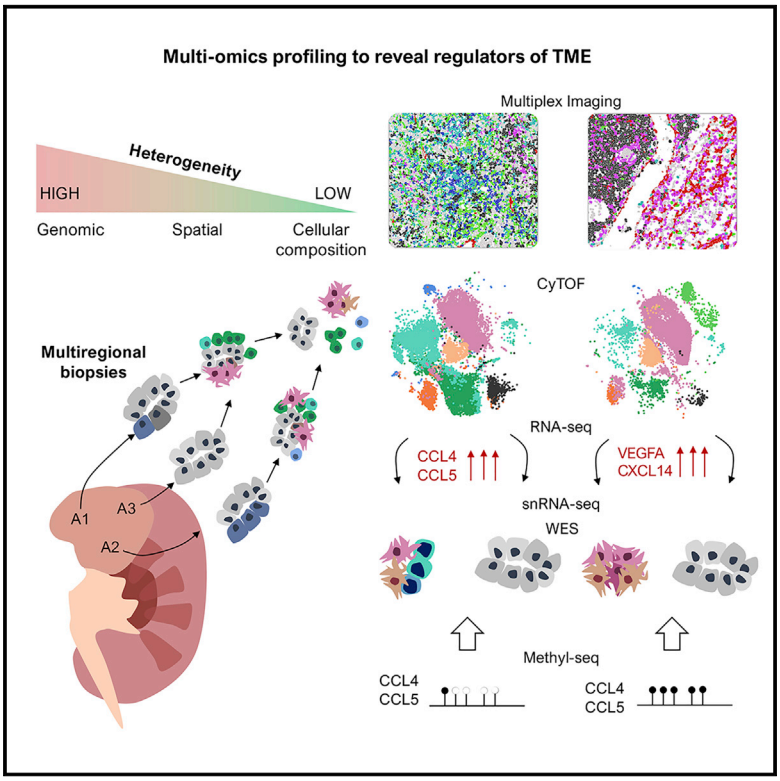
Cell Reports

Article

Multiregional single-cell proteogenomic analysis of ccRCC reveals cytokine drivers of intratumor spatial heterogeneity

Graphical abstract



Authors

Natalia Miheecheva, Ekaterina Postovalova, Yang Lyu, ..., Ravshan Ataullakhanov, Nathan Fowler, James J. Hsieh

Correspondence

nathan.fowler@bostongene.com (N.F.), jhsieh@wustl.edu (J.J.H.)

In brief

Using an integrated approach, including multiomic profiling and multiplex imaging, Miheecheva et al. identify heterogeneity within the tumor microenvironment (TME) spatial architecture of primary ccRCC tumors. Molecular analysis shows correlation of TME tumor-expressed cytokines with treatment response, which may drive clinical heterogeneity of ccRCC patient outcomes.

Highlights

- Integrated multiomic profiling and multiplex imaging of multiregional ccRCC biopsies
- Heterogeneity in TME spatial architecture is detected across different tumor regions
- MxIF analysis reveals 14 distinct cellular neighborhoods in multiregion ccRCC tumors
- TME tumor-expressed cytokines are correlated with ccRCC clinical outcomes

Miheecheva et al., 2022, Cell Reports *40*, 111180
August 16, 2022
https://doi.org/10.1016/j.celrep.2022.111180

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Multiregional single-cell proteogenomic analysis of ccRCC reveals cytokine drivers of intratumor spatial heterogeneity

Natalia Miheecheva,[1] Ekaterina Postovalova,[1] Yang Lyu,[2] Akshaya Ramachandran,[2] Alexander Bagaev,[1] Viktor Svekolkin,[1] Ilia Galkin,[1] Vladimir Zyrin,[1] Vladislav Maximov,[1] Yaroslav Lozinsky,[1] Sergey Isaev,[1] Pavel Ovcharov,[1] Diana Shamsutdinova,[1] Emily H. Cheng,[3] Krystle Nomie,[1] Jessica H. Brown,[1] Maria Tsiper,[1] Ravshan Ataullakhanov,[1] Nathan Fowler,[1,4,*] and James J. Hsieh[2,*]

[1]BostonGene Corporation, University Office Park III, 95 Sawyer Road, Waltham, MA 02453, USA
[2]Molecular Oncology, Division of Oncology, Department of Medicine, Washington University, St. Louis, MO 63110, USA
[3]Human Oncology and Pathogenesis Program and Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[4]Lead contact
*Correspondence: nathan.fowler@bostongene.com (N.F.), jhsieh@wustl.edu (J.J.H.)
https://doi.org/10.1016/j.celrep.2022.111180

SUMMARY

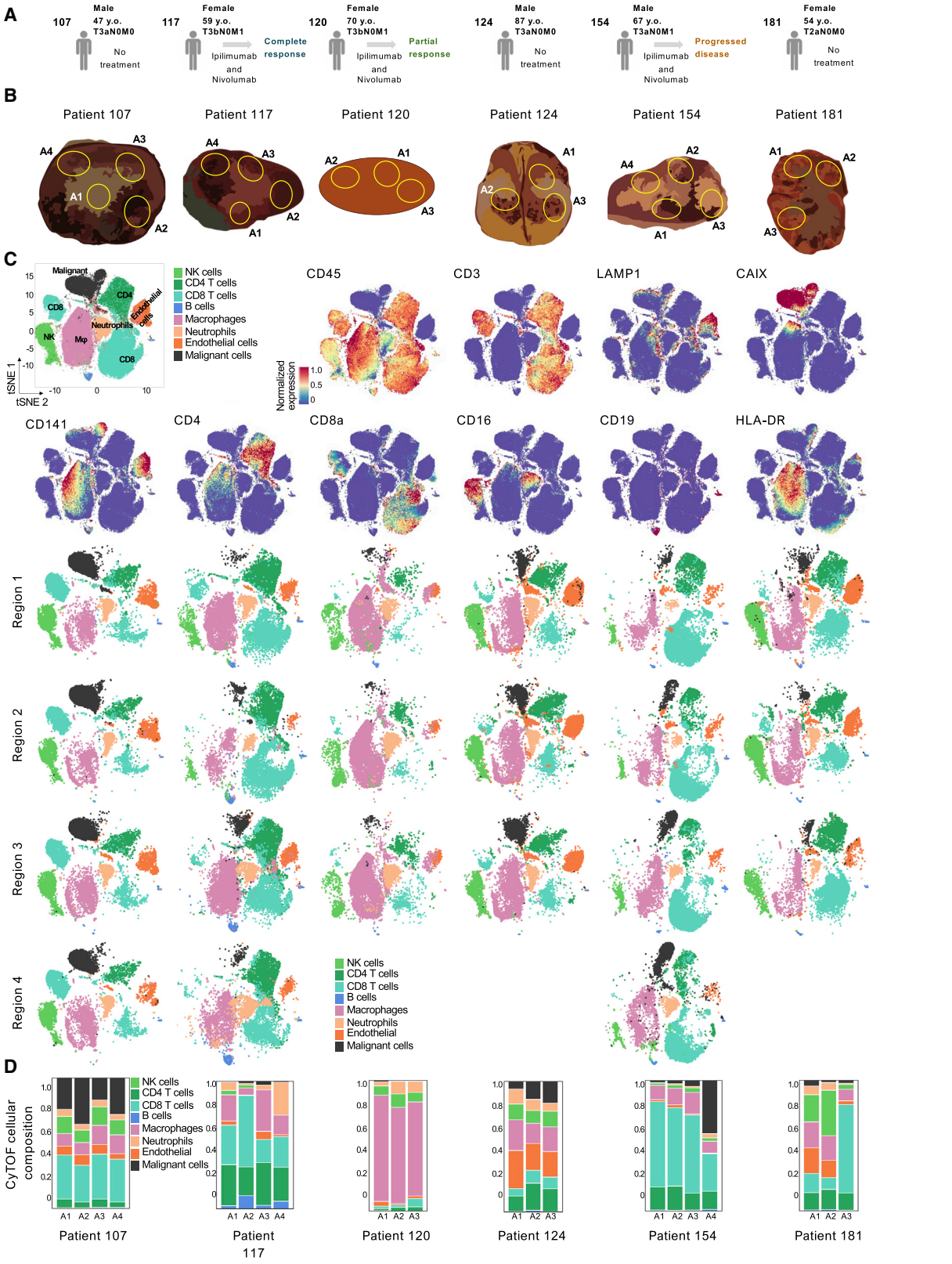
Intratumor heterogeneity (ITH) represents a major challenge for anticancer therapies. An integrated, multidimensional, multiregional approach dissecting ITH of the clear cell renal cell carcinoma (ccRCC) tumor microenvironment (TME) is employed at the single-cell level with mass cytometry (CyTOF), multiplex immunofluorescence (MxIF), and single-nucleus RNA sequencing (snRNA-seq) and at the bulk level with whole-exome sequencing (WES), RNA-seq, and methylation profiling. Multiregional analyses reveal unexpected conservation of immune composition within each individual patient, with profound differences among patients, presenting patient-specific tumor immune microenvironment signatures despite underlying genetic heterogeneity from clonal evolution. Spatial proteogenomic TME analysis using MxIF identifies 14 distinct cellular neighborhoods and, conversely, demonstrated architectural heterogeneity among different tumor regions. Tumor-expressed cytokines are identified as key determinants of the TME and correlate with clinical outcome. Overall, this work signifies that spatial ITH occurs in ccRCC, which may drive clinical heterogeneity and warrants further interrogation to improve patient outcomes.

INTRODUCTION

The intricate interplay between kidney cancer cells and the surrounding tumor microenvironment (TME) contributes to the marked genetic intratumor heterogeneity (ITH) of clear cell renal cell carcinoma (ccRCC), which is thought to underlie tumor evolution, metastasis, and clinical responses to various therapies (Díaz-Montero et al., 2020; Krishna et al., 2021). Recent studies show the utilization of whole-exome sequencing (WES) to uncover evolution patterns, the mutational profile underlying ITH, clonal architecture, and interpatient tumor differences (Hsieh et al., 2017; Motzer et al., 2020a; Ricketts et al., 2018; Turajlic et al., 2018). Notably, multiregion DNA sequencing revealed that two common ccRCC aberrations, del(3p) and ampl(5q), are the first genetic events in kidney cancer development (Ricketts and Linehan, 2018) and suggests that clonal architecture divides ccRCC tumors into three evolutionary trajectories correlated with clinical outcome (Turajlic et al., 2018). Although immunotherapy has dramatically improved the clinical outcomes of ccRCC patients, and the combination of ipilimumab and nivolumab is FDA approved for frontline metastatic ccRCC (Gao and McDermott, 2018), the differences observed in patient responses cannot be explained by genetic heterogeneity alone (Banchereau et al., 2021; Motzer et al., 2020b).

Uncovering the underlying molecular and cellular mechanisms of this observed clinical heterogeneity, particularly the role of the TME, may lead to improved patient outcomes. Sequencing technologies cannot reveal spatial architecture to decode cell-to-cell interactions, limiting our understanding of the complex tumor ecosystem in kidney cancer (Bi et al., 2021; Braun et al., 2021; Krishna et al., 2021). Single-cell pathology has recently shown promise for understanding the clinical behavior of breast and colorectal cancers (Jackson et al., 2020; Schürch et al., 2020). Furthermore, how genetic and epigenetic (DNA methylation) events control the formation of individual tumor ecosystems remains unclear, with some links established between *BAP1* mutations and the formation of an immune-enriched TME and *PBRM1* mutation enrichment in non-inflamed clusters (Liu et al., 2020; Motzer et al., 2020b).

Here, we utilized diverse methods to characterize the ITH, microenvironment composition, spatial architecture, and cellular neighborhoods of multiregion ccRCC tumor tissues to uncover



(legend on next page)

specificities in tumor behavior linked with patient response and the formation of complex ecosystems across different regions of the same tumor. Integration of diverse methods for analysis provides further insights into the epigenetic control of transcription through DNA methylation and gene expression levels measured at bulk and single-cell levels.

RESULTS

Multiregional single-cell proteogenomic assessment of ccRCC heterogeneity

To comprehensively analyze the ITH, the TME, and the tumor ecosystem at cellular, molecular, and spatial levels of ccRCC, primary tumors collected from six patients with varying clinical outcomes and therapeutic interventions (Figures 1A and 1B) were biopsied at multiple locations and subjected to cytometry by time of flight (CyTOF), multiplex immunofluorescence (MxIF) in conjunction with artificial intelligence (AI) algorithm-based digital pathology, WES, RNA sequencing (RNA-seq), single-nucleus RNA-seq (snRNA-seq), and whole genome bisulfite sequencing (WGBS), enabling integrated multiregion multiproteogenomic analysis (Figures 1B and S1A–S1C). The regions selected were significantly distant from one another, allowing geographically distinct regions to be analyzed separately (Figures 1B and S1B). Patient 117 (Pt117), Pt120, and Pt154 had metastatic disease and were treated with ipilimumab and nivolumab, while Pt107, Pt124, and Pt181 had localized disease resulting in a nephrectomy. The ccRCC tumors were primarily grade 3 (33%) or 4 (50%) and tumor stage T3a (50%) or T3b (33%), except for one T2a. Among the three patients treated with immune checkpoint blockade (ICB), Pt117 achieved a complete response, Pt120 was a partial responder, and Pt154 was a non-responder with disease progression. Machine learning (ML) platforms (STAR Methods) were employed for all analyses, and integration of all methods and the determination of the drivers of tumor behavior and clinical responses to therapy were performed (Figure S1A).

CyTOF analysis revealed striking similarity in immune cell composition across different regions

To determine whether heterogeneity of the TME manifests in alterations of immune cell composition, multiregional CyTOF (Figure S1D; Table S1) was performed (Chevrier et al., 2017). Interestingly, CyTOF analyses revealed similarity in cellular composition across different regions within the same patient's tumor, while individual patient tumors were distinctly different from one another, with notable heterogeneity in immune cell proportions between patients (Figure 1C). For example, T cell and myeloid populations ranged from 3% to 83% and 6% to 81% among patients, respectively (Figure 1D). Remarkably, CyTOF analyses showed that each patient's tumor regions clustered together based on cellular contents (Figure 1C), indicating low ITH in immune cell composition but large interpatient differences. These findings also suggest that each patient has his or her own immune composition signature, as characterization of the major immune populations showed intratumor similarities in cellular composition (Figure 1C); only slight differences were shown among multiregion tumors of a few patients. The tumor regions of Pt154 had different relative tumor cell contents, with the highest number of tumor cells in the A4 region (41% of total cell composition; Figures 1C and 1D). While region A3 of Pt181 primarily comprised CD8 T cells (68%), the other two regions had much lower levels of CD8 T cells (15% and 9%) (Figures 1C and 1D). In addition, differences in the percentages of neutrophils (Pt117 and Pt181) and endothelial cells (Pt181) were observed among the regions. Lastly, Pt117 had varied cell types in the tumor regions and the highest proportion of B cells in all of the tumor regions compared with the other patients (Figures 1C and 1D). Overall, these analyses showed marked interpatient cellular heterogeneity but not strong ITH in cellular composition in each tumor region of the same patient (Figures S2A and S2B).

In-depth cellular phenotypic analysis identified numerous cell subpopulations conserved across regions

Further characterization of the subpopulations and TME within each tumor region using phenograph-based clustering of major cell populations (Figure S1D; Table S1) identified 14 $CD4^+$ T cells, 13 $CD8^+$ T cells, and 10 myeloid subpopulations across the ccRCC tumors (Figures 2A–2C), showing marked heterogeneity in the immune cells comprising the TME across patients. Based on the expression of CD45RA, CCR7, IL-7R, CD27, IL-2R (CD25), FoxP3, CD57, and granzyme B (GRB), subpopulations of conventional memory (TCM), effector memory (TEM), effector (TEF), cytotoxic $GRZB^+$ (Cyto), and regulatory T cells (Tregs) were identified (Figures 2B, 2C, S2C–S2E). PD-1, CD69, and CD38 markers primarily defined level 2 subpopulations for CD8 and CD4 T cells (Figure 2C).

The subpopulations for CD8 and CD4 T cells and monocytes were quantified per tumor region for each patient (Figure 2D). As expected, percentages of naive cells were negligible for tumor tissue (Ruf et al., 2015). TCM ($CCR7^+$) and TEM ($CCR7^-$) cells were observed in all patients (Figure 2D), representing one-half of the T cell population. In different regions, from 4% to 80% of CD8 T cells and up to 10% of CD4 T cells were classified as cytotoxic (GRB, $CD57^+$), suggesting the presence of active immune processes within tumor tissue. Among all T cells, approximately 50% of CD4 and CD8 T cells expressed PD-1 and were distributed across all subtypes except TCM. The percentage of $CD69^+$ T cells was increased across all patients, the majority of which may represent tissue-resident T cells. Up to 5% of TEFs expressed high $PD\text{-}1^+$ (Figure S2D) and $CD69^+$ ($TEX\ PD\text{-}1^+$), which are a potential exhausted population of both CD4 and

Figure 1. Multiregion ccRCC tumor immune cell composition determined by CyTOF

(A) Schematic of treatment history.

(B) Primary ccRCC tumor images (reproduced) of the selected biopsy locations (A1–A4).

(C) t-distributed stochastic neighbor embedding (t-SNE) projections of the major cell populations across the tumor regions for each patient (n = 6).

(D) Bar graphs of each patient's major cell-type differences per region (21 samples from six patients). See also Figures S1 and S2 and Table S1.

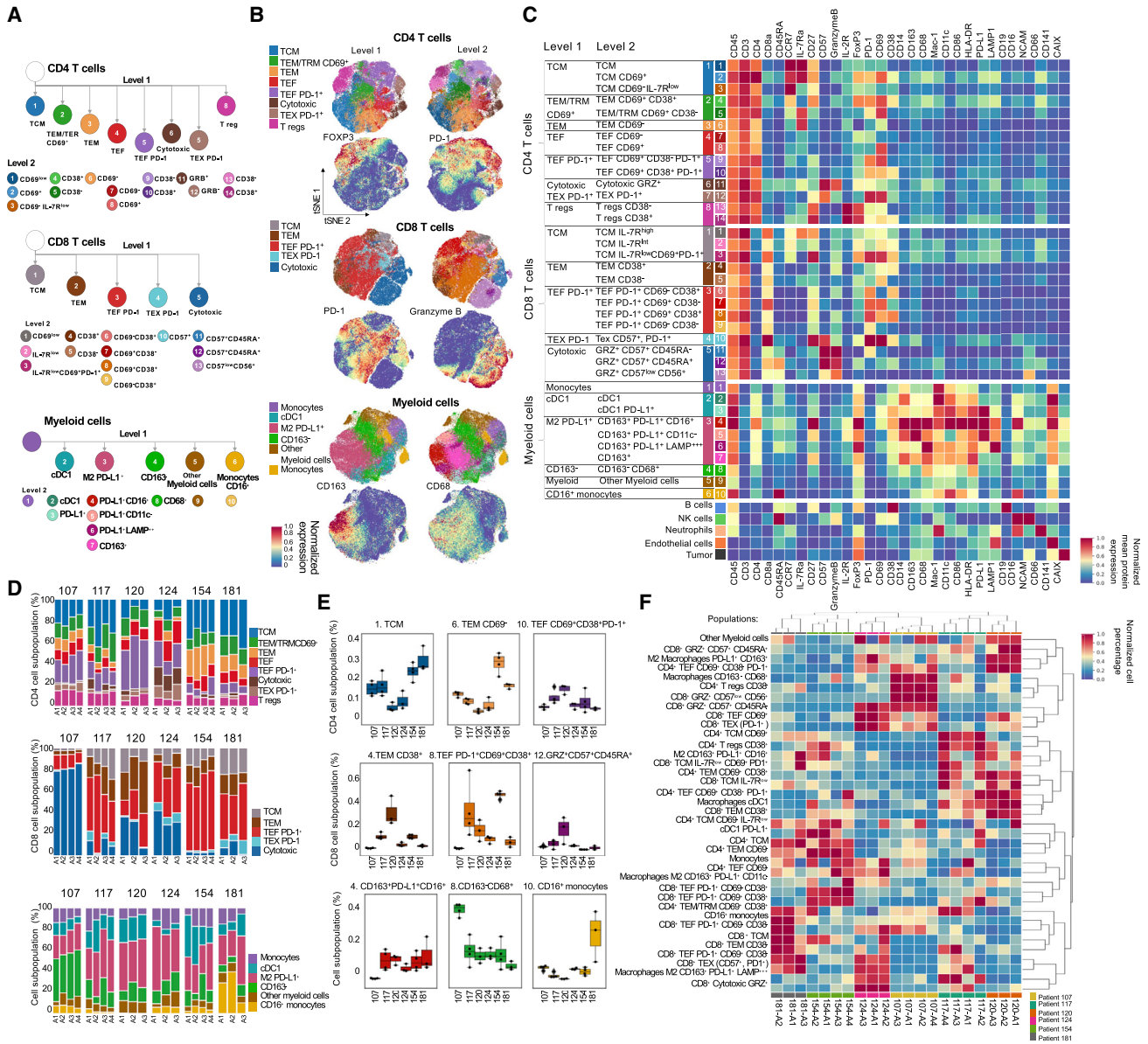


Figure 2. CyTOF analysis revealed a diverse set of cellular subpopulations conserved across regions

(A) CyTOF analysis of major cell populations and their subpopulations.

(B) t-SNE plots of level 1 and 2 subpopulations of CD4 and CD8 T cells and myeloid cells.

(C) Heatmap of the 30 markers across all cells depicting level 1 and 2 cell subpopulations.

(D) Bar graphs of each patient's relative content of level 1 CD4 and CD8 T cells and myeloid cells per region (21 samples from six patients).

(E) Frequencies of the most variable CD4 and CD8 T cells and macrophage subpopulations per patient (n = 6). In the boxplots, the upper whisker indicates the maximum value or 75th percentile + 1.5 interquartile range (IQR); the lower whisker indicates the minimum value or 25th percentile − 1.5 IQR.

(F) Clustermap with normalized [0–1] percentages of CD4 and CD8 T cells and myeloid subpopulations; dendrograms represent hierarchical clustering of subpopulations (rows) and regions (columns) (21 samples from six patients). See also Figures S1 and S2 and Table S1.

CD8 T cells. Interestingly, the proportion of CD4 Tregs was similar across patients (about 15%).

The percentages of CD4 memory T cell subsets were similar within a patient's tumor regions but differed compared with other patients (Figures 2D and 2E). For example, $CD69^-$ $CCR7^+$ $IL\text{-}7R^+$ $CD4^+$ TCMs were different in tumor regions for Pt120 and Pt124 (Figure 2D). Interestingly, all patients possessed $CD38^-$ Tregs, while $CD38^+$ Tregs were not found in Pt107 and Pt124 (Figure S2E). CD4 memory T cells were present in all patients, but CD8 memory T cell subpopulations were different between all patients (Figure 2D). Notable differences were also seen in the proportion of the TEF $PD\text{-}1^+$ $CD69^+$ $CD38^+$ CD8 T cells (Figure 2E) and different subsets of cytotoxic GRZ^+ $CD57^+$ $CD45RA^-$, GRZ^+ $CD57^+$ $CD45RA^+$, and GRZ^+ $CD57^{low}$ $CD56^+$ CD8 T cells (Figure S2E).

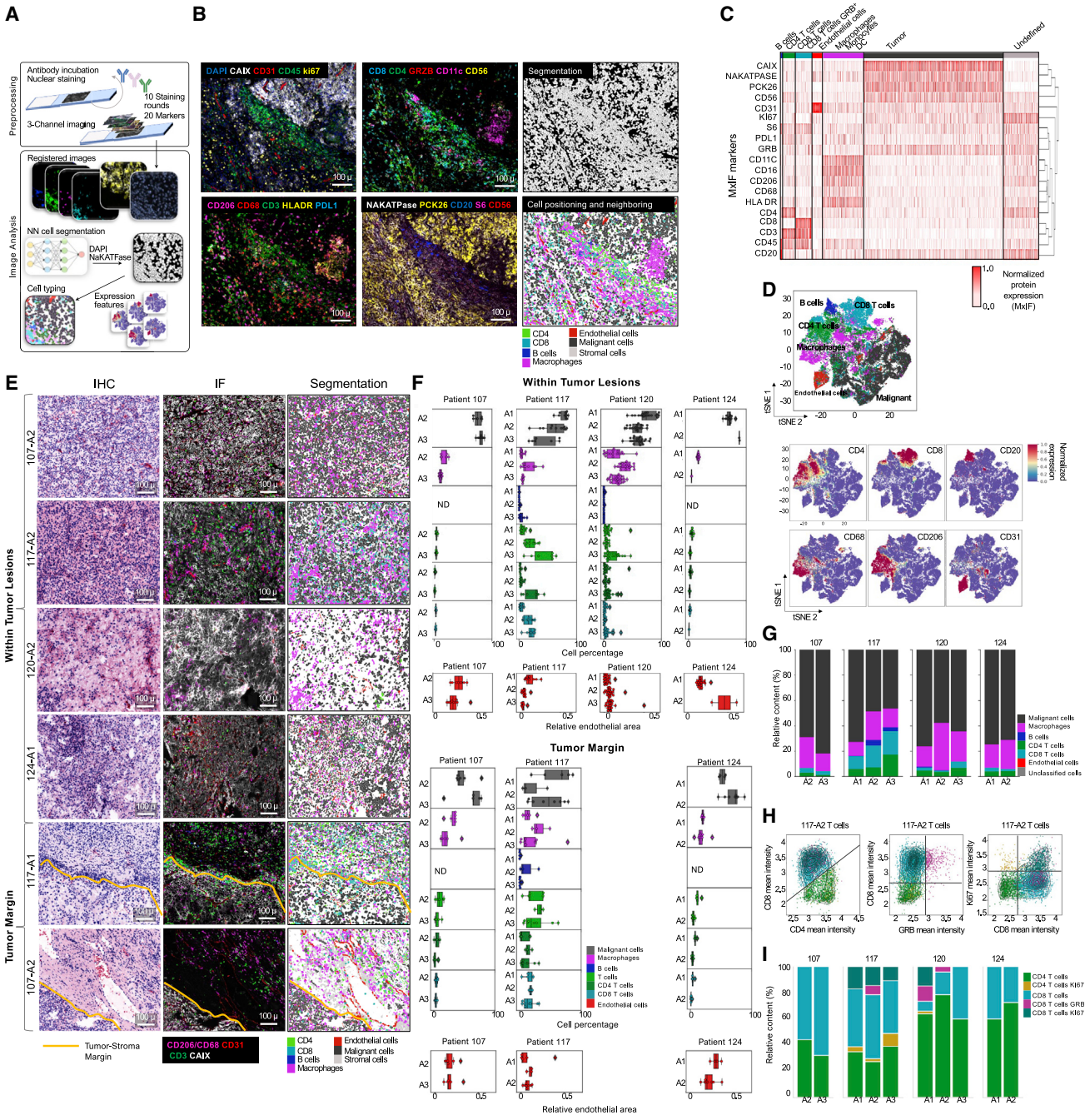


Figure 3. MxIF showed intratumor similarity across regions but substantial heterogeneity among patients

(A) Schematic overview of the automated ML-based cell segmentation pipeline (STAR Methods).

(B) Pseudo-color images of selected ccRCC intratumor regions. Scale bar, 100 μm.

(C) Heatmap of 19 markers across all cell segments measured by MxIF (77 ROIs, 10 samples from four patients).

(D) t-SNE plot of cell segments derived from MxIF analysis; cells are colored according to six major cell populations and expression levels of markers (10 samples from four patients).

(E) Representative images of ccRCC intratumor and tumor margin areas showing comparison of cell typing with MxIF imaging analysis (5 samples from four patients). Scale bar, 100 μm.

(legend continued on next page)

Myeloid populations were subclassified primarily based on the expression of CD14, CD163, CD68, Mac-1, CD11c, CD86, HLA-DR, PD-L1, CD141, and LAMP1, and included tissue classical ($CD16^-$) and non-classical monocytes ($CD16^+$), classical dendritic cells (cDCs) ($CD141^+$), and various macrophage populations predominantly expressing PD-L1 (Figures 2D and 2E). Populations of macrophages expressing CD163 or CD68 together represented more than 50% of the total number of myeloid cells.

A population of $LAMP^+$ macrophages, known cytokine producers, was also found (Figure 2C). Interestingly, $CD163^+$ $PD\text{-}L1^+$ macrophages (M2 PD-L1) had similar proportions across the majority of patients (Figure 2D). In contrast, cell types such as TCM CD4 T cells, TEM $CD38^+$ CD8 T cells, and CD16 monocytes showed interpatient differences (Figure 2E). Indeed, there was low interregion ITH, but marked interpatient heterogeneity, across these multiregion ccRCC tumors (Figure 2F). Interestingly, ICB-responsive Pt117 and Pt120 clustered together, suggesting similar TME signatures.

As our tumor dissociation method adapted for CyTOF was optimized at analyzing the tumor immune microenvironment, the tumor cell and stromal populations were low, hindering our ability to truly assess the tumor ecosystem. To provide spatial resolution of the tumor ecosystem and direct cell-to-cell communications, we used MxIF imaging of intact tumor tissue to analyze the different tumor regions.

Large-scale multiplex imaging analysis showed low ITH in tumor immune composition across regions but substantial heterogeneity among patients

Spatial ITH of ccRCC tumors was investigated (Figures 3A and 3B) with MxIF (Figures S1E and S3; Table S2). We employed ML-based segmentation (STAR Methods), resulting in the single-cell proteomic dataset of 860,000 intact cells from 10 regions of four patients (Figures 3C and 3D). Using a clustering approach similar to that applied in the CyTOF analysis, cellular populations per patient were identified based on the expression patterns of antibody-based markers: malignant ccRCC cells ($CAIX^+$, $PCK26^+$, $NAKATPASE^+$), B cells ($CD45^+$ $CD20^+$), various T cell populations ($CD45^+$ $CD3^+$), endothelial cells ($CD31^+$), macrophages ($CD11c^+$ HLA DR^+ $CD206^+/CD68^+$), monocytes ($CD11c^+$ HLA DR^+ $CD16^+$), and dendritic cells (DCs; CD11c) (Figures 3C and 3D). The ML-based segmentation pipeline identified tumors with marked spatial tumor heterogeneity in immune infiltration (Figure 3E). Cells were typed at two major tumor locations: intratumor regions and tumor margins (Figure S4). For example, in Pt117, intratumor and tumor margin regions were highly infiltrated by $CD4^-$ and $CD8^-$ T cells; Pt107 and Pt124 had much lower T cell density at the tumor margins (Figures 3E and 3F).

Pt107 and Pt124 showed the highest vascularization with endothelial cells calculated as a surface mask (Figures 3E and 3F). Notably, most regions within the same patients had no significant differences in the area of endothelial cells at either intratumor regions or margins, whereas in Pt124 there was a difference at the endothelial cell content between two tumor regions (Figure 3F). Overall, differences in the relative contents of macrophages, $CD4^-$ and $CD8^-$ T cells, and B cells were noted among intratumor regions in all patients. For Pt117, the greatest amount of T cells was found in the A3 region (37% of total cells versus 17% in the 117-A1 and 24% in the 117-A2 region; Figure 3G). Notably, Pt120 had a high proportion of macrophages and moderate amounts of T cells in all three regions. Tumor margin cellular composition analysis of all three regions of Pt117 detected significant infiltration by immune cells, especially $CD4^-$ and $CD8^-$ T and B cells (Figure 3F).

To further characterize the cellular subpopulations within each tumor region, additional functional markers were analyzed, including GRB, the immune activation marker, and the proliferation marker Ki67. Proliferating $Ki67^+$ $CD4^-$ and $CD8^-$ T cells were identified in three regions of Pt117 and one region (A1) of Pt120; cytotoxic GRB^+ CD8 T cells were identified in select regions of Pt117 and Pt120 (Figures 3H and 3I). In Pt117, no significant differences in the relative contents of proliferating $Ki67^+$ CD4 and CD8 T cells were found between all three tumor regions. Regions of Pt117 and Pt120 differed in the presence of T cell subpopulations, with five detected in regions A2 and A1, respectively (Figure 3I). Interestingly, activation markers, such as proliferation and GRB expression, were more prevalent in Pt117 and Pt120, who experienced responses to ICB.

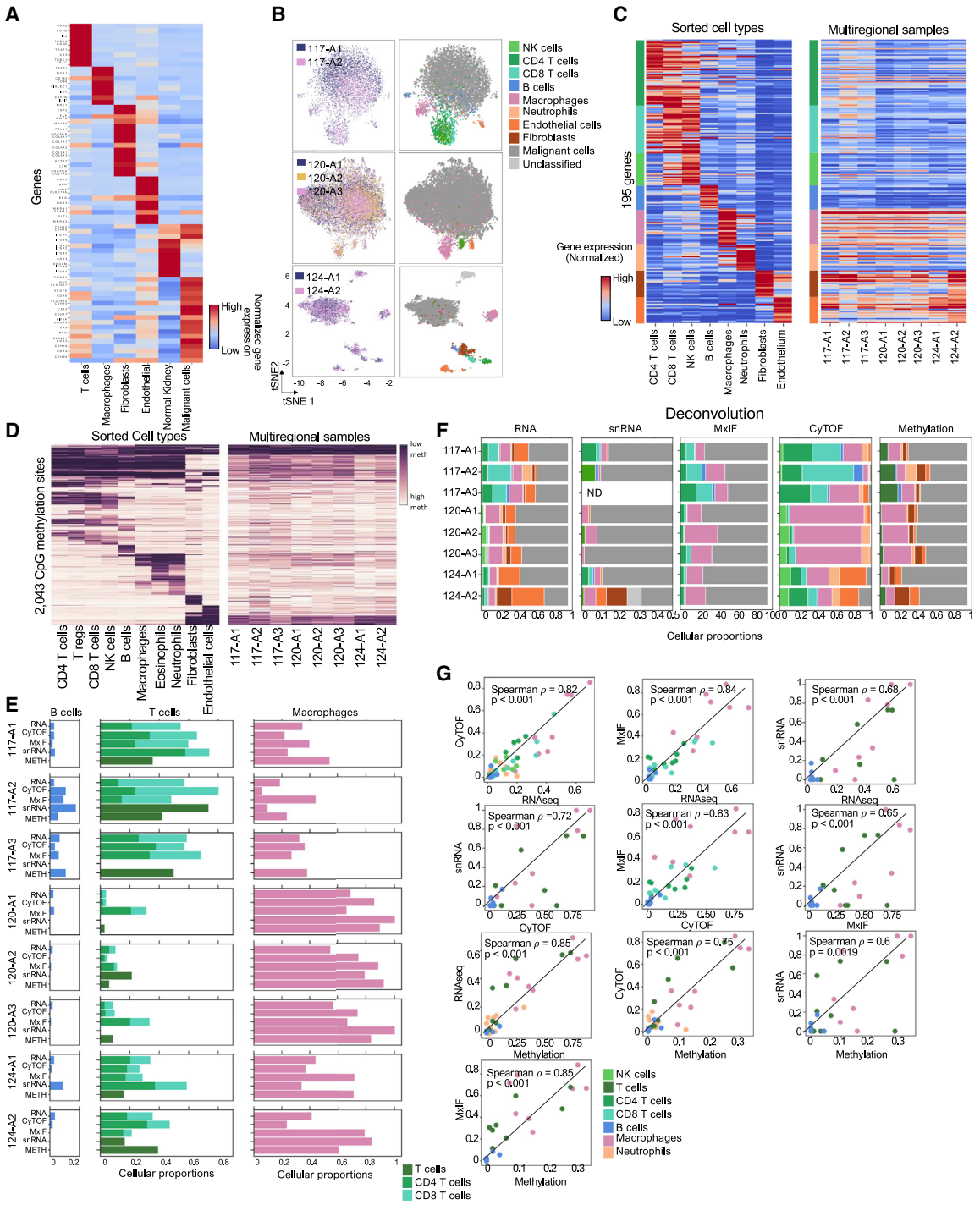
Overall, based on MxIF imaging analyses of cellular compositions and the TME, only minimal differences in the major cellular populations among intratumor regions were found in the ccRCC samples. Altogether, low ITH in immune composition was observed using both CyTOF and MxIF.

Multiple technologies correlated in determining the cellular composition of the TME

Surprising similarities identified in cellular composition among distant tumor regions by CyTOF and MxIF led us to perform a comprehensive inter-method comparison of the TME cellular composition using orthogonal technologies. snRNA-seq analysis of the same biopsies identified six cellular populations out of 27,000 sequenced cells (Figure 4A), including major TME cell types. In all samples, malignant ccRCC cells comprised more than 70% of all cells (Figure 4B), which is similar to the MxIF analysis, where malignant cells comprised approximately 50%–80% of the tumor regions (Figure 3G).

Further assessment of cellular heterogeneity used two additional methodologies that can differentiate the cell types present in a tumor: deconvolution of bulk RNA-seq using a custom algorithm (Zaitsev et al., 2022) (Figure 4C; Table S3) and

(F) Boxplots showing the relative cell-type percentage for tumor lesions and margins (10 samples from four patients); endothelial cell content is presented as a relative area of the CD31 marker mask. In the boxplots, the right whisker indicates the maximum value or 75th percentile + 1.5 interquartile range (IQR); the left whisker indicates the minimum value or 25th percentile − 1.5 IQR.
(G) Relative content of cell types (segments) identified per region per patient (10 samples from four patients).
(H) T cell subpopulation analysis for A2 region of Pt117; a single dot represents a distinct event.
(I) Relative percentages of T cell subpopulations per region per patient (10 samples from four patients). See also Figures S3 and S4 and Table S2 related to STAR Methods.



(legend on next page)

deconvolution from DNA methylation data (Figure 4D) by MethylCIBERSORT. Although the relative percentages of immune cell composition were largely concordant among the five methods (Figure 4E), the cumulative immune cell percentages varied across the different methods, as the immune TME of Pt117 and Pt120 mainly consisted of T cells and macrophages, respectively (Figure 4F). Among all methods, the highest correlations were between CyTOF and bulk RNA-seq deconvolution ($\rho = 0.82$ and $p < 0.001$) and MxIF and bulk RNA-seq deconvolution ($\rho = 0.84$ and $p < 0.001$) (Figure 4G). Moreover, MxIF and CyTOF ($\rho = 0.83$ and $p < 0.001$), as well as bulk RNA-seq deconvolution and methyl-seq deconvolution ($\rho = 0.85$ and $p < 0.001$), showed a strong correlation in TME composition prediction (Figure 4G). In addition to the immune cell content within the TME, the percentage of endothelial cells, inferred by RNA-seq and MxIF, was also strongly correlated (Figure 4H); however, endothelial cell prediction via deconvolution from bulk RNA-seq and methyl-seq did not correlate ($\rho = 0.38$ and $p = 0.35$). In contrast, fibroblasts predicted by RNA-seq and methylation deconvolution were strongly correlated (Figure 4H), demonstrating that these two technologies could assist in the estimation of fibroblasts, a marker of poor clinical outcome, including in ccRCC (Bagaev et al., 2021).

Marked ITH was observed in the cellular neighborhood and spatial architecture across different tumor regions

Cellular composition analysis alone provides no spatial information regarding the intricate cellular communication among different cell types within a complex tumor ecosystem that dictates tumor behavior and treatment responses. To decode these cellular interplays, we analyzed cell-to-cell interactions using MxIF and categorized distinct cellular neighborhoods encompassing distinct sets of cells (Figure 5A). To perform community analysis, we analyzed all available regions that were imaged for a particular sample, including individual regions of interest (ROIs) (15 ROIs) for Pt107 and Pt117 and full slide images for Pt120 and Pt124 (Figure S1C). For each sample, the cells analyzed were assigned to a neighborhood, resulting in a total of 14 neighborhoods: five ccRCC malignant cell -enriched, six T cell enriched, one macrophage enriched, one B cell enriched, and one endothelial cell enriched. Based on their biology, 11 of these cellular neighborhoods were grouped into malignant ccRCC-cell- and T-cell-enriched neighborhoods (Figures 5B and 5C). Proximity neighborhoods with a high density of ccRCC malignant cells interacting with macrophages and blood vessels, as well as immune cells, were identified (N1–N5) (Figures 5C and 5D). Additional T-cell-enriched neighborhoods (N6–N11) were also found, including T cells enriched at the tumor margin (N9) (Figures 5C and 5D).

The presence or absence of these different neighborhoods was determined across tumor regions within the tumor (Figures 5E and 5F) and at the margin (Figure S5A), showing marked interpatient heterogeneity in the immune cell interactions and neighborhoods comprising the TME among all patients (Figure 5F). Indeed, within the intratumor regions, the macrophage-enriched Pt120 had approximately 10 diverse neighborhood regions, primarily consisting of malignant ccRCC cells, macrophages, and CD4 and CD8 T cells, which were different from those of the other three patients (Figure 5F). Pt107 had the largest neighborhoods of interacting blood vessels (N3) and macrophages (N1, N2), with ccRCC malignant cells in both tumor regions (N1–N3) (Figure 5F). The variability among the tumors was associated with immune infiltration and immune enrichment in each tumor. Unlike cellular composition, which displayed little ITH, marked ITH in cellular neighborhoods was found in tumor regions from the same patient both in the tumor lesion and at the tumor margins, which is most strikingly apparent in Pt117 (Figures 5F and S5A). In the A1 region within the tumor lesion of Pt117, the tumor tissue had the neighborhoods of interacting blood vessels (N3) and macrophages intermixed (N1, N2), with ccRCC malignant cells that were not found in the A2 and A3 regions (Figure 5F); those regions were similar, whereas region A1 was distinct according to cellular neighborhoods (Figure S5B). In region A2 a T-cell- (N6, N7, N9–N11) and macrophage-enriched (N12) neighborhood was identified within the tumor and at the tumor margin, which was not as pronounced in the other regions. Notably, Pt117 had the highest levels of B cells within the tumor in regions A2 and A3 and in the tumor margin in region A2 (Figures 1C, 1D, 5E, 5F, and S5A), resulting in the presence of tertiary lymphoid structures (TLSs; N13), ectopic lymphoid organs that often arise in highly inflammatory environments and are linked to better patient clinical outcomes (Dieu-Nosjean et al., 2016). B-cell repertoire (BCR) analysis of total RNA-seq showed that regions A2 and A3 of Pt117 had a high fraction of BCR reads in bulk RNA-seq (Figure S5C), confirming TLS presence in N13. Further analysis showed that expanded major immunoglobulin heavy chain (IGH) and immunoglobulin kappa chain (IGK) clonotype families, comprising several BCR clones that arose from hypermutation, caused a high fraction of BCR reads (Figures S5D and S5E). Intriguingly, Pt117 had a complete response to ICB, highlighting the potential role of both B cells and TLSs in this response (Helmink et al., 2020; Simonaggio et al., 2021). Pt117 and Pt120, who both responded to immunotherapy, possessed the T-cell-enriched N7 and N9 neighborhoods.

Figure 4. Correlations of immune cell composition of the TME across methods

(A) Heatmap of 71 cell-type-specific markers based on snRNA-seq analysis of region A2 of Pt124.

(B) snRNA-seq t-SNE plots showing cell populations per patient (7 samples from three patients).

(C) Left: heatmap of the cell-type-specific gene expression of 195 genes selected for RNA-seq deconvolution. Right: heatmap of the cell-type-specific gene expression of 195 genes (Zaitsev et al., 2022) for each patient and region (8 samples from three patients).

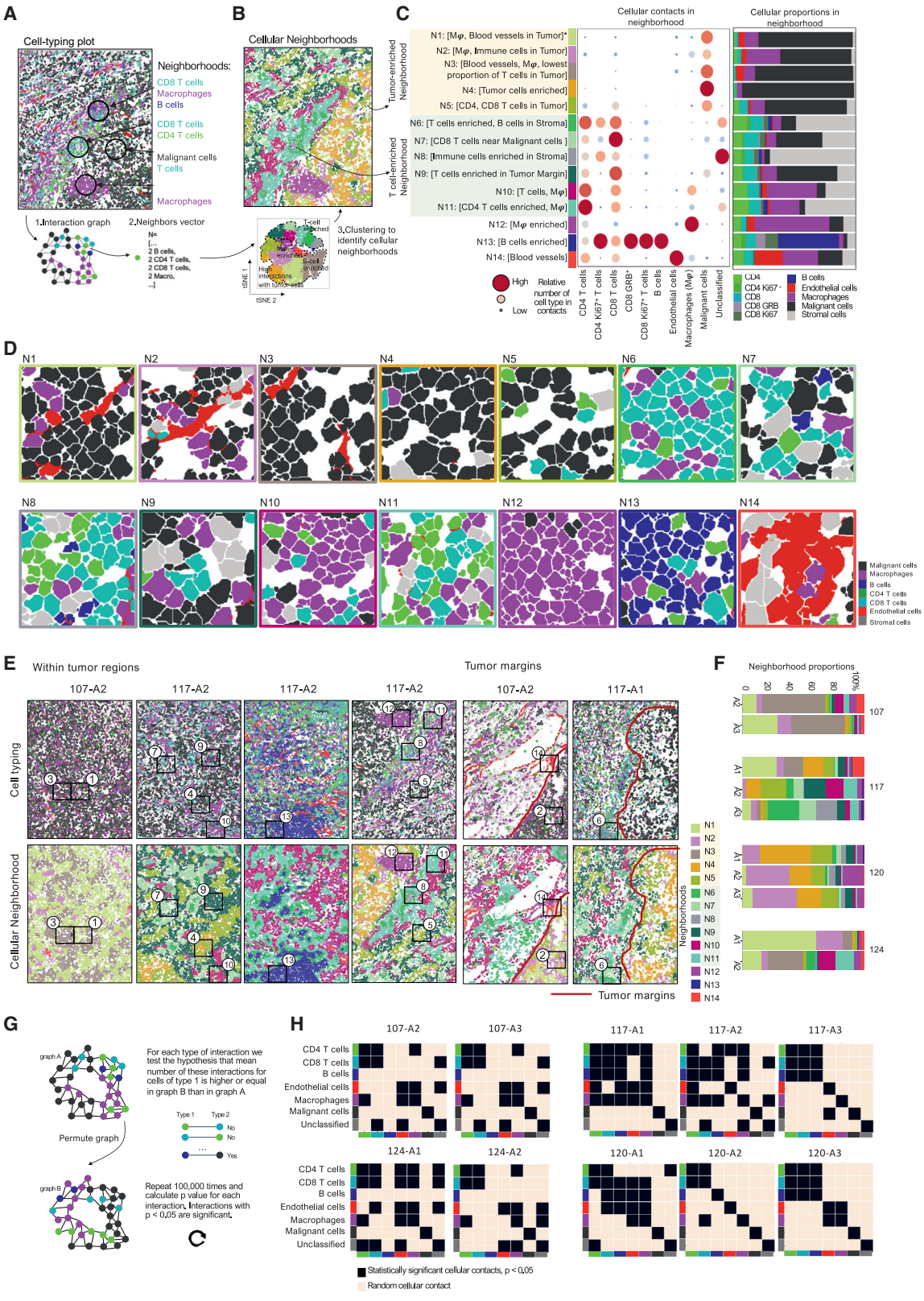
(D) Left: heatmap of the methylation levels of 2,043 CpG methylation sites selected for deconvolution per cell type. Right: heatmap of the methylation levels of 2,043 CpG methylation sites for each patient and region listed (8 samples from three patients).

(E) Relative contents of B and T cells and macrophages derived from each analysis method for all patients and regions (8 samples from three patients).

(F) Relative cellular percentages per patient and region derived from each analysis method (8 samples from three patients).

(G) Spearman correlation in the enumeration of relative immune cell contents among the analysis methods (n = 5).

(H) Spearman correlation of the percentage of endothelial cells and fibroblasts calculated using the listed methods (n = 5). See also Table S3.



(legend on next page)

To further support the significance of the identified neighborhoods, the non-randomness of cellular interactions was calculated using a permutation test (Jackson et al., 2020) (Figures 5G, 5H, and S5F). Interestingly, despite the neighborhood heterogeneity of individual regions within tumors, cellular interactions remained similar across regions (Figures 5H and S5F). Different regions in Pt107 and Pt124 had a low proportion of immune cells (Figure 5H). Pt117 had similar direct contacts in regions A1 and A2 within the tumor lesion, and A2 and A3 were similar in the margin areas (Figures 5H and S5F). Analysis of cellular neighborhoods supports the marked heterogeneity observed in all three regions of Pt117 in both tumor lesions and margin areas (Figures 5H and S5F), showing that the formation of TLSs (N13) emphasizes the observed heterogeneity. Regions in Pt120 that were enriched with macrophages showed cellular interactions between macrophages and endothelial cells (N3), as well as macrophages and T cells (N1–N2) (Figures 5F and 5H). Together, these results suggest somewhat consistent tumor biology and cell behavior among regions, supporting the idea that malignant ccRCC cells influence the TME.

Subclonal genetic heterogeneity was associated with methylation patterns and TME heterogeneity

To further uncover the underlying genetics that may influence the observed heterogeneity among patients, the mutation, chromosome, and methylation landscapes of the tumor regions were studied (Figures 6A–6C). Mutations from bulk RNA-seq (Figure 6A) were called to highlight expressed mutated proteins and support mutation calls from WES, and somatic mutations expressed in single tumor cells were identified using snRNA-seq (Figures 6A and S6A). The proportion of tumor cells (purity) measured by various methods correlated poorly (Figure S6B). In Pt124, a decrease in 3p gene expression and an increase in 5q gene expression were observed in tumor cells (Figure S6C), which correspond to a 3p deletion and 5q amplification. Mutations found at both DNA and RNA levels may represent the core set of alterations driving tumor ecosystem formation and spatial heterogeneity with the caveat of the inability to find expressed mutated transcripts from mutations driving nonsense-mediated decay or associated with losses of the entire gene.

The mutations were divided into three groups based on their clonality and presence in the different tumor regions (Figure 6A): truncal, clonal in at least one region and present in all regions; region specific, clonal in at least one region and potentially subclonal in the other regions; and subclonal, present in any region (Figure S6D). Truncal mutations were mostly found by all types of analysis, whereas subclonal mutations were typically found only by WES (Figure 6A). For each patient, we constructed a phylogenetic tree with truncal and region-specific events (Figure 6D; STAR Methods) and analyzed associations between genetic events and TME composition inferred by multiple platforms (Figure 6D). In Pt117, truncal nonsense mutations in *BAP1* and *VHL* were identified in all three regions by WES and RNA-seq-based mutation calling, as well as mutations in *TET3* and *USP20*. Region-specific mutations in *HSPA6* and *POLK* were found in regions A1 and A3, respectively, with a subclonal frameshift mutation in *SET2D* and a missense *TP53* mutation in region A1 found by WES and RNA-seq mutation calling (Figure 6A). Pt117 also displayed a hypermethylated genome in region A1 (Figure 6C), with methylation at *PPP2R2C* in all three regions and *EPO* methylation only in region A1 (Figure 6D), a region primarily composed of macrophages and T cells as well as macrophage-enriched neighborhoods (N1, N2) (Figure 6D). Notably, the truncal *BAP1* mutation may underlie the observed T cell enrichment in these regions (Brugarolas et al., 2020). Pt120 showed a different mutation profile compared with Pt117 and Pt124. Truncal *CREBBP* and *PBRM1* missense mutations were identified in all three regions with all three sequencing platforms (WES, RNA-seq, and snRNA-seq), as well as *TSC1* splice-site mutations in all regions, which are similar in composition and architecture (Figure 6A). Interestingly, the subclonal *TSC1* in-frame insertion mutation was not found in the A3 region of Pt120 by WES; instead, RNA-seq-based mutation calling led to its identification (Figure 6A). *TACC1* methylation was also observed in all three regions (Figure 6D). Moreover, the *PBRM1* mutations identified in Pt120 may underlie the high levels of angiogenesis and lower T cell infiltration as reported (Braun et al., 2021) (Figures 6A and 6D). Pt124 had a truncal missense *VHL* mutation and *PBRM1* and *RHOB* mutations in region A1 (Figure 6A). WES analysis also identified *SETD5* (region A1), *KDM5C*, and *STAG2* (region 2) subclonal mutations, with hypermethylation in region A1 (Figures 6A and 6C). Region A1 of Pt117 and Pt124 had the highest whole-genome hypermethylation levels (Figure S6E) and had mutations in *SETD2* and *SETD5* (Figure 6A), respectively, which are lysine methyltransferases whose disruption may cause genome-wide epigenetic changes (Cancer Genome Atlas Research Network et al., 2013) and genome hypermethylation. This hypothesis is supported by an analysis of the TCGA: phs000178 cohort that showed that patients with *SETD2* or *PBRM1* mutations have higher genome methylation levels

Figure 5. Cellular neighborhoods and cell-to-cell interactions among the different tumor regions of the same patient

(A) Cell-typing plot of cell-to-cell interactions; schematic overview of the graph-network-based unsupervised clustering process.

(B) Top: cell-typing plot of cellular neighborhoods (different colors). Bottom: t-SNE plot of cells clustered into neighborhoods (different colors) based on cell-to-cell interactions graph (10 samples from four patients).

(C) Left: bubble heatmap of cell types identified in each neighborhood. Right: bar plot of the relative cell percentage per neighborhood (10 samples from four patients).

(D) Cell-typing plots of the cell-to-cell interactions across the patients for each neighborhood.

(E) Top: cell-typing plots of representative tumor regions of Pt107 and Pt117 showing immune cell composition and distribution of immune cells. Bottom: proximity neighborhood plots for the same regions.

(F) Bar graphs showing the proportions of proximity neighborhoods identified by MxIF for tumor regions in Pt107, Pt117, Pt120, and Pt124 (10 samples from four patients).

(G) Schematic diagram of the permutation test utilized to assess significance of the computed cellular contacts.

(H) Heatmaps showing statistically significant (non-random) cellular contacts for the listed regions per patient (10 samples from four patients). See also Figure S5.

CellPress
OPEN ACCESS

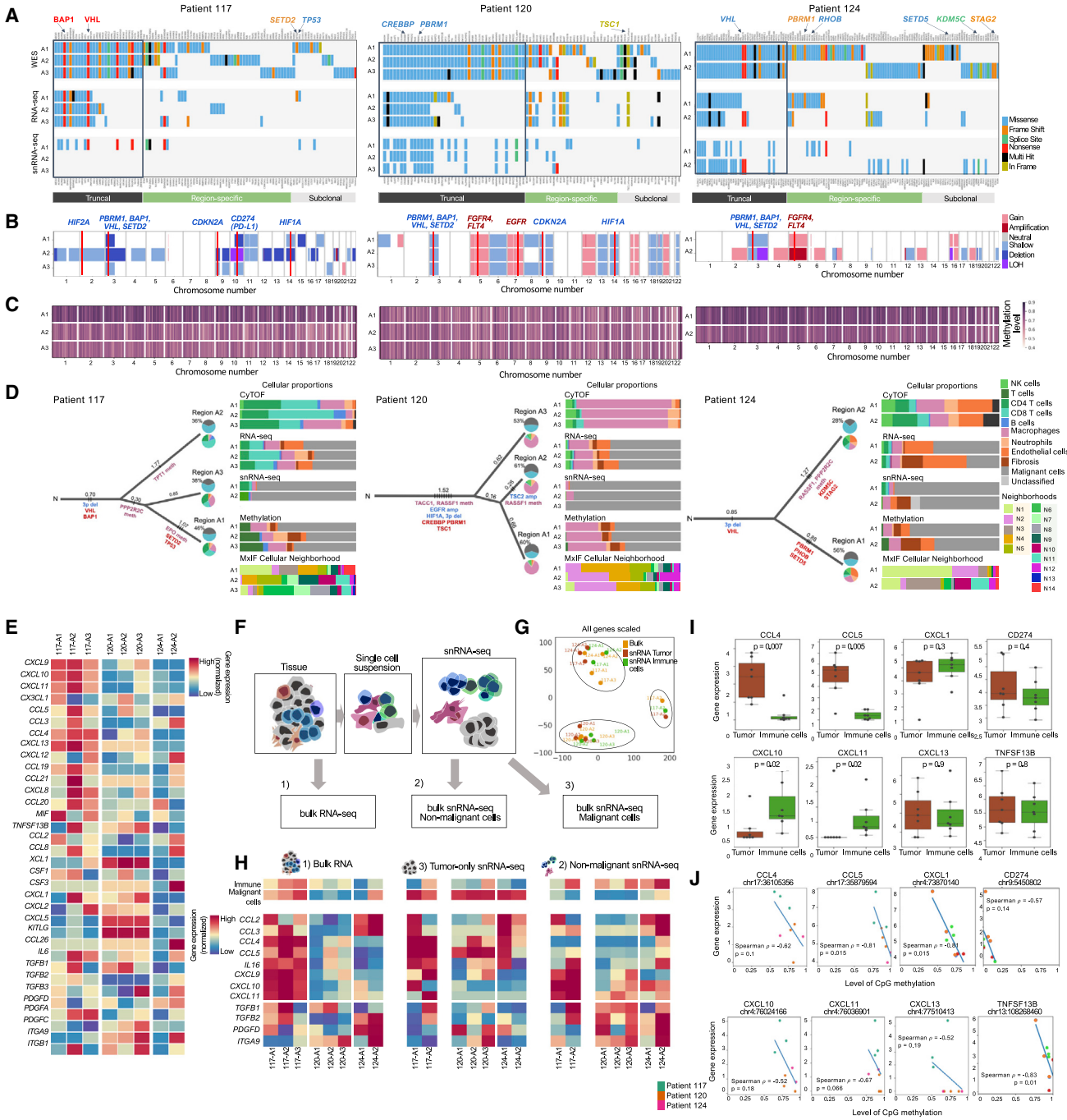


Figure 6. Association of genetic heterogeneity with TME composition and production of inflammatory cytokines by malignant cells

(A) Oncoplots of identified variants called by WES, RNA-seq, and snRNA-seq per patient (8 samples from three patients).

(B) Copy number gains and losses per region per patient (8 samples from three patients).

(C) Methylation heatmap of the entire genome colored by 1-Mb bins per patient (8 samples from three patients).

(D) Trees of main mutational events that led to the genetic diversity of regions per patient. Branches represent potential driver mutations; the nodes/leaves show the cellular composition of regions. Pie charts represent the ratio between malignant and non-malignant cells; charts show the composition of non-malignant cells calculated by CyTOF. Top four bar plots represent the cellular composition of regions calculated by CyTOF (8 samples from three patients), RNA-seq (8 samples from three patients), snRNA-seq (7 samples from three patients), and methylation (8 samples from three patients). Bottom bar plot represents MxIF cellular neighborhood per patient (77 ROIs; 8 samples from three patients).

(E) Heatmap of gene expression of cytokines derived from RNA-seq (8 samples from three patients).

(F) snRNA-seq pseudo-bulk and bulk RNA-seq analysis scheme.

(legend continued on next page)

(Figure S6F). The A1 region of Pt124 had the highest endothelial cell content (Figure 6D), which may be linked to the identified *PBRM1* mutation (Carril-Ajuria et al., 2019). Moreover, Pt124 had a 5q amplification, which was consistent with snRNA copy number alteration (CNA) findings (Figures 6B and S6C). Mutations found in *VHL*, *PBRM1*, *SETD2*, and *BAP1* cause heterogeneity among patients and affect individual therapy response for the treatments (Hsieh and Cheng, 2020).

TME-associated cytokines are influenced by the methylation patterns of malignant cells

To understand the driving factors underlying the observed inter-patient heterogeneity in TME cellular populations and tumor architecture, individual cytokine expression was assessed at the transcriptional level using bulk RNA-seq (Figure 6E). Pt117, who had large percentages of CD4 and CD8 T cell populations, as well as B cells that formed TLSs (Figures 1C, 1D, 5E, S5D, and S5E), had high expression of the pro-inflammatory cytokines CXCL9, CXCL10, and CXCL11 (Figure 6E). Region A2 of Pt117, with the highest immune infiltration, showed that the majority of cytokines were associated with T cells and B cells (e.g., CCL4 [median-scaled score = 1.7], CXCL9 [median-scaled score = 2.5], CXCL10 [median-scaled score = 2.1], CXCL11 [median-scaled score = 1.9], and CXCL13 [median-scaled score = 3]). CCL5 expression was also higher in immune-enriched Pt117. All regions of Pt120, which were macrophage enriched as shown by CyTOF and MxIF, demonstrated the highest expression of the granulocyte-attracting chemokine CXCL5 of all the patients. In regions A1 and A2 of Pt124, noted by high endothelial cell content (Figure 3F) and the presence of blood vessel communities (Figure S5A), the highest expression of platelet-derived growth factor-D (PDGF-D) (Figure 6E), a potent mitogenic factor known to be expressed in the kidney and whose expression may be inhibited by TGF-β signaling, was observed.

Based on the notable differences in cytokine expression across all the patients determined by bulk RNA-seq analysis, the cells controlling this differential cytokine expression were narrowed down to immune microenvironment or tumor cell populations based on snRNA-seq analyses (Figures 6F, 6G, and 6H). First, we confirmed that the bulk RNA-seq and pseudo-bulk snRNA-seq (all reads from snRNA-seq mixed together to imitate bulk RNA-seq) (Figure 6F) expression analyses correlated, demonstrating that comparative analyses could be implemented. Samples from the same patient clustered together, after scaling of all bulk samples and pseudo-bulk samples from tumor and immune cells (Figure 6G). Based on the snRNA-seq results, CXCL10 and CXCL11 were predominantly expressed by immune cells, whereas other cytokines such as CCL4 and CCL5 were expressed by the malignant ccRCC cells (Figure 6I). Several studies have demonstrated that epithelial cells in colorectal cancer are also a major source of CCL4 (De la Fuente López et al., 2018). The expression of cytokines by ccRCC malignant cells suggests that this expression may modulate the composition and activity of specific TME and tumor spatial architecture via context-specific cytokines, similar to previously identified CCL5-CCR5 interactions in ccRCC (Gelbrich et al., 2017; Li et al., 2018; Yoo et al., 2015; Zhou et al., 2020). Lastly, the expression of several cytokines, including CCL5, inversely correlated with CpG methylation at their respective promoter regions (Figure 6J), linking their methylation with their expression. CCL5 and CXCL9 co-expression revealed immunoreactive tumors with prolonged survival and response to checkpoint blockade (Dangaj et al., 2019). Intriguingly, in immune-enriched Pt117, the CCL4 and CCL5 promoters had the lowest methylation level. DNA methylation is a principal epigenetic mechanism negatively regulating CCL5, in agreement with prior evidence in lung and colon cancer (Dangaj et al., 2019; Li et al., 2014).

Cytokines correlate with the formation of cellular neighborhoods and treatment outcome

To investigate if the differences in cytokine expression were linked to the spatial tumor architecture as well as the formation of distinct neighborhoods, we performed multiomics integrated correlation analysis between proximity neighborhoods measured by MxIF, the relative content of immune cell types by CyTOF, and gene expression analysis using RNA-seq data (Bagaev et al., 2021). Six highly correlated groups were identified (Figure 7A). Group 1, the largest, consisting of immune inflammation in ccRCC samples and B-cell- (TLS) and T-cell-enriched neighborhoods (Figure S7D; ROI 1), correlated with the relative content of $CD8^+$ TEM, $CD8^+$ TEM PD-1^+, B cells, and CD8 TCM, measured by CyTOF, as well as the expression of T-cell traffic, B and T cells, checkpoint inhibition, and effector cell gene signatures (Figures 7A and 7B). Specific T cell neighborhoods also independently correlated with RNA expression of CCL4 (r = 0.66) and CCL5 (r = 0.88) (Figure S7A).

Group 2 was associated with CD4 T helper cells; gene expression signatures in this cluster associated with T cell trafficking, T cells, and Th1 and Th2 signatures correlated with the relative content of CD4 TEM, CD4 TEM $CD69^+$, CD4 TEF PD-1^+, and CD4 Treg (CyTOF) (Figures 7A and 7B). The macrophage-enriched group 3 (Figure S7C) included correlations between T cells and CD4 T cells in particular, and the gene expression signatures associated with macrophage and DC trafficking as well as pro-tumor cytokines (Figures 7A and 7B). Another correlation was found in group 4 between blood-vessel-enriched proximity neighborhoods (N1 and N3) and the expression of angiogenesis and endothelial gene signatures (Figure S7A), which coincided with the relatively

(G) Principal component analysis (PCA) of scaled gene expression of bulk RNA-seq (8 samples from three patients) and snRNA-seq (7 samples from three patients) samples clustered by patients.
(H) Heatmaps of the scaled expression of 12 cytokines in bulk RNA-seq (8 samples from three patients) and tumor-only and non-malignant parts of snRNA-seq (7 samples from three patients) expression.
(I) Boxplots showing the expression of eight cytokines from malignant cells (tumor-only snRNA-seq) and immune cells (non-malignant snRNA-seq) (7 samples from three patients). In the boxplots, the upper whisker indicates the maximum value or 75th percentile + 1.5 interquartile range (IQR); the lower whisker indicates the minimum value or 25th percentile − 1.5 IQR.
(J) Correlation plots of cytokine expression with single CpG methylation in the promoter region (8 samples from three patients). See also Figure S6.

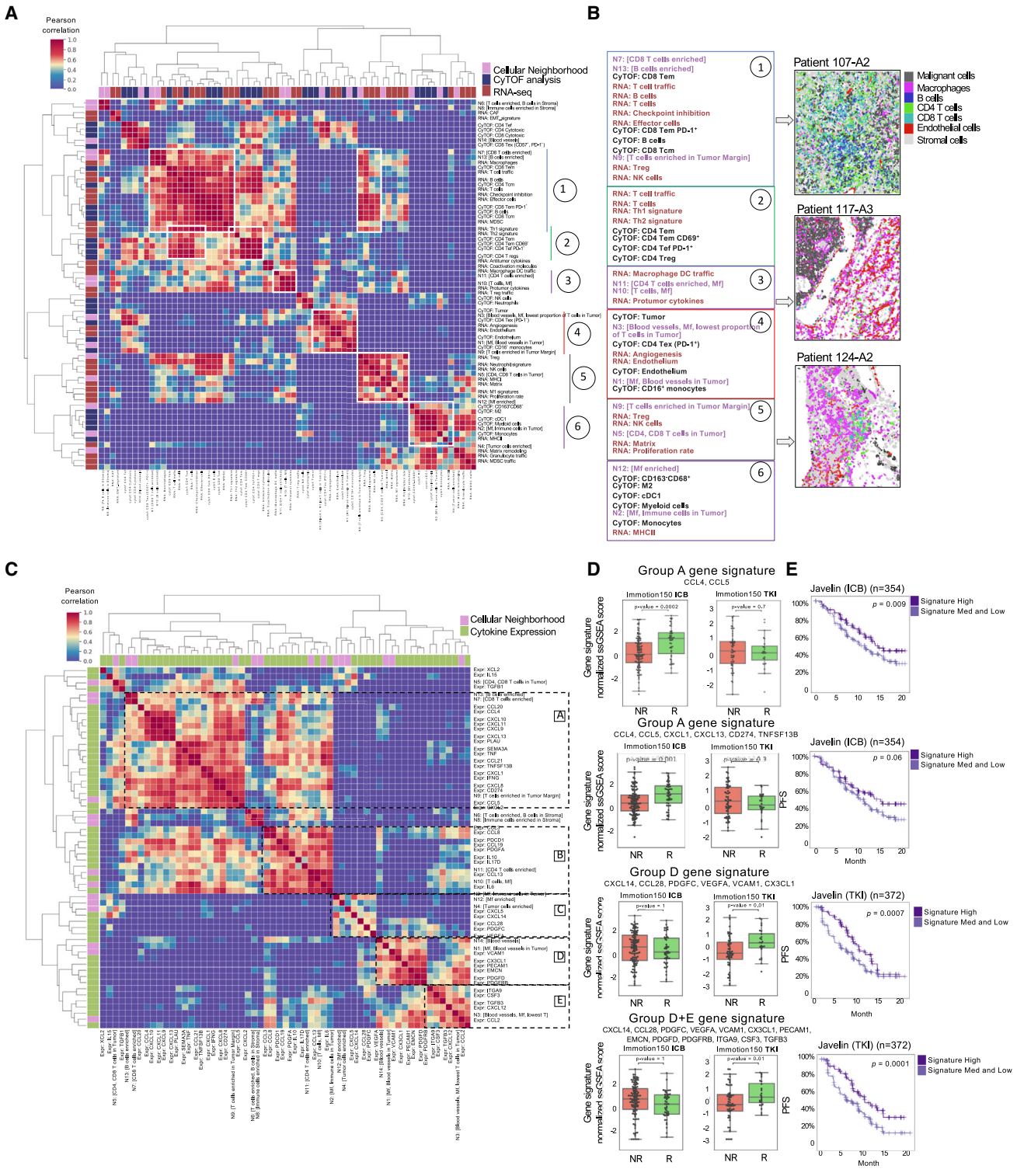


Figure 7. Intraregional TME and cellular neighborhoods are formed by the context of specific cytokine expression

(A) Clustermap of pairwise correlation between proximity neighborhoods by MxIF (eight samples from three patients), relative content of immune cell types by CyTOF (eight samples from three patients), and gene expression signature by RNA-seq data (eight samples from three patients). Colors represent Pearson correlation coefficient. Main correlation clusters are highlighted in white and supplemented with the numbers. Row and column annotation represents the analysis methods used.

(legend continued on next page)

high content of endothelial cells by CyTOF (Figures 7A and 7B). Group 5 identified an association between T cell neighborhoods N5 and N9 and Treg, natural killer (NK), matrix, and proliferation gene signatures. Neighborhoods N2 and N12, comprising myeloid cells, correlated with high percentages of macrophages, monocytes, and DCs (group 6, Figures 7A and 7B).

Cytokine expression was also correlated with neighborhood formation in these different tumors, resulting in five highly correlated groups of cytokines, important ligands/receptors, and neighborhoods (Figure 7C). In group A, CD8 T cells near malignant cell neighborhoods (N7), T cells enriched in tumor margin neighborhoods (N9), and B-cell-enriched neighborhoods (N13) correlated with the expression of diverse cytokines such as CCL4 ($r = 0.66$ for correlation with N7), CCL5 ($r = 0.88$ for correlation with N9), and CXCL13 ($r = 0.7$ for correlation with N13) (Figures 7C and S7A). Expression of CCL4 and CCL5 corresponded to the presence of tumor infiltrating lymphocytes (TILs), which is consistent with previous studies (Dangaj et al., 2019). Notably, N7, N9, and N13 were identified only in patients who responded to ICB (Figure 5F). Blood-vessel-enriched areas (N1, N3, N14) correlated with VCAM1 ($r = 0.84$ for correlation with N1) and VEGFA expression ($r = 0.7$ for correlation with N3) in group D as well as TGF-B3 ($r = 0.62$ for correlation with N3), ITGA9 ($r = 0.53$ for correlation with N3), and CXCL12 ($r = 0.8$ for correlation with N3) in group E (Figures 7C and S7A).

We further linked ccRCC clinical outcome and survival with specific cytokines correlating with the different TME neighborhoods and communities in three different cohorts: TCGA-KIRC: phs000178 ($n = 511$), the JAVELIN: https://doi.org/10.1038/s41591-020-1044-8 RCC clinical trial cohort (Motzer et al., 2020a) comparing avelumab (anti-PD-L1) plus axitinib ($n = 354$, ICB) with sunitinib ($n = 372$, tyrosine kinase inhibitor [TKI]), and the IMmotion150: EGAS00001002928 cohort (McDermott et al., 2018) comparing atezolizumab (anti-PD-L1) ± bevacizumab ($n = 169$, ICB) with sunitinib (TKI) ($n = 89$). A group A gene signature was identified that correlated with response to immunotherapy in the IMmotion150: EGAS00001002928 ICB cohort ($p = 0.001$ to $p = 0.0002$; Figure 7D) and could somewhat stratify survival in the JAVELIN: https://doi.org/10.1038/s41591-020-1044-8 ICB cohort by tumor-only derived cytokines CCL4 and CCL5 ($p = 0.02$). Notably, CCL5 expression alone significantly stratified poor patient survival in the TCGA: phs000178 cohort ($p = 0.02$) yet was associated with response in IMmotion150: EGAS00001002928 ICB ($p < 0.001$, Figure S7B) and survival in JAVELIN: https://doi.org/10.1038/s41591-020-1044-8 ICB cohorts ($p = 0.02$, Figure S7B). CCL4 correlated with response in the IMmotion150: EGAS00001002928 ICB cohort (Figure S7B). High CXCL13 expression was associated with immunotherapy response in the IMmotion150: EGAS00001002928 cohort and the worst survival in TCGA patients ($p = 0.02$) (Figure S7B). In contrast, the group D cytokines and the combined groups D and E gene signature significantly correlated with response to TKI ($p = 0.01$, 0.01) in the IMmotion150: EGAS00001002928 (TKI) and stratified survival to TKI ($p = 0.0007$, $p = 0.0001$) in the JAVLEIN: https://doi.org/10.1038/s41591-020-1044-8 (TKI) cohorts (Figures 7D and 7E).

DISCUSSION

Significant ITH heterogeneity in ccRCC has impeded the application of precision medicine strategies for kidney cancer patients and may be responsible for the clinical differences in patient responses to treatments. We investigated the downstream effects of this genetic heterogeneity on the tumor ecosystem, comprising malignant ccRCC cells and their TME, using a multiplatform hierarchical approach to dissect tumor behavior at the epigenetic, genetic, transcriptomic, proteomic, and spatial levels of biopsies collected from multiple tumor regions. Notably, we did not observe marked ITH in immune microenvironment cellular composition as measured by CyTOF and MxIF, suggesting that subclonal genetic ITH does not dramatically alter the type of cells comprising the TME. Instead, surprisingly, we found significant ITH in the architecture of the TME across the different tumor regions of the individual patients, specifically in the presence of diverse cellular neighborhoods. The formation of the various neighborhoods and dissimilarities in TME spatial architecture have been linked to differences in the antitumor activity of targeted and immunotherapy agents, potentially through variations in drug distribution and residence of immune cells within the tumor regions due to these differences in spatial organization (Schürch et al., 2020). In our limited cohort, the only patient (Pt117) who achieved a complete response to ICB had marked spatial ITH, but all the tumor regions were immune enriched, including the residence of a TLS, suggesting that in this very limited sample size the presence of immune-enriched neighborhoods and the lack of stroma-enriched areas determined response to ICB. Evaluation of the cellular populations in the context of the TME spatial architecture in larger cohorts will be essential to validate the application of ITH in clinical practice.

The underlying genetic and molecular drivers that sculpt the formation of these distinct communities and neighborhoods were also investigated at the epigenetic, genetic, and expression levels. A truncal nonsense mutation in *BAP1* was identified in the immune-enriched tumor regions of Pt117 consisting of high levels of CD4 and CD8 T cells and B cells (TLS). *BAP1* mutations have been previously associated with increased immune infiltration in

(B) Composition of the most important functional clusters in (A) and cell-typing plots corresponding to clusters showing the inflammation type within the tumor region and tumor margin area.

(C) Clustermap of correlation between different proximity neighborhoods by MxIF (eight samples from three patients) and cytokine expression by RNA-seq (eight samples from three patients). Colors represent Pearson correlation coefficient. Main correlation clusters are highlighted by black dashed lines and supplemented with the letters. Row and column annotation represents the analysis methods used.

(D) Boxplots of the gene-signature-normalized single-sample gene set enrichment analysis (ssGSEA) score in non-responders and responders in ICB ($n = 354$) and tyrosine kinase inhibitor (TKI) ($n = 372$) JAVELIN: https://doi.org/10.1038/s41591-020-1044-8 cohorts.

(E) Association between progression-free survival and activity of cytokine groups A, D, and D + E (shown in [D]) in ICB ($n = 354$) and TKI ($n = 372$) JAVELIN: https://doi.org/10.1038/s41591-020-1044-8 cohorts. Genes from cytokine groups were analyzed as a signature. See also Figure S7.

kidney cancer (Wang et al., 2018) and immunogenic microenvironments in various cancers (Shrestha et al., 2019; Souri et al., 2019), potentially through the upregulation of different chemokines and cytokines, including CXCL13. Here, we found that CXCL13 expression correlated with B-cell-enriched neighborhoods (TLS) in the *BAP1*-mutant Pt117, further linking *BAP1* mutations to diverse immune-enriched microenvironments, including TLS formation. Further, whole-genome methylation was higher in the A1 regions of Pt117 and Pt124 and possessed frameshift mutations in chromatin remodeling genes (*SETD2* and *SETD5*), which have been shown to correlate with higher average methylation (Figure S7F) (Cancer Genome Atlas Research Network et al., 2013).

Interestingly, this work supports the idea that the ccRCC TME composition and architecture is shaped by different genetic drivers that act through changes in oncogenic signaling pathways to alter chemokine/cytokine signaling, directly influencing patient therapeutic outcomes. We found that CCL4 and CCL5 cytokines are directly expressed by malignant cells, potentially due to changes in the methylation profile of their promoters, and that this expression correlated with the formation of spatial CD8 T cells near malignant cells (N7) or at the tumor margin (N9) and the presence of high percentages of PD-1$^+$ CD8 T cells. Although we were not able to identify an underlying genetic regulation of methylation changes, this finding suggests that examining cytokine-mediated cell trafficking to the tissue is crucial to elucidating the formation of specific tumor spatial organization patterns that have an impact on clinical outcomes.

Overall, this work highlights the importance of understanding all facets of tumor biology to meaningfully identify the underlying factors responsible for different responses and outcomes to therapies. While notable ITH at the genomic and tumor cellular composition levels was not observed, variations in spatial tumor architecture may be a contributing factor to heterogeneous responses to therapy of an individual patient. Indeed, chemokine-based signatures, the building blocks of these diverse neighborhoods, architectures, and tumor ecosystems, and specifically, the expression of T cell trafficking-associated cytokines controlled by methylation, directly correlated with ccRCC patient response to immunotherapy across large cohorts, most likely promoting antitumor activity by stimulating immune infiltration to create immune-enriched tumor ecosystems. These results demonstrate that MxIF analysis of an entire tissue section is a powerful method to study the morphology of a resected ccRCC tumor and can further be applied to other cancer types to better understand ITH as a predictor of variability in clinical outcomes. In cancer types with high ITH (e.g., prostate cancer) or low tumor purity (e.g., pancreatic cancer), tumor resection and analysis of the entire tissue section are necessary for efficient spatial analysis. In other solid cancer types, such as breast cancer, lung adenocarcinoma, gastric cancer, colon cancer, and ovarian cancer, two tissue samples with core sizes of 2 mm are sufficient to perform cell-to-cell interaction analysis and extrapolate the MxIF analysis to the entire tissue (Jackson et al., 2020).

In future studies, coupling this MxIF analysis with CyTOF will enable investigation of not only the spatial architecture, and thus the location of immune cells, but also the functions of the immune cells. Further combination of MxIF and CyTOF methods with proteogenomic and transcriptomic analyses, as demonstrated in this study, allows for an even more comprehensive understanding of a patient's tumor and can provide additional insights into the highly varied treatment outcomes associated with a specific cancer. Dissection of the ITH of the complex tumor ecosystem resulting from this integrated approach can be a powerful biomarker of response to therapies, ultimately helping to improve precision medicine strategies for patients with ccRCC and other cancer types.

Limitations of the study

The main limitation of this work is the small number of patients included in the study. The results from the limited number of samples that show spatial ITH in ccRCC require further validation in meta-cohorts to demonstrate the contribution of the spatial architecture to the clinical heterogeneity seen in clinical outcomes. While the results showing the correlation of cytokines in the different TME neighborhoods with ccRCC clinical outcome and survival were validated in three large cohorts (Figure 7), future prospective studies will be essential to demonstrate the usefulness of this technology in clinical decision-making. Another limitation is the use of only primary ccRCC samples in this study. Evaluating metastatic samples will help us understand if a difference in tumor heterogeneity is observed between primary and metastatic samples, and if the tumor biology of the metastatic samples plays a role in clinical response and survival outcomes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - CyTOF sample preparation
  - CyTOF data analysis
  - MxIF imaging
  - MxIF imaging data analysis
  - Sequencing
  - Bioinformatics
  - RNA-seq analysis
  - snRNA-seq analysis
  - WGBS analysis
  - Building phylogenetic trees
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2022.111180.

ACKNOWLEDGMENTS

The results published here are in whole or in part based upon data generated by The Cancer Genome Atlas (TCGA) managed by the NCI and NHGRI. Information about TCGA can be found at http://cancergenome.nih.gov/.

AUTHOR CONTRIBUTIONS

J.H., A.B., R.A., and N.F. conceived and jointly supervised the study. A.R. and Y.L. conceived of the experiments. N.M., E.P., V.S., Y.L., A.R., V.Z., V.M., I.G., Y.L., S.I., and P.O. conducted the overall data analyses and sequencing data processing. V.Z. performed CyTOF data analysis. V.S., I.G., E.P., and P.O. conducted MxIF data processing and analysis. Y.L. and S.I. performed snRNA-seq data processing. R.A., A.B., N.M., and E.P. created and validated the cytokine gene expression signatures. E.P. and D.S. performed a histological assessment of the tissue slides. N.M., E.P., A.B., K.N., J.B., and I.G. contributed to figure and table generation for the manuscript. J.H., E.P., N.M., K.N., J.B., A.B., and M.T. wrote and revised the manuscript and prepared the figures. All authors participated in the final review of the text.

DECLARATION OF INTERESTS

J.H. is a medical oncologist in the Division of Oncology at Washington University School of Medicine in St. Louis and has received consulting fees and research funding from BostonGene. N.F. is the chief medical officer of BostonGene Corp. and a professor at The University of Texas MD Anderson Cancer Center. This research was funded by BostonGene Corp., and all BostonGene authors were employees thereof at the time the study was performed.

Received: December 22, 2021
Revised: June 23, 2022
Accepted: July 19, 2022
Published: August 16, 2022

REFERENCES

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461. https://doi.org/10.1038/nature12787.

Bagaev, A., Kotlov, N., Nomie, K., Svekolkin, V., Gafurov, A., Isaeva, O., Osokin, N., Kozlov, I., Frenkel, F., Gancharova, O., et al. (2021). Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. Cancer Cell *39*, 845–865.e7. https://doi.org/10.1016/j.ccell.2021.04.014.

Banchereau, R., Leng, N., Zill, O., Sokol, E., Liu, G., Pavlick, D., Maund, S., Liu, L.-F., Kadel, E., Baldwin, N., et al. (2021). Molecular determinants of response to PD-L1 blockade across tumor types. Nat. Commun. *12*, 3969. https://doi.org/10.1038/s41467-021-24112-w.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. *37*, 38–44. https://doi.org/10.1038/nbt.4314.

Bergmann, E.A., Chen, B.-J., Arora, K., Vacic, V., and Zody, M.C. (2016). Conpair: concordance and contamination estimator for matched tumor–normal pairs. Bioinformatics *32*, 3196–3198. https://doi.org/10.1093/bioinformatics/btw389.

Bi, K., He, M.X., Bakouny, Z., Kanodia, A., Napolitano, S., Wu, J., Grimaldi, G., Braun, D.A., Cuoco, M.S., Mayorga, A., et al. (2021). Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. Cancer Cell *39*, 649–661.e5. https://doi.org/10.1016/j.ccell.2021.02.015.

Bolotin, D.A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I.Z., Putintseva, E.V., and Chudakov, D.M. (2015). MiXCR: software for comprehensive adaptive immunity profiling. Nat. Methods *12*, 380–381. https://doi.org/10.1038/nmeth.3364.

Braun, D.A., Street, K., Burke, K.P., Cookmeyer, D.L., Denize, T., Pedersen, C.B., Gohil, S.H., Schindler, N., Pomerance, L., Hirsch, L., et al. (2021). Progressive immune dysfunction with advancing disease stage in renal cell carcinoma. Cancer Cell *39*, 632–648.e8. https://doi.org/10.1016/j.ccell.2021.02.013.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527. https://doi.org/10.1038/nbt.3519.

Brugarolas, J., Rajaram, S., Christie, A., and Kapur, P. (2020). The evolution of angiogenic and inflamed tumors: the renal cancer paradigm. Cancer Cell *38*, 771–773. https://doi.org/10.1016/j.ccell.2020.10.021.

Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature *499*, 43–49. https://doi.org/10.1038/nature12222.

Carril-Ajuria, L., Santos, M., Roldán-Romero, J.M., Rodriguez-Antona, C., and de Velasco, G. (2019). Prognostic and predictive value of PBRM1 in clear cell renal cell carcinoma. Cancers *12*, 16. https://doi.org/10.3390/cancers12010016.

Cavalcante, R.G., and Sartor, M.A. (2017). annotatr: genomic regions in context. Bioinformatics *33*, 2381–2383. https://doi.org/10.1093/bioinformatics/btx183.

Chakravarthy, A., Furness, A., Joshi, K., Ghorani, E., Ford, K., Ward, M.J., King, E.V., Lechner, M., Marafioti, T., Quezada, S.A., et al. (2018). Pan-cancer deconvolution of tumour composition using DNA methylation. Nat. Commun. *9*, 3220. https://doi.org/10.1038/s41467-018-05570-1.

Chevrier, S., Levine, J.H., Zanotelli, V.R.T., Silina, K., Schulz, D., Bacac, M., Ries, C.H., Ailles, L., Jewett, M.A.S., Moch, H., et al. (2017). An immune atlas of clear cell renal cell carcinoma. Cell *169*, 736–749.e18. https://doi.org/10.1016/j.cell.2017.04.016.

Dangaj, D., Bruand, M., Grimm, A.J., Ronet, C., Barras, D., Duttagupta, P.A., Lanitis, E., Duraiswamy, J., Tanyi, J.L., Benencia, F., et al. (2019). Cooperation between constitutive and inducible chemokines enables T cell engraftment and immune attack in solid tumors. Cancer Cell *35*, 885–900.e10. https://doi.org/10.1016/j.ccell.2019.05.004.

Davidson-Pilon, C. (2019). lifelines: survival analysis in Python. J. Open Source Softw. *4*, 1317. https://doi.org/10.21105/joss.01317.

De la Fuente López, M., Landskron, G., Parada, D., Dubois-Camacho, K., Simian, D., Martinez, M., Romero, D., Roa, J.C., Chahuán, I., Gutiérrez, R., et al. (2018). The relationship between chemokines CCL2, CCL3, and CCL4 with the tumor microenvironment and tumor-associated macrophage markers in colorectal cancer. Tumour Biol. *40*. https://doi.org/10.1177/1010428318810059.

der Auwera, G.A.V., and O'Connor, B.D. (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra (O'Reilly Media, Inc.).

Díaz-Montero, C.M., Rini, B.I., and Finke, J.H. (2020). The immunology of renal cell carcinoma. Nat. Rev. Nephrol. *16*, 721–735. https://doi.org/10.1038/s41581-020-0316-3.

Dieu-Nosjean, M.-C., Giraldo, N.A., Kaplon, H., Germain, C., Fridman, W.H., and Sautès-Fridman, C. (2016). Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. Immunol. Rev. *271*, 260–275. https://doi.org/10.1111/imr.12405.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

Favero, F., Joshi, T., Marquard, A.M., Birkbak, N.J., Krzystanek, M., Li, Q., Szallasi, Z., and Eklund, A.C. (2015). Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann. Oncol. *26*, 64–70. https://doi.org/10.1093/annonc/mdu479.

Gao, X., and McDermott, D.F. (2018). Ipilimumab in combination with nivolumab for the treatment of renal cell carcinoma. Expert Opin. Biol. Ther. *18*, 947–957. https://doi.org/10.1080/14712598.2018.1513485.

Gassen, S.V., Callebaut, B., Van Helden, M.J., Lambrecht, B.N., Demeester, P., Dhaene, T., and Saeys, Y. (2015). FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. Cytometry A. *87*, 636–645. https://doi.org/10.1002/cyto.a.22625.

Gelbrich, N., Ahrend, H., Kaul, A., Brandenburg, L.-O., Zimmermann, U., Mustea, A., Burchardt, M., Gümbel, D., and Stope, M.B. (2017). Different cytokine and chemokine expression patterns in malignant compared to those in nonmalignant renal cells. Anal. Cell. Pathol. *2017*, 7190546. https://doi.org/10.1155/2017/7190546.

Goldman, M.J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A.N., et al. (2020). Visualizing and interpreting cancer genomics data via the Xena platform. Nat. Biotechnol. *38*, 675–678. https://doi.org/10.1038/s41587-020-0546-8.

Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., Black, S., and Nolan, G.P. (2018). Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. Cell *174*, 968–981.e15. https://doi.org/10.1016/j.cell.2018.07.010.

Helmink, B.A., Reddy, S.M., Gao, J., Zhang, S., Basar, R., Thakur, R., Yizhak, K., Sade-Feldman, M., Blando, J., Han, G., et al. (2020). B cells and tertiary lymphoid structures promote immunotherapy response. Nature *577*, 549–555. https://doi.org/10.1038/s41586-019-1922-8.

Hsieh, J.J., and Cheng, E.H. (2020). Exploiting the circuit breaker cancer evolution model in human clear cell renal cell carcinoma. Cell Stress *4*, 191–198. https://doi.org/10.15698/cst2020.08.227.

Hsieh, J.J., Purdue, M.P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., Heng, D.Y., Larkin, J., and Ficarra, V. (2017). Renal cell carcinoma. Nat. Rev. Dis. Primers *3*, 17009. https://doi.org/10.1038/nrdp.2017.9.

Jackson, H.W., Fischer, J.R., Zanotelli, V.R.T., Ali, H.R., Mechera, R., Soysal, S.D., Moch, H., Muenst, S., Varga, Z., Weber, W.P., and Bodenmiller, B. (2020). The single-cell pathology landscape of breast cancer. Nature *578*, 615–620. https://doi.org/10.1038/s41586-019-1876-x.

Kennedy, D.W., White, N.M., Benton, M.C., Fox, A., Scott, R.J., Griffiths, L.R., Mengersen, K., and Lea, R.A. (2018). Critical evaluation of linear regression models for cell-subtype specific methylation signal from mixed blood cell DNA. PLoS One *13*, e0208915. https://doi.org/10.1371/journal.pone.0208915.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.-R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

Krishna, C., DiNatale, R.G., Kuo, F., Srivastava, R.M., Vuong, L., Chowell, D., Gupta, S., Vanderbilt, C., Purohit, T.A., Liu, M., et al. (2021). Single-cell sequencing links multiregional immune landscapes and tissue-resident T cells in ccRCC to tumor topology and therapy efficacy. Cancer Cell *39*, 662–677.e6. https://doi.org/10.1016/j.ccell.2021.03.007.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572. https://doi.org/10.1093/bioinformatics/btr167.

Lake, B.B., Chen, S., Hoshi, M., Plongthongkum, N., Salamon, D., Knoten, A., Vijayan, A., Venkatesh, R., Kim, E.H., Gao, D., et al. (2019). A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. Nat. Commun. *10*, 2832. https://doi.org/10.1038/s41467-019-10861-2.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Li, H., Chiappinelli, K.B., Guzzetta, A.A., Easwaran, H., Yen, R.-W.C., Vatapalli, R., Topper, M.J., Luo, J., Connolly, R.M., Azad, N.S., et al. (2014). Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. Oncotarget *5*, 587–598. https://doi.org/10.18632/oncotarget.1782.

Li, J., Byrne, K.T., Yan, F., Yamazoe, T., Chen, Z., Baslan, T., Richman, L.P., Lin, J.H., Sun, Y.H., Rech, A.J., et al. (2018). Tumor cell-intrinsic factors underlie heterogeneity of immune cell infiltration and response to immunotherapy. Immunity *49*, 178–193.e7. https://doi.org/10.1016/j.immuni.2018.06.006.

Liu, X.-D., Kong, W., Peterson, C.B., McGrail, D.J., Hoang, A., Zhang, X., Lam, T., Pilie, P.G., Zhu, H., Beckermann, K.E., et al. (2020). PBRM1 loss defines a nonimmunogenic tumor phenotype associated with checkpoint inhibitor resistance in renal carcinoma. Nat. Commun. *11*, 2135. https://doi.org/10.1038/s41467-020-15959-6.

Lun, A.T.L., Riesenfeld, S., Andrews, T., Dao, T.P., Gomes, T., and participants in the 1st Human Cell Atlas Jamboree; and Marioni, J.C. (2019). EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biol. *20*, 63. https://doi.org/10.1186/s13059-019-1662-y.

McDermott, D.F., Huseni, M.A., Atkins, M.B., Motzer, R.J., Rini, B.I., Escudier, B., Fong, L., Joseph, R.W., Pal, S.K., Reeves, J.A., et al. (2018). Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. Nat. Med. *24*, 749–757. https://doi.org/10.1038/s41591-018-0053-3.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The ensembl variant effect predictor. Genome Biol. *17*, 122. https://doi.org/10.1186/s13059-016-0974-4.

Melsted, P., Booeshaghi, A.S., Gao, F., Beltrame, E., Lu, L., Hjorleifsson, K.E., Gehring, J., and Pachter, L. (2019). Modular and efficient pre-processing of single-cell RNA-seq. Preprint at bioRxiv. https://doi.org/10.1101/673285.

Motzer, R.J., Robbins, P.B., Powles, T., Albiges, L., Haanen, J.B., Larkin, J., Mu, X.J., Ching, K.A., Uemura, M., Pal, S.K., et al. (2020a). Avelumab plus axitinib versus sunitinib in advanced renal cell carcinoma: biomarker analysis of the phase 3 JAVELIN Renal 101 trial. Nat. Med. *26*, 1733–1741. https://doi.org/10.1038/s41591-020-1044-8.

Motzer, R.J., Banchereau, R., Hamidi, H., Powles, T., McDermott, D., Atkins, M.B., Escudier, B., Liu, L.-F., Leng, N., Abbas, A.R., et al. (2020b). Molecular subsets in renal cancer determine outcome to checkpoint and angiogenesis blockade. Cancer Cell *38*, 803–817.e4. https://doi.org/10.1016/j.ccell.2020.10.011.

Pachynski, R.K., Kim, E.H., Miheecheva, N., Kotlov, N., Ramachandran, A., Postovalova, E., Galkin, I., Svekolkin, V., Lyu, Y., Zou, Q., et al. (2021). Single-cell spatial proteomic revelations on the multiparametric MRI heterogeneity of clinically significant prostate cancer. Clin. Cancer Res. *27*, 3478–3490. https://doi.org/10.1158/1078-0432.CCR-20-4217.

Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D.E., and Gfeller, D. (2017). Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. Elife *6*, e26476. https://doi.org/10.7554/eLife.26476.

Reinius, L.E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S.E., Greco, D., Söderhäll, C., Scheynius, A., and Kere, J. (2012). Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS One *7*, e41361. https://doi.org/10.1371/journal.pone.0041361.

Ricketts, C.J., and Linehan, W.M. (2018). Multi-regional sequencing elucidates the evolution of clear cell renal cell carcinoma. Cell *173*, 540–542. https://doi.org/10.1016/j.cell.2018.03.077.

Ricketts, C.J., De Cubas, A.A., Fan, H., Smith, C.C., Lang, M., Reznik, E., Bowlby, R., Gibb, E.A., Akbani, R., Beroukhim, R., et al. (2018). The cancer genome atlas comprehensive molecular characterization of renal cell carcinoma. Cell Rep. *23*, 313–326.e5. https://doi.org/10.1016/j.celrep.2018.03.075.

Ruf, M., Moch, H., and Schraml, P. (2015). Interaction of tumor cells with infiltrating lymphocytes via CD70 and CD27 in clear cell renal cell carcinoma. Oncoimmunology *4*, e1049805. https://doi.org/10.1080/2162402X.2015.1049805.

Saunders, C.T., Wong, W.S.W., Swamy, S., Becq, J., Murray, L.J., and Cheetham, R.K. (2012). Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics *28*, 1811–1817. https://doi.org/10.1093/bioinformatics/bts271.

Schürch, C.M., Bhate, S.S., Barlow, G.L., Phillips, D.J., Noti, L., Zlobec, I., Chu, P., Black, S., Demeter, J., McIlwain, D.R., et al. (2020). Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive. Cell *182*, 1341–1359.e19. https://doi.org/10.1016/j.cell.2020.07.005.

Shen, R., and Seshan, V.E. (2016). FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. Nucleic Acids Res. *44*, e131. https://doi.org/10.1093/nar/gkw520.

Shrestha, R., Nabavi, N., Lin, Y.-Y., Mo, F., Anderson, S., Volik, S., Adomat, H.H., Lin, D., Xue, H., Dong, X., et al. (2019). BAP1 haploinsufficiency predicts a distinct immunogenic class of malignant peritoneal mesothelioma. Genome Med. *11*, 8. https://doi.org/10.1186/s13073-019-0620-3.

Simonaggio, A., Epaillard, N., Pobel, C., Moreira, M., Oudard, S., and Vano, Y.-A. (2021). Tumor microenvironment features as predictive biomarkers of response to immune checkpoint inhibitors (ICI) in metastatic clear cell renal cell carcinoma (mccRCC). Cancers *13*, E231. https://doi.org/10.3390/cancers13020231.

Souri, Z., Wierenga, A.P.A., van Weeghel, C., van der Velden, P.A., Kroes, W.G.M., Luyten, G.P.M., van der Burg, S.H., Jochemsen, A.G., and Jager, M.J. (2019). Loss of BAP1 is associated with upregulation of the NFkB pathway and increased HLA class I expression in uveal melanoma. Cancers *11*, 1102. https://doi.org/10.3390/cancers11081102.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics *30*, 3310–3316. https://doi.org/10.1093/bioinformatics/btu548.

Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. *9*, 5233. https://doi.org/10.1038/s41598-019-41695-z.

Turajlic, S., Xu, H., Litchfield, K., Rowan, A., Horswell, S., Chambers, T., O'Brien, T., Lopez, J.I., Watkins, T.B.K., Nicol, D., et al. (2018). Deterministic evolutionary trajectories influence primary tumor growth: TRACERx renal. Cell *173*, 595–610.e11. https://doi.org/10.1016/j.cell.2018.03.043.

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28*, 2184–2185. https://doi.org/10.1093/bioinformatics/bts356.

Wang, T., Lu, R., Kapur, P., Jaiswal, B.S., Hannan, R., Zhang, Z., Pedrosa, I., Luke, J.J., Zhang, H., Goldstein, L.D., et al. (2018). An empirical approach leveraging tumorgrafts to dissect the tumor microenvironment in renal cell carcinoma identifies missing link to prognostic inflammatory factors. Cancer Discov. *8*, 1142–1155. https://doi.org/10.1158/2159-8290.CD-17-1246.

Wingett, S.W., and Andrews, S. (2018). FastQ Screen: a tool for multi-genome mapping and quality control. F1000Res. *7*, 1338. https://doi.org/10.12688/f1000research.15931.2.

Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. *19*, 15. https://doi.org/10.1186/s13059-017-1382-0.

Wu, H., Kirita, Y., Donnelly, E.L., and Humphreys, B.D. (2019). Advantages of single-nucleus over single-cell RNA sequencing of adult kidney: rare cell types and novel cell states revealed in fibrosis. J. Am. Soc. Nephrol. *30*, 23–32. https://doi.org/10.1681/ASN.2018090912.

Yoo, K.H., Lee, D.-G., Won, K.Y., Lim, S.-J., Park, Y.-K., and Chang, S.-G. (2015). Expression of CC chemokine receptor 5 in clear cell renal cell carcinoma and its clinical significance. Oncol. Lett. *9*, 2085–2089. https://doi.org/10.3892/ol.2015.3048.

Zaitsev, A., Chelushkin, M., Dyikanov, D., Cheremushkin, I., Shpak, B., Nomie, K., Zyrin, V., Nuzhdina, K., Lozinsky, Y., Zotova, A., et al. (2022). Precise reconstruction of the tumor microenvironment using bulk RNA-seq and a unique machine learning algorithm trained on artificial transcriptomes. Cancer Cell *40*, 879–894.

Zheng, S.C., Breeze, C.E., Beck, S., Dong, D., Zhu, T., Ma, L., Ye, W., Zhang, G., and Teschendorff, A.E. (2019). EpiDISH web server: epigenetic dissection of intra-sample-heterogeneity with online GUI. Bioinformatics, btz833. https://doi.org/10.1093/bioinformatics/btz833.

Zhou, Q., Qi, Y., Wang, Z., Zeng, H., Zhang, H., Liu, Z., Huang, Q., Xiong, Y., Wang, J., Chang, Y., et al. (2020). CCR5 blockade inflames antitumor immunity in BAP1-mutant clear cell renal cell carcinoma. J. Immunother. Cancer *8*, e000228. https://doi.org/10.1136/jitc-2019-000228.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-human CD45 (clone HI30)	Fluidigm	Cat# 3089003; RRID: AB_2661851
Purified anti-human CD3 (clone UCHT1)	BioLegend	Cat# 300402; RRID: AB_314056
Anti-human CD57 (clone HCD5)	Fluidigm	Cat# 3142007B; RRID: AB_2868399
Anti-human CD69 (clone FN50)	Fluidigm	Cat# 3144018; RRID: AB_2687849
Anti-human CD4 (clone RPA-T4)	Fluidigm	Cat# 3145001; RRID: AB_2661789
Purified anti-human CD8 (clone SK1)	BioLegend	Cat# 344727; RRID: AB_2563762
Anti-human CD11C (clone Bu15)	Fluidigm	Cat# 3147008; RRID: AB_2687850
Anti-human CD16 (clone 3G8)	Fluidigm	Cat# 3148004; RRID: AB_2661791
CD25 IL-2R (clone 02A3)	Fluidigm	Cat# 3149010B; RRID: AB_2756416
Anti-human CD107a/LAMP1 (clone H4A3)	Fluidigm	Cat# 3151021D; RRID: AB_2858230
Anti-human CD66b (clone 80H3)	Fluidigm	Cat# 3152011; RRID: AB_2661795
Anti-human CD45RA (clone HI100)	Fluidigm	Cat# 3153001B; RRID: AB_2802108
Anti-human CD163 (clone GHI/61)	Fluidigm	Cat# 3154007B; RRID: AB_2661797
Mouse anti-human CAIX (clone 303123)	R&D	Cat# MAB2188; RRID: AB_2066530
Anti-human CD86 (clone IT2.2)	Fluidigm	Cat# 3156008; RRID: AB_2661798
Anti-human CD27 (clone L128)	Fluidigm	Cat# 3158010B; RRID: AB_2858231
Anti-human CD197 (CCR7) (clone G043H7)	Fluidigm	Cat# 3159003; RRID: AB_2714155
Anti-human CD14 (clone M5E2)	Fluidigm	Cat# 3160001B; RRID: AB_2687634
Anti-human FoxP3 (clone 259D/C7)	Fluidigm	Cat# 3162024A
Anti-human CD127 (IL-7Ra) (clone A019D5)	Fluidigm	Cat# 3165008B; RRID: AB_2868401
Anti-human CD141 (Thrombomodulin) (clone M80)	Fluidigm	Cat# 3166017B; RRID: AB_2892693
Anti-human CD38 (clone HIT2)	Fluidigm	Cat# 3167001B; RRID: AB_2802110
Anti-human CD19 (clone HIB19)	Fluidigm	Cat# 3169011B; RRID: AB_2893034
Anti-human HLA-DR (clone L243)	Fluidigm	Cat# 3170013B; RRID: AB_2888929
Anti-human CD68 (clone Y1/82A)	Fluidigm	Cat# 3171011B; RRID: AB_2687637
Anti-human/mouse granzyme B (clone GB11)	Fluidigm	Cat# 3173006B; RRID: AB_2811095
Anti-human CD279 (PD-1) (clone EH12.2H7)	Fluidigm	Cat# 3174020B; RRID: AB_2868402
Anti-human CD274 (PD-L1) (clone 29E.2A3)	Fluidigm	Cat# 3175017B; RRID: AB_2687638
Anti-human CD56 (NCAM) (clone NCAM16.2)	Fluidigm	Cat# 3176008; RRID: AB_2661813
Anti-human CD11b (Mac-1) (clone ICRF44)	Fluidigm	Cat# 3209003; RRID: AB_2687654
Anti-CD56 (clone 56C04, same as 123A8)	ThermoFisher Scientific	Cat# MS-1149-P1ABX; RRID: AB_64047
Anti-CD4 (clone EPR6855)	Abcam	Cat# ab181724; RRID: AB_2864377
Anti-NaKATPase (clone EP1845Y)	Abcam	Cat# ab167390; RRID: AB_2890241
PCK26 (clone PCK-26)	Sigma	Cat# C5992; RRID: AB_2134432
S6 (clone 5G10)	Cell Signaling Technology	Cat# 2217; RRID: AB_331355
CD45 (clone 2B11 + PD7/26)	Agilent	Cat# M070101-2; RRID: AB_2750582
CD20 (clone EP459Y)	Abcam	Cat# ab166865
CD206 (clone 5C11)	Sigma-Aldrich	Cat# WH0004360M2; RRID: AB_1840543
CD31 (clone 89C2)	Cell Signaling Technology	Cat#3528; RRID: AB_2160882

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PDL1 (clone E1L3N)	Cell Signaling Technology	Cat# 13684; RRID: AB_2687655
HLA DR (clone LN3)	Lab Vision	Cat# MS-133-PABX; RRID: AB_63218
CAIX (Polyclonal)	Thermo Fisher Scientific	Cat# PA1-16592; RRID: AB_568476
Granzyme B (clone 2C5)	Santa Cruz Biotechnology	Cat # sc-8022; RRID: AB_2232723
CD11c (clone D3V1E)	Cell Signaling Technology	Cat# 45581; RRID: AB_2799286
CD8 (clone C8/144B)	Agilent	Cat# M7103; RRID: AB_2075537
CD68 (clone KP1)	Lab Vision	Cat# MS397PABX; RRID: AB_720551
CD3e (clone SP7)	Thermo Fisher Scientific	Cat# MA5-14524; RRID: AB_10982026
Anti-human CD16 (clone DJ130c)	Santa Cruz Biotechnology	Cat# sc-20052; RRID: AB_626925
Anti- Ki67 (clone SP6)	abcam	Cat# ab16667; RRID: AB_302459
Chemicals, peptides, and recombinant proteins		
ACK Lysis Buffer	Lonza, Basel, Switzerland	Lonza™ 10548E
Fc-receptor blocking solution	Invitrogen, Waltham, MA, USA	AB_468582
Foxp3/Transcription Factor Staining Buffer Set	eBioscience, San Diego, CA, USA	00-5523-00
Ir-intercalator	Fluidigm, San Francisco, CA, USA	201192A
IDT KAPA Hyper libraries with xGen Exome Research Panel v1.0	IDT, Skokie, IL	Rxn 10005151
TruSeq Stranded Total RNA	Illumina	20020596
Critical commercial assays		
Tumor Dissociation Kit	Miltenyi Biotec, Bergisch Gladbach, Germany	130-095-929
ACCEL-NGS® METHYL-SEQ DNA LIBRARY KIT	Swift biosciences	30024
Deposited data		
Microarray reference data	Reinius et al., 2012 Kennedy et al., 2018	GEO: GSE35069
IMmotion150	McDermott et al., 2018	EGAS00001002928
JAVELIN 101	Motzer et al., 2020a	Supplemental of PMID
TCGA-KIRC	N/A	phs000178
CyTOF data	This paper	https://github.com/BostonGene/Multiregional_ccRCC/tree/main/data/cytof
RNAseq data	This paper	https://github.com/BostonGene/Multiregional_ccRCC/tree/main/data/snRNA
snRNAseq data	This paper	https://github.com/BostonGene/Multiregional_ccRCC/tree/main/data/snRNA
WES data	This paper	https://github.com/BostonGene/Multiregional_ccRCC/tree/main/data/mutations
Methylation data	This paper	https://github.com/BostonGene/Multiregional_ccRCC/tree/main/data/methylation
Imaging data (MxIF)	This paper	https://github.com/BostonGene/Multiregional_ccRCC/tree/main/data/mxif
human reference genome GRCh38	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/grc/human
Software and algorithms		
FlowJo (FlowJo X 10.0.7r)	N/A	https://docs.flowjo.com/flowjo/getting-acquainted/10-1-release-notes/10-0-7-release-notes/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FlowSOM version 1.20.0	Gassen et al., 2015	https://bioconductor.org/packages/release/bioc/html/FlowSOM.html
ImageJ (ImageJ bundled with Java 1.8.0_172)	N/A	https://imagej.nih.gov/ij/
Pytorch Python library with UnetResNet-101	N/A	https://github.com/sshuair/torchsat/tree/develop
FastQC v0.11.5	N/A	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
FastQ Screen v0.11.1	Wingett and Andrews, 2018	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
RSeQC v3.0.0	Wang et al., 2012	http://rseqc.sourceforge.net/
MultiQC v1.6	Ewels et al., 2016	https://github.com/ewels/MultiQC
OptiType	Szolek et al., 2014	https://github.com/FRED-2/OptiType
Conpair algorithm	Bergmann et al., 2016	http://github.com/nygenome/conpair
FilterByTile/BBMap v37.90	N/A	https://jgi.doe.gov/data-and-tools/software-tools/bbtools/bb-tools-user-guide/bbmap-guide/
BWA v0.7.17	Li and Durbin, 2009	https://github.com/lh3/bwa/releases/tag/v0.7.17
Picard's v2.6.0 MarkDuplicates	"Picard Toolkit", 2019. Broad Institute, GitHub Repository.	http://broadinstitute.github.io/picard/
GATK v3.8.1	der Auwera and O'Connor, 2020	https://gatk.broadinstitute.org/hc/en-us
Strelka v2.9	Saunders et al., 2012	https://github.com/Illumina/strelka
Variant Effect Predictor v92.1	McLaren et al., 2016	https://uswest.ensembl.org/info/docs/tools/vep/index.html
FACETS 0.5.14	Shen and Seshan, 2016	https://github.com/mskcc/facets
Sequenza v2.1.2	Favero et al., 2015	https://github.com/cran/sequenza
Kallisto v0.42.4	Bray et al., 2016	https://pachterlab.github.io/kallisto
MIXCR v.2.1.7	Bolotin et al., 2015	https://mixcr.readthedocs.io/en/master/
ssGSEA	Subramanian et al., 2005	http://software.broadinstitute.org/gsea
DropletUtils package	Lun et al., 2019	https://bioconductor.org/packages/release/bioc/html/DropletUtils.html
scanpy python package	Wolf et al., 2018	https://scanpy.readthedocs.io/en/stable/
Harmony	Korsunsky et al., 2019	https://github.com/immunogenomics/harmony
UMAP package	Becht et al., 2018	https://github.com/lmcinnes/umap
cellSNP tool	N/A	https://github.com/single-cell-genetics/cellSNP/tree/master/test
bismark v0.22.1	Krueger and Andrews, 2011	https://github.com/FelixKrueger/Bismark/blob/master/CHANGELOG.md
annotatr R package	Cavalcante and Sartor, 2017	https://bioconductor.org/packages/release/bioc/html/annotatr.html
MethylCIBERSORT R package	Chakravarthy et al., 2018	https://zenodo.org/record/1284582
EpiDISH R package	Zheng et al., 2019	https://github.com/sjczheng/EpiDISH
python (version 3.7.5)	Python	https://www.python.org/downloads/release/python-375/
R (version 4.0.2)	N/A	https://cran.r-project.org/
CamDavidsonPilon/lifelines: v0.14.6 (Version v0.14.6)	Davidson-Pilon, 2019	https://github.com/CamDavidsonPilon/lifelines/tree/v0.14.6
K-means algorithm	N/A	https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Phenograph	N/A	https://github.com/jacoblevine/PhenoGraph
Other		
Kassandra	Zaitsev et al., 2022	https://science.bostongene.com/kassandra/

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the Lead Contact, Nathan Fowler (nathan.fowler@bostongene.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Data: Detailed code, along with associated datasets and documentation are available at https://github.com/BostonGene/Multiregional_ccRCC.

Code: All original code has been deposited at https://github.com/BostonGene/Multiregional_ccRCC and is publicly available as of the date of publication.

General statement: Any additional information required to reanalyze the data reported in this paper is available from the lead contact, Nathan Fowler, upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

This study was approved by the Washington University Institutional Review Board. Informed consent was obtained from all patients who provided samples for this work. Primary ccRCC tumors were collected from 6 patients. The detailed characteristics, including gender and age, of the patients and samples are provided in Figure S1C. There were 3 female and 3 male patients, with ages ranging from 47 to 87 years old. One-half of the patients had metastatic diseases and were treated with ipilimumab and nivolumab, while the other half had localized diseases and only received a nephrectomy. Overall, the ccRCC tumors were primarily grades 3 (33%) or 4 (50%) and tumor stages T3a (50%) or T3b (33%) except one T2a. Among the three patients treated with ipilimumab and nivolumab, one patient achieved a complete response, one patient was a partial responder, and one patient was a non-responder with disease progression.

METHOD DETAILS

CyTOF sample preparation

Single cell dissociation

Overall, fresh tumor tissue (100 mg) was minced into 3 mm pieces and mixed with enzymes obtained from human tumor cells using the Tumor Dissociation Kit (Miltenyi Biotec, Bergisch Gladbach, Germany). The automated cell dissociator (gentleMACS^{TM} Dissociator, Miltenyi Biotec) was used to gently agitate the cells for disassociation for 1 h at 37 °C. Next, the red blood cells were removed using the ACK Lysis Buffer (Lonza, Basel, Switzerland), and the remaining cell suspension was maintained in RPMI for CyTOF antibody staining.

CyTOF antibody staining

After dissociation, 1 × 10^5–1 x 10^6 tumor cells were incubated with Fc-receptor blocking solution (Invitrogen, Waltham, MA, USA) for 10 min at room temperature. Next, the cells were incubated with surface marker antibodies at the optimized concentrations (Table S2) for 1 h on ice. The cells were then stained with 2.5 uM Cisplatin solution for 1 min at room temperature to identify the dead cells. The cells were permeabilized using Foxp3/Transcription Factor Staining Buffer Set (eBioscience, San Diego, CA, USA) for 30 min on ice. The permeabilized cells were then incubated with intracellular marker antibodies at the optimized concentrations (Table S2) for 1 h on ice. The cells were stored in 4% PFA solution in 4 °C for up to 2 weeks until the CyTOF analysis. The cells were stained with 125 uM Ir-intercalator (Fluidigm, San Francisco, CA, USA) at 1:3,000 dilution overnight before the CyTOF analysis. Before acquisition, the samples were washed in MilliQ water and resuspended in 500 mL 1X solution of EQ Beads (Fluidigm). The samples were analyzed in the CyTOF2/Helios (Fluidigm) mass cytometer by the The Immunomonitoring Laboratory (IML) Core in the Center for Human Immunology and Immunotherapy Programs (CHiiPs) at Washington University in Saint Louis.

CyTOF data analysis

Gating and quality control

After the CyTOF experimentation, the data (117-A1/A2/A3/A4, 120-A1/A2/A3, 124-A1/A2/A3, 107-A1/A2/A3/A4, 154-A1/A2/A3/A4, 181-A1/A2/A4) were analyzed using a gating strategy. Only viable cells were selected for subsequent analysis. Gating manipulation was performed with FlowJo (FlowJo X 10.0.7r). After removing the beads, "live cell" events were selected by one-stage gating according to their viability (cisplatin) and DNA signal.

Major cell population identification

All samples underwent bootstrapped clusterization, and the clusters were automatically annotated by correlation with the reference sample (124-A1), which had been annotated manually. CyTOF signals were subjected to standard transformation (arcsine (x/5)). Clusterization was performed with FlowSOM version 1.20.0 (Gassen et al., 2015), using the following parameters: xdim = 10, ydim = 10; the number of clusters set to 30; and all other options set to default. Reference populations were acquired by clustering of 124-A1 using 30 markers, resulting in 30 identified clusters that were then annotated into 9 cell types using the following list of markers: CD8 T cells ($CD45^+$, $CD3^+$, $CD8^+$, $CD4^-$, $CD56^-$), CD4 T cells ($CD45^+$, $CD3^+$, $CD4^+$, $CD8^-$, $CD56^-$), NK cells ($CD45^+$, $CD56^+$, $CD16^+$), neutrophils ($CD45^+$, $CD66^+$, $CD16^+$), macrophages/monocytes ($CD45^+$, HLA-DR^+, $CD11c^+$, $CD66b^-$), endothelium ($CD45^-$ $CD107a^+$ $CAIX^-$), B cells ($CD45^+$ $CD19^+$ $CD38^-$), plasmatic cells ($CD45^+$ $CD19^+$ $CD38^+$) and malignant ccRCC cells ($CAIX^+$). All unidentified cells and cells with non-specific annotation were grouped as "other".

All samples were clustered with FlowSOM 30 times, each time with a 90% random cell subset, with all markers except viability and DNA1/DNA2. For all samples, automated annotation was used. For each of the bootstrapped clusters, the vector of median marker values across all cells was correlated with corresponding vectors in reference clusters. The tested cluster was annotated by a reference cluster with the highest Pearson correlation. If the highest correlation coefficient was less than 0.6 or two reference clusters had a high correlation with the tested cluster, but the difference between correlation coefficients was lower than 0.05, the tested cluster was set as "unidentified". As each cell participated in several runs of clusterization and were annotated several times, the final cell type of each event was defined as the most frequent type in all clusterization runs.

Cell subpopulation identification

To identify cell subpopulations, the following steps were conducted: 1) batch correction of signals by quantile normalization within each cell type; 2) bootstrapped clusterization of cell sets from all samples of each major population; 3) additional clustering of clusters to select the most stable groups; and 4) annotation of the clusters of cells.

Subpopulations were defined for four major populations (macrophages, CD4 T cells, CD8 T cells, and B cells). Clustering was performed on a subset of markers as follows: for CD4 and CD8 T cells - CD45RA, CCR7, IL-2R, IL-7Ra, FoxP3, CD57, GRB, CD27, CD69, PD-1 CD38; for macrophages and monocytes - CD14, CD11c, Mac-1, HLA-DR, CD16, CD163, CD68, CD4, CD38, PD-L1, CD86, LAMP1, and CD69.

For batch correction, quantile normalization within a cell type was performed. For each marker in every sample and every population, only cases with raw CyTOF signals >20 (linear scale) were considered. Expression levels were normalized by the 95th percentile of the sample/marker/population combination. After that, the levels were multiplied by the average 95th percentile for marker/population combination. Cells were pooled across patients after normalization. Normalized data were subjected to hyperbolic transformation (arcsine (x/5)) and bootstrapped FlowSOM clustering. FlowSOM version 1.20.0 was used for clusterization, with options xdim = 10, ydim = 10, the number of metacluster set to 30, and the metaclustering strategy and all other options set to default. We performed 200 iterations of clustering with 50% subsampling without replacement.

To select stable clusters across subsampled clusters, FlowSOM obtained 200 × 30 = 6,000 clusters (from 200 iterations) that were subsequently clustered by k-means, k = 20. Metaclusters were correlated to each other using a vector of median expression calculated across all cells. If the Pearson correlation was >0.95, the metaclusters were merged together. Clusters containing less than 2% of cells were merged to the most similar one based on the Pearson correlation analysis of markers. Finally, clusters were annotated with biologically meaningful labels.

MxIF imaging

Tissue preparation for histology and MxIF imaging

Fresh renal tissue specimens (kidney after resection and core needle biopsy) were handled carefully and appropriately fixed after dissection with 10% formalin for 48 h at room temperature. Tissue specimens were dehydrated by immersing specimens in a series of ethanol solutions of increasing concentration until water-free ethanol: 70% (1x), 80% (1x), 95% (1x), 100% (2x) for 1 h each. The specimens were then immersed in fresh-xylene (or xylene substitute) three times for 1.5 h each and in paraffin wax (58–60 ºC) twice for 2 h each. Next, the tissue specimens were embedded into paraffin blocks.

Histological examination

Samples were obtained and stored in 10% neutral-buffered formalin until examination. The tissues were embedded in wax, sectioned, and stained using VIRTUAL HE. Virtually stained slides were evaluated by a pathologist, and the fibrotic area was defined as an eosinophilic area of densely packed fibers. The fibrosis ratio for each specimen was calculated with ImageJ software. The tumor area and intratumoral fibrotic areas were calculated using the area selection tool. Then, the percentage of fibrosis was assessed using the ratio of fibrotic area to tumor area.

MxIF staining

Tissue sections (5 μm thick) were cut from FFPE ccRCC tissue blocks using a microtome and then mounted onto Superfrost™ Ultra Plus adhesion slides for morphological examination and multiplexed imaging analysis. Slide tissue sections were baked at 60 ºC in an oven for 1 h. Tissue sections were deparaffinized with 2 washes of fresh-xylene and were then rehydrated with washes of ethanol 100% (2x), 95% (2x), 70% (2x), 50% (2x), 1X PBS (1x) and 0.3% Triton X-100 in 1X PBS (1x) and subjected to a two-step antigen retrieval process. Next, the tissue sections underwent repeated cycles of staining, imaging and signal removal. The sections were stained with antibodies directly conjugated with either Cy3 or Cy5 dye at a previously optimized concentration (Table S3). All antibody mixes used for seven incubation rounds were incubated at room temperature for 1 h in a humid chamber. After incubation during all rounds, the slides were washed in 1X PBS for 5 min (3X). The tissue sections were then stained with DAPI solution (1 ug/mL) for 15 min. The slides were washed with 1X PBS, and the coverslip was added immediately using mounting media.

MxIF imaging data analysis

Cell segmentation

Cell segmentation, which is the estimation of a position and shape of single cell on a fluorescent image slide, was performed using the UNet semantic segmentation neural network and implemented with the Pytorch Python library with UnetResNet-101 (https://github.com/sshuair/torchsat/tree/develop). Watershed post-processing of the identified cell masks was performed to reduce the under-segmented cell counts using the OpenCV2 Python library. Training of the network was performed with manual image segmentation, which consisted of three channels: DAPI, NaKaATPase, and S6, with a combination of binary cross entropy and DICE scores as loss weighted equally, as described in Pachynski et al. (Pachynski et al., 2021). For each segmented cell, the expression of 20 markers (19 antibodies plus DAPI; Table S2, Figures S1E, S3) was measured in addition to the spatial location of each cell.

Cell typing

Cell subtype assignment was performed using Phenograph cell clusterization by standardized mean marker intensity in a given cell contour. Cell subtype assignment was guided by tSNE projection plots, allowing us to join similar cell subpopulation clusters. Cell typing was visualized by drawing cell contours and coloring them according to cell types.

Overall, 8 major cell populations were obtained using the following list of markers: B cells ($CD45^+$, $CD19^+$, $CD3^-$), CD4 T cells ($CD45^+$, $CD3^+$, $CD4^+$, $CD8^-$, $CD56^-$), CD8 T cells ($CD45^+$, $CD3^+$, $CD8^+$, $CD4^-$, $CD56^-$), endothelial cells ($CD31^+$, $CD45^-$), myeloid cells ($CD45^+$, $HLADR^+$, $CD68^+$, $CD206^+$, $CD16^+$, $CD11c^+$), malignant and epithelial cells ($CAIX^+$, $NAKATPASE^+$, PCK26). All unidentified cells and cells with non-specific annotation were grouped as ''other''.

Cell subpopulation identification

T cell subpopulations were identified by applying linear thresholds in two-dimensional plots for various marker combinations (Figure 3H). CD4 and CD8 T cells were separated based on marker intensities in plots for CD4 vs CD8 markers. Granzyme B-positive cells were identified in the CD8 vs Granzyme B marker plot. Ki67-positive cells were identified in the Ki67 vs CD8 marker plot. Thresholds were identified manually based on data distribution for each tumor region.

Community analysis

To perform community detection, mutual arrangements of cells using Delaunay triangulation for the set of their centroids to obtain neighborhood graphs were reconstructed. The upper limit for edge length was set to 200 pixels. For each cell, we identified neighbors as cells on the edges, obtaining neighborhood information in the form of a vector of cell classes and selected mask percentages for each cell. Then, using the resulting graph, the DeepGraphInfomax graph convolutional neural network (from Pytorch Geometric python package) was trained in an unsupervised manner to obtain 32-dimensional embedding of cell neighborhoods in vector form. After this transformation, neighborhood vectors were clustered using the k-means algorithm, obtaining 14 different neighborhood clusters. For each community cluster, the percentage of cells within a cluster and mask percentage was calculated to facilitate community descriptions. Clusters were grouped based on the dominant cell type proportion and morphological structures into two major groups: tumor-enriched neighborhoods and T cell-enriched neighborhoods, as well as three separated clusters of macrophage-, B cell-, and blood vessel-enriched neighborhoods. Cell communities were visualized by drawing cell contours and coloring them according to each community.

Cell-to-cell interactions

Cellular neighborhood analysis was conducted using a graph neural network-encoding approach and K-Means clustering (Goltsev et al., 2018) to identify individual neighborhoods encompassing similar types, proportions, densities, and distributions of cells. To assess the significance of contacts between different cell types, cell neighborhood graphs obtained during community analysis were reused. Then, the cell type labels in the graph were permuted multiple times to count and compare the number of contacts between permuted cell types with original distribution as described in Goltsev et al. (Goltsev et al., 2018). The proportion of permuted cases with numbers of contacts greater than in the original showed significance. Results were presented as a p value for each cell type pair for each sample. Cell-to-cell interactions were visualized on a schematic interaction graph of cell contacts.

Sequencing

WES

Each tumor and blood sample as the normal had a single enriched library constructed. Exome libraries were captured by IDT KAPA Hyper libraries with xGen Exome Research Panel v1.0. The libraries were sequenced on a NovaSeq S4 with at least 100x coverage for tumor and 30x for normal.

RNA-seq

TruSeq Stranded Total RNA (Illumina) was used for RNA library construction. The libraries were sequenced on a NovaSeq S4 with 100M reads/sample and a sequencing depth of 100 M reads (2 × 150 pairs).

WGBS

For WGBS, the ACCEL-NGS® METHYL-SEQ DNA LIBRARY KIT (Swift biosciences) was used, with a sequencing depth of 25X and 2 × 150 pairs snRNA-seq.

snRNA-seq

We conducted single-nucleus RNA-seq (snRNA-seq) on the same biopsies using the whole tissue by lysing the cell membrane and isolating nuclei, potentially leading to the recovery of each cell from the tissue (Lake et al., 2019). In brief, 100 mg frozen tumor was used for single nuclei dissociation. The dissociation method was based on a previously published protocol (Wu et al., 2019). Fresh nuclei were used for 10 × 5' Single Cell Library (5′ GEX only) construction. Libraries were sequenced on a NovaSeq S4 with 50K read pairs/cell.

Bioinformatics

NGS data quality control and analysis

Quality control of all NGS samples was performed using FastQC v0.11.5 (available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), FastQ Screen v0.11.1 (Wingett and Andrews, 2018), RSeQC v3.0.0 (Wang et al., 2012), MultiQC v1.6 (Ewels et al., 2016). HLA genotypes were calculated from RNA-seq or WES using OptiType (Szolek et al., 2014). Sample correspondence was checked using HLA comparison and the Conpair algorithm (Bergmann et al., 2016).

WES analysis: Alignment, variant calling, and mutation clonality identification

Low quality reads were filtered using FilterByTile/BBMap v37.90 and aligned to human reference genome GRCh38 (GRCh38.d1.vd1 assembly) using BWA v0.7.17 (Li and Durbin, 2009). Duplicate reads were removed using Picard's v2.6.0 MarkDuplicates (''Picard Toolkit'', 2019. Broad Institute, GitHub Repository; http://broadinstitute.github.io/picard/; Broad Institute), indels were realigned by IndelRealigner and recalibrated by BaseRecalibrator and ApplyBQSR; tools were taken from GATK v3.8.1 (der Auwera and O'Connor, 2020).

Somatic single nucleotide variations (sSNVs), small insertions and deletions were all detected using Strelka v2.9 (Saunders et al., 2012). All variants, insertions and deletions were annotated using Variant Effect Predictor v92.1 (McLaren et al., 2016).

To calculate the clonality status of each mutation, the mathematical approach that identifies clonality status from VAF and mutation copy number was employed.

Copy number alteration (CNA) analysis

FACETS v0.5.14 (Shen and Seshan, 2016) and a customized version of Sequenza v2.1.2 (Favero et al., 2015) were used as CNA callers with refinement and tuning to obtain better quality and speed of calculations. FACETS was used to create a common input file for both the CNA callers. Then, Sequenza generated a resulting file with coverage for each position in.pileup format. After the coverage extraction stage, Sequenza and FACETS were run in parallel. Both the CNA callers generated the genome/chromosome plots and output tables with segments and respective statistics (depth.ratio and BAF - for Sequenza; log-ratio and log-odds-ratio for FACETS) with collected and analyzed copy numbers.

RNA-seq analysis

RNA-seq processing

RNA-seq reads were aligned using Kallisto v0.42.4 (Bray et al., 2016) to GENCODE v23 transcripts with default parameters. The protein-coding transcripts, IGH/K/L- and TCR-related transcripts were retained, and the noncoding RNA, histone- and mitochondria-related transcripts were removed, resulting in 20,062 protein coding genes. Gene expression was quantified as transcripts per million (TPM) and log2-transformed (Goldman et al., 2020).

Deconvolution of bulk RNA-seq

The recently-described machine learning algorithm, Kassandra, was used to predict cell percentages from bulk RNA-seq (Zaitsev et al., 2022). The model consisted of a two-level hierarchical ensemble that used LightGBM as building blocks. The model was trained on artificial RNA-seq mixtures of different cell types (T cells, B cells, NK, macrophages, cancer-associated fibroblasts, and endothelium), obtained from multiple datasets of sorted cells. All datasets satisfied the following criteria: isolated from human tissue, poly-A or total RNA-Seq performed with read length higher than 31 bp, have at least 4 million coding read counts, passed quality control by FASTQC, no contamination detected (<2%). The model was trained to predict the percentage of RNA belonging to particular cell types. The predicted percentages of RNA were later converted into percentages of cells using the methodology described by Racle et al. (Racle et al., 2017).

T cell receptor/B cell receptor repertoire profiling

MIXCR v.2.1.7 (Bolotin et al., 2015) was used to analyze the RNA-seq samples. Single clonotypes were grouped into clones with a distinct VDJ combination and identical CDR3 nucleotide sequences. For B cells, the clones were further aggregated into clone groups if the VDJ combination was the same and if the CDR3 nucleotide sequences differed no more than 1 nt.

Gene signature score calculations

Correlation analysis was performed between gene expressions of single cytokines measured by RNAseq and proximity communities based on MxIF. Five high correlation groups were identified (A, B, C, D, and E). Gene signature scores were calculated using the

single-sample gene set enrichment analysis (ssGSEA) algorithm in R package GSVA (Subramanian et al., 2005). Raw scores were medium scaled to (−2, 2) or to (−3, 3) range.

snRNA-seq analysis

snRNA-seq processing

FASTQ files were processed using the kallisto bustools pipeline described in (Melsted et al., 2019). Empty droplets were filtered out manually using rank filtration implemented in DropletUtils package (Lun et al., 2019). The distribution of genes per cell for each sample was bimodal. As the left mode was supposedly made of unfiltered droplets, we filtered out cells from the left mode and kept cells from the right mode.

The scanpy python package (Wolf et al., 2018) was used to remove low-quality cells with high mitochondrial read count content (>20%). An unsupervised approach to phenotype cells was employed as follows: data were jointly visualized by Uniform Manifold Approximation and Projection (UMAP) (Becht et al., 2018), and Leiden graph-based clustering was performed as described in (Traag et al., 2019). Batch correction was performed using Harmony (Korsunsky et al., 2019). T cells were identified by *CD3E* and *CD3D* gene expression (CD4 T cells also express the *CD4* gene, CD8 T cells by *CD8A* and *CD8B*), B cells by *CD19* or *MS4A1*, macrophage by *CD163* and *CSF2RA*, malignant cells by *CA9*, endothelial cells by *PECAM1*, fibroblasts by *COL1A1*, and normal kidney cells by *SLC12A1*.

Pseudo-bulk creation from snRNA-seq

The expression in each cell was normalized to reach 10,000 UMI counts, and the expression vectors of all cells in each sample were summed. In Figure 7D, two gene signatures were used to show differences between malignant and non-malignant pseudo-bulk: 'Malignant' signature - MYC, IFI27, APOC1, IRGM; 'Immune cells' signature - PDGFRB, MYO1F, CD6, IL10RA, BRCC3, IKZF1, FGFR2.

snRNA-seq CNA analysis

CNA analysis was performed for sample 124-A2 only because this sample had the highest data quality. Gene expression from the 3p and 5q arms was averaged and scaled among cells. The resulting scores for 3p and 5q genes showed a lower 3p score in tumor cells (showing 3p deletion) and higher 5q gene score in tumor cells (showing 5q amplification).

snRNA-seq mutation calling

The coverage of mutated positions from WES was assessed using the cellSNP (https://github.com/single-cell-genetics/cellSNP/tree/master/test) tool applied to VCF files obtained from bulk DNA sequencing. Mutation information was mapped back onto the visualization tSNE plots.

WGBS analysis

WGBS preprocessing

Raw reads were aligned to reference genome hg38 using bismark v0.22.1 (Krueger and Andrews, 2011). Cytosines with read coverage less than 10 were excluded from the assay. Cytosines were annotated using the annotatr R package (Cavalcante and Sartor, 2017). CpG islands were identified based on the annotatr package, with genes region coordinates derived from biomart and human enhancer coordinates derived from the FANTOM5 database (Andersson et al., 2014).

Methylation deconvolution

Illumina 450K microarray data were used as a reference: GSE35069 (Kennedy et al., 2018; Reinius et al., 2012) for immune subtypes and methylation of TCGA samples with the highest purity for malignant cells. Differentially methylated cytosines (DMCs) were found with MethylCIBERSORT (Chakravarthy et al., 2018) R package with the following parameters: minimal difference between average methylation of groups was 20%, FDR cutoff of 0.01, and the maximum number of DMCs was 150. The cell composition ratio was obtained by the EpiDISH R package (Zheng et al., 2019) with the CIBERSORT method.

Whole genome methylation profile and tumor whole genome hypermethylation level

Firstly, CpGs were aggregated in genomic windows of 10^6 nucleotides with 50% overlap with neighboring windows. Then, the window methylation was considered as the average methylation of CpGs related to windows. The genome hypermethylation level was calculated as the mean methylation of all windows. To calculate the whole genome methylation level in TCGA data, mean methylation of all probes in the sample were calculated.

Correlation of methylation with gene expression

We calculated spearman correlation of mean DNA methylation of gene promoter CpG island with expression of gene. As most cytokines had no CpG islands in the promoter region, we calculated the correlation of cytokine expression with methylation of single CpGs in the promoter region, and then CpG with the best correlation was chosen.

Building phylogenetic trees

To build phylogenetic trees for each patient, we calculated the number of events that happened in each region of patients. We considered each event to happen only one time in patients and not to happen independently. We calculated three type of events:

1. Hypermethylation of promoter (>10% beta-value) of genes which promoter is not hypermethylated in normal TCGA samples (6689 genes)
2. Protein affecting somatic mutations (376 individual mutations)

3. Copy number alteration in chromosome bands (747 bands overall)

Events of each type were calculated separately, normalized (divided by maximum number events of this type) and then summed up. The root of the patient's tree was chosen as the common event for all samples in the patient. Branching point was chosen as the maximum common event for pairs of samples of the patient.

QUANTIFICATION AND STATISTICAL ANALYSIS

For statistical analyses and plotting, python (version 3.7.5) and R (version 4.0.2) were used. The statistical information is detailed in the text, figure legends, and figures. Significance values correspond to p values as follows: '-' > 0.05, * <0.05, ** <0.01, *** <0.001. To compare non-categorical values between groups we used the Mann-Whitney U test. To calculate correlation we used Spearman's correlations coefficient. Other statistical tests that were used are reported in the figure legends. In the boxplots, the upper whisker indicates the maximum value or 75th percentile +1.5 IQR; the lower whisker indicates the minimum value or 25th percentile 1.5 IQR. Survival differences were assessed using the log rank test CamDavidsonPilon/lifelines: v0.14.6 (Davidson-Pilon, 2019).