sipuleucel-T immunotherapy for castration-resistant prostate cancer," *New England Journal of Medicine,* vol. 363, pp. 411-22, 2010.

[134] K. Boudadi et al., "Ipilimumab plus nivolumab and DNA-repair defects in AR-V7-expressing metastatic prostate cancer," *Oncotarget,* vol. 9, pp. 28561-71, 2018.

[135] P. Sharma et al., "Initial results from a phase II study of nivolumab (NIVO) plus ipilimumab (IPI) for the treatment of metastatic castration-resistant prostate cancer (mCRPC; CheckMate 650)," *Journal of Clinical Oncology,* vol. 37, no. 142, 2019.

[136] C. Drake et al., "Prostate cancer as a model for tumour immunotherapy," *Nature Reviews Immunology,* vol. 10, pp. 580-93, 2010.

[137] M. Haffner et al., "Comprehensive Evaluation of Programmed Death-Ligand 1 Expression in Primary and Metastatic Prostate Cancer," *American Journal of Pathology,* vol. 188, no. 6, pp. 1478-1485, 2018.

[138] M. Sanda et al., "Molecular characterization of defective antigen processing in human prostate cancer," *Journal of the National Cancer Institute,* vol. 87, pp. 280-5, 1995.

[139] W. Su et al., "The Polycomb Repressor Complex 1 Drives Double-Negative Prostate Cancer Metastasis by Coordinating Stemness and Immune Suppression," *Cancer Cell,* vol. 36, pp. 139-55 , 2019.

[140] A. Erlandsson et al., "M2 macrophages and regulatory T cells in lethal prostate cancer," *Prostate,* vol. 79, pp. 363-9, 2019.

[141] A. Calcinotto et al., "IL-23 secreted by myeloid cells drives castration-resistant prostate cancer," *Nature,* vol. 559, pp. 363-9, 2018.

[142] B. Benzon et al., "Correlation of B7-H3 with androgen receptor, immune pathways and poor outcome in prostate cancer: an expression-based analysis," *Prostate Cancer and Prostatic Diseases,* vol. 20, pp. 28-35, 2017.

[143] M. Mercader et al., "Early effects of pharmacological androgen deprivation in human prostate cancer," *BJU International,* vol. 99, pp. 60-7, 2007.

[144] H. Yano et al., "Intratumoral regulatory T cells: markers, subsets and their impact on anti-tumor immunity," *Immunology,* vol. 157, pp. 232-47, 2019.

[145] H. Akaike et al., "Information Theory and an Extension of the Maximum Likelihood Principle," *Springer Series in Statistics Breakthroughs in Statistics,* p. 610–624, 1992.

[146] L. N. Sanchez-Pinto et al., "Comparison of Variable Selection Methods for Clinical Predictive Modeling," *International Journal of Medical Informatics,* vol. 116, p. 10–17, 2018.

[147] E. Vittinghoff and C. Mcculloch, "Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression," *American Journal of Epidemiology,* vol. 165, no. 6, p. 710–718, December 2007.

[148] M. Ammirante et al., "B-cell-derived lymphotoxin promotes castration-resistant prostate cancer," *Nature,* vol. 464, pp. 302-5, 2010.

[149] M. Mercader et al., "T cell infiltration of the prostate induced by androgen withdrawal in patients with prostate cancer," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 98, no. 25, pp. 14565-14570, 2001.

[150] K. Hagihara et al., "Neoadjuvant sipuleucel-T induces both Th1 activation and immune regulation in localized prostate cancer," *Oncoimmunology,* vol. 8, 2019.

[151] L. Fong et al., "Activated lymphocyte recruitment into the tumor microenvironment following preoperative sipuleucel-T for localized prostate cancer," *Journal of the National Cancer Institute,* vol. 106, 2014.

[152] J. Fu et al., "STING Agonist Formulated Cancer Vaccines Can Cure Established Tumors Resistant to PD-1 Blockade," *Science Translational Medicine,* vol. 7, no. 283, 2015.

[153] V. Bronte et al., "Boosting antitumor responses of T lymphocytes infiltrating human prostate cancers," *Journal of Experimental Medicine,* vol. 201, no. 8, pp. 1257-1268, 2005.

[154] M. A. Elmonem et al., "Immunomodulatory Effects of Chitotriosidase Enzyme," *Enzyme Research,* pp. 1-9, 2016.

[155] S. Wada et al., "Cyclophosphamide Augments Antitumor Immunity: Studies in an Autochthonous Prostate Cancer Model," *Cancer Research,* vol. 69, no. 10, p. 4309–4318, December 2009.

[156] J. Machiels et al., "Cyclophosphamide, Doxorubicin, and Paclitaxel Enhance the Antitumor Immune Response of Granulocyte/Macrophage-Colony Stimulating Factor-Secreting Whole-Cell Vaccines in HER-2/Neu Tolerized Mice," *Cancer Research,* vol. 61, no. 9, May 2001.

[157] L. A. Emens et al., "Timed Sequential Treatment With Cyclophosphamide, Doxorubicin, and an Allogeneic Granulocyte-Macrophage Colony-Stimulating Factor–Secreting Breast Tumor Vaccine: A Chemotherapy Dose-Ranging Factorial Study of Safety and Immune Activation," *Journal of Clinical Oncology,* vol. 27, no. 35, pp. 5911-5918, October 2009.

[158] M. Scurr et al., "Low-Dose Cyclophosphamide Induces Antitumor T-Cell Responses, Which Associate with Survival in Metastatic Colorectal Cancer," *Clinical Cancer Research,* vol. 23, no. 22, November 2017.

[159] C. M. Huijts et al., "Phase 1 Study of Everolimus and Low-Dose Oral Cyclophosphamide in Patients with Metastatic Renal Cell Carcinoma," *Cancer Immunology, Immunotherapy,* vol. 68, no. 2, p. 319–329, September 2018.

[160] J. Stultz and L. Fong, "How to turn up the heat on the cold immune microenvironment of metastatic prostate cancer," *Prostate Cancer and Prostatic Diseases,* vol. 24, no. 3, pp. 697-717, 2021.

[161] M. Bilusic et al., "Immunotherapy of Prostate Cancer: Facts and Hopes," *Clinical Cancer Research,* vol. 23, no. 22, 2017.

[162] K. Sfanos et al., "Human prostate-infiltrating CD8+ T lymphocytes are oligoclonal and PD-1+," *Prostate,* vol. 69, no. 15, pp. 1694-1703, 2009.

[163] L. Vargas Roditi et al., "Single-Cell proteomics defines the cellular heterogeneity of localized prostate cancer," *bioRxiv,* 26 January 2021.

[164] J. Sutherland et al., "Activation of thymic regeneration in mice and humans following androgen blockade," *Journal of Immunology,* vol. 175, no. 4, pp. 2741-2753, 2005.

[165] P. Gannon et al., "Characterization of the intra-prostatic immune cell infiltration in androgen-deprived prostate cancer patients," *Journal of Immunological Methods,* vol. 348, no. 1-2, pp. 9-17, 2009.

[166] M. Morse and D. McNeel, "Prostate cancer patients on androgen deprivation therapy develop persistent changes in adaptive immune responses," *Human Immunology,* vol. 71, no. 5, pp. 496-504, 2010.

[167] B. Greenstein et al., "Reappearance of the thymus in old rats after orchidectomy: inhibition of regeneration by testosterone," *Journal of Endocrinology,* vol. 110, no. 3, pp. 417-422, 1986.

[168] A. Roden et al., "Augmentation of T cell levels and responses induced by androgen deprivation," *Journal of Immunology,* vol. 173, no. 10, pp. 6098-6108, 2004.

[169] H. Kissick et al., "Androgens alter T-cell immunity by inhibiting T-helper 1 differentiation," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 111, no. 27, pp. 9887-9892, 2014.

[170] C. Wilson et al., "Enhanced production of B lymphocytes after castration," *Blood,* vol. 85, no. 6, pp. 1535-1539, 1995.

[171] K. Windmill and V. Lee, "Effects of castration on the lymphocytes of the thymus, spleen and lymph nodes," *Tissue Cell,* vol. 30, no. 1, pp. 104-111, 1998.

[172] C. Sorrentino et al., "Androgen deprivation boosts prostatic infiltration of cytotoxic and regulatory T lymphocytes and has no effect on disease-free survival in prostate cancer patients," *Clinical Cancer Research,* vol. 17, no. 6, pp. 1571-1581, 2011.

[173] Y. Pu et al., "Androgen receptor antagonists compromise T cell response against prostate cancer leading to early tumor relapse," *Science Translational Medicine,* vol. 8, no. 333, 2016.

[174] S. Tang et al., "Increased CD8+ T-cell function following castration and immunization is countered by parallel expansion of regulatory T cells," *Cancer Research,* vol. 72, no. 8, pp. 1975-1985, 2012.

[175] L. Brady et al., "Inter- and intra-tumor heterogeneity of metastatic prostate cancer determined by digital spatial gene expression profiling," *Nature Communications,* vol. 12, no. 1, pp. 1-16, 2021.

[176] M. He et al., "Transcriptional mediators of treatment resistance in lethal prostate cancer," *Nature Medicine,* vol. 27, no. 3, pp. 426-433, 2021.

[177] J. Fox et al., "Mass cytometry reveals species-specific differences and a new level of complexity for immune cells in the prostate," *American journal of clinical and experimental urology,* vol. 7, no. 4, pp. 281-296, 2019.

[178] A. Oliver et al., "Tissue-dependent tumor microenvironments and their impact on immunotherapy responses," *Frontiers in Immunology,* vol. 9, no. JAN, p. 70, 2018.

[179] D. Hanahan and R. Weinberg, "Hallmarks of cancer: The next generation," *Cell,* vol. 144, no. 5, pp. 646-674, 2011.

[180] J. Clark and C. Cooper, "ETS gene fusions in prostate cancer," *Nature Reviews Urology,* vol. 6, no. 8, pp. 429-439, 2009.

[181] G. Petrovics et al., "Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome," *Oncogene,* vol. 24, no. 23, pp. 3847-3852, 2005.

[182] S. Chen et al., "Single-cell analysis reveals transcriptomic remodellings in distinct cell types that contribute to human prostate cancer progression," *Nature Cell Biology,* vol. 23, no. 1, pp. 87-98, 2021.

[183] F. Triebel et al., "LAG-3, a novel lymphocyte activation gene closely related to CD4," *Journal of Experimental Medicine,* vol. 171, no. 5, pp. 1393-1405, 1990.

[184] F. Balkwill et al., "TNF-α in promotion and progression of cancer," *Cancer and Metastasis Reviews,* vol. 25, no. 3, pp. 409-416, 2006.

[185] A. Montfort et al., "The TNF Paradox in Cancer Progression and Immunotherapy," *Frontiers in Immunology,* vol. 1818, 10 July 2019.

[186] E. Perez-Ruiz et al., "Prophylactic TNF blockade uncouples efficacy and toxicity in dual CTLA-4 and PD-1 immunotherapy," *Nature,* vol. 569, no. 7756, pp. 428-432, 2019.

[187] Y. Zhang et al., "Overexpression of a novel candidate oncogene KIF14 correlates with tumor progression and poor prognosis in prostate cancer," *Oncotarget,* vol. 8, no. 28, pp. 45459-45469, 2017.

[188] D. Wishart et al., "DrugBank: A knowledgebase for drugs, drug actions and drug targets," *Nucleic Acids Research,* vol. 36, 2008.

[189] D. Robinson et al., "Integrative clinical genomics of advanced prostate cancer," *Cell,* vol. 161, no. 5, pp. 1215-1228, 2015.

[190] D. Quigley et al., "Genomic Hallmarks and Structural Variation in Metastatic Prostate Cancer," *Cell,* vol. 174, no. 3, pp. 758-769, 2018.

[191] B. Dong et al., "Single-cell analysis supports a luminal-neuroendocrine transdifferentiation in human prostate cancer," *Communications Biology,* vol. 3, no. 1, pp. 1-15, 2020.

[192] W. Karthaus et al., "Regenerative potential of prostate luminal cells revealed by single-cell analysis," *Science,* vol. 368, no. 6490, pp. 497-505, 2020.

[193] H. Song et al., "Single-cell analysis of human primary prostate cancer reveals the heterogeneity of tumor-associated epithelial cell states," *bioRxiv,* 2020.

[194] E. Shenderov et al., "Genomic and clinical characterization of pulmonary-only metastatic prostate cancer: A unique molecular subtype," *Prostate,* vol. 79, no. 13, pp. 1572-1579, 2019.

[195] M. Haffner et al., "Genomic and phenotypic heterogeneity in prostate cancer," *Nature Reviews Urology,* vol. 18, no. 2, pp. 79-92, 2021.

[196] H. Beltran et al., "Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer," *Nature Medicine,* vol. 22, no. 3, pp. 298-305, 2016.

[197] E. Bluemn et al., "Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling," *Cancer Cell,* vol. 32, no. 4, pp. 474-489.e6, 2017.

[198] M. Labrecque et al., "Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer," *Journal of Clinical Investigation,* vol. 129, no. 10, pp. 4492-4505, 2019.

[199] A. Obradovic et al., "T-cell infiltration and adaptive treg resistance in response to androgen deprivation with or without vaccination in localized prostate cancer," *Clinical Cancer Research,* vol. 26, no. 13, pp. 3182-3192, 2020.

[200] J. Grosso et al., "Functionally Distinct LAG-3 and PD-1 Subsets on Activated and Chronically Stimulated CD8 T Cells," *The Journal of Immunology,* vol. 182, no. 11, pp. 6659-6669, 2009.

[201] J. Grosso and C. Drake, "Current immunotherapeutic strategies in prostate cancer," *Surgical Oncology Clinics of North America,* vol. 16, no. 4, pp. 861-871, 2007.

[202] E. Lipson et al., "Relatlimab (RELA) plus nivolumab (NIVO) versus NIVO in first-line advanced melanoma: Primary phase III results from RELATIVITY-047 (CA224-047)," *Journal of Clinical Oncology,* vol. 39, p. 9503, 2021.

[203] C. Zahm et al., "PD-1 and LAG-3 blockade improve anti-tumor vaccine efficacy," *OncoImmunology,* vol. 10, no. 1, 2021.

[204] C. Fares et al., "Mechanisms of Resistance to Immune Checkpoint Blockade: Why Does Checkpoint Inhibitor Immunotherapy Not Work for All Patients?," *American Society of Clinical Oncology Educational Book,* vol. 39, pp. 147-164, 2019.

[205] R. Zappasodi et al., "Rational design of anti-GITR-based combination immunotherapy," *Nature Medicine,* vol. 25, no. 5, pp. 759-766, 2019.

[206] L. Vence et al., "Characterization and Comparison of GITR Expression in Solid Tumors," *Clinical Cancer Research,* vol. 25, no. 21, pp. 6501-6510, 2019.

[207] K. Heinhuis et al., "Safety, Tolerability, and Potential Clinical Activity of a Glucocorticoid-Induced TNF Receptor–Related Protein Agonist Alone or in Combination With Nivolumab for Patients With Advanced Solid Tumors: A Phase 1/2a Dose-Escalation and Cohort-Expansion Clinic," *JAMA Oncology,* vol. 6, no. 1, pp. 100-107, 2020.

[208] B. Tran et al., "Dose escalation results from a first-in-human, phase 1 study of glucocorticoid-induced TNF receptor–related protein agonist AMG 228 in patients with advanced solid tumors," *Journal for ImmunoTherapy of Cancer,* vol. 6, no. 1, p. 93, 2018.

[209] D. Killock et al., "GITR agonism — combination is key," *Nature Reviews Clinical Oncology,* vol. 16, no. 7, p. 402, 2019.

[210] R. Geva et al., "First-in-human phase 1 study of MK-1248, an anti–glucocorticoid-induced tumor necrosis factor receptor agonist monoclonal antibody, as monotherapy or with pembrolizumab in patients with advanced solid tumors," *Cancer,* vol. 126, no. 22, pp. 4926-4935, 2020.

[211] F. Balkwill et al., "Tumour necrosis factor and cancer," *Nature Reviews Cancer,* vol. 9, no. 5, pp. 361-371, 2009.

[212] J. E. Hawley et al., "Association between immunosuppressive cytokines and PSA progression in biochemically recurrent prostate cancer treated with intermittent hormonal therapy," *The Prostate,* vol. 80, no. 4, pp. 336-344, 2020.

[213] C. Sweeney et al., "Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer," *New England Journal of Medicine,* vol. 373, no. 8, pp. 737-746, 2015.

[214] C. Ager et al., "Longitudinal immune profiling reveals unique myeloid and T-cell phenotypes associated with spontaneous immunoediting in a prostate tumor model," *Cancer Immunology Research,* vol. 9, no. 5, pp. 529-541, 2021.

[215] P. Thall et al., "New statistical strategy for monitoring safety and efficacy in single-arm clinical trials," *Journal of Clinical Oncology,* vol. 14, no. 1, pp. 296-303, 1996.

[216] A. Obradovic et al., "PISCES: A pipeline for the Systematic, Protein Activity-based Analysis of Single Cell RNA Sequencing Data," *bioRxiv,* 1 June 2021.

[217] E. Chen et al., "Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool," *BMC Bioinformatics,* vol. 14, 2013.

[218] A. Abeshouse et al., "The Molecular Taxonomy of Primary Prostate Cancer," *Cell,* vol. 163, no. 4, pp. 1011-1025, 2015.

[219] T. Simpson et al., "Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma," *Journal of Experimental Medicine,* vol. 210, pp. 1695-1710, 2013.

[220] Z. Freeman et al., "A conserved intratumoral regulatory T cell signature identifies 4-1BB as a pan-cancer target," *Journal of Clinical Investigation,* vol. 130, pp. 1405-1416, 2020.

[221] J. Arriaga et al., "A MYC and RAS co-activation signature in localized prostate cancer drives bone metastasis and castration resistance," *Nature Cancer,* vol. 1, no. 11, pp. 1082-1096, November 2020.

[222] M. LaFleur et al., "Inhibitors of the PD-1 Pathway in Tumor Therapy," *Journal of immunology,* vol. 200, pp. 375-383, 2018.

[223] M. De Simone et al., "Transcriptional landscape of human tissue lymphocytes unveils uniqueness of tumor-infiltrating T regulatory cells," *Immunity,* vol. 45, no. 5, pp. 1135-1147, 2016.

[224] G. Plitas et al., "Regulatory T cells exhibit distinct features in human breast cancer," *Immunity,* vol. 45, no. 5, pp. 1122-1134, 2016.

[225] C. Zheng et al., "Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing," *Cell,* vol. 169, no. 7, pp. 1342-1356, 2017.

[226] A. M. Magnuson et al., "Identification and validation of a tumor-infiltrating Treg transcriptional signature conserved across species and tumor types," *Proceedings of the National Academy of Sciences,* vol. 115, no. 45, pp. E10672-E10681, 2018.

[227] M. Stubbington et al., "T cell fate and clonality inference from single-cell transcriptomes," *Nature Methods,* vol. 13, pp. 329-332, 2016.

[228] S. Loontiens et al., "Purification of high-quality RNA from a small number of fluorescence activated cell sorted zebrafish cells for RNA sequencing purposes," *BMC Genomics,* vol. 20, no. 228, 2019.

[229] D. Aggen et al., "Blocking IL1 Beta Promotes Tumor Regression and Remodeling of the Myeloid Compartment in a Renal Cell Carcinoma Model: Multidimensional Analyses," *Clinical Cancer Research,* vol. 27, no. 2, pp. 608-621, 15 January 2021.

[230] M. Uhlén et al., "Proteomics. Tissue-based map of the human proteome," *Science,* vol. 347, no. 6220, p. 1260419, 23 January 2015.