Supplementary Materials for

Intratumoral CD103 CD8+ T cells Predict Response to PD-L1 Blockade

Romain Banchereau[‡], Avantika S Chitre[‡], Alexis Scherl[†], Thomas D. Wu[†], Namrata S. Patil[†], Patricia de Almeida, Edward E. Kadel, III, Shravan Madireddi, Amelia Au-Yeung, Chikara Takahashi, Ying-Jiun Chen, Zora Modrusan, Jacqueline McBride, Rhea Nersesian, Ehab A. El-Gabry, Mark D. Robida, Jeffrey C. Hung, Marcin Kowanetz, Wei Zou, Mark McCleland, Patrick Caplazi, Shadi Toghi Eshghi, Hartmut Koeppen, Priti S. Hegde, Ira Mellman, W. Rodney Mathews, Thomas Powles, Sanjeev Mariathasan*, Jane Grogan*, William E. O'Gorman*

*Corresponding Author. Emails: ogormanw@gene.com (W.E.O.); jane@arsenalbio.com (J.G.); mariathasan.sanjeev@gene.com (S.M.)

This file includes:

Other Supplementary Materials for this manuscript include the following:

Data File S1 (Microsoft Excel Format). Extended Table of Genes from Differential Expression Analysis Comparing Inflamed and Excluded Tumors in IMvigor210.

Data File S2 (Microsoft Excel Format). *ITGAE* and *CD8A* Gene Expression by Individual for Patients Included in the Clinical Trials.

Supplementary Materials & Methods

Bulk RNA Sequencing & Bioinformatics Analyses

The pathologic diagnosis of each case was confirmed by review of hematoxylin and eosin (H&E) stained slides and all samples that advanced to nucleic acid extraction contained a minimum of 20% tumor cells. H&E images were marked for macro-dissection by a pathologist. RNA (High Pure FFPET RNA Isolation Kit, Roche, Basel, Switzerland) and DNA (QIAamp DNA FFPE Tissue Kit, Qiagen, Hilden, Germany) were then extracted from the macro-dissected sections. Whole transcriptome profiles were generated using TruSeq RNA Access technology (Illumina, San Diego, California). RNAseq reads were first aligned to ribosomal RNA sequences to remove ribosomal reads. The remaining reads were aligned to the human reference genome (NCBI Build 38) using GSNAP 4.5 version '2013-10-10', allowing a maximum of two mismatches per 75 base sequence (parameters: '-M 2 -n 10 -B 2 -i 1 -N 1 -w 200000 -E 1 --pairmax-rna=200000 --clip-overlap). To quantify gene expression levels, the number of reads mapped to the exons of each RefSeq gene was calculated in a strand-specific manner using the functionality provided by the R package GenomicAlignments[86] (Bioconductor). Count data was normalized to transcripts-per-million (TPM) and $\log_2$-transformed. Gene expression data for ITGAE and CD8A by individual are included in Data File S2. To identify genes associated with infiltration status, we grouped patients into "inflamed" and "excluded" phenotypes as below. Differentially expressed genes between these two groups were determined using the limma R package[87] and are shown in Data File S1. Pathway enrichment analysis was conducted on the Reactome[88] database using the ReactomePA R Package.[89] Top correlates of *ITGAE* were identified by Pearson correlation.

Immunohistochemical Staining of Paraffin-Embedded Tumor Tissues

Formalin-fixed paraffin-embedded (FFPE) tumor tissue was stained prospectively for CD8 (rabbit anti-human CD8 monoclonal antibody, Clone SP16, Catalog No. M3160, Spring

Bioscience, Pleasanton, California) and PD-L1 (proprietary diagnostic anti-human PD-L1 monoclonal antibody, SP142) by immunohistochemistry (IHC) as previously described[45]. For determining CD8+ T cell infiltration status, tumors were categorized into three different immunophenotypic classifications based on the prevalence of CD8+ cells in addition to the level of infiltration with respect to malignant epithelial cells. Briefly, a cut off <10 CD8+ T cells was used to designate the tumors into the “desert” category. If CD8+ cells were seen exclusively in the stroma immediately adjacent to the main tumor mass, it was classified as an “excluded” phenotype. Tumors were categorized as “inflamed” if CD8+ cells were seen in direct contact of the malignant epithelial cells in forms such as a spillover of stromal infiltrates or diffuse infiltration of CD8+ cells in tumor cell aggregates or sheets of tumor cells. Samples were scored for PD-L1 expression on tumor-infiltrating immune cells (IC), which included macrophages, dendritic cells and lymphocytes. Specimens were scored as IHC IC0, IC1, IC2, or IC3 if <1%, ≥1% but <5%, or ≥5% of IC were PD-L1 positive respectively. PD-L1 scores in patients with multiple specimens from different time points or samples were based on the highest score.

FFPE tissue was stained for CD103 (rabbit anti-human CD103 monoclonal antibody, Clone EPR4166(2), Catalog No. ab129202, Abcam, Cambridge, United Kingdom) using a fully automated procedure on the BenchMark ULTRA (Roche Tissue Diagnostics, Oro Valley, Arizona) staining platform. Antigen retrieval was performed for 64 minutes using ULTRA CC1 buffer (Catalog No. #950-224, Roche Tissue Diagnostics). The primary antibody was diluted to 3.0 µg/ml in a PBS-based dilution buffer and applied for 16 minutes at 36°C. Secondary antibody application and chromogenic visualization were performed at instrument default times using the OptiView DAB IHC Detection Kit (Catalog No. 760-700, Roche Tissue Diagnostics). Samples were counterstained for 4 minutes with hematoxylin II (Catalog No. 790-2208, Roche Tissue Diagnostics) and post-counterstained for 4 minutes with bluing reagent (Catalog No. 760-2037, Roche Tissue Diagnostics).

Images were scanned and the frequency of CD103+ events out of total cellular events quantified using an automated analysis developed with the 2019a version of the Matlab software package (MathWorks, Natick, Massachusetts). Regions of interest (ROI's) were defined by a pathologist. Within the ROI border, cell nuclei marked with hematoxylin or CD103/DAB were identified by intensity thresholding and simple morphological filtering. The distribution of data was plotted as a histogram and $CD103^{high}$ and $CD103^{low}$ groups were defined based on the right tail of the normal distribution (frequency of 2% or greater).

Tissue Dissociation

Fresh surgical tumor samples were shipped overnight to our institution and handled immediately upon arrival. Tumor tissues were digested using collagenase D (0.5 mg/ml) (MilliporeSigma, St.Louis, Missouri) and DNase (0.1 mg/ml) (MilliporeSigma) for 15 minutes in a 37°C rotating incubator. Tissue samples were then subjected to mechanical dissociation using a gentleMACS Octo Dissociator (Miltenyi Biotec, Bergisch Gladback, Germany), followed by an additional incubation at 37°C for 10 minutes. Samples were filtered and washed prior to downstream applications. Dissociated cell suspensions were either directly stained for mass cytometry analysis or enriched for CD3+ T cells by flow cytometry for use in single-cell RNA gene expression and TCR V(D)J sequencing.

Peripheral Blood Mononuclear Cell Isolation

Peripheral blood mononuclear cells (PBMCs) from healthy human donors were isolated using 50mL Leucosep™ tubes (Greiner Bio-One International, Kremsmünster, Austria) and Ficoll-Paque™ PLUS (GE Healthcare Life Sciences, Malborough, Massachussetts). Whole blood drawn into sodium heparin blood collection tubes was diluted with phosphate-buffered saline (PBS) without calcium or magnesium (Lonza, Walkersville, Maryland) and was centrifuged for

15 minutes at 800 x g at room temperature (RT). PBMCs were harvested and then washed with PBS and subsequently centrifuged for 10 minutes at 250 x g at RT before further processing.

Staining of Healthy Reference PBMC Sample

Washed PBMCs were resuspended in an appropriate volume of PBS to obtain a cell concentration of 10^7 cells/mL. Cells were incubated with a viability reagent, Cell-ID™ Cisplatin (Fluidigm, South San Francisco, California) at a final concentration of 5 µM for 5 minutes on ice. Cisplatin was quenched by washing once with 5x volume of MaxPar® Cell Staining Buffer (Fluidigm) and centrifuged at 300 x g, then resuspended to a final concentration of 30 million cells/mL in staining buffer. To start antibody labeling, 3 million cells were transferred to Falcon® 5 mL 12 x 75 mm tubes (Corning, Corning, New York) and incubated with 5 µL of Human TruStain FcX™ (BioLegend, San Diego, California) for 10 minutes on ice to block Fc receptor binding. Healthy PBMCs were stained with CD45-198Pt (in-house conjugation of Clone HI30, BioLegend) as a reference to incorporate into each tumor sample as our internal reference control and tumor cells were stained with CD45-89Y (Fluidigm) for 30 minutes on ice. Cells were washed twice with 4 mL cell staining buffer before being prepared for surface staining.

Staining of Cells for Mass Cytometry Analysis

Differentially labeled CD45+ samples were combined together as one sample. 20% of CD45-198Pt labeled healthy PBMCs were incorporated into 3 million tumor cells labeled with CD45-89Y in a Falcon® 5 mL 12 x 75 mm tube (Corning). A healthy PBMC sample labeled with only CD45-198Pt was also used as a reference control sample. A master surface antibody cocktail with all metal-conjugated antibodies (50µL of total staining reagent volume) was added to samples for cell staining and incubated for 30 minutes on ice. See Table S2 for a list of the metal conjugated monoclonal antibodies used in these studies. Cells were washed once with 4 mL of cell staining buffer before being prepared for intracellular staining by using the FoxP3

Staining Buffer Set (eBioscience, San Diego, California). Cells were then resuspended in 1 mL of fixation/permeabilization solution for 45 minutes on ice. After incubation, cells were washed with 3 mL of permeabilization buffer and centrifuged at 800 x g for 5 minutes and resuspended in 50 µL of permeabilization buffer. Cells were then stained for intracellular markers by addition of 50 µL of the master intracellular antibody cocktail for 30 minutes on ice. Lastly, the cells were washed with 4 mL cell of staining buffer and fixed in a 1 mL solution overnight at 4°C containing Cell-ID™ Intercalator-Ir in 1.6% EMS paraformaldehyde (Electron Microscopy Sciences, Hatfield, Pennsylvania) diluted with PBS. For NSCLC6, an antibody detecting CD39 (Clone A1, BioLegend) was also included in the staining panel.

Acquisition on CyTOF® Instrument

After overnight fixation, cells were washed with 3 mL of MaxPar® cell staining buffer and centrifugated at 800 x g for 5 minutes. After aspiration of the wash buffer and resuspension of the cell pellet, another round of wash was performed with 4 mL of MaxPar® Water (Fluidigm). Cells were resuspended in 1mL of MaxPar® Water and counted. After obtaining cell counts, 3 mL of MaxPar® Water was added and cells were pelleted one final time prior to instrument acquisition. Before introduction into the Helios™, a CyTOF® System (Fluidigm), pelleted cells were resuspended with 1X MaxPar® Water containing EQ™ Four Element Calibration Beads (Fluidigm) then filtered using a 12 x 75 mm tube with a 35 µm nylon mesh cell-strainer cap (Corning).

Data Processing and Analysis of CyTOF® Data

All FCS files from each indication were normalized together using the MATLAB® (MathWorks) normalizer and analyzed using FlowJo® software (FlowJo, LLC, Ashland, Oregon).

Aggregated UMAP Analysis of Mass Cytometry Data

Protein marker expression intensities from mass cytometry analysis were aggregated from multiple samples and transformed using the inverse hyperbolic sine function. Dimensionality reduction was applied to the transformed expression matrix using the uniform manifold approximation and projection (UMAP) package in R[53] with the following default parameters: min_dist :0.1, n_neighbors: 15, n_components: 2, and metric: euclidean. For each aggregated UMAP analysis, individual samples were downsampled to equal cell numbers as follows: 8,000 cells per sample in CD45+ and CD8+ populations for the aggregated NSCLC analysis, 20,000 and 12,000 cells per sample in CD45+ and CD8+ populations for the aggregated Endometrial analysis respectively, and ~7,500 cells per sample in the analysis of the representative tumor and adjacent tissue NSCLC sample. UMAP coordinates were appended to ".fcs" files as additional channels for integration with manual gating analysis in FlowJo (FlowJo, LLC).

Flow-Based Sorting of T cells for Single Cell RNA & TCR Clone Sequencing

The workflow for single-cell RNA sequencing (scRNAseq) and TCR V(D)J sequencing (scTCRseq) library preparation, sequencing, and data processing was performed as published previously.[44] After enzymatic dissociation of tissues, cell suspensions were stained with antibodies against EpCAM (Clone 9C4, Catalog No. 324208, BioLegend), CD45 (Clone 2D1, Catalog No. 368516, BioLegend), and CD3 (Clone SK7, Catalog No. 564001, BD Biosciences, San Jose, California) along with viability markers Calcein Blue AM (ThermoFisher Scientific, Waltham, Massachusetts) and 7-AAD (ThermoFisher Scientific). Purified T cell populations were obtained by first gating for live cells (Calcein Blue$^+$, 7-AAD$^-$) and subsequently sorting for CD3+CD45+EpCAM- events via fluorescence-activated cell sorting (FACS) using a Becton Dickinson FACSAria Fusion Cell Sorter comprised of four lasers (405 nm, 488 nm, 561 nm, 638 nm) (Becton Dickinson, Franklin Lakes, New Jersey).

Single-Cell RNA & TCR V(D)J Library Preparation

T cells enriched by FACS were used to prepare libraries for scRNAseq and scTCRseq using the Chromium Single Cell 5' Library and Gel Bead Kit v1 (10x Genomics, Pleasanton, California) according to manufacturer's instructions. Briefly, the cell density and viability of sorted T cells from tumors and adjacent tissues were determined with a Vi-CELL XR cell counter (Beckman Coulter, Pasadena, California). Approximately 10,000 FACS-sorted CD3+ immune cells (90-95% viability) were encapsulated into droplets, aiming to achieve ~6,000 cells per sample. After Gel Bead-in-Emulsion reverse transcription (GEM-RT), droplets were disrupted and barcoded cDNA was purified with DynaBeads MyONE SILANE (Thermofisher Scientific) and cleaned up. This was followed by 14 cycles of PCR amplification. The cDNA was used to construct both 5' gene expression libraries as well as TCR V(D)J target enriched libraries. For TCR target enrichment, cDNA was first amplified with two PCR reactions with 10 cycles each using nested primers provided in the Chromium Single Cell V(D)J Enrichment Kit, Human T cell (10x Genomics), and then proceeded to TCR library generation. Libraries for both scRNAseq and V(D)J were profiled using the Bioanalyzer High Sensitivity DNA Kit (Agilent Technologies, Santa Clara, California) and quantified with a Kapa Library Quantification Kit (Kapa Biosystems, Wilmington, Massachusetts). Single-cell RNAseq libraries were sequenced in one lane of HiSeq4000 (Illumina); single-cell TCR V(D)J libraries were pooled and then sequenced in one lane of HiSeq2500 (Illumina). Sequencing was done according to the manufacturer's specifications (10x Genomics).

Processing of scRNAseq Data

scRNAseq data were processed using the Cell Ranger software (10x Genomics). Illumina base call (BCL) 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.mtx file containing counts of transcripts for each

gene in each single cell, which we parsed using a script in R to create a count matrix for each sample. Single cells were identified by unique barcodes. In total, scRNAseq data were obtained on 30,611,692 transcripts in 26,951 CD8+ T cells.

Processing of scTCRseq Data

TCRseq data for each sample were processed using Cell Ranger software with the command 'cellranger vdj' using a custom reference set of 30,727 genes, based on human reference genome GRCh38 and RefSeq gene models.[90] For each sample, Cell Ranger generated an output file, filtered_contig_annotations.csv, containing TCR α-chain and β-chain CDR3 nucleotide sequences for single cells that were identified by barcodes. Clonotypes were re-grouped across tumor samples for each patient, requiring that they share all reported α-chain and β-chain CDR3 consensus nucleotide sequences in common. Clonotypes were assigned an expansion pattern based on their clone sizes in the tumor.

Cluster Visualization of scRNAseq Data for Combined CD8+ T cells

Count matrices for each sample were processed using Seurat[91] using the SCTransform procedure with default parameters to perform a regularized negative binomial regression based on the 3,000 most variable genes. Normalized datasets for each patient were combined using the FindIntegrationAnchors and IntegrateData functions in Seurat with the default value of 30 dimensions. The resulting datasets for the 26,951 CD8+ T cells from the 6 patients were then combined using the FindIntegrationAnchors and IntegrateData functions in Seurat with the default value of 30 dimensions. The integrated dataset was scaled and processed under principal components analysis using the ScaleData and RunPCA functions in Seurat. Two-dimensional map coordinates were generated using the RunUMAP procedure. Cluster definitions were generated using data for total T cells (including both CD8+ and non-CD8 populations) from multiple tissue sampling sites (tumor, non-diseased adjacent tissue, and

blood).[44] Briefly, cluster analysis was performed using the FindNeighbours procedure and the FindClusters procedure at a resolution of 1.6, larger than the default value of 0.8, in order to obtain a finer resolution of cell subtypes, yielding 8 clusters of CD8+ T cells. The FindMarkers procedure in Seurat was run on each cluster to obtain biomarker genes upregulated in that cluster. Data were computationally clustered specifically for this study and only annotations, not the clusters themselves, were used from Wu et al., 2020. Our analysis focused on tumor CD8+ T cells which, being a single cellular lineage from one tissue in the context of this broader analysis, are inherently more closely related to one another and thus the clusters may not appear entirely distinct in the UMAP.

Integration of scRNAseq and scTCRseq Data

Clonal expansion pattern information from TCRseq is applied to clones, whereas cluster assignments from RNAseq are applied to individual cells, based on their transcriptional profiles. Therefore, a given clone may contain a diverse set of cells belonging to different phenotypic clusters. To integrate the data, we assigned expansion patterns from clones to their constituent cells. Clones were assigned a primary cluster, based on the cluster with the largest representation of cells in the clone. In cases of ties, in which the two largest representative clusters had equal counts, we assigned no primary cluster to that clone.

Software & Data Representation of scRNAseq and scTCRseq Analyses

Data were collected using Cell Ranger software (10x Genomics) v.2.2.0.[90] Data were analysed using Perl v.5.18.2, R v.3.6.0, and the following packages and versions in R for analysis: Seurat, 3.0.2[91]; SingleR, 1.0.1[92]; RankProd, 3.11.0[93]; GSEABase, 1.47.0; limma, 3.41.15[87] annotate, 1.63.0; homologene, 1.4.68.19.3.27; mouse4302.db, 3.2.3; cocor, 1.1-3[94]; and iNEXT, 2.0.19[95]. Two-dimensional gene expression maps were generated using coordinates from the UMAP algorithm[53] as implemented in the umap-learn module v.0.3.10 in Python v.3.6.0, run

under R using reticulate v.1.13. Figures were produced using the following packages and versions in R: colorspace, 1.4-1[96] RColorBrewer, 1.1-2; pheatmap, 1.0.12; and superheat, 0.1.0[97].

Supplementary Figure Titles & Legends

Figure S1. Extended Data for Mass Cytometry Analysis & Aggregate UMAPs for Endometrial Tumor Samples. (A) Resultant UMAP clusters derived from total CD45+ cells from endometrial patient tumors (N=3) as analyzed by mass cytometry (left). Immune subpopulations were defined by manual gating and immune subtypes projected onto the UMAP. Expression pattern of CD103 across CD45+ cells overlaid onto the UMAP (right). (B) Frequency of immune cell subpopulations amongst CD45+ cells across nine patient tumor samples (N=6 NSCLC, solid dots; N=3 Endometrial, open dots) (left). Frequency of cells expressing CD103 within the indicated immune subsets (right). (C) UMAPs of CD8+ T cells for each individual patient sample depicting variation in cellular distribution amongst immune subpopulations for NSCLC (top two rows) or Endometrial (bottom row) tumors. Aggregate UMAPs were generated separately for each tumor indication. (D) Expression of proliferation and dysfunction markers in Endometrial tumors (N=3) overlaid onto the aggregate UMAP of CD8+ T cells. (E) Correlation of the percentage of PD-L1-expressing cells out of total CD45+ cells to the frequency of CD8+ T cells (top), CD103+ cells within CD8+ T cells (middle), or proliferating (Ki-67+) CD103+ CD8+ T cells (bottom) as analyzed by mass cytometry on freshly procured tumors. Data from NSCLC tumors is indicated by solid dots and endometrial tumors depicted by open dots. The Spearman correlation coefficient is displayed on each plot alongside the resultant p-value. *$p<0.05$, **$p<0.01$, n.s.= not significant.

Figure S2. Comparison of CD103+ and CD103- CD4+ T cells and NK cells. (A) Frequency of positivity for indicated markers of proliferation, activation, and dysfunction between CD103+ (red) and CD103- (blue) subsets of CD4+ T cells in either tumor (left) or adjacent (right) tissues in NSCLC (closed circles) and Endometrial (open circles) tumors. (B) Parallel analysis as done in (A) for CD103+ (red) and CD103- (blue) subsets of NK cells. Statistical significance was determined using a Two-Way ANOVA with a Bonferroni Correction for multiple comparisons. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$, n.s.= not significant.

Figure S3. CD39 Further Enriches for Ki-67 and Immune Checkpoint Expression Within the CD103+ CD8+ Compartment. (A) CD8+ T cells from tumor and adjacent tissues from NSCLC (N=1, NSCLC6) were hand gated by manual analysis, merged, and clustered to generate a single UMAP. Data is displayed in a tissue-specific manner, with tumor (left) and adjacent (right) represented as individual UMAPs. (B) Expression of CD103 (top) and CD39 (bottom) overlaid onto the UMAP for CD8+ T cells shown in tumor (left) as well as non-diseased adjacent (right) tissues. (C) Bivariate plots depicting CD103 (x-axis) and CD39 (y-axis) within CD8+ T cells. (D) Percentage of cells expressing the indicated markers for double negative (blue), CD103 single-positive (purple), and CD103+CD39+ (orange) subpopulations. (E) Expression of indicated proteins overlaid onto the merged UMAP of CD8+ T cells from (A) for tumor (top) and adjacent (bottom) tissues.

Figure S4. Extended Data for Matched scRNAseq and scTCRseq Analysis of Tumor Samples. (A) Gene expression of the indicated individual genes in tumor CD8+ T cells overlaid onto the UMAP. Individual cells are colored on a scale of gray (0) to red (1) according to the quantile of their expression. (B) Expression of indicated genes associated with a naive/stem-like phenotype (y-axis) plotted according to the size of a given clonotype (rank ordered on the x-axis). (C) TCR clone composition and phenotype sharing across clusters. Clonotypes are grouped by their primary cluster assignment on the x-axis, as indicated by dot color (Green: 8.2-T_{EM}, Blue: 8.3a-T_{RM}, Purple: 8.3b-T_{RM}, Red: 8.3c-T_{RM}). Heatmap demonstrates the fraction of each clonotype that falls into each cluster phenotype, depicted as a gradient from 0% (white) to 100% (black). Within each primary cluster, clones are rank ordered according to gradation of cell fraction. The sum clone size of each clonotype is indicated at the top.

Figure S5. CD8+ T cell Phenotypes in the Bladder Tumor Microenvironment. (A) Heatmap depicting cross-referencing of CD8+ T cell cluster definitions from our dataset (x-axis) with those determined by Oh et al. (y-axis) in a bladder tumor dataset with color intensity indicating extent of correlation. (B) Cluster

definitions, as determined by our analysis, projected onto the UMAP of CD8+ T cells in bladder tumors from Oh, et al. (C) Frequency of various indicated clusters in CD8+ T cells from bladder tumors.

Figure S6. Extended Data for Survival Analyses of *ITGAE* and *CD8A* in Clinical Trials of Immunotherapy Treatment. (A) Kaplan-Meier curves showing the probability of overall survival of patients from OAK (NSCLC) treated with either atezolizumab (top panels) or chemotherapy (docetaxel) (bottom panels) categorized by quartiles (Q1: lowest, Q4: highest) of transcriptional expression of either *ITGAE* (left panels) or *CD8A* (right panels). (B) Kaplan-Meier curves showing the probability of overall survival of patients from IMvigor210 (bladder) treated with atezolizumab categorized by quartiles (Q1: lowest, Q4: highest) of transcriptional expression of either *ITGAE* (left panel) or *CD8A* (right panel). Statistical significance was determined by Cox proportional hazards modeling. p-values are displayed within each panel. Kaplan-Meier curves showing the probability of OS of patients from IMvigor211 (bladder) treated with either atezolizumab (top panels) or chemotherapy (docetaxel, paclitaxel, or vinflunine) (bottom panels) categorized by median (high in red, low in blue) (C) or quartiles (Q1: lowest, Q4: highest) (D) of transcriptional expression of either *ITGAE* (left panels) or *CD8A* (right panels). p-values and hazard ratios (where applicable) are displayed within each panel. Statistical significance for all analyses was determined by Cox-proportional hazards modeling.

Table S1. Patient Characteristics & Tumor Metadata by Individual for Tissue Samples Analyzed by Mass Cytometry, scRNAseq, and scTCRseq.

Sample Code	NSCLC1	NSCLC2	NSCLC3	NSCLC4	NSCLC5	NSCLC6	Endo1	Endo2	Endo3
Matched scRNAseq	NA	NA	NA	Y	Y	Y	Y	Y	Y
Donor ID	D37070	D37075	D37079	D37084	D37073	D39903	D40221	D40222	D40220
Tumor Indication	NSCLC	NSCLC	NSCLC	NSCLC	NSCLC	NSCLC	Endome-trial	Endome-trial	Endome-trial
Age	72	83	51	60	55	67	62	59	61
Gender	F	M	M	F	M	F	F	F	F
Ethnicity	White	White	White	White	White	White	Native Hawaiian & Other Pacific Islander	White	White
Tumor Stage	Ia	IIa	III	I	III	Not Available	I	I	I
Histology Subtype	Adenoca-rcinoma	Squamo-us Cell Carcino-ma	Squamo-us Cell Carcino-ma (Keratiniz-ing)	Squamo-us Cell Carcino-ma (Nonkera-tinizing)	Not Available	Not Available	Adenoca-rcinoma	Endomet-rioid Carcino-ma	Adenoca-rcinoma
T (Area of Cancer)	T1b	T2b	T2a	T2a	T4	Not Available	T1a	T1a	T1a
N (Lymph Node Spread)	N0	N0	N2	N0	N1	Not Available	N0	N0	N0
M (Metastasis)	MX	MX	MX	MX	MX	Not Available	MX	MX	MX

Table S2. Antibody Information for Mass Cytometry Panel Used to Assess Tumor Phenotypes.

Mass	Metal	Target	Source	Vendor	Clone	Catalog No.	Lot Number
89	Y	CD45	Fluidigm		HI30	3089003B	2291710
113	In	EpCAM	Custom	Biolegend	9C4		In-house
115	In	CD57	Custom	Biolegend	HCD57		In-house
140	Ce	EQ Beads					
141	Pr	Perforin	Custom	Abcam	B-D48		In-house
142	Nd	CCR4	Custom	R&D	205410		In-house
143	Nd	CD127/ CD137/CCR8	Fluidigm/ Custom	Biolegend	A019D5/ 4B4-1/ L263G8	3143012B	2191516 In-house
144	Nd	Granzyme B	Custom	Biolegend	GB11		In-house
145	Nd	CD4	Fluidigm		RPA-T4	3145001B	1931719
146	Nd	CD8	Fludigm		RPA-T8	3146001B	1671716
147	Sm	CD11c	Fluidigm		Bu15	3147008B	0731704
148	Nd	CD56	Custom	Miltenyi	REA196		In-house
149	Sm	Granzyme A	Custom	Biolegend	CB9		In-house
150	Nd	CD103	Custom	Biolegend	Ber-ACT8		In-house
151	Eu	ICOS	Custom	eBioscience	ISA-3		In-house
152	Sm	CD155 (PVR)/ CD39	Custom	Biolegend	SKII.4/A1		In-house
153	Eu	TIM3	Custom	R&D	344823		In-house
154	Sm	CD3	Fluidigm		UCHT1	3154003B	1351723

155	Gd	CD27	Fluidigm		L128	3155001B	1031712
156	Gd	CXCR3	Fluidigm		G025H7	3156004B	1681613
157	Gd	CD14	Custom		M5E2		In-house
158	Gd	OX40	Custom	Genentech	1A7.gr.1		In-house
159	Tb	CD226	Custom	BD	DX11		In-house
160	Gd	Tbet	Fluidigm		4B10	3160010B	1061601
161	Dy	CTLA4	Custom	Biolegend	BNI3		In-house
162	Dy	Foxp3	Fluidigm		PCH101	3162011A	0731702
163	Dy	Eomes	Custom	eBioscience	WD1928		In-house
164	Du	CD161	Fluidigm		HP-3G10	3164009B	3271610
165	Ho	CD19	Fluidigm		HIB19	3165025B	3121611
166	Er	NKG2D	Fluidigm		ON72	3166016B	0091703
167	Er	CCR7	Fluidigm		G043H7	3167009A	0241815
168	Er	Ki-67	Fluidigm		Ki-67	3168007B	0861703
169	Tm	CD25	Fluidigm		2A3	3169003B	1931713
170	Er	CD45RA	Fluidigm		HI100	3170010B	0981521
171	Yb	PD-L1	Custom	Genentech	14D3		In-house
172	Yb	CD28	Custom	Biolegend	CD28.2		In-house
173	Yb	TIGIT	Fluidigm		MBSA43	3999999-5	1461701
174	Yb	HLA-DR	Fluidigm		L243	3174001B	0041719

175	Lu	PD-1	Fluidigm		EH12.2H7	3175008B	2431709
176	Yb	CD38	Custom	Biolegend	HIT2		In-house
191	Ir	Nucleic Acid	Fluidigm			201192A	2451506A
192	Pt	Cisplatin	Fluidigm			201064	0671809
193	Ir	Nucleic Acid	Fluidigm			201192A	2451506A
195	Pt	Cisplatin	Fluidigm			201064	0671809
209	Bi	CD16	Fluidigm		3G8	3209002B	2031601