Quantifying the single-cell spatial landscape of cancer

By

Katie E. Blise

A DISSERTATION

Presented to the Department of Biomedical Engineering

and the Oregon Health & Science University School of Medicine

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

March 2024

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2D	two-dimensional
3D	three-dimensional
anti-CD40	agonistic CD40
αSMA	alpha smooth muscle actin
ADCC	antibody-dependent cellular cytotoxicity
APC	antigen presenting cell
AUC	area under the receiver operating characteristic curve
CAF	cancer-associated fibroblast
CAIX	carbonic anhydrase IX
CODEX	co-detection by indexing
CTLA-4	cytotoxic T-lymphocyte-associated protein 4
DC	dendritic cell
DFS	disease-free survival
EN	elastic net
ER	estrogen receptor
FFPE	formalin-fixed paraffin-embedded
FISH	fluorescence in situ hybridization
GrzB	granzyme B
GVAX	granulocyte-macrophage colony-stimulating factor vaccine
H&E	hematoxylin and eosin
HER2	human epidermal growth factor receptor 2
HNSCC	head and neck squamous cell carcinoma
HPV	human papillomavirus
IA	immune aggregate
ICB	immune checkpoint blockade
IFN	T-cell secreted interferon
IHC	immunohistochemistry
ILC1	innate lymphoid cells type 1
IQR	interquartile range
KIR	killer-cell immunoglobulin-like receptor
KL	Kullback-Leibler
mAb	monoclonal antibodies
MIBI-TOF	multiplexed ion beam imaging by time-of-flight
mIHC	multiplex immunohistochemistry
ML	machine learning
NAP	normal-adjacent pancreas
NK	natural killer

OHSU	Oregon Health & Science University
OS	overall survival
PanCK	pan-cytokeratin
PCA	principal component analysis
PD-1	programmed death protein 1
PD-L1	programmed death-ligand 1
PDAC	pancreatic ductal adenocarcinoma
PFS	progression free survival
PR	progesterone receptor
RCN	recurrent cellular neighborhood
ROC	receiver operating curve
ROI	region of interest
RT	room temperature
SEM	standard error of the mean
SHAP	Shapley Additive exPlanations
SMOTE	Synthetic Minority Over-sampling Technique
T	tumor
TAS	tumor-adjacent stroma
TiME	tumor-immune microenvironment
TLS	tertiary lymphoid structures
TME	tumor microenvironment
TNBC	triple negative breast cancer
Treg	regulatory T cell

Acknowledgements

I am extremely grateful for everyone who supported me in the pursuit of my degree.

First, thank you to my mentor, Dr. Jeremy Goecks, for providing me with outstanding training and mentorship throughout my degree. Your feedback, patience, and unwavering support allowed me to grow as a scientist over the past five years, and I am grateful for the time and energy you put into helping me learn.

To Dr. Lisa Coussens, thank you for taking a chance on me and welcoming me into your lab. You always made time for me, even when there seemed to be no time in your schedule, and you have continually advocated for me and my work. Thank you for your support—you have made me a better scientist.

Thank you to Dr. Katelyn Byrne for your selfless support and guidance over the past two years. This work would not have been possible without your investment. Thank you for being willing to dive head-first into the computational analyses and for teaching me more immunology than I could have ever imagined.

I would also like to extend my gratitude to the additional members of my dissertation advisory committee, Dr. Laura Heiser and Dr. Reid Thompson. Your guidance and feedback have helped shape my research and training for the better.

Thank you to the members of the Goecks and Coussens labs for providing me with ideas and feedback on my work. To Sam Sivagnanam, thank you for imparting so much of your wisdom on me. None of this work would have been possible without you.

Finally, I would like to thank my community for your steady support throughout this journey. Thank you to my parents and the rest of my family for always believing in me. To Claire, Anna, and Annie, thank you, thank you, thank you. You encouraged me, fed me, and kept things fun. You all are my home. Thank you to Julia, Molly, Anna, and Leslie for countless video calls, phone calls, and trips to Oregon. You've been there for me since before the start, and I am so grateful for each of you. To Ian, thank you for your friendship throughout this wild ride. I've appreciated having someone to turn to who is right there with me in this thing. To Lisa, Leah, Emery, and Shea, thank you for the outdoor adventures, dinner parties, and friendships. You all provided me with fun escapes from work and helped keep me grounded in this experience. Finally, to Whitney, you are my biggest cheerleader. Thank you for your endless encouragement and for never giving up on me. I am so lucky.

Abstract

Cancer is among the leading causes of death worldwide, and there is a critical need to improve treatment strategies and clinical outcomes for patients. In addition to neoplastic cells, tumors also consist of varying types and quantities of stromal, endothelial, and immune cells, making them highly heterogeneous entities. This diverse ecosystem of cells, termed the tumor microenvironment (TME), is now widely appreciated for its important role in shaping tumor progression, response to therapy, and clinical outcome. Propelled by recently developed multiplex tissue imaging assays, which generate massive datasets involving phenotypic and spatially resolved maps of cells residing in a tissue sample, research has demonstrated that both the cellular composition and spatial organization of the TME are important determinants of clinical outcome for cancer patients. However, given the sheer magnitude of these datasets and TME biological complexity, the exact relationships among TME phenotypic, functional, and spatial aspects and various clinical endpoints are still poorly understood. Thus, computational analyses that interrogate the single-cell spatial landscape of TMEs are desperately needed to improve the clinical care of cancer patients. This dissertation aims to address this need by (1) quantitatively assessing single-cell spatial proteomics datasets generated by multiplex immunohistochemistry for three tumor types and (2) applying statistical and machine learning approaches to identify potential candidate biomarkers of various clinical parameters. First, head and neck squamous cell carcinomas (HNSCCs) were analyzed for their cellular heterogeneity, longitudinal changes following therapy, and spatial architecture. Immune, neoplastic, and mesenchymal spatial organizations associated with progression-free survival; these TME spatial aspects may be used as potential candidate biomarkers for future HNSCC patient stratification. Breast TMEs were then assessed specifically for their natural killer (NK) cell function and

spatial organization in the context of human epidermal growth factor receptor 2 (HER2) disease. NK cell proximity with neoplastic cells correlated with the expression of various functionality biomarkers on NK and neoplastic cells, as well as HER2 status. These results reveal potential immunosuppressive mechanisms in $HER2^+$ breast TMEs, which could be targeted by new therapeutic interventions to improve outcomes for these patients. Finally, pancreatic ductal adenocarcinomas (PDACs) were assessed for T cell phenotypes, functions, and spatial organizations within the TME in the context of neoadjuvant immunotherapy. T cell exhaustion was reduced following immunotherapy use, and the presence of spatial neighborhoods of effector T cells located specifically in immune aggregates was associated with improved disease-free survival following immunotherapy. These results reveal important immunological features of the TME that could be leveraged to improve treatment strategies and clinical outcome for PDAC patients. Collectively, the work presented in this dissertation provides a roadmap for future quantitative analyses of highly complex spatial proteomics datasets. Importantly, the computational algorithms used here are applicable across single-cell spatial proteomics datasets and can be used to quantify the single-cell spatial landscape of additional TMEs, as well as other diseased or healthy tissues. Finally, future studies may leverage the immunological results presented herein to design new therapeutic interventions and biomarkers for patient stratifications to improve clinical outcomes for cancer patients.

Chapter 1: Introduction

1.1 Tumor microenvironment

1.1.1 Brief overview

Cancer is characterized by the dysregulated proliferation of genetically altered cells, leading to the growth of malignant—and often deadly—tumors [1]. In many ways, cancer can be thought of as an evolutionary process whereby genetic alterations in mutated cells provide a survival advantage to the mutated cells over normal cells. However, it has become increasingly clear that beyond the genetic changes within the cancerous cells, the environment in which these cells reside also plays a significant role in determining how the cancer will progress [2].

This environment, known as the tumor microenvironment (TME), comprises a diverse collection of cells [3, 4]. Beyond the malignant neoplastic cells, TMEs contain structure-related cells, including mesenchymal and vascular endothelial cells. These cells are the building blocks for the stromal tissues that hold tumors together and the blood vessels that contribute to angiogenesis and hypoxia. TMEs also contain a wide assortment of immune cells, which play varying pro- or anti-tumorigenic roles within the TME. Beyond these broad cell lineages, each cell may express an array of functional markers that further differentiates it from cells of the same canonical classification. Combined with its phenotypic lineage, these functional markers help determine the specific role or ability of the cell in the TME. Finally, the spatial location of each individual cell within the TME further dictates the function of the cell, as cells communicate and affect each other's abilities through spatial proximity [5]. Overall, TMEs are heterogeneous and complex ecosystems composed of thousands to millions of cells [6]. Each cell descends from a specific phenotypic lineage, serves a particular function, and is located within a coordinated spatial network of neighboring cells within the TME (**Figure 1.1**).

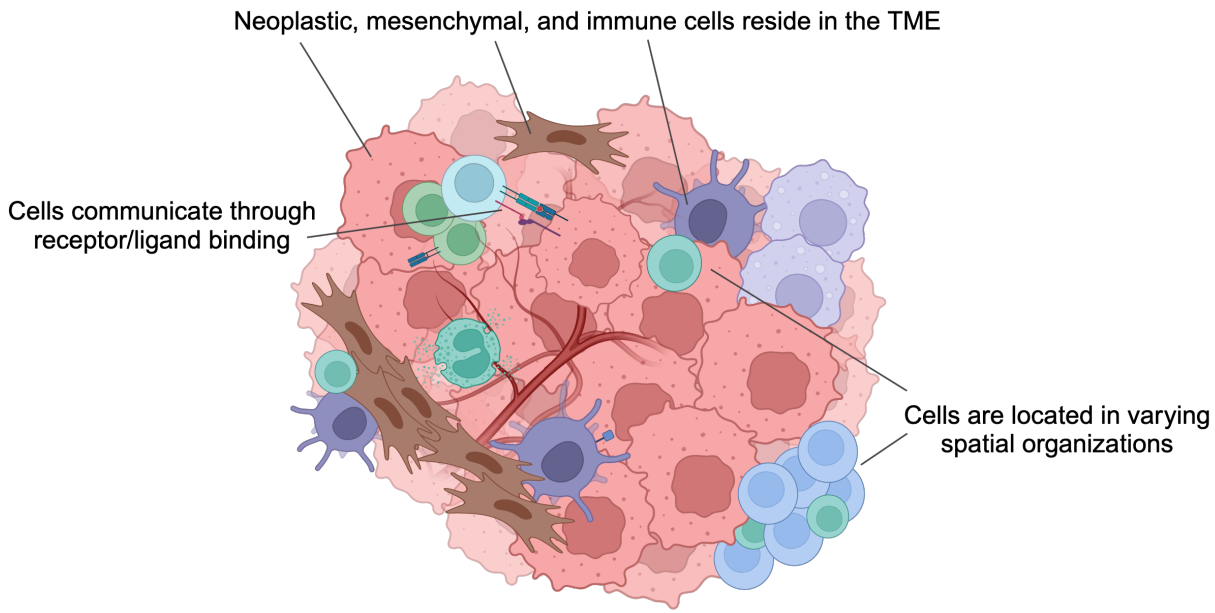


Figure 1.1: TME overview. Many types of cells expressing various functional markers are located in distinct spatial arrangements throughout the TME. This figure was created using BioRender.com.

1.1.2 TME immune contexture

Understanding the types, functions, and locations of the leukocytes, or immune cells, residing in the TME is of particular interest to cancer researchers. This is because tumors form, progress, and metastasize throughout the body, in large part, due to a failure in immunosurveillance [7]. Under homeostatic conditions, a healthy immune system continually surveilles the body for abnormal cells, tissue damage, and pathogens. Upon detection of these aberrations, immune cells destroy and/or remove the unhealthy cells, thus preventing their continued presence in the body. In the context of cancer, genetically mutated cells avoid detection and/or destruction by the immune system, resulting in the progression of neoplastic tissues into malignant tumors. Although not always attributed to the same mechanism, the general skirting of neoplastic cells past immune detection and destruction is now considered a "hallmark" of cancer [2]. As TMEs contain numerous immune populations, understanding which of these immune cells are capable of mounting an anti-tumor response with or without

therapeutic priming, versus which immune cells are complicit in, or even enable, tumor progression, can help inform future targeted therapies that leverage the immune system of the body to eradicate the tumor [8, 9].

All types of leukocytes with varying functions have been found within the TME, including cells from both adaptive and innate lineages. T cells are a critical adaptive immune cell population present in the TME [10]. From a rudimentary perspective, canonical $CD8^+$ T cells and the majority of $CD4^+$ T helper cells predominately exert anti-tumor functions in the TME and help identify and kill neoplastic cells [11, 12], while regulatory T cells (Tregs)—which are characterized by the expression of both CD4 and FOXP3—contribute to an immunosuppressive TME [13, 14]. However, in reality, the T cell phenotypic landscape is much more heterogeneous than this [15-18]—many $CD8^+$ T cells and $CD4^+$ T helper cells in the TME are dysfunctional and/or exhausted due to chronic stimulation and signals from nearby cells [19-22], and certain subsets of Tregs counterintuitively possess anti-tumor functions [23]. Additionally, the degree of T cell infiltration into TMEs varies based on tumor type and signals from the collection of cells present in the TME, among other factors [24, 25]. Thus, deeper investigations into the amounts and phenotypes of T cells present, their spatial locations and cellular neighbors within the TME, and their associated functions are needed to better understand how to harness the full cytotoxic potential of T cells against tumors.

B cells are another adaptive immune population present in TMEs [26]. While certain subsets of B cells may be pro-tumorigenic [27-29], B cells are often located in spatial clusters of immune cells, including tertiary lymphoid structures (TLS) [30], which are typically associated with anti-tumor functions [31-33]. Interrogation of the cells present in immune aggregates and TLS may enable improved understanding of how B cells contribute to anti-tumor immunity.

Bridging the adaptive and innate immune systems are natural killer (NK) cells [34]. NK cells are potent cytotoxic leukocytes, but they are often present in limited numbers within the TME [35, 36]. Additional research into the types of TMEs NK cells infiltrate and how their surrounding cellular neighbors in the TME impact their function may reveal new mechanisms to further leverage the cytotoxic powers of NK cells within the TME.

While sometimes overlooked given their lack of memory of and specificity for neoplastic cells, cells of the innate immune system play pivotal roles in tumor progression [37]. These myeloid cells include dendritic cells (DCs), macrophages/monocytes, and granulocytes (neutrophils, basophils, eosinophils, and mast cells). DCs typically contribute to anti-tumor immunity in the TME by aiding in the detection of neoplastic cells. DCs present neoplastic antigen to T cells to help activate an effector T cell response against the tumor, and they are considered a major antigen presenting cell (APC) in the TME [38, 39]. Macrophages are also capable of presenting antigen to T cells and contributing to anti-tumor immunity [40]. However, more often, and especially in late-stage tumors, macrophages polarize away from an anti-tumorigenic "M1-like" phenotype toward an "M2-like" phenotype, which is characterized by many pro-tumorigenic functions [41]. Unlike their M1-like counterparts, M2-like macrophages create a highly immunosuppressive TME and contribute to neoplastic cell proliferation, angiogenesis within the TME, and tumor metastasis [42]. Finally, granulocytes also fulfill varied and often opposing roles in the TME depending on the phenotype of the granulocyte, tumor type, tumor stage, and spatial proximity to other cells within the TME [43-45]. Granulocytes are characterized by the presence of granules inside the cell, which are released into the surrounding tissue as a defense mechanism. Neutrophils, which comprise a major subtype of granulocytes, are typically found to contribute to an immunosuppressive TME through the promotion of tumor

metastasis, among other functions [46, 47]. Comprising a much smaller subset of granulocytes found in the TME, the roles of basophils and eosinophils in the context of cancer are poorly understood [44, 48, 49]. Prior studies on the functions of basophils and eosinophils report conflicting results [43, 50], again highlighting the complexity of immune cell function in the TME. Finally, mast cells, which are tissue-resident granulocytes, have been shown to recruit other immune cells, spanning immunoreactive and immunosuppressive cells, to the TME, and they can also contribute to increased angiogenesis in the TME [51]. Thus, similar to the lymphocytes of the adaptive immune system, myeloid cells have diverse functions in the TME and can exert both pro- and anti-tumorigenic functions depending on a variety of circumstances.

In summary, control versus progression of a tumor is due, in part, to the immune cells present in the TME. However, immune cells are highly heterogeneous in both their phenotypes and functions, making it challenging to fully assess the concerted role of the immune system for a given TME. Further, the cells surrounding the immune cells have a profound impact on the abilities of the immune cells present in the TME. Analyses that are capable of isolating and identifying the phenotype of each immune cell present in a TME, deducing its associated function, and also contextualizing its spatial location among other cells in the TME are needed to gain clarity on how TME immune contexture works to slow or accelerate tumor growth.

1.1.3 Clinical relevance

Just as the phenotypic composition and spatial organization of cells in the TME play a critical role in impacting tumor progression, TME cellular contexture also contributes to patient survival and regulating response to therapy [6, 8, 52-55]. From a purely compositional perspective, increased densities of $CD8^+$ T cells have been associated with improved survival across several tumor types, including breast [56], colorectal [57], clear cell ovarian [58], and

prostate [59]. Conversely, presence of M2-like macrophages has been implicated in shorter survival times for head and neck [60], bladder [61], and breast [62] cancer patients. Incorporating both cellular density and spatial localization, the Immunoscore has been used over the past decade to aid in colorectal tumor staging [63]. Higher Immunoscores, which reflect increased presence of specific $CD8^+$ T cells in the tumor core and invasive margin are associated with reduced rates of recurrence, while lower Immunoscores are associated with reduced $CD8^+$ T cell infiltration and increased rates of recurrence [64]. Finally, the spatial organization of the TME alone has been associated with response to therapy and patient survival in several tumor types. For example, increased spatial proximity between cells involved in the programmed cell death protein 1 (PD-1)/programmed death-ligand 1 (PD-L1) immunoregulatory checkpoint has been correlated with improved response to therapies targeting this checkpoint in Merkel cell carcinoma [65] and metastatic melanoma [66, 67], and specific TME cellular architectures were found to associate with survival for breast cancer patients [68] and several cancer types [52].

Beyond derivation of potential molecular biomarkers, interrogation of TME cellular contexture has also helped identify new targets for therapeutic interventions for cancer patients. Immune checkpoint blockades (ICBs) leverage knowledge of basic T cell immunology and TME immune contexture specifically, as they target various T cell regulatory pathways, such as the PD-1 and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) pathways [69]. PD-1 and CTLA-4 are receptors expressed on the surface of T cells, and when bound to their corresponding ligands, the effector function of the T cell is inhibited [70, 71]. This immunoregulatory mechanism, which exists under homeostatic circumstances to prevent autoimmunity, is hijacked by the cancer to block T cells from killing its own neoplastic cells. ICBs disrupt these inhibitory receptor/ligand bindings and remove the “brake” from T cells, thus

reinvigorating their ability to detect and kill neoplastic cells in largely antigen-dependent manners [72, 73]. ICBs have demonstrated unparalleled success for some patients, highlighting the power of harnessing the natural immune system of the body to fight tumors [74-76]. However, ICB success stories are limited, and the vast majority of cancer patients do not respond to these therapies [77], underlining our still naive understanding of TME complexity.

As cancer and subsequent metastases are oftentimes lethal, there is a critical need to identify superior treatment strategies, as well as identify improved molecular biomarkers of recurrence likelihood, therapeutic response, and patient survival. Studies that interrogate TME cellular composition and spatial architecture have already enabled more personalized treatment regimens and better patient stratification approaches for certain cancers [78]. However, many cancer patients still have poor clinical outcomes, highlighting the need for continued investigation of these TMEs.

1.1.4 Head and neck squamous cell carcinoma

Head and neck squamous cell carcinomas (HNSCCs) arise from the mucosal epithelium lining the oral cavity, pharynx, and larynx [79]. Risk factors for developing this cancer include alcohol and tobacco use, as well as infection with human papillomavirus (HPV) [80]. Roughly 890,000 new HNSCC cases are diagnosed per year, and 450,000 people die from HNSCC per year [80]. Global incidence is continuing to rise, and there is an urgent need to identify new treatments and improved molecular biomarkers of survival, response to therapy, and tumor progression. Investigation of head and neck TME cellular landscapes presents an opportunity to elucidate new targeted therapies and clinical biomarkers for HNSCC patients.

HNSCC TMEs are remarkably heterogeneous, and the immune contexture of HNSCCs varies greatly depending on the HPV status of the tumor [81]. HPV-positive [HPV(+)] tumors

tend to have more tumor infiltrating lymphocytes than HPV-negative [HPV(-)] tumors, which are instead characterized by highly immunosuppressive TMEs [81]. This difference in TME immune contexture is thought to contribute to the discrepancy in survival outlook for patients with HPV(+) versus HPV(-) tumors. Patients with HPV(+) HNSCC tend to have better clinical outcomes—although still poor—with a 3-year overall survival (OS) rate of 34%, as compared to HPV(-) HNSCC patients, whose 3-year OS rate is only 19% [82].

Current standard of care for primary HNSCC involves surgery and adjuvant radiotherapy and/or chemotherapy [83]. ICBs that target PD-1 have also been approved for treatment of recurrent or metastatic HNSCC [84-86], yet prognosis has hardly improved—especially for HPV(-) HNSCC patients [87]. Thus, there is an even a more critical need to develop improved therapies for HPV(-) HNSCC patients in particular. While infiltrating immune cells are often present in limited numbers in HPV(-) HNSCC TMEs, understanding which immune cells do exist and how they may be suppressed by neighboring cells could inform new therapies that can convert the traditionally immunosuppressive TME into an immunoreactive one.

1.1.5 Breast cancer

Although not always as lethal as HNSCC, breast cancer is the most commonly diagnosed cancer worldwide, with over 2.2 million new cases diagnosed per year [88]. Breast cancers are traditionally classified based upon their expression of estrogen receptor (ER), progesterone receptor (PR), and/or human epidermal growth factor receptor 2 (HER2) [89]. This molecular subtyping helps inform treatment regimens, such as hormone therapy to target ER^+ and PR^+ tumors and anti-HER2 antibodies for $HER2^+$ tumors. However, nearly 700,000 people still die from breast cancer per year [88], and an improved understanding of breast TME heterogeneity within molecular subtypes is urgently needed to guide more personalized treatment strategies.

Breast TMEs are heterogeneous entities, and their immune contexture and its associated prognostic value varies depending on the molecular classification of the tumor. Lymphocytes infiltrating breast tumors tend to be $CD8^+$ T cells, whose presence is associated with improved prognosis specifically for $HER2^+$ disease and triple negative breast cancer (TNBC) [90], which is defined by the lack of ER, PR, and HER2 positivity. Immunosuppressive lymphocytes, including Tregs and $CD4^+$ T helper type 2 (Th2) cells are also present in breast TMEs and associate with poor prognosis [90]. Finally, breast TMEs contain varying amounts of innate immune cells, including DCs, macrophages, and NK cells.

Although less common than tumor infiltrating lymphocytes, NK cells are of particular importance in $HER2^+$ breast tumors. This is because HER2-targeted therapies, such as the anti-HER2 antibody Trastuzumab, have been shown to elicit an originally unintended NK cell-mediated anti-tumor response in $HER2^+$ breast cancers [91]. Trastuzumab has been reported to increase the number of tumor-infiltrating NK cells in $HER2^+$ breast TMEs, as well as trigger NK cell cytotoxicity through an antibody-dependent cellular cytotoxicity (ADCC) mechanism [92, 93]. However, despite this positive effect of Trastuzumab, inhibitory signals from neighboring cells in the TME often dampen NK cell cytotoxicity [94, 95]. Improving understanding of how to effectively harness the NK cell-mediated anti-tumor response following Trastuzumab use in $HER2^+$ breast tumors could greatly improve clinical outcomes for these patients. This is especially important, as $HER2^+$ breast cancer patients have a much higher risk of recurrence and metastasis than their ER^+ or PR^+ counterparts [96]. Thus, investigations into NK cell phenotype, function, and spatial organization within $HER2^+$ breast TMEs specifically are critically needed.

1.1.6 Pancreatic ductal adenocarcinoma

While not as common as HNSCC or breast cancer, pancreatic ductal adenocarcinoma (PDAC) is one of the deadliest cancers, with a 12% 5-year survival rate [97]. There are nearly 500,000 new cases diagnosed per year, and there are an almost equivalent number of deaths per year (over 460,000) resulting from PDAC [98]. Current standard of care includes surgery and adjuvant chemotherapy. However, this line of treatment is largely ineffective, and the vast majority of patients experience recurrence in as few as 7 months [99]. Additionally, PDAC patients often present with late-stage disease and are ineligible for surgery due to the existence of distant metastases and/or due to the proximity of critical vasculature surrounding the pancreas [97]. The lack of feasibility and efficacy of existing treatments and the dismal survival outlook for PDAC patients highlight the desperate need to identify new therapeutic interventions for these patients.

The PDAC TME presents several opportunities to target with new therapeutic strategies. The TME is characterized by a dense stroma and extracellular matrix which is thought to preclude effector lymphocytes from entering the tumor core. While intuitively, disruption of the stromal barrier could be hypothesized to enable the passage of lymphocytes into the TME, initial attempts to deplete stromal fibroblasts through therapeutic intervention instead contributed to increased tumor vascularization and metastasis [100-103]. Up until recently, therapies that targeted the immune system were also largely unsuccessful [104]. This is due in part to the large proportion of myeloid cells present, which promote a highly immunosuppressive TME [105-107]. Additionally, PDAC is characterized by relatively few genetic mutations, thus limiting T cells' natural ability to infiltrate and detect the cancer [108]. Altogether, these TME attributes are main contributors to the failure of ICB use in PDAC.

More recently, therapies that act on T cells through mechanisms differing from ICBs have shown the ability to promote T cell priming against PDAC. Two of these therapeutic interventions are peptide vaccination [109] and agonistic CD40 (anti-CD40) monoclonal antibodies [110]. Both strategies indirectly prime T cells through the direct stimulation of APCs. Peptide vaccination involves administration of a highly personalized vaccine that informs T cells of the neoantigens specific to the tumor cells from a patient, which APCs present to T cells [109]. This treatment demonstrated remarkable success in a small clinical trial [109] and deserves further attention in larger clinical trials. However, personalized vaccine development took over two months from time of surgery to time of vaccination, which is costly for most PDAC patients, whose recurrence timeline usually occurs on the scale of months. Anti-CD40 therapy also induces T cell responses by binding to the CD40 receptor expressed on the cell surface of APCs [111]. This mimics the binding of the CD40 receptor on the APC with its corresponding CD40 ligand. When this binding occurs, the APC becomes licensed to activate T cells without additional signals needed. Anti-CD40 therapy with or without combination chemotherapy and/or ICB has shown promise in preclinical models and is now being tested in clinical trials [110, 112, 113].

For one of the first times, peptide vaccination and anti-CD40 therapy demonstrate the feasibility of eliciting a T cell response against PDAC. However, optimization of these therapies to maximize clinical outcome still requires extensive investigation into the T cell response powered by these therapies. Thus, deep interrogations of T cell phenotype, function, and spatial localization within the PDAC TME with and without therapies that prime T cells are urgently needed to continue improving clinical outcome for patients with this devastating disease.

1.2 Multiplex immunohistochemistry

1.2.1 Brief overview

As discussed above, studies that investigate the composition and spatial organization of the cells in the TME have the potential to improve understanding for many cancer patients. However, most traditional assays fail to comprehensively evaluate both the composition and the spatial organization of the tissue concurrently. Common molecular assays, such as RNA-sequencing [114] and flow cytometry [115], assess the molecular makeup of the tissue, but these assays disrupt the spatial context of the tissue sample. Immunohistochemistry (IHC), which involves the staining of antibodies on tissue sections on a slide, retains cellular locations, and it has been the gold standard of pathology over the last twenty years [116]. However, IHC can only stain up to a few antibodies per tissue slide, thus failing to capture the detailed cellular phenotypic information of the tissue. Until recently, the available methods for interrogating the TME required a tradeoff between knowing the phenotypic composition of the TME or the spatial landscape of the TME. Advances in single-cell spatial proteomic imaging technologies have resulted in several novel multiplex tissue imaging assays, including cyclic immunofluorescence [117], imaging mass cytometry [118], co-detection by indexing technology (CODEX) [119], multiplexed ion beam imaging [120], and multiplex immunohistochemistry (mIHC) [121], which overcome the aforementioned challenge and deeply phenotype cells while also preserving the cellular spatial arrangement of the tissue.

Given the critical role that immune cells play in detecting and destroying neoplastic cells in the TME, mIHC was designed specifically to interrogate the immune contexture of solid tumors [121, 122]. mIHC builds on the foundation provided by standard IHC while employing a cyclical antibody staining protocol (**Figure 1.2**). It sequentially assays the same tissue slide many times and can stain for over 25 antibodies, all while maintaining the spatial context of the

tissue sample. The resulting imaging data is processed by a computational analysis pipeline to identify cellular locations in the form of Cartesian coordinates, and the resulting protein expression levels are hierarchically gated to classify each cell into neoplastic, mesenchymal or a specific leukocyte subtype of either lymphoid or myeloid lineage. Further, the expression of additional proteins is also used to determine cells that are positive or negative for various functional markers, such as the PD-1 immunoregulatory protein. Thus, the data produced from the mIHC assay and its downstream computational image processing pipeline provides an unrivaled picture of the TME that includes both cellular phenotypic and spatial information. Given the value of understanding the single-cell spatial landscape of the TME, analysis of this mIHC data has the potential to identify promising biomarker candidates and enable increased personalized treatment strategies for cancer patients.

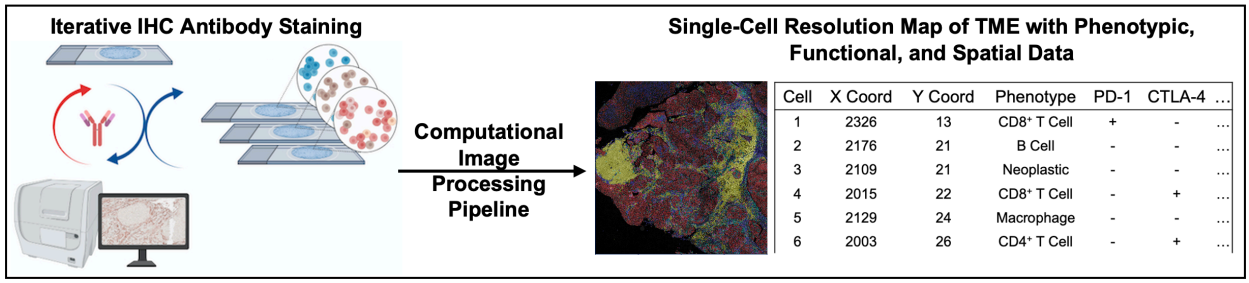


Figure 1.2: mIHC proteomics platform and computational image processing pipeline yield spatially resolved maps of TMEs.

1.2.2 Challenges to mIHC data interpretation

As mIHC provides single-cell resolution maps of the TME, resulting mIHC datasets are incredibly large. For example, a single tissue region sampled from the HNSCC TME with area 2500 x 2500 μm^2 contains on average 21,000 cells. Thus, on average, the processed mIHC data matrix corresponding to one tissue region has 21,000 rows in the table, with each row corresponding to data for one cell. Oftentimes in an mIHC experiment, several tissue regions per TME are assayed, resulting in hundreds of thousands of cells assayed for a single cancer patient.

When considering the resulting data from an entire cohort of cancer patients whose tissue samples are all assayed with mIHC, the dataset suddenly consists of phenotypic, functional, and spatial data for millions of cells. The incredible size of mIHC datasets alone presents a challenge for researchers attempting to evaluate the types and roles of cells present in the TME and how they are spatially organized.

Beyond quantification and description of TME cellular contexture using mIHC datasets, incorporation of accompanying tissue and clinical metadata provides interesting, but challenging, opportunities to assess the clinical relevance of TME biological complexities. Examples of tissue metadata include anatomic site from which the tumor was collected, histopathologic information about the TME sample, and tumor stage, among others. Clinical metadata include parameters such as therapy administered, response to therapy, clinical subtype of the tumor, timeline to recurrence, and survival duration. Ultimately, linking TME biology to these various parameters has the potential—and is necessary—to improve treatment strategies and clinical outcomes for cancer patients. However, identifying clinically meaningful relationships in the data is particularly challenging when several tissue or clinical parameters are varied across the dataset. For example, some datasets encompass tumor specimens spanning each stage, collected from varying anatomic sites, at multiple timepoints, and following different therapeutic interventions. When considering a dataset like this—which also possesses data for millions of cells—it becomes increasingly challenging to interpret biological and clinical meaning.

Finally, given the novelty of mIHC datasets—and multiplex tissue imaging datasets in general—methods for interpreting their biological meaning or clinical relevance remain to be standardized. While metrics derived from traditional IHC analyses to assess basic cellular phenotypes can be applied to mIHC data, no standard metrics exist for quantifying the spatial

layout of cells in the TME. Overall, analysis of the single-cell spatial landscape of the TME has great potential to revolutionize the treatment and care of cancer patients; however, it remains unclear how to best evaluate these types of datasets.

1.3 Quantitative analyses of the TME

1.3.1 Algorithms to quantify TME single-cell spatial landscape

As biological datasets continue to grow in size and complexity, computational approaches are needed to effectively analyze them. This is particularly true for the single-cell spatial proteomics datasets of the TME resulting from multiplex tissue imaging assays, where the data includes phenotypic, functional, and spatial information for millions of cells, often in addition to various accompanying metadata.

Assessment of the single-cell spatial landscape of the TME necessitates metrics that calculate the abundance of cells possessing distinct phenotypes and functions, as well as metrics that quantify the cells' spatial organization in the TME. Quantification of the phenotypic and functional composition of the TME can often still be performed using simple metrics such as cell densities, ratios, or proportions present in the TME [121]. Demonstrating the clinical utility of these metrics, proportions of cells positive for various biomarkers are used to determine sensitivity to treatment with ICB. For example, the U.S. Food and Drug Administration approved the use of an ICB targeting the PD-1/PD-L1 checkpoint in patients with non-small cell lung cancer when at least 50% of the neoplastic cells present express PD-L1 [123-125]. While this biomarker is easy to calculate, it does not always accurately predict ICB response [125, 126], suggesting that additional TME factors, such as cellular spatial localization, must also be considered.

Quantifying the spatial arrangement of cells within the TME is more complex than assessing the composition of the TME. Dozens of methods have been used to quantify TME cellular spatial organization, and the choice of method is partially dependent upon the magnification level desired for the analysis [52]. Evaluation of full surgical resections enables detection of the broad histopathologic location each cell resides in, including the tumor core, invasive margin, and more distal stromal tissue surrounding the tumor (**Figure 1.3A**) [107]. More granular spatial analyses include metrics borrowed from ecology [127], which were designed to understand the overlap of different species within a given habitat. For example, metrics such as the Morisita-Horn index and the Getis-Ord geospatial hotspot statistic are ecological metrics that have been used to assess the spatial distribution of immune and neoplastic cells within various TMEs [128, 129]. Respectively, these metrics calculate the spatial colocalization and extent of spatial clustering versus dispersion of two populations within a given space; they do not account for the actual distances between single cells within the outlined space (**Figure 1.3B**).

Highly granular distance-based metrics are particularly valuable in the context of studying TME biology, as cells require spatial proximity to communicate [5]. For example, immune cells secrete chemokines and cytokines, which are small proteins that diffuse through the TME and signal to other nearby cells, impacting their function [130]. While one study found T-cell-secreted interferon (IFN)-γ secreted by activated $CD8^+$ T cells could travel over 800 μm in the TME [131], most studies report chemokine and cytokine signaling distances of 30 to 250 μm [132-136]. Direct signaling mechanisms, which involve the binding of receptor ligand pairs expressed on cell surfaces, require even more proximity between cells. These cells must be directly adjacent to one another to facilitate binding of their proteins, and thus assessing

distances in the 15 to 30 μm range is more appropriate to capture these types of interactions [137, 138], such as the binding of PD-1 with its corresponding PD-L1 ligand [65]. Thus, spatial metrics that integrate the distance between cells into the calculation are important for understanding how cells affect one another in the TME.

Among the spatial analyses that involve a user-defined distance parameter are a recently developed "mixing score" metric [139], cell-cell interaction measures [65], and cellular neighborhood analyses [140] (**Figure 1.3C**). The mixing score was developed to specifically interrogate immune cell spatial mixing or compartmentalization with neoplastic cells within the TME, and it was found to associate with survival in TNBC [139]. Although the mixing score is a regional measure of spatial organization, it still incorporates knowledge of which types of cells are in contact with one another in the given region. Cell-cell interaction measures strip spatial analyses to their core, as they simply count the frequency of two cell types located within a specified distanced threshold from one another in the TME. Interactions can then be normalized to the densities of the cells present to avoid skewing the interaction counts by cells that are present in greater abundances. When interaction frequencies are calculated across all pairings of cell phenotypes present, it becomes easy to determine which cell types tend to be located most proximally to one another versus cell types that are rarely present in close proximity. Spatial interactions can also be calculated for cell phenotypes expressing specific functional markers. For example, studies interrogating the PD-1/PD-L1 immunoregulatory checkpoint found that tumors containing more spatial interactions involving PD-1^+ cells and PD-$L1^+$ cells were associated with improved response to ICBs targeting this checkpoint across multiple cancers [65, 67, 141]. Finally, cellular neighborhood analyses comprise an emerging sophisticated technique used to quantify TME spatial organization. Neighborhoods are defined as spatial groupings of

cells surrounding a given "seed" cell within the TME. Neighborhoods can be calculated by either identifying a set number of nearest cells to the seed cell or by identifying all cells within a specified radius around the seed cell. Then, based upon the composition of cells residing in the neighborhood, neighborhoods can be clustered using statistical techniques to identify recurring spatial groupings of cells that are present across multiple TMEs. Recurrent cellular neighborhoods were originally used to quantify colorectal TMEs, and the frequency of PD-1^{+} CD4^{+} T cells in a specific cellular neighborhood were found to associate with survival in these patients [140]. Cellular neighborhoods can be used to characterize the spatial organization of larger tissue structures within the TME, such as TLS or tumor nests.

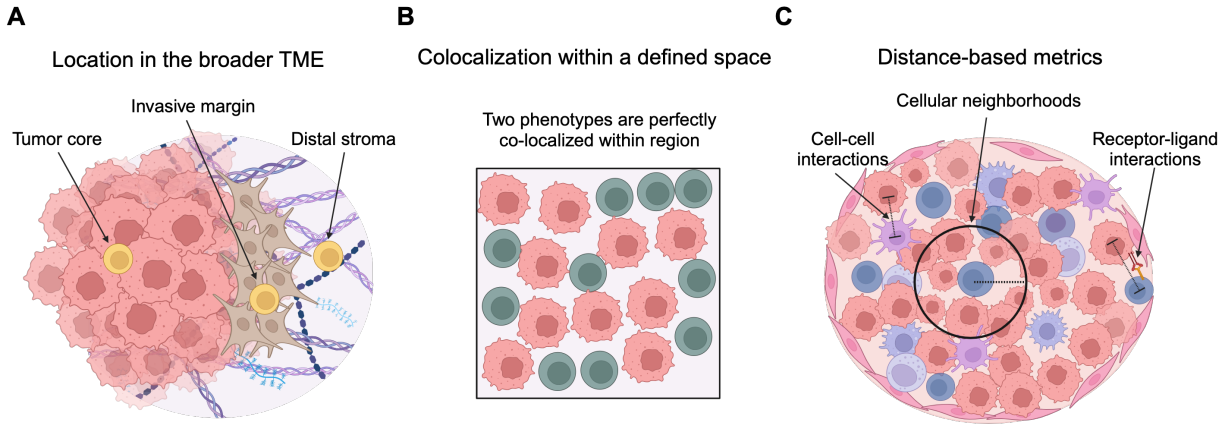


Figure 1.3: Overview of TME spatial quantifications. TME spatial organization can be defined in terms of **A)** location in the broader TME context, **B)** ecological measures of colocalization within a defined space, and **C)** distance-based metrics requiring proximity between cells. This figure was created using BioRender.com.

While each of the methods described above quantify various aspects of the TME single-cell spatial landscape, it remains unclear which of these aspects are clinically meaningful, and thus, which of the methods should become standardized for future analyses.

1.3.2 Machine learning to link TME biology to clinical parameters

Understanding the relationship between TME biology and clinical parameters is needed to ultimately improve treatment strategies and outcome for cancer patients. For example, analysis of longitudinal TME data, such as paired tissue biopsies collected at multiple time points from the same patient, may enable a deeper understanding of how a tumor progresses over time. Knowing the clinical subtypes of tumors facilitates deeper interrogation of TME cellular and spatial heterogeneity between, and within, subtypes, which may ultimately lead to increasingly personalized therapeutic approaches. Comparison of TMEs exposed to varying therapies may improve understanding of how and where specific therapies impact the TME, as well as how they may be combined with other therapies to improve response. Finally, linking clinical outcome data, such as progression- or disease-free survival and overall survival, to TME biology provides important insight into the types of TMEs associated with improved or worsened outcomes and identifies potential candidate biomarkers for future patient stratifications.

For some of these analyses, simple statistical correlation and difference tests suffice. For example, comparison of the average density of a given cell phenotype present in samples collected in primary versus recurrent tumors can be performed using a traditional statistical difference test. This analysis would test the hypothesis that the samples collected at the two timepoints differed in their densities of the cell phenotype assessed, and this knowledge could aid in understanding tumor progression. While traditional statistical approaches work well in relatively simple circumstances, these approaches have several limitations, especially in the context of large and complex datasets [142].

First, traditional statistical approaches require prior knowledge about the data to generate a testable hypothesis. In the case of high-dimensional biological datasets, such as single-cell spatial proteomics datasets, it is not always clear which features should be tested for their

relationship to a given clinical parameter. Second, traditional statistical approaches are often parametric, meaning they assume the shape of the underlying distribution of the data. Complex biological datasets do not always contain features that follow a normal distribution, and thus parametric statistical methods may be inappropriate for use with these datasets. Finally, traditional statistical approaches fail to evaluate the downstream impact of the combination of many features, each potentially weighted with varying degrees of importance. In the context of single-cell spatial proteomics datasets and TME complexity, often several biological features contribute to a given clinical parameter. Therefore, it is necessary to consider the role of combinations of data features, which traditional statistical methods fail to do. Thus, more advanced calculations are needed when analyzing highly complex datasets, such as single-cell spatial proteomics datasets.

Machine learning (ML), a form of artificial intelligence, offers a solution to the aforementioned data complexity challenges and is becoming increasingly used to analyze high-dimensional biological datasets [143]. At its core, ML involves the use of statistical algorithms to learn patterns from data, and it excels at analyzing large and complex datasets for several reasons. First, unlike traditional statistical approaches, ML approaches are largely unbiased; the methods are data-driven and are able to detect previously unknown patterns, requiring little prior knowledge about the dataset. Second, ML approaches are flexible, as they are not limited to parametric models and do not always assume the shape of a distribution. Finally, ML methods are capable of weighing and combining data features, to identify which groupings of features most contribute to a given output. Thus, given the complexity of single-cell spatial proteomics datasets, where TME biology can be quantified in seemingly countless ways and the clinical relevance of the biological features is unclear, ML offers a promising approach for learning

which TME biological features are associated with clinical parameters for these datasets [144, 145].

Several types of ML algorithms exist, including unsupervised and supervised algorithms, and the selection of the appropriate method depends upon the task in question. Unsupervised ML algorithms learn patterns in data without the help of labeled data. These types of algorithms are particularly useful in exploratory data analyses, where the goal is to identify similarities and differences among the data without requiring the algorithm to bin samples according to a label. Two types of unsupervised learning methods include dimensionality reduction techniques and clustering techniques. Dimensionality reduction techniques are particularly valuable when working with high-dimensional datasets, where the number of features is far more than the number of samples. Dimensionality reduction methods combine data features through linear or non-linear algorithms to reduce the overall number of features. The smaller number of resulting features can be more easily visualized in a two- or three-dimensional projection (**Figure 1.4A**), and various metadata or clinical annotations can be overlaid with the projection to identify patterns in the data. Clustering techniques are another form of unsupervised learning, which are used to bin data samples into distinct clusters, such that samples within the same cluster possess similar characteristics. In the context of single-cell spatial proteomics datasets, clustering methods are often used to group tumor samples based upon their phenotypic compositions [107]. This results in clusters that contain tumors composed of similar cell types, and clusters can also be manually labeled with various metadata or clinical annotations to determine if the clusters represent clinically meaningful differences (**Figure 1.4B**).

In contrast to unsupervised learning algorithms, supervised ML leverages labeled data to train an ML model. Classification tasks encompass one main type of supervised learning.

Classification involves training models to learn which combinations and weights of inputted data features associate with each of the resulting labels (**Figure 1.4C**). Models can then be used to make predictions on new data without the need for labels. The predictive potential of supervised learning makes it an attractive technique for detecting clinical meaning, where models could be trained, for example, based upon single-cell spatial features of the TME to predict tumor progression, response to therapy, or survival [144, 145]. Additionally, quantitative analyses can be used to interpret how the model is making its predictions [146]. In the example where a model is trained to predict survival duration based upon compositional and spatial features of the TME, interpretation of the predictions of the model would provide researchers with an improved understanding of which combinations of TME features were related to either improved or worsened survival. These features may represent candidate biomarkers of survival, and they could be used to stratify future cancer patients.

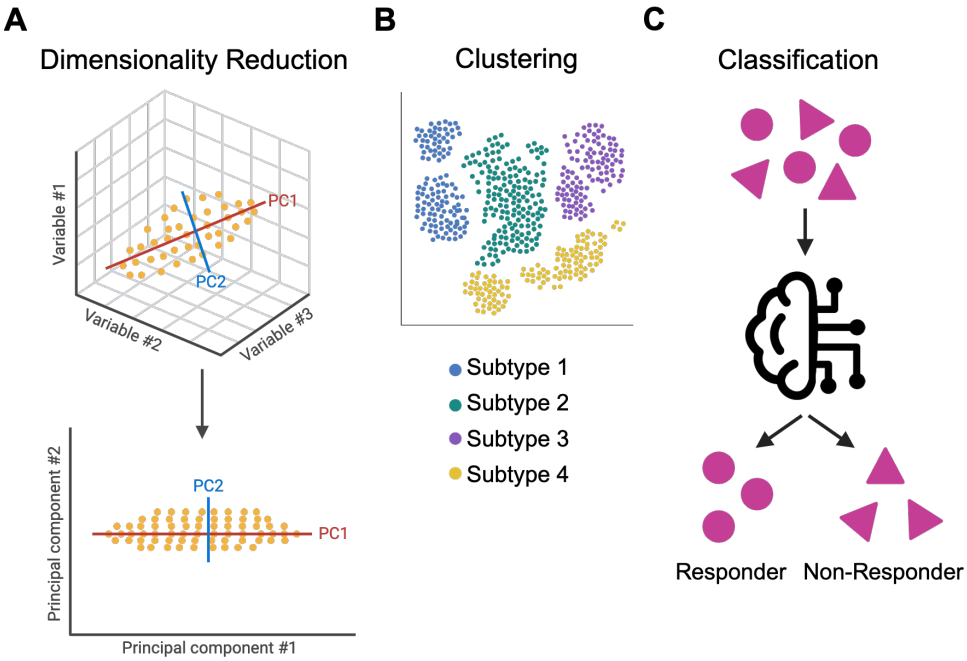


Figure 1.4: Overview of ML approaches. A) Dimensionality reduction techniques enable visualization of data in low-dimensional projections. **B)** Clustering methods group samples into distinct clusters based upon shared characteristics, which can be annotated by various metadata, such as tumor subtype. **C)** Classification tasks train a model to accurately predict data labels, such as response to therapy. This figure was created using BioRender.com.

Pursuit of computational analyses such as those described above—both to quantify the single-cell spatial landscape of the TME, as well as to link TME biology to clinical parameters—has the potential to reveal important TME biology that could be leveraged to design novel therapeutic interventions and improve clinical outcomes for cancer patients.

1.4 Contributions

The work presented in this dissertation attempts to addresses the challenges and opportunities described above. TME cellular phenotypes, functions, and spatial organizations were quantified for three extremely common and/or aggressive tumor types from mIHC single-cell spatial proteomics datasets, and clinically relevant TME biological features were then identified through various statistical techniques. Chapter 2 describes a detailed quantitative assessment of HPV(-) HNSCC tumor heterogeneity, progression over time, and spatial architecture. Immune cell and alpha smooth muscle actin (αSMA^{+}) mesenchymal cell spatial organization were found to correlate with progression-free survival. These results provide novel insight into how the immune system and stromal cells contribute to clinical outcome for HPV(-) HNSCC patients and reveal potential candidate biomarkers for progression-free survival for future HPV(-) HNSCC patients. Chapter 3 presents a functional and spatial assessment of NK cells within the TMEs of $HER2^{+}$ and $HER2^{-}$ breast cancer patients. This study identifies new relationships among NK cell functionality, NK cell spatial neighbors, and HER2 status, which may be leveraged to improve treatment for $HER2^{+}$ breast cancer patients. Finally, Chapter 4 builds upon the single-cell spatial algorithms utilized in Chapters 2 and 3 and leverages supervised ML approaches to perform a deep quantitative assessment of T cell phenotype, function, and spatial organization in the PDAC TME with and without treatment with anti-CD40

therapy. This study reveals important aspects of T cell biology impacted by anti-CD40 therapy in the PDAC TME, and it identifies novel TME immunological properties associated with improved disease-free survival following anti-CD40 therapy. Overall, the research in this dissertation provides a framework for quantitatively analyzing complex single-cell spatial proteomics datasets, and, importantly, the results highlight various immunological aspects of the TME which may be leveraged in future studies to improve clinical outcome for cancer patients.

Chapter 2: Quantifying the single-cell spatial landscape of head and neck cancer

This chapter has been formatted for inclusion in this dissertation from the manuscript "Single-cell spatial architectures associated with clinical outcome in head and neck squamous cell carcinoma" by Katie E. Blise, Shamilene Sivagnanam, Grace L. Banik, Lisa M. Coussens, and Jeremy Goecks, published in *npj Precision Oncology* (2022) [147]. The author of this dissertation is the primary author of this manuscript and used single-cell gated data generated by LMC's laboratory to conduct computational experiments to generate all figures in this manuscript except Main Figures 2.1a, 2.3a, and 2.4e and Supplementary Figures 2.1a, 2.1b, and 2.3a.

2.1 Abstract

There is increasing evidence that the spatial organization of cells within the tumor-immune microenvironment (TiME) of solid tumors influences survival and response to therapy in numerous cancer types. Here, we report results and demonstrate the applicability of quantitative single-cell spatial proteomics analyses in the TiME of primary and recurrent human papillomavirus (HPV)-negative head and neck squamous cell carcinoma (HNSCC) tumors. Single-cell compositions of a nine patient, primary and recurrent (n=18), HNSCC cohort is presented, followed by deeper investigation into the spatial architecture of the TiME and its relationship with clinical variables and progression free survival (PFS). Multiple spatial algorithms were used to quantify the spatial landscapes of immune cells within TiMEs and demonstrate that neoplastic tumor-immune cell spatial compartmentalization, rather than mixing, is associated with longer PFS. Mesenchymal (αSMA^{+}) cellular neighborhoods describe distinct immune landscapes associated with neoplastic tumor-immune compartmentalization and improved patient outcomes. Results from this investigation are concordant with studies in other tumor types, suggesting that trends in TiME cellular heterogeneity and spatial organization may be shared across cancers and may provide prognostic value in multiple cancer types.

2.2 Introduction

Tumor microenvironments, comprising both neoplastic tumor cells and recruited stromal cells of various lineages, including a diverse assemblage of immune, mesenchymal, and vascular cells, play a key role in both de novo progression of tumors and regulating response to therapies [6, 8, 33]. Numerous studies have reported that, in addition to the types and quantities of cells present in the tumor immune-microenvironment (TiME), the spatial organization of the TiME is prognostic for survival and response to therapy in multiple cancer types [65-68, 139-141, 148-150]. Metrics that quantify this spatial organization can range from simple density ratios within specific tumor regions [151], such as the Immunoscore [63], a now commonly used biomarker for colorectal tumor staging, to more complex measures that account for the precise locations of specific cells relative to other cells, such as mixing scores [139] and cellular neighborhood measures [140]. These more advanced spatial quantifications are a result of emerging single-cell multiplex tissue imaging modalities [68, 121, 122, 139, 140, 152], which provide detailed phenotypic and effector proteomic markers for each cell, while maintaining the spatial architecture of the tissue assayed. Knowing the precise locations of cells in the TiME enables a deeper understanding of how cells interact within the tumor, as both direct and indirect cell signaling mechanisms require cells to be near, if not directly adjacent to one another [5]. This understanding can aid treatment decisions, as many therapies require spatial proximity of specific cell types for efficacy [66]. Given that single-cell imaging technologies are still relatively new, there is much to be discovered regarding how the spatial organization of cells within the TiME relates to clinical outcome and may be used for patient stratification decisions for therapy.

Head and neck squamous cell carcinoma (HNSCC) is the sixth leading form of cancer worldwide [153], and it accounts for more than 10,000 deaths per year in the US alone [154]. While patients harboring human papillomavirus (HPV) within neoplastic cells tend to exhibit a better prognosis, their HPV-negative [HPV(-)] counterparts typically exhibit T cell suppressive TiMEs and have a significantly greater risk of recurrence and shorter 3-year survival [155-157]. There is a critical need to improve understanding of HNSCC TiMEs to enable better patient stratification for therapy, as well as identify new targets that could be leveraged for therapeutic intervention to improve outcome, particularly for patients with HPV(-) tumors who currently lack promising therapeutic options. We previously developed a multiplex immunohistochemistry (mIHC) imaging platform to aid studies investigating the immune contexture of solid tumors and their response to therapies at the single-cell level [121, 122]. Using a sequential antibody staining protocol, detection of 12-30 proteins can be enumerated at single-cell resolution across a single formalin-fixed paraffin-embedded (FFPE) tissue section. This enables single-cell phenotyping of discrete leukocyte lineages, and importantly, reveals their spatial relationships with other cells in the tissue section. Utilizing this mIHC approach on a small cohort of eighteen HPV(-) primary and matched recurrent HNSCC tumor samples collected from nine patients, we previously reported immune contextures associated with disease recurrence, most notably that myeloid inflamed profiles in primary tumors exhibited shorter progression free survival (PFS) compared to lymphoid inflamed profiles [121, 122].

In this study, we have significantly extended our prior analysis of this cohort, focusing on tumor heterogeneity and compositional changes from primary to recurrent tumors, in addition to using multiple spatial algorithms to quantify the spatial organization of the TiMEs. We then correlated these spatial features with PFS and identified TiME architectures that may be

important for therapeutic decision making. Overall, we found increased neoplastic tumor-immune cell spatial compartmentalization in primary tumors to be associated with longer PFS. These tumors also contained alpha smooth muscle actin (αSMA^{+}) cells with more organized structure located near T and B cells, as well as near cells involved in antigen presentation. Our results are concordant with those from other studies, indicating that the features identified herein are likely shared and prognostic across cancer types.

2.3 Results

One to three regions of 2500^2 μm^2 from each of the nine patients primary and matched recurrent tumor resections (n=18) were analyzed, for a total of 47 regions (**Table 2.1**, **Figure 2.1a**). For this study, we utilized a gating strategy with thirteen lineage or functional protein biomarkers to classify cells as neoplastic tumor cells, stromal cells (mesenchymal), or one of seven different leukocyte subtypes spanning lymphoid and myeloid lineages (**Table 2.2, Supplementary Figure 2.1a**). We investigated tumor heterogeneity both within and across patient samples, quantified the cellular spatial relationships within the TiME using a mixing score and performed a neighborhood clustering method to describe the association between TiME spatial architecture, clinical features, and PFS.

Table 2.1: Patient and tumor characteristics.

Patient ID	Anatomic Site of Resection	Primary Tumor TNM Stage	Therapy Following Primary Resection	Gender	Race	Alcohol History	Smoking History	HPV Status	Progression Free Survival (days)
1	Oral Cavity	2	Cisplatin + Radiation	Male	White	Yes	Yes	Negative	804
2	Oral Cavity	4	Cetuximab + Radiation	Male	White	Yes	Yes	Negative	123
3	Oral Cavity	1	Cisplatin + Radiation	Female	White	Yes	No	Negative	1447
4	Oral Cavity	1	Cisplatin + Radiation	Female	White	No	No	Negative	246

5	Oropharynx	4	Cetuximab + Radiation	Female	White	Yes	No	Negative	202
6	Oropharynx	4	Cisplatin + Radiation	Female	Asian	No	No	Negative	188
7	Larynx	4	Cisplatin + Cetuximab + Radiation	Female	White	Yes	Yes	Negative	409
8	Larynx	3	Cisplatin + Radiation	Female	White	Yes	Yes	Negative	1033
9	Larynx	4	Cisplatin + Radiation	Female	White	Yes	Yes	Negative	83

Table 2.2: mIHC cell phenotype classification.

Cell Phenotype	**Antibody Markers**
$CD8^+$ T Cell	$CD45^+$ $CD20^-$ $CD3^+$ $CD8^+$
$CD4^+$ T Helper	$CD45^+$ $CD20^-$ $CD3^+$ $CD8^-$ $FOXP3^-$
$CD4^+$ Regulatory T Cell	$CD45^+$ $CD20^-$ $CD3^+$ $CD8^-$ $FOXP3^+$
B Cell	$CD45^+$ $CD20^+$
Macrophage	$CD45^+$ $CD20^-$ $CD3^-$ $CD66B^-$ $CD68^+$
Granulocyte	$CD45^+$ $CD20^-$ $CD3^-$ $CD66B^+$
Antigen Presenting Cell	$CD45^+$ $CD20^-$ $CD3^-$ $CD66B^-$ $CD68^-$ $MHCII^+$
Other Immune	$CD45^+$ $CD20^-$ $CD3^-$ $CD66B^-$ $CD68^-$ $MHCII^-$ $CD8^-$ $FOXP3^-$
αSMA^+ Mesenchymal	$CD45^-$ $PANCK^-$ αSMA^+
Neoplastic Tumor	$CD45^-$ $PANCK^+$
Functional Markers	PD-1, PD-L1, Ki-67

2.3.1 Single-cell proteomic analyses reveal varying degrees of tumor heterogeneity

To quantify how cellular composition varied across tumor regions, we assessed tumor heterogeneity at multiple levels, including intra-tumoral, intra-patient, and inter-patient cellular heterogeneity by calculating Kullback-Leibler (KL) divergences for each region, performing hierarchical clustering, and conducting a principal component analysis (PCA). The density of

each cell type per region was measured for all eighteen tumor specimens by taking the count of each cell type divided by the measured tissue area in mm^2 (**Figure 2.1b,c**). We then calculated the coefficient of variation per cell type for each tumor, and averaged these values to quantitatively describe the cell types contributing most to intra-tumoral heterogeneity within the cohort. The coefficient of variation is defined as the standard deviation divided by the mean, and it provides a normalized measure of variability for comparison across cell types with large differences in densities. On average, B cells exhibited the greatest coefficient of variation across the cohort relative to other cell types (**Table 2.3**). This is likely due to the fact that B cells were frequently observed to be spatially clustered together, resulting in regions of either high B cell density or low B cell density despite being collected from the same tumor (**Supplementary Figure 2.1b**).

Table 2.3: Coefficient of variation.

Cell Phenotype	Average Coefficient of Variation
B Cell	0.658
Macrophage	0.563
αSMA^+ Mesenchymal	0.545
Neoplastic Tumor	0.541
Granulocyte	0.523
Antigen Presenting Cell	0.519
Other Immune	0.402
$CD8^+$ T Cell	0.384
$CD4^+$ T Helper	0.376
$CD4^+$ Regulatory T Cell	0.311

To further quantify and assess tumor heterogeneity both within and across patients, we calculated the KL divergence of each tumor region from five average cell type distributions. KL

divergence is a relative measure of how similar two distributions are, with larger values reflecting less similarity between the distributions and smaller values reflecting more similarity between the distributions. This measure has been used previously to quantify tumor heterogeneity [68]. By calculating and comparing the divergences of each tumor region from multiple average cell type distributions, we were able to assess heterogeneity within and across tumors and patients. Overall, we observed that heterogeneity was lower across regions from the same tumor and tumors from the same patient (primary or recurrent), while higher across tumors from different patients. This is evidenced by smaller intra-tumoral and intra-patient KL divergence values for the majority of tumor regions (**Figure 2.1d**).

The cellular distributions used to calculate the five KL divergence values per tumor region were (1) the average cellular distribution across all regions sampled from the same tumor ["Intra-Tumor (P or R only)"]; (2) the average cellular distribution across the patient's primary and recurrent tumors ["Intra-Patient (P and R)"]; (3) the average cellular distribution across all tumors in the cohort collected from the same timepoint ["Inter-Patient (P or R only)"]; (4) the average cellular distribution across all tumors in the cohort resected from the same anatomic site ["Inter-Patient (Same Anatomic Site)"], and; (5) the average cellular distribution across all tumors collected from all patients in the cohort, regardless of primary or recurrent status or anatomic site ["Inter-Patient (all)"]. By comparing the relative KL divergence values to each other, we found tumor regions to be more similar to regions sampled from the same tumor and patient than regions collected from tumors of other patients. Notably, we found no significant difference between "Inter-Patient (Same Anatomic Site)" and "Inter-Patient (all)," indicating that tumor regions diverged by the same degree from regions sampled at the same anatomic site as they did from regions sampled at all three anatomic sites (oral cavity, oropharynx, larynx) of the

head and neck region (**Figure 2.1d**). Finally, given the large proportion of neoplastic tumor cells comprising the TiME for many of the tumor regions, we assessed KL divergence using only the distribution of immune cells present and found similar results (**Supplementary Figure 2.1c**). This indicates that immune cell composition is more similar within regions from the same patient than across regions collected from different patients.

To further investigate intra-patient heterogeneity, we performed unsupervised hierarchical clustering on the 47 tumor regions based on their normalized density composition (**Figure 2.1e**). We found that two patients (3, green; 7, pink) contained all tumor regions clustering together, independent of primary or recurrent state. These patients also had the smallest intra-patient KL divergence values (**Figure 2.1d**), indicating that the cell densities of these patients' primary and recurrent tumors were similar to each other. Three patients (2, orange; 4, red; 9, yellow) contained nearly all regions clustered together. The remaining four patients' tumors exhibited greater degrees of intra-patient heterogeneity, as demonstrated by the distance between primary and recurrent tumor regions on the clustered heatmap (**Figure 2.1e**). Overall, we found that regions evaluated from the same patient tended to cluster together more than regions evaluated from different patients (**Figure 2.1e**), indicating increased heterogeneity between patients as compared to between samples from the same patient. We also examined whether tumor regions clustered by the anatomic resection site and found that the clusters formed did not group by site. These results, in addition to those of our KL divergence analyses, indicate that anatomic site was likely not the main contributor of cellular heterogeneity in this cohort. PCA results also supported these observations (**Figure 2.1f,g**).

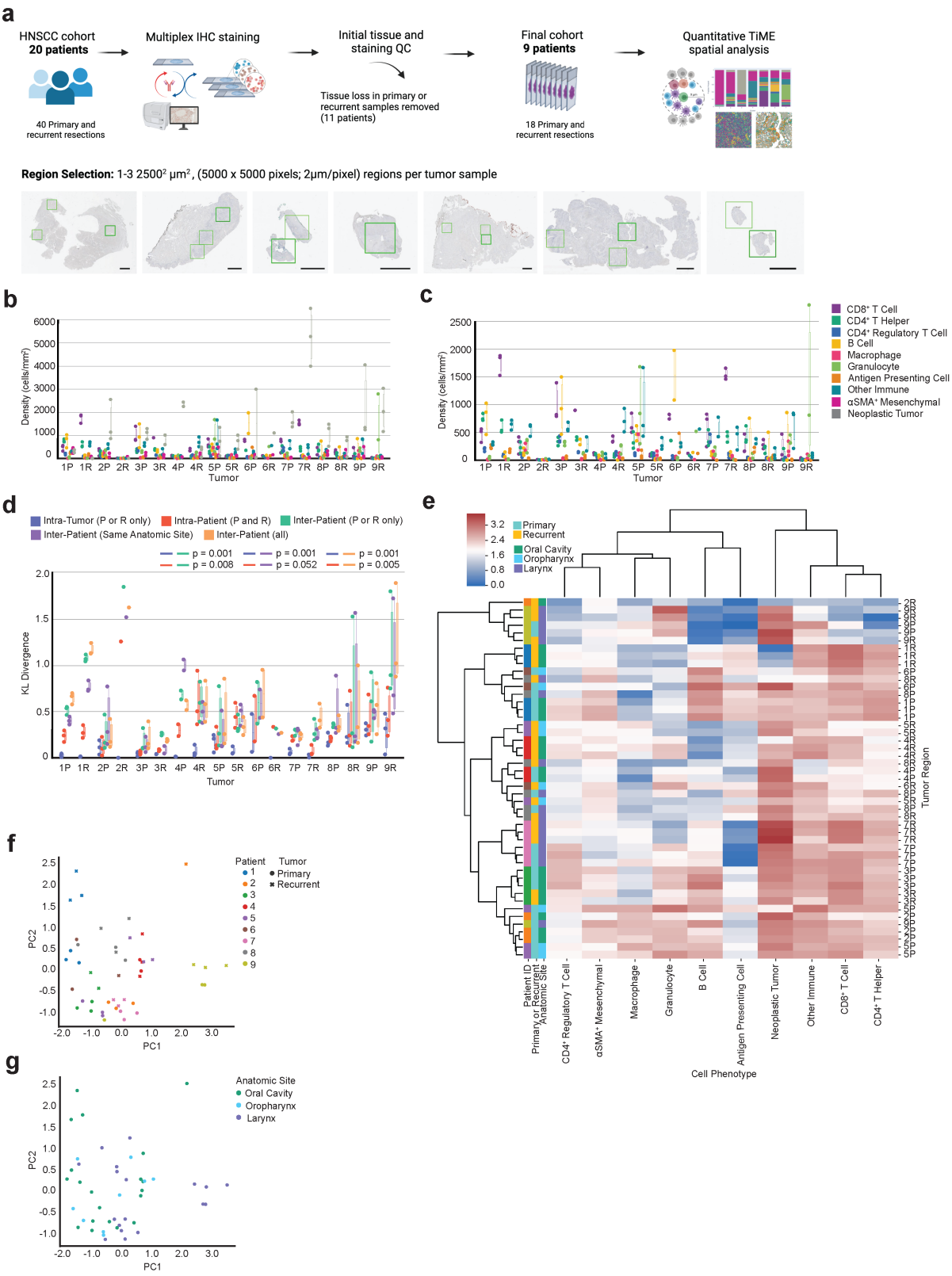


Figure 2.1: Heterogeneity across patients and tumor regions. a, Cohort and tissue region selection overview. One to three regions of 2500^2 µm^2 were assayed with mIHC per eighteen tumor resections and are represented by the green boxes in the tissue images. Black scale bar = 2500 µm. **b**, Density (cells/mm^2) of each cell

type present per individual primary (P) and recurrent (R) tumor. Each dot represents a single tumor region (n=47). **c**, Density (cells/mm^2) of each immune cell type present per individual primary (P) and recurrent (R) tumor. Each dot represents a single tumor region (n=47). **d**, Box plot of the Kullback-Leibler divergences from a single tumor region's cellular distribution compared to: the tumor's average cellular distribution [Intra-Tumor (P or R only)], the patient's average cellular distribution [Intra-Patient (P and R)], the cohort's average cellular distribution across tumors of the same timepoint [Inter-Patient (P or R only)], the cohort's average cellular distribution across tumors of the same anatomic site [Inter-Patient (Same Anatomic Site)], the cohort's average cellular distribution across all tumors from all patients [Inter-Patient (all)]. P-values calculated using a one-way ANOVA multi-group significance test followed by a Tukey honestly significant difference post-hoc test. **e**, Heatmap of cellular composition across tumor regions. Rows are individual tumor regions that are ordered by the hierarchical clustering of their cellular composition. Columns are the cell types used as clustering features. Composition was normalized using a log10+1 transformation before clustering. Leftmost three columns are color coded by patient, tumor timepoint, and anatomic site. **f**, Principal component analysis on cellular density following a log10+1 transformation. Each point (n=47) represents one tumor region and is colored by patient. The shape of each point denotes primary or recurrent status. **g**, Principal component analysis on cellular density following a log10+1 transformation. Each dot (n=47) represents one tumor region and is colored by anatomic resection site.

2.3.2 TiME cellular composition altered by therapy

Multiple studies have reported differences in TiME cellular makeup [158, 159] and tumor clonal diversity [160] between primary and recurrent tumors. To assess whether any immune contexture changes occurred following post-operative therapy in our cohort, we analyzed the cellular composition of primary tumors as compared to their recurrent tumors. All patients received a combination therapy of cisplatin and/or cetuximab accompanied by radiation following surgical resections of their primary tumors. We used the average density of each cell type present across regions for a given tumor and compared primary tumor composition to their matched recurrent tumor composition. While we did not observe any significant differences in cell density between primary and recurrent tumors ($p>0.112$), we did find that all patients experienced a decrease in the density of B cells from their primary to recurrent tumors (**Figure 2.2a,b**). This result is supported by a recent study that found that a large cohort of HNSCC patients experienced a decline in B cells from primary to recurrent tumors, and this reduction was dependent upon patients receiving adjuvant chemoradiotherapy [161]. We found no common differences across all patients for any other cell type, although eight of the nine patients saw a

decrease in the density of neoplastic tumor cells from their primary to recurrent tumors (**Figure 2.2b**).

We then assessed whether patients exhibited similar compositional differences in primary and recurrent TiMEs by performing unsupervised hierarchical clustering on the normalized average difference in cellular composition for each cell type. This resulted in two groups of patients (**Figure 2.2c**). Interestingly, the two patients that received cetuximab (2, orange; 5, purple), rather than cisplatin, clustered together within one of these groups and were the only two patients to experience a decrease in the density of every cell type following therapy. Across the cohort these patients had the greatest decrease in the density of macrophages, granulocytes, and other $CD45^+$ immune cells present from their primary tumors to their recurrent tumors following therapy (**Figure 2.2b,c)**. Interestingly, the one patient who received both cisplatin and cetuximab (7, pink) was present in the other cluster from the two patients who received only cetuximab. This was the only patient to experience an increase in the density of neoplastic tumor cells (**Figure 2.2b,c**). This patient also experienced the second largest increase in $CD8^+$ T cells as well as the greatest decrease in $CD4^+$ regulatory T cells, potentially indicating a pro-inflammatory response to—or despite—increased neoplastic tumor cell density. Altogether, these results indicate that shared trends in TiME composition changes from primary to recurrent tumors specific to therapy exist, and regardless of therapy, all patients exhibited a decrease in B cells from primary to recurrent tumors.

Approximately half (n=4) of the patients in the cohort experienced an increase in $CD8^+$ T cell density while the other half (n=5) experienced a decrease in $CD8^+$ T cell density following therapy. This was the only cell type that increased in density for nearly half of the cohort and decreased for the other half. To determine if there was a survival advantage for patients that

experienced this increase, we split our cohort into short-term or long-term survivor groups using median PFS and observed that all patients who experienced an increase in $CD8^+$ T cell density from their primary to recurrent tumors were long-term survivors (**Figure 2.2d**). Interestingly, the density of $CD8^+$ T cells in the primary tumor alone did not associate with PFS ($p=0.829$). Prior research has revealed that increased $CD8^+$ T cell abundance in the primary tumor is associated with better outcomes in HNSCC [162-166]. However, these studies largely included HPV-positive HNSCCs, which is more often associated with greater densities of $CD8^+$ T cells and improved survival [155-157], thus unsurprising that our results differ. However, our results are concordant with a recent study in HNSCC that reported longer survival was associated with patients who had experienced an increase in $CD8^+$ tumor-infiltrating lymphocytes from their primary to recurrent tumor [167]. Another study in HNSCC found a similar trend between increased $CD8^+$ T cell infiltration, longer survival, presence of specific neoantigens, and increased cytolytic activity in recurrent tumors [168]. Notably, the four patients in our cohort that experienced the greatest decrease in $CD8^+$ T cell density in recurrence had TNM stage 4 primary tumors, while patients that experienced an increase in $CD8^+$ T cell density in recurrence included TNM stages 1, 2, 3, and 4.

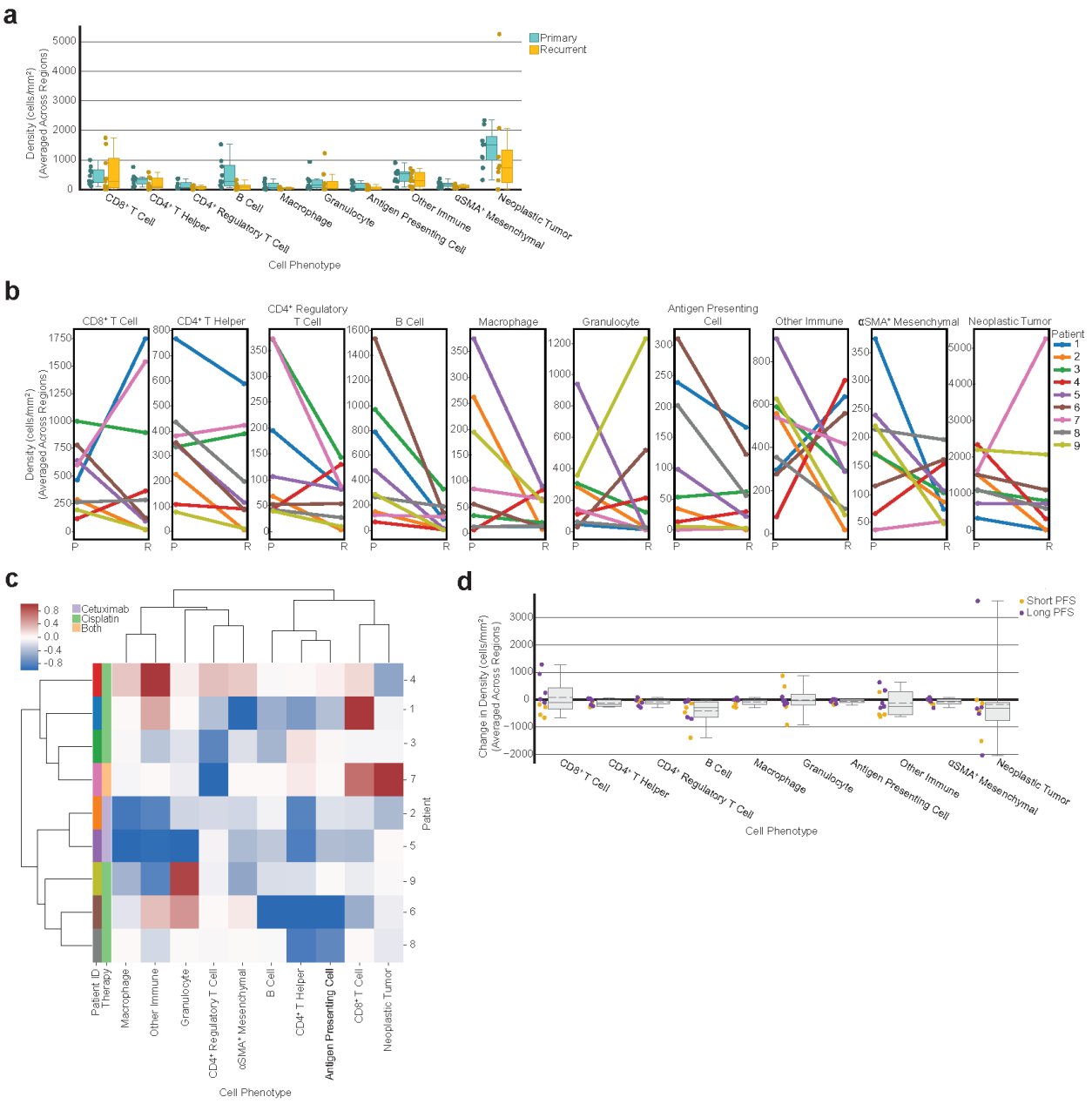


Figure 2.2: Tumor cellular composition changes following therapy. a, Box plot showing the average density of each cell type split by primary and recurrent status. Each dot represents the density of that cell type for one tumor, averaged across regions (n=9 primary tumors, n=9 recurrent tumors). Boxes = quartile 1 (Q1) to quartile 3 (Q3); whiskers = smallest and largest datapoints within 1.5*interquartile range (IQR) +/- Q3/Q1; solid line = median. **b**, Line plots showing the change in cellular composition from primary to recurrent tumors for each cell type. Each dot represents the density of that cell type for one tumor, averaged across regions. Lines are colored by patient. **c**, Heatmap of average change in cellular composition from primary to recurrent tumors for each patient. Rows are individual patients that are ordered by the hierarchical clustering of their change in TiME cellular composition (averaged across regions). Columns are the cell types used as clustering features. Compositional change was normalized [-1,1] before clustering (see Methods). Leftmost two columns are color coded by patient followed by therapy administered. **d**, Box plot showing the average change in density of each cell type for each patient (n=9) colored by short-term or long-term progression free survival, splitting on median progression free survival. Boxes = Q1 to Q3; whiskers = smallest and largest datapoints within 1.5*IQR +/- Q3/Q1; solid line = median; dotted line = mean.

2.3.3 Quantifying the spatial organization of neoplastic and immune cells

Prior studies reported the abundance of various cell types, including $CD8^+$ T cells [166], $CD4^+$ regulatory T cells [169], and macrophages [170] to be associated with survival in HNSCC. We analyzed the average density of each cell type across primary tumors for their correlations with PFS, but found no significant association with PFS for any single cell type ($p>0.159$). Given the prognostic potential of TiME cellular spatial organization as has been reported in other cancer types [65-68, 140, 141, 148-150], we quantified the spatial organization of cells within tumor regions and examined the association of the spatial features with clinical outcome.

We first deployed a mixing score, used previously to analyze immune cell spatial compartmentalization in triple negative breast cancers [139]. The mixing score measures the enrichment of neoplastic tumor-immune cell proximity relative to immune-immune cell proximity within a set distance. We quantified the number of immune and neoplastic tumor cells within 15 μm of each other, divided by the number of immune cells within 15 μm from another immune cell. Each region was labeled as mixed or compartmentalized using the median mixing score value for all primary tumors as the threshold (**Figure 2.3a,b**; see Methods). This threshold classified tumor regions as mixed if at least one neoplastic tumor cell was within 15 μm from an immune cell for approximately every ten immune cells within 15 μm from another immune cell. Tumor regions were considered compartmentalized when this ratio was smaller. Regions with fewer than 250 $CD45^+$ immune cells per 800^2 μm^2 present were labeled as cold, utilizing the same immune cell density threshold from the original study [139]. Of the 47 total tumor regions, 25 were mixed, 20 were compartmentalized, and two were cold.

After calculating the mixing score for each region, we examined the spatial heterogeneity of our cohort. Our tumor compositional heterogeneity analyses revealed that intra-tumoral and intra-patient heterogeneity was less than inter-patient heterogeneity. To determine whether this observation held for spatial organization heterogeneity, we compared each region's mixing score to five groups of average mixing scores, which were computed from the same five groups as the analysis in Figure 1d: Intra-Tumor (P or R only), Intra-Patient (P and R), Inter-Patient (P or R only), Inter-Patient (Same Anatomic Site), and Inter-Patient (all). Contrary to our analysis in Figure 1d, the mixing score is only one feature, not a distribution of features, thus we used the difference in mixing scores, rather than the KL divergence. In addition, we calculated the absolute values of these differences as a way to normalize the data in order to capture the degree of difference in spatial organization, allowing us to subsequently test for differences across the five levels of heterogeneity. We found there to be less intra-tumoral heterogeneity than inter-patient heterogeneity (**Supplementary Figure 2.2a**). This result indicates that, in terms of neoplastic tumor-immune cell mixing, tumor regions resemble regions sampled from the same tumor more than regions sampled from tumors of other patients.

We then considered whether tumor regions sampled from the same anatomic site differed in their spatial organization and found that of the three anatomic sites, tumor regions from the oral cavity contained significantly different average mixing scores than tumor regions from the larynx ($p=0.016$, **Figure 2.3c**). No significant differences were found in average mixing scores between the oral cavity and the oropharynx or the larynx and the oropharynx. Regions from larynx tumors did exhibit a greater range of mixing scores than oral cavity or oropharynx (**Figure 2.3c**) indicating that larynx tumors exhibit greater spatial heterogeneity in terms of

neoplastic tumor-immune cell proximity. We found no significant difference in the mixing scores of primary versus recurrent tumors (Wilcoxon signed rank test, p=0.441).

2.3.4 Spatial compartmentalization associated with longer progression free survival

To investigate how spatial mixing correlated with patient outcome, we averaged the mixing scores across regions and assigned a final mixing score and spatial label for each tumor. Patients with more compartmentalization between neoplastic cells and immune cells in their primary tumors exhibited significantly longer PFS than those with greater mixing between these cell types (p=0.032, **Figure 2.3d**). We then examined how the average mixing score of the tumors related to the TNM stage and anatomic site of each primary tumor. Four of the five mixed primary tumors were TNM stage 4, and one was TNM stage 1. The four compartmentalized primary tumors were TNM stages 1, 2, 3, and 4. All anatomic sites were present in both mixed and compartmentalized spatial architecture groups (mixed: 2 oral cavity, 1 oropharynx, 2 larynx; compartmentalized: 2 oral cavity, 1 oropharynx, 1 larynx), indicating no single anatomic site had predominantly mixed or compartmentalized spatial architecture.

2.3.5 Spatial architecture associated with cellular composition

We next explored how cellular composition related to spatial organization, in an effort to explain the association found between mixing score and PFS. By coloring the initial PCA on TiME composition by mixing score, we found that tumor regions from the same mixing group clustered together (**Figure 2.3e**). Due to this association between cellular composition and spatial organization, we wondered whether certain cell types would be more frequent in mixed or compartmentalized tumors. Given how the mixing score is calculated, we hypothesized that compartmentalized tumor regions would have greater densities of immune cells than mixed

regions, while mixed tumor regions would have greater densities of neoplastic tumor cells than compartmentalized regions, and found this to be the case (**Figure 2.3f**). Namely, compartmentalized tumor regions had greater densities of $CD4^+$ T helper cells (p=0.016), B cells (p=0.048), antigen presenting cells (APCs) (p=0.037), and other $CD45^+$ immune cells (p=0.048) than mixed tumor regions. Given the role of $CD4^+$ T helper cells, B cells, and other $MHCII^+$ immune cells in antigen presentation, these results could indicate enhanced antigen presentation in compartmentalized tumor regions as compared to mixed regions. Conversely, mixed tumor regions contained greater densities of neoplastic tumor cells than compartmentalized regions (p<0.001).

We further examined associations between functional phenotypes of leukocytes and mixing scores for each tumor region. Mixed regions contained a greater proportion of lymphocytes, including $CD8^+$ T cells (p=0.023), $CD4^+$ T helper cells (p=0.017), and B cells (p=0.016), expressing the immunoregulatory protein programmed death ligand (PD)-1 than compartmentalized regions (**Figure 2.3g**). PD-1 is recognized as an indicator of T cell antigen experience, whereas its expression on B cells has been reported to suppress T cell effector function [171, 172], indicating a suppressive and potentially dysfunctional immune environment in mixed tumors. On the contrary, compartmentalized regions contained a greater proportion of APCs expressing the proliferation marker Ki-67 than mixed regions (p=0.032, **Figure 2.3h**), supporting the notion that antigen presentation is a key feature of compartmentalized regions. We performed a bootstrapping analysis to confirm the robustness of these results and demonstrate that no one tumor was biasing the functional marker results (**Supplementary Figure 2.2b,c,d,e**).

In addition to identifying differences between spatial architectures and their respective immune and neoplastic tumor cell densities and functional marker expressions, we found that

tumor regions with more compartmentalization between neoplastic cells and immune cells also contained greater densities of αSMA^+ mesenchymal cells, as compared to those with higher mixing ($p=0.002$, **Figure 2.3f**). This result was intriguing because these cells were not included when computing the mixing score, yet there is a clear association between αSMA^+ cell density and the tumor's spatial organization (**Figure 2.3i**).

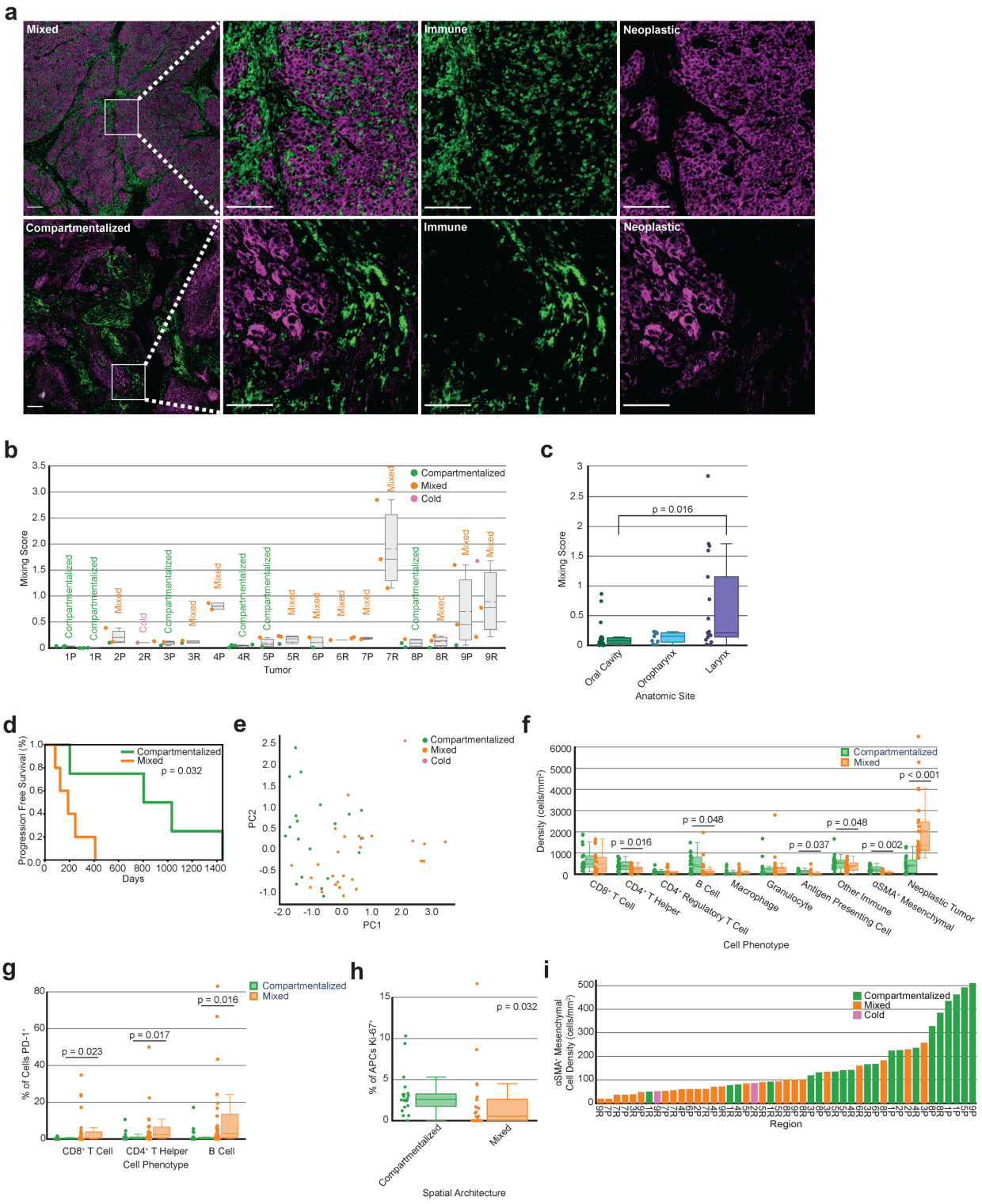


Figure 2.3: Mixing score quantifies the spatial organization of tumors. a, mIHC images of a representative mixed tumor region (top) versus a compartmentalized tumor region (bottom). Leftmost panel shows tumor regions with neoplastic tumor cells (purple) and CD45^{+} immune cells (green); white scale bar = 200 µm. Remaining panels show zoomed in areas of mixing (top) and compartmentalization (bottom), first with both cell populations present and then separated; white scale bar = 100 µm. **b**, Box plot showing the mixing scores across all

primary and recurrent tumors (n=18). Each dot (n=47) represents the mixing score for one tumor region and is colored according to its spatial architecture. The average spatial architecture designation for the overall tumor is printed above each box. Boxes = Q1 to Q3; whiskers = smallest and largest datapoints within 1.5*IQR +/- Q3/Q1; solid line = median; dotted line = mean. **c**, Box plot showing the mixing score of each tumor region split by the anatomic site of its resection. Boxes = Q1 to Q3; whiskers = smallest and largest datapoints within 1.5*IQR +/- Q3/Q1; solid line = median. P-value calculated using a one-way ANOVA multi-group significance test followed by a Tukey honestly significant difference post-hoc test. **d**, Kaplan-Meier curve of progression free survival for patients split by the mixing score of their primary tumors. Patients were split on the median value. P-value calculated using the log-rank test. **e**, Principal component analysis on cellular density following a log10+1 transformation. Each dot (n=47) represents one tumor region and is colored according to the region's spatial architecture. **f**, Box plot showing the density of each cell type split by the tumor region's spatial architecture. Each dot represents the density of that cell type for one region (n=47 per cell type). Boxes = Q1 to Q3; whiskers = smallest and largest datapoints within 1.5*IQR +/- Q3/Q1; solid line = median; dotted line = mean. Statistical significance calculated using independent one-tailed t-tests for cell types whose differences follow a normal distribution and non-parametric one-tailed t-tests (Mann-Whitney U Test) for cell types whose differences do not follow a normal distribution. P-values were corrected using the Benjamini-Hochberg procedure. **g**, PD-1 expression on $CD8^+$ T cells, $CD4^+$ T helper cells, and B cells by spatial architecture. Each dot represents the percentage of each cell type positive for PD-1 for a single tumor region (n=45, excluding cold regions). Boxes = Q1 to Q3; whiskers = smallest and largest datapoints within 1.5*IQR +/- Q3/Q1; solid line = median. P-values calculated using a one-tailed Mann-Whitney U test and corrected using the Benjamini-Hochberg procedure. **h**, Ki-67 expression on APCs by spatial architecture. Each dot represents the percentage of cells positive for Ki-67 for a single tumor region (n=45, excluding cold regions). Boxes = Q1 to Q3; whiskers = smallest and largest datapoints within 1.5*IQR +/- Q3/Q1; solid line = median. P-value calculated using a Mann-Whitney U test and corrected across all cell types using the Benjamini-Hochberg procedure. **i**, Bar chart showing the density of αSMA^+ mesenchymal cells present per tumor region. Bars are ordered by αSMA^+ cell density and are colored according to the region's spatial architecture.

2.3.6 αSMA^+ mesenchymal spatial cellular neighborhoods reveal spatial landscapes associated with progression free survival advantage

Given the relationship between mixing score and αSMA^+ mesenchymal cell density, we deployed a cellular neighborhood clustering analysis to identify which cell types were spatially proximal to αSMA^+ cells across tumors in order to gain a better understanding of whether these cells were contributing to the neoplastic tumor-immune spatial compartmentalization observed. This analysis first involved calculating neighborhoods, which were defined as physical groupings of cells within a set distance threshold from a seed cell (**Figure 2.4a**). Each cell within the distance threshold was deemed a neighbor of the seed cell, contributing to that seed cell's neighborhood's composition. After identifying neighborhoods for each seed cell present across all tumor regions, neighborhoods were grouped with K-means clustering, using the fraction of

each cell type present in the neighborhoods as the clustering features. This revealed clusters of αSMA$^+$ cell neighborhoods with similar cellular makeups across all tumor regions.

We applied a neighborhood clustering analysis with αSMA$^+$ cells as the seed cells with a distance threshold of 30 μm, as this produced neighborhoods with an average of approximately ten neighbor cells. A recent study involving cellular neighborhood analyses employed a method that selected the ten nearest spatial neighbors of the seed cell, regardless of the distance between the seed cell and its neighbors [140]. By setting a distance threshold of 30 μm, our method required the cells be close, if not directly touching, while still capturing enough neighbors to cluster on. Any αSMA$^+$ cell that did not have any neighbors was removed from downstream clustering analyses. Clustering results yielded seven groups, each different in their average composition of αSMA$^+$ neighborhoods (**Figure 2.4b**). Clusters 1 and 2 contained mostly other αSMA$^+$ cells comprising the neighborhoods; in fact, cluster 1 was almost exclusively made up of αSMA$^+$ cells. Cluster 3 contained the greatest proportion of neoplastic tumor cell neighbors. Clusters 4, 5, 6, and 7 were all comprised of roughly 75% immune cells as neighbors, although they differed in the types of immune cells present. Cluster 4 was defined by a majority of other CD45$^+$ immune cells not explicitly defined within our gating strategy. To elucidate marker expression within these other CD45$^+$ cells, we performed a post-hoc t-distributed stochastic neighbor embedding (t-SNE) analysis on cells classified as 'other immune,' and found them to contain a large population of CD163$^+$ cells and a smaller population of mast cells (tryptase$^+$) (**Supplementary Figure 2.3a**). Cluster 5 contained primarily CD8$^+$ T cells as the dominant immune population. Cluster 6 consisted of CD4$^+$ T helper cells, B cells, and the greatest proportion of APCs of any cluster. Finally, cluster 7 was defined by its large proportion of granulocytes.

We confirmed that no single tumor region, entire tumor, or patient dominated any of the αSMA$^+$ neighborhood clusters by examining the percent contribution of each of the seven clusters per 47 tumor regions, eighteen tumors, and nine patients. While clusters were present in varying degrees across tumor regions, our results indicated that no region (**Supplementary Figure 2.3b**), tumor (**Supplementary Figure 2.3c**), or patient (**Supplementary Figure 2.3d**) was solely responsible for giving rise to any of the clusters. We also confirmed that all seven clusters were present in tumors collected from each of the three anatomic sites (**Supplementary Figure 2.3e**), as well as each of the four TNM stages (**Supplementary Figure 2.3f**).

To identify groups of patients with primary tumors of similar αSMA$^+$ cellular neighborhoods, we performed unsupervised hierarchical clustering on the normalized average αSMA$^+$ cellular neighborhood composition across the nine primary tumors. This resulted in two groups of patients, differing in proportional compositions of αSMA$^+$ cellular neighborhoods (**Figure 2.4c**). On average, both groups had roughly 20% of their αSMA$^+$ cells assigned to cluster 2 and roughly 10% of their αSMA$^+$ cells assigned to cluster 5 (**Figure 2.4d**). However, the two groups differed in that group 1 (blue) included patients with αSMA$^+$ cells predominantly assigned to clusters 1 and 6, meaning their αSMA$^+$ cells were primarily surrounded by CD4$^+$ T helper cells, B cells, and other αSMA$^+$ cells. On the contrary, group 2 (orange) included patients with more of their αSMA$^+$ cells assigned to clusters 3, 4, and 7, meaning their αSMA$^+$ cells were mostly surrounded by neoplastic tumor cells, other immune cells, and granulocytes. Consistent with our results, visualization of tissue regions illustrates primary tumors in group 1 containing more αSMA$^+$ stromal cells overall, frequently neighboring CD4$^+$ T helper cells and B cells, with greater MHCII positivity (**Figure 2.4e**). Primary tumors in group 2 contained less structured αSMA$^+$ stromal cells, fewer neighboring immune cells, and less MHCII positivity, differences

that were strikingly apparent in the mIHC stained tissue images (**Figure 2.4e**). Despite both groups containing nearly equal proportions of $CD8^+$ T cells neighboring αSMA^+ cells, group 1 contained higher densities of $CD4^+$ T cells and increased MHCII positivity as compared to group 2 (**Figure 2.4e**).

To determine if the composition of αSMA^+ cellular neighborhood groups was correlated with clinical outcome, we performed a survival analysis on the two groups of patients. We found that patients in group 1 had significantly longer PFS than patients in group 2 ($p=0.047$, **Figure 2.4f**). Patients in group 1 had primary tumors annotated as TNM stages 1, 2, 3, and 4, while patients in group 2 had four primary tumors annotated as TNM stage 4 and one annotated as TNM stage 1. Tumors from all three anatomic sites were represented in both groups. An analysis of the proportions of αSMA^+ neighborhood clusters present in each of the nine primary tumors revealed the varying degrees to which each of the seven clusters were present in each of the tumors (**Figure 2.4g**). Notably, we found there to be a positive correlation between the presence of clusters 1 and 6 ($r = +0.69$) as well as clusters 4 and 7 ($r = +0.66$). We found negative correlations between the presence of clusters 1 and 4 ($r = -0.88$), clusters 3 and 6 ($r = -0.69$), and clusters 6 and 7 ($r = -0.65$). Finally, we found the two groups resulting from hierarchical clustering to be associated with mixing status. Specifically, group 1 consisted of 75% compartmentalized tumors, and group 2 consisted of 80% mixed tumors. Overall, these results describe interesting spatial relationships between immune cells, mesenchymal stroma, and neoplastic tumor cells, indicating increased antigen presentation and immune activity associated with compartmentalization and progression free survival.

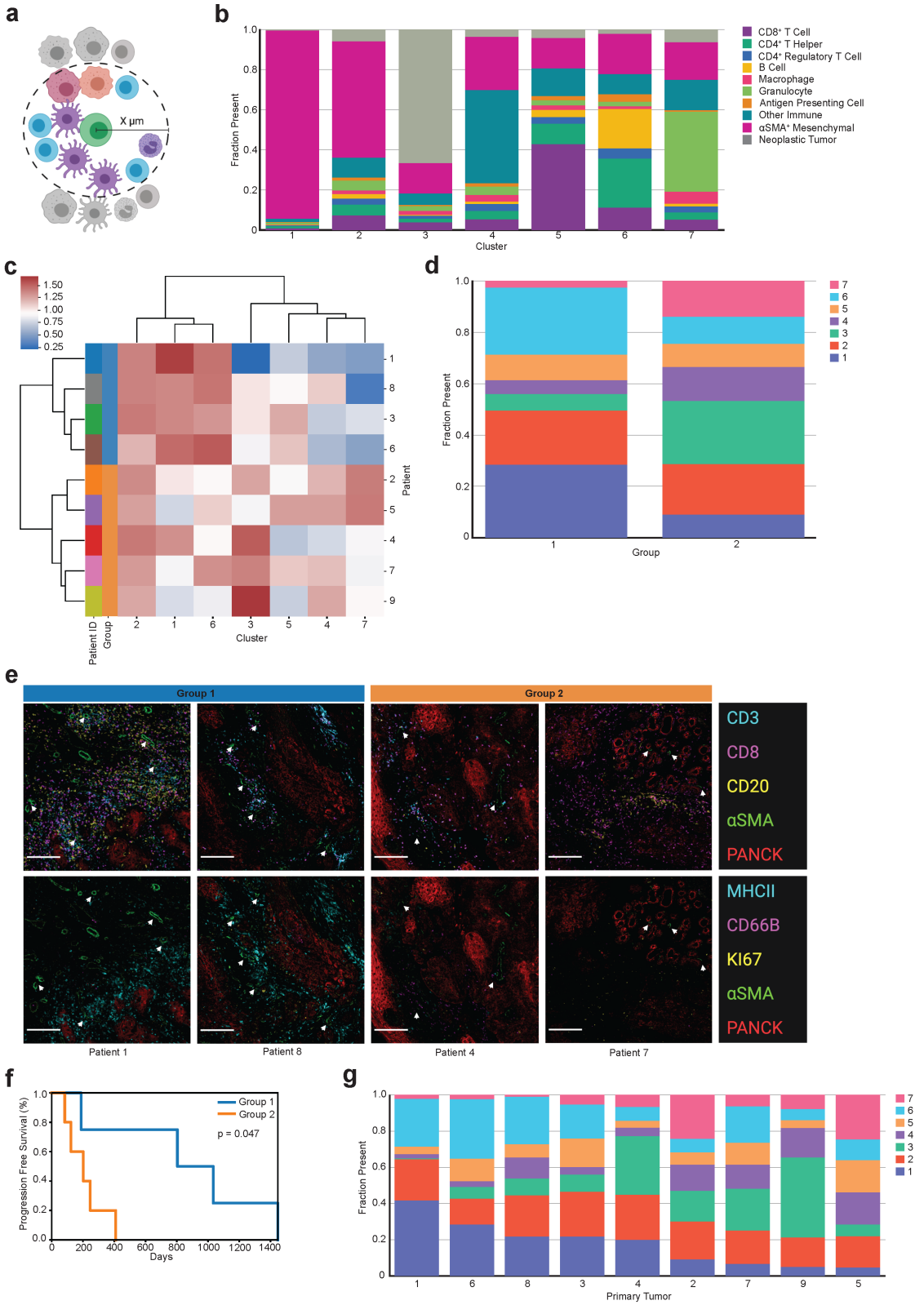


Figure 2.4: αSMA$^+$ mesenchymal cellular neighborhood clustering. a, Cellular neighborhoods were defined by drawing a circle of a specified radius around each seed cell (green) of a designated phenotype. Cells whose centers were inside the circle were considered neighbors of that seed cell. This figure was created using

BioRender.com. **b**, Stacked bar chart showing the average cellular composition of each αSMA$^+$ mesenchymal cell neighborhood cluster (n=7). Bars are colored by cell type and represent the average fraction (out of 1.0) of each cell type present in the neighborhoods belonging to each cluster. **c**, Heatmap of αSMA$^+$ cell neighborhood clusters present averaged across primary tumors. Rows are primary tumors that are ordered by the hierarchical clustering of their average of αSMA$^+$ neighborhood cluster presence. Columns are the αSMA$^+$ cell neighborhood clusters used as clustering features. Percent (out of 100) of αSMA$^+$ neighborhood clusters was normalized using a log10+1 transformation before clustering. Leftmost column is color coded by patient. **d**, Stacked bar chart showing the average proportion (out of 1.0) of αSMA$^+$ cell neighborhood clusters present in each of the two hierarchically clustered groups of primary tumors. **e**, Panel of mIHC images containing two merged pseudo-colored images of one region from four patients visually illustrates representative regions from patients in group 1 and group 2. The top panel of images are merged pseudo-colored stains containing CD3, CD8, CD20, αSMA, and PANCK. The bottom panel of images are the same regions as the top panel with merged pseudo-colored stains containing MHCII, CD66B, Ki-67, αSMA, and PANCK. White arrows point out representative αSMA$^+$ cells. White scale bar = 200 μm. **f**, Kaplan-Meier curve of progression free survival for patients grouped together in the hierarchical clustering of their primary tumor αSMA$^+$ mesenchymal cell neighborhood abundance. P-value calculated using the log-rank test. **g**, Stacked bar chart showing the average proportion (out of 1.0) of αSMA$^+$ cell neighborhood clusters present in each of the primary tumors (n=9). Tumors are ordered by descending proportions of cluster 1.

2.4 Discussion

The significant role that TiME cellular composition plays in tumor progression and response to therapy has been accepted for over a decade [173]. However, recent findings powered by single-cell proteomics imaging technologies have found that the spatial organization of the cells present in the TiME also plays a critical role [65-68, 139-141, 148-150]. TiME spatial quantifications are just beginning to provide novel insights into tumor biology, and thus it is still unclear exactly which spatial features are important in dictating response to therapy or clinical outcome, as well as whether these features are shared across cancer types, and how they could be leveraged for therapeutic decisions and patient stratification. Here, we leveraged single-cell spatial proteomics data generated by our mIHC immunoassay-based imaging platform to quantitatively assess the TiME of nine matched primary and recurrent HPV(-) HNSCCs in order to demonstrate the use of spatial features in disease prognosis. Our results on this cohort of nine patients provide insight into the heterogeneity and spatial landscape of HPV(-) HNSCCs, and highlight possible TiME spatial landscapes that may impact clinical outcome across cancer types.

We found concordance between the results from our heterogeneity and composition analyses and those from several other studies. Our results are similar to a study that clustered HNSCC biopsies based on their neoplastic and immune gene signatures as determined by RNA-sequencing, and reported that samples from the same patient were more similar to each other than samples from different patients [174]. This trend has also been observed across cancer types, including melanoma [175], hepatocellular carcinoma [176], pancreatic ductal adenocarcinoma [177], and breast cancer [178]. Moreover, we found that patients who experienced an increase in $CD8^+$ T cell density from their primary to recurrent tumors were associated with improved PFS, which has previously been reported in head and neck cancer [167, 168]. Conversely, patients experiencing the greatest decrease in $CD8^+$ T cell density in recurrence had TNM stage 4 primary tumors. This could indicate that later staged primary tumors are better equipped to evade immune attack in recurrence, and a therapy to elicit an anti-tumor immune response may be beneficial.

Despite the fact that HPV(-) HNSCCs often contain limited immune infiltrates [155], our analyses revealed significant differences in the immune cell spatial organization within these tumors that were associated with progression free survival, highlighting the importance of considering the spatial context of the TiME. Most strikingly, patients whose primary tumors contained more compartmentalization between their neoplastic tumor cell and immune cell populations demonstrated longer PFS. This correlation was also identified in a similar analysis of triple negative breast cancer patients [139], indicating that a compartmentalized spatial architecture may play a favorable role in survival across cancer types. Our analyses of TiME composition and functional marker expression indicate there is likely more antigen presentation and less immunoregulation in regions of compartmentalization rather than in regions of mixing.

This favorable immune landscape in compartmentalized tumors may contribute to improved survival. On the contrary, four of the five of the mixed primary tumors were classified as TNM stage 4, which may contribute to shortened survival for these patients. However, given that compartmentalized tumors were not all classified as an early TNM stage and instead ranged from TNM stage 1 to 4, it is difficult to determine the exact association between neoplastic tumor and immune cell mixing versus TNM stage, and further investigation is warranted to understand these correlations and impact on prognosis.

Prior studies have identified a mesenchymal HNSCC subtype [179, 180]; our αSMA^+ mesenchymal cellular neighborhood analyses highlight the importance of considering how these cells are organized within the TiME, beyond simply considering their presence in the tumor. This could reveal more precise mesenchymal subtypes for improved stratification for patient care. αSMA is a common marker for cancer-associated fibroblasts (CAFs) [153, 181], whose presence in tumors, including HNSCC, tends to be associated with tumor progression, metastasis, and resistance to therapy [182-187]. However, despite compartmentalized tumors having increased αSMA^+ cell density, patients with compartmentalized primary tumors demonstrated longer PFS, which indicates that the spatial organization of αSMA^+ cells may be related to their function.

We found αSMA^+ cells neighboring immune cells in both mixed and compartmentalized primary tumors, but the types of immune cells near αSMA^+ cells differed. This is interesting, as emerging research has found that CAFs, which are often defined by their expression of αSMA, can modulate immune cell function within the TiME. With our mIHC platform we were able to identify the differences in the types of immune cells neighboring αSMA^+ cells and relate these differences to survival, supporting recent research on the impact of CAFs on various immune cell populations. A recent study in melanoma found CAFs to be instrumental in aiding tertiary

lymphoid structure (TLS) development [188]; TLS are defined by their large concentration of B cells surrounded by T cells, and their presence in tumors is associated with improved patient outcome [31-33]. In compartmentalized tumors, we noticed the αSMA^+ cells to be more structured and located neighboring dense pockets of T cells and B cells, perhaps indicating the formation of TLS. Although this antibody panels did not include a biomarker for high endothelial venules to confirm presence of TLS, our results support the notion that CAF-lymphocyte interactions can be clinically beneficial.

Conversely, patients with predominantly mixed primary tumors and shorter PFS contained αSMA^+ cells primarily located near granulocytes and other $CD45^+$ immune cells, many of which were likely $CD163^+$ myelomonocytic cells or mast cells. CD163 is a marker for scavenger receptor activity and is commonly used to demark pro-tumor type tumor-associated macrophages and monocytes. Supporting these findings, a study in oral squamous cell carcinoma found the presence of CAFs to be associated with increased presence of $CD163^+$ macrophages and worse survival [189]. Another study reported CAFs to be correlated with an increase in monocyte expression of CD163, which in turn suppressed T cell proliferation and increased neoplastic cell proliferation in breast cancer [190]. Mast cells mediate innate and acquired immune response as a part of the myeloid lineage. They have been reported to facilitate neo-vascularization and tumor dissemination in HNSCC, and found to be correlated with increased angiogenesis in advanced HNSCC [191]. However, interactions between mast cells and CAFs is largely unknown; our results indicate this interaction may be associated with increased neoplastic density and worse PFS. Of note, a recent study in melanoma and pancreatic adenocarcinoma found neutrophils, a subclass of granulocytes, to exert pro-tumor effects when in the presence of

CAFs [192], and another study found the combined presence of neutrophils and CAFs to be associated with shortened survival in gastric adenocarcinoma [193].

Finally, patients with predominantly mixed primary tumors and shorter PFS were found to contain fewer overall stromal cells and greater proportions of those αSMA$^+$ cells neighboring neoplastic tumor cells. It has been reported that neoplastic tumor cells in direct contact with CAFs move along tracks laid by the CAFs in the extracellular matrix, promoting tumor growth [194, 195]. It is possible that αSMA$^+$ cells in mixed tumors may provide avenues for neoplastic tumor cells to transit, thus leading to more advanced tumor progression. Overall, spatial analyses herein deepened our understanding of neoplastic tumor and immune cell organization relative to each other, and how this organization is related to αSMA$^+$ cells working in tandem with many cells, beyond single cell-cell interactions, to impact TiME organization, and ultimately, clinical outcome. The conclusions from our spatial analyses and their relationship to clinical variables are summarized in **Figure 2.5**.

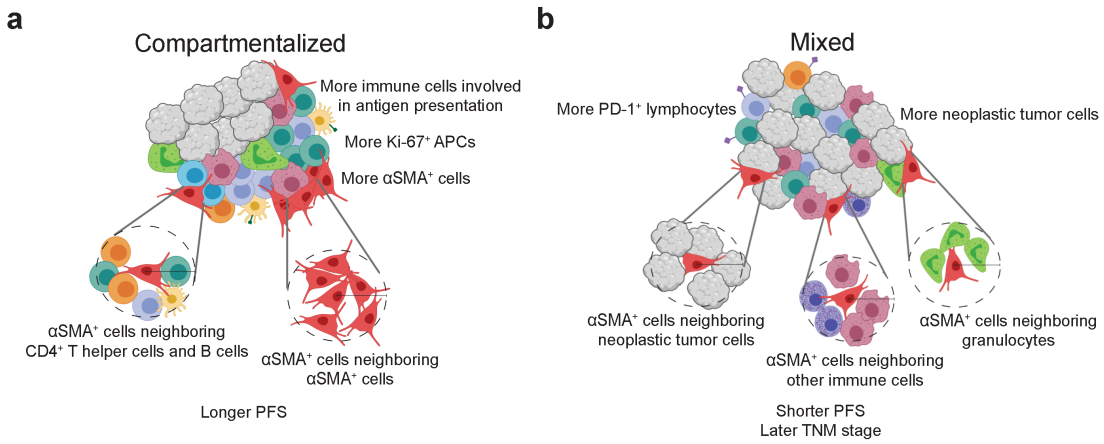


Figure 2.5: Proposed model of primary HPV(-) HNSCC tumor-immune microenvironments. a, Depiction of a tumor with a compartmentalized spatial architecture. These tumors have decreased mixing between immune cells and neoplastic tumor cells and tend to contain greater immune cell density, specifically cells involved in antigen presentation, as well as increased density of αSMA$^+$ mesenchymal cells. The αSMA$^+$ cells present tend to be neighbored by CD4$^+$ T helper cells and B cells, as well as other αSMA$^+$ cells. Compartmentalized primary tumors were found to be associated with longer progression free survival. **b**, Depiction of a tumor with a mixed spatial architecture. These tumors contain increased mixing between immune cells and neoplastic tumor cells and tend to contain increased neoplastic tumor cells, as well as increased PD-1-positive lymphocytes. The αSMA$^+$ cells present

tend to be neighbored by neoplastic tumor cells, other immune cells, many of which are $CD163^+$ or are mast cells, and granulocytes. Mixed primary tumors were found to be associated with shorter progression free survival and later TNM stage. This figure was created using BioRender.com.

The concordance found among results herein and those of the aforementioned studies provides orthogonal support for our conclusions and indicates that single-cell heterogeneity and spatial organization of tumors may share similarities across different types of cancer and across different molecular assays. Our spatial analyses demonstrate the use of various algorithms to quantify the spatial landscape of tumors using single-cell imaging data. Further, our computational methods provide a framework for future single-cell imaging analyses, as they are applicable to any multiplex imaging assay, including mIHC, co-detection by indexing technology (CODEX) [140], cyclic immunofluorescence [117, 152, 196], and imaging mass cytometry [68]. Our results highlight several spatial architectures that may help guide precision medicine approaches for HPV(-) HNSCC patients, including architectures that may help stratify patients who may have shorter PFS, and thus warrant more aggressive therapy or clinical follow-ups. While our results provide evidence that the spatial organization of HPV(-) HNSCC tumors correlates with clinical outcome, future studies with larger cohorts will be needed to evaluate the strength and validity of our observations. Studies with greater representation of anatomic site and stage are also needed to assess the prognostic value of the spatial features identified in this HNSCC cohort. Despite these limitations, this study demonstrates practical analysis strategies that elucidate spatial architecture features for potential use in precision medicine.

2.5 Methods

2.5.1 Multiplex immunohistochemistry data generation

mIHC is an immunohistochemical-based imaging platform that evaluates sequentially stained immune lineage epitope-specific antibodies for immunodetection on FFPE tissue sections

[121, 122]. Images were stained and processed as described in our previous report [122]. A table of antibodies, species, vendor, and concentration used in staining are previously reported in Table 1 of Banik et al [122]. Briefly, sequentially stained images were co-registered in MATLAB. AEC signal from each antibody stain was extracted and normalized, and the mean intensity of each single cell for each marker was quantified in Cell Profiler. Watershed based nuclei segmentation on hematoxylin staining was used to identify single cells in FIJI. Using a hierarchical gating strategy, single cells were phenotyped using image gating cytometry in FCS Express 7 Image Cytometry RUO (**Supplementary Figure 2.1a**). A threshold was set on the scatterplot of mean intensity for each marker within the gating strategy, validated by visual live rendering of masked cell objects within the selected gate on extracted marker signals. Cartesian coordinates of each phenotyped cell were maintained relative to the tissue region.

We applied the mIHC pipeline to analyze matched primary and recurrent FFPE tissue specimens from nine HPV(-) HNSCC patients. Each patient underwent surgical resection of their primary tumor prior to beginning a regimen of chemotherapy and radiation therapy. Upon recurrence, the patient underwent another surgical resection of their recurrent tumor. Tissue specimens for each patient were obtained from the Oregon Health & Science University Knight Biolibrary and were deidentified and coded with a unique identifier prior to analysis. Patient demographic and clinical data including HPV status, tobacco and alcohol use, treatment regimens, and survival outcomes were collected. All HNSCC tumors were staged according to the 8th edition AJCC/UIC TNM classification and cohort characteristics are shown in **Table 2.1** as reported in Banik et al [122]. All studies involving human tissue were approved by institutional IRB (protocol #809 and #3609), and written informed consent was obtained.

2.5.2 Tumor heterogeneity analyses

The Kullback-Leibler divergence was calculated using the entropy function from the Scipy Python package [197]. The distribution of each individual tumor region's cellular composition was compared to the average of five different cellular composition distributions: the average distribution for the region's tumor, the average distribution for the patient's primary and recurrent tumors, the average distribution across all tumors of the same timepoint in the cohort (primary or recurrent), the average distribution across all tumors resected from the same anatomic site in the cohort, and the average distribution across all tumors from all patients in the cohort (primary and recurrent combined). Log base 2 was used for the calculation. Unsupervised hierarchical clustering of each tumor region was performed using the log10+1 normalized density of each cell type present as the features. Euclidean distance was used to determine distances between observations, and the Ward method was used for the linkage.

2.5.3 TiME compositional change clustering analysis

Unsupervised hierarchical clustering of each patient was performed using the normalized change in density of each cell type as the features. Euclidean distance was used to determine distances between observations, and the Ward method was used for the linkage. Normalization in the change in density was computing by first calculating the absolute value of the raw change in density for each cell type. These values were then normalized to a range of [0,1]. Finally, the values that were originally negative (decreasing change), were flipped to be negative values again, such that all values ranged from [-1,1] with zero representing no change.

2.5.4 Mixing score analysis

Keren et al. developed a mixing score to quantify the ratio of neoplastic and immune cell spatial interactions [139]. This score is defined as the number of interactions between neoplastic

tumor cells and immune cells divided by the number of interactions between immune cells and another immune cell within a tumor region. We defined there to be an interaction between two cells if their centers [the (x,y) coordinates provided by mIHC] were within 15 μm from one another. We used the median mixing score value (0.107) for all primary tumors as the threshold to distinguish between mixed and compartmentalized spatial organization groups. Tumor regions with a mixing score of greater than 0.107 were defined as mixed. Tumor regions with a mixing score of less than 0.107 were defined as compartmentalized. We used a density threshold of less than 250 immune cells per 800^2 μm^2 to define tumor regions as cold. We chose this threshold to match that used by Keren et al.

2.5.5 Functional marker bootstrapping analyses

Bootstrapping analyses involved the following steps. First, one tumor region per eighteen tumor samples was randomly selected. The regions were then split into two groups based on their mixed or compartmentalized spatial architecture designation, and the average proportion of the specified cell population expressing the specified functional marker was calculated for each group. Finally, this process was repeated 100 times, yielding 100 values representing the proportion of the specified cell population expressing the specified functional marker for mixed tumor regions and 100 values representing the proportion of the specified cell population expressing the specified functional marker for compartmentalized tumor regions. These values were then compared between the mixed and compartmentalized groups for differences.

2.5.6 Cellular neighborhood clustering analyses

Cellular neighborhoods were defined by drawing a circle of a specified radius around all seed cells of a given phenotype. Cells whose centers were inside the circle were considered

neighbors of that seed cell and contributed to that cell's neighborhood. All neighborhoods across all tumor regions were then clustered using scikit-learn's MiniBatchKMeans function [198] to perform K-means clustering according to their normalized cellular composition. The elbow method was used to determine the number of clusters to form. Each resulting cluster was comprised of neighborhoods with a similar cellular makeup. Unsupervised hierarchical clustering of each primary tumor was performed using the log10+1 normalized proportion (out of 100) of αSMA^+ cell neighborhood clusters present in the tumor as the features. Euclidean distance was used to determine distances between observations, and the Ward method was used for the linkage.

2.5.7 Statistics

Independent t-tests were used to determine statistically significant differences for independent samples whose differences followed a normal distribution. Mann-Whitney U tests were used to determine statistically significant differences for independent samples whose differences did not follow a normal distribution. Paired t-tests were used to determine statistically significant differences for paired samples whose differences followed a normal distribution. Wilcoxon signed rank tests were used to determine statistically significant differences for paired samples whose differences did not follow a normal distribution. One-way ANOVA tests were used to determine statistically significant differences for multi-group comparisons. If the ANOVA result was significant, a Tukey honestly significant difference post-hoc test was conducted to determine which groups were significantly different from one another. A Benjamini-Hochberg correction was used to account for multiple hypothesis testing in analyses that involved systematically testing multiple variables. P-values less than 0.05 were

considered statistically significant. All statistical calculations were performed with the Scipy and statsmodels packages using Python software [197, 199].

2.5.8 Survival analyses

Kaplan-Meier curves were generated and a log-rank test was performed using the lifelines package with Python software [200]. P-values less than 0.05 were considered statistically significant.

2.5.9 Data availability

All of the data produced by our mIHC computational image processing pipeline, including protein abundance, cell phenotype, and cell location information saved in the form of a matrix, in addition to survival data, is available for download on Zenodo at https://doi.org/10.5281/zenodo.5540356.

2.5.10 Code availability

All computational analyses in this study were performed using Python software, version 3.6.5. The code created to produce the results of this study is available at https://github.com/kblise/HNSCC_mIHC_paper.

Acknowledgements

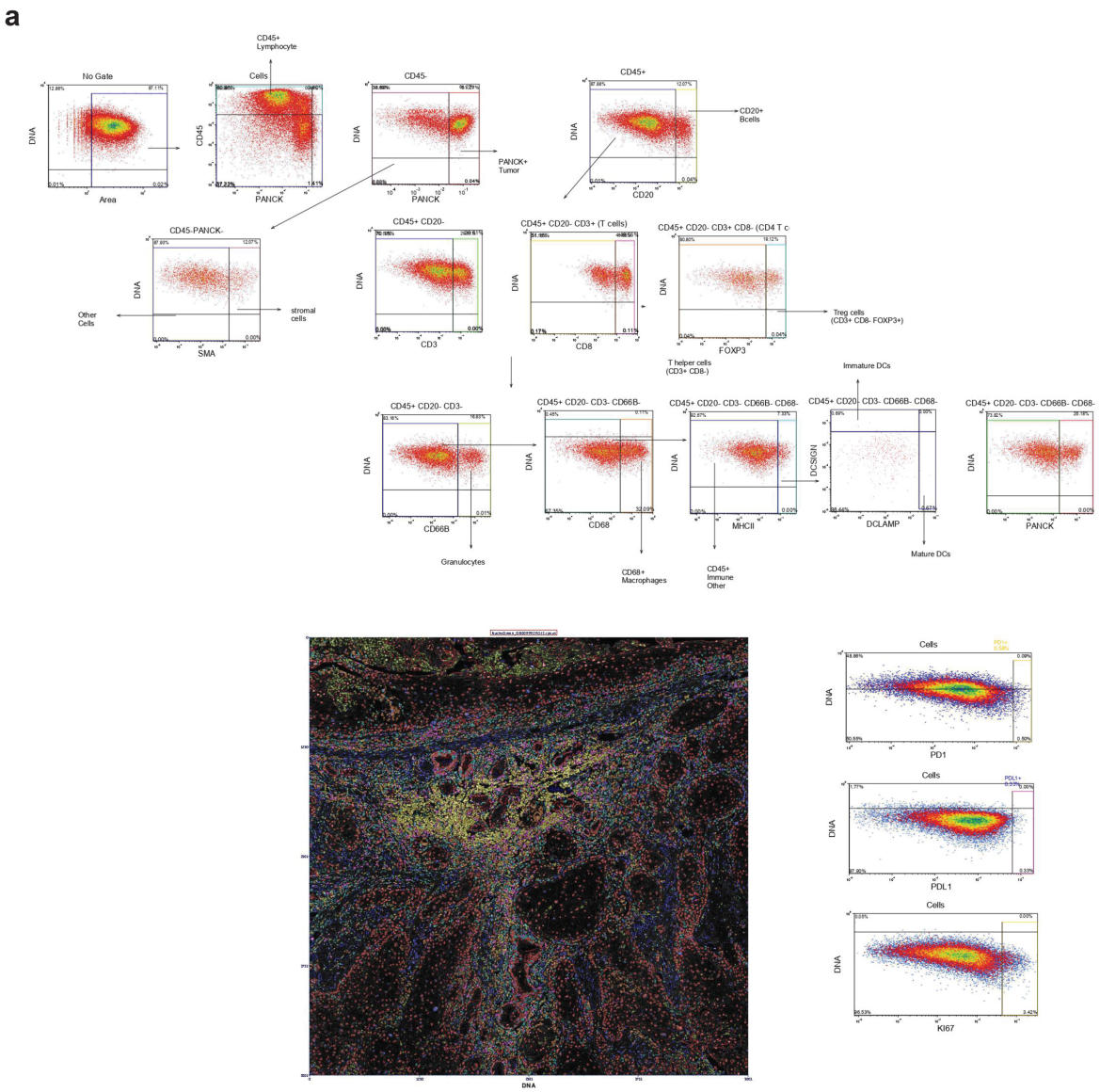
The authors thank Dr. Courtney Betts, Justin Tibbitts, Teresa Beechwood, and Meghan Lavoie for regulatory and technical assistance. The authors thank the OHSU Knight Biostatistics Shared Resource for guidance with statistical calculations. The authors also acknowledge and thank the patients who donated tissue samples for this study. K.E.B. acknowledges funding from the National Cancer Institute (NCI) of the National Institutes of Health under award number

T32CA254888. L.M.C. acknowledges funding from the NCI awards U01CA224012, R01CA223150, R01CA226909, and R21HD099367 and funding from the Brenden-Colson Center for Pancreatic Care at OHSU. J.G. acknowledges funding under award number U24CA231877. L.M.C. and J.G. acknowledge funding from the OHSU Knight Cancer Institute, NCI award number U2CCA233280, and the Prospect Creek Foundation to the OHSU SMMART (Serial Measurement of Molecular and Architectural Responses to Therapy) Program.

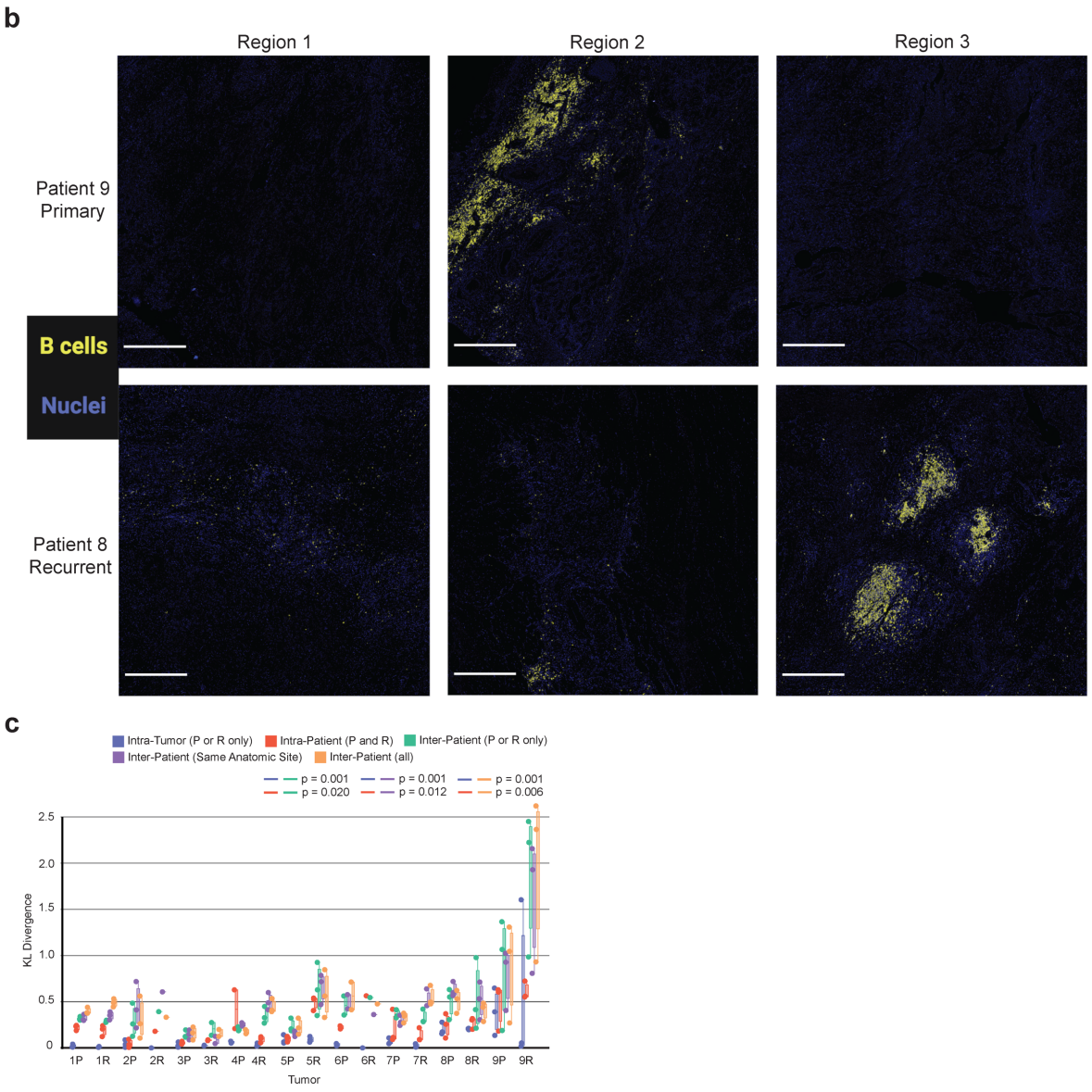
Author contributions

K.E.B., S.S., L.M.C., and J.G. conceived the project and designed the overall study. G.L.B. collected the data using the multiplex immunohistochemistry experimental assay and provided clinical expertise. S.S. performed computational analyses for image processing to produce the single-cell spatial dataset that was analyzed in the study. K.E.B. designed the computational experiments, developed software for the study, and analyzed all results with support from S.S. K.E.B., S.S., L.M.C., and J.G. wrote the original and revised manuscript. L.M.C. and J.G. supervised the project. All authors read, edited, and approved the final manuscript.

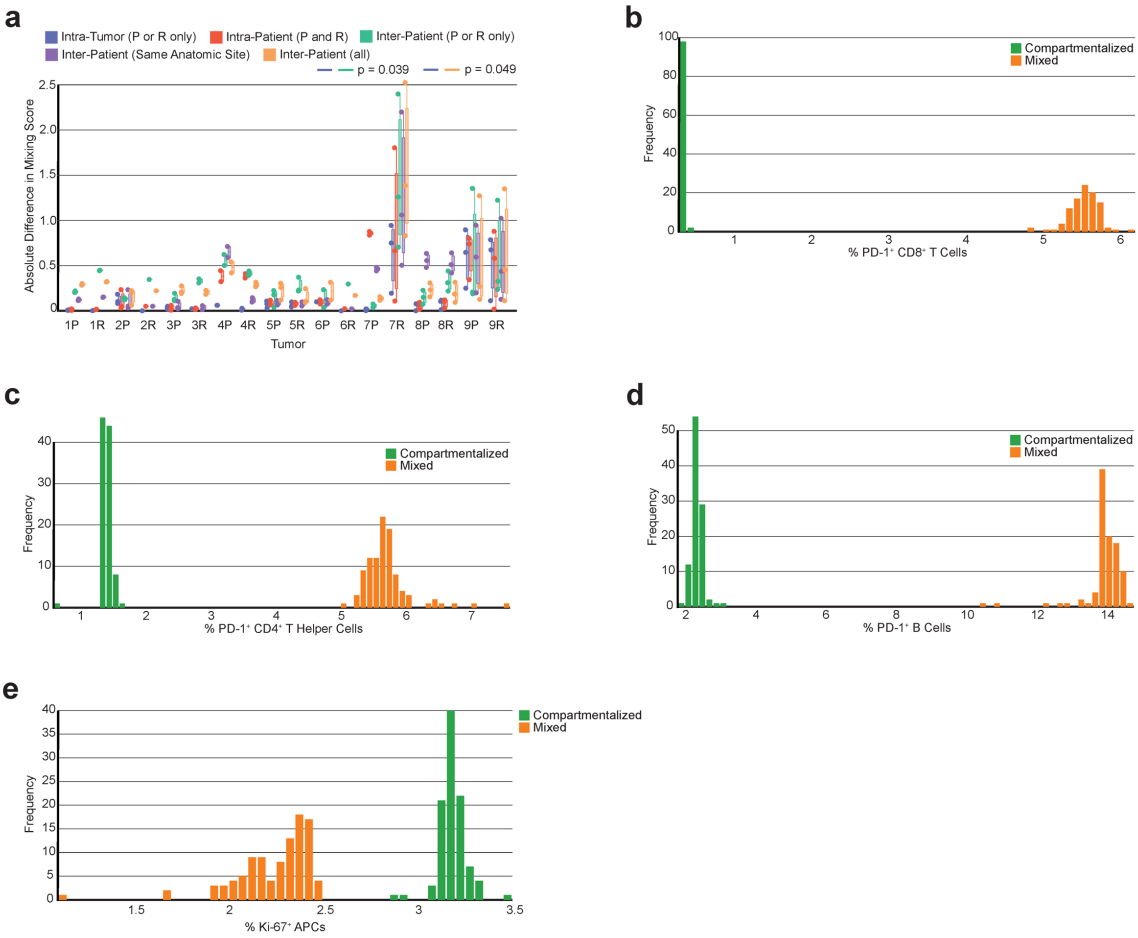
2.6 Supplementary Data



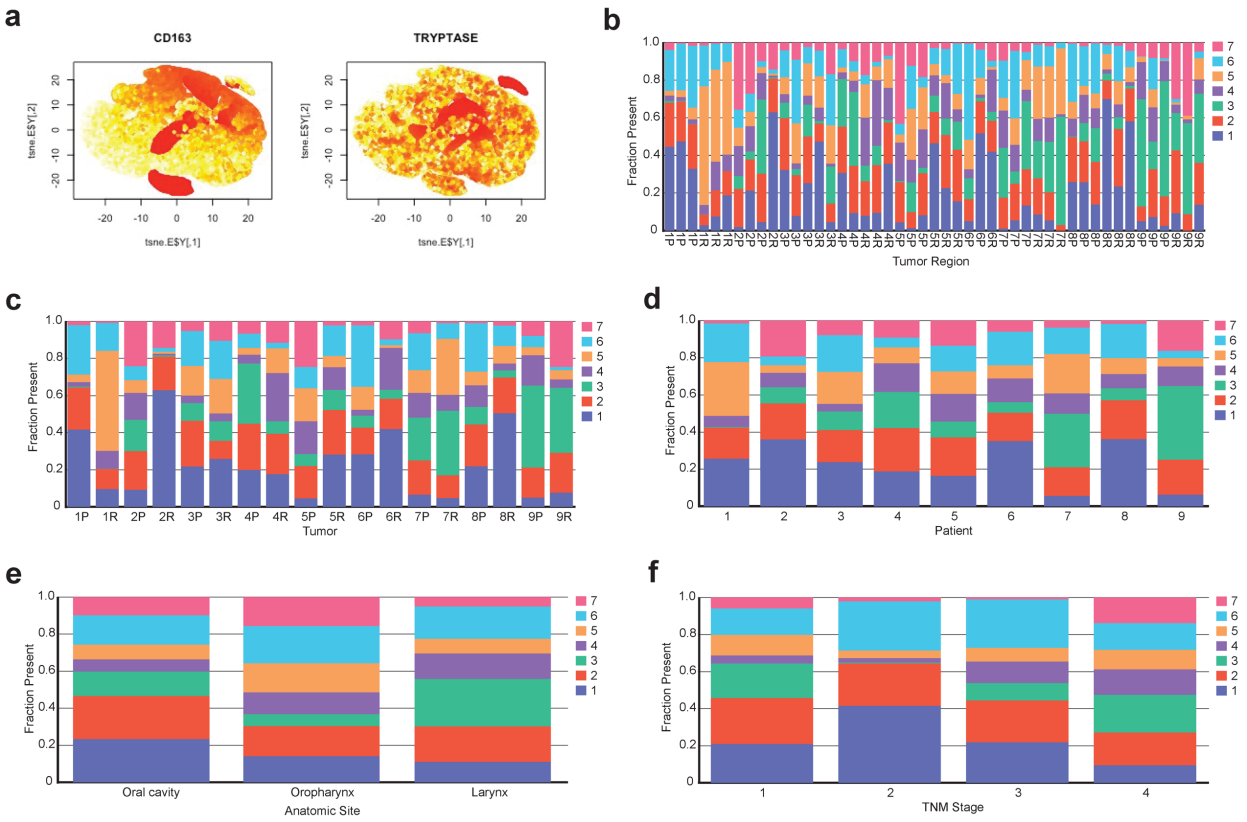
Supplementary Figure 2.1: Tumor heterogeneity assessed by mIHC assay. a, Gating template used to determine cell phenotypes and classification of ten populations using image gating cytometry in FCS Image Cytometry 7 RUO. Scatter plots show single cell mean intensity distributions rescaled from 0-1 on a log10 scale with gated populations in colored boxes. Image plot shows a semi-transparent pseudo-colored mask of each cell object colored by the classified population overlaid on signal extracted nuclei image. The markers used for identification of cell phenotypes are shown in Table 1.



Supplementary Figure 2.1 (continued): **b**, mIHC images of six tumor regions selected from two different tumor resections (patient 9 primary and patient 8 recurrent). Blue pseudo-color indicates all cell nuclei present in the region (Hematoxylin^{+}) and yellow indicates B cells. In both tumors, two regions have low densities of B cells, whereas one region has a high density and spatial concentration of B cells, despite the three regions being sampled from the same tumor. **c**, Box plot of the Kullback-Leibler divergences from a single tumor region's immune cell distribution compared to: the tumor's average immune cell distribution [Intra-Tumor (P or R only)], the patient's average immune cell distribution [Intra-Patient (P and R)], the cohort's average immune cell distribution across tumors of the same timepoint [Inter-Patient (P or R only)], the cohort's average immune cell distribution across tumors of the same anatomic site [Inter-Patient (Same Anatomic Site)], the cohort's average immune cell distribution across all tumors from all patients [Inter-Patient (all)]. P-values calculated using a one-way ANOVA multi-group significance test followed by a Tukey honestly significant difference post-hoc test.



Supplementary Figure 2.2: Leukocyte functionality compared to mixing score. a, Box plot of the absolute difference from a single tumor region's mixing score compared to: the tumor's average mixing score [Intra-Tumor (P or R only)], the patient's average mixing score [Intra-Patient (P and R)], the cohort's average mixing score across tumors of the same timepoint [Inter-Patient (P or R only)], the cohort's average mixing score across tumors of the same anatomic site [Inter-Patient (Same Anatomic Site)], the cohort's average mixing score across all tumors from all patients [Inter-Patient (all)]. P-value calculated using a one-way ANOVA multi-group significance test followed by a Tukey honestly significant difference post-hoc test. **b**, Histogram showing the results of a bootstrapping analysis on the percentage of $CD8^+$ T cells expressing PD-1. Bars are colored by spatial architecture. **c**, Histogram showing the results of a bootstrapping analysis on the percentage of $CD4^+$ T helper cells expressing PD-1. Bars are colored by spatial architecture. **d**, Histogram showing the results of a bootstrapping analysis on the percentage of B cells expressing PD-1. Bars are colored by spatial architecture. **e**, Histogram showing the results of a bootstrapping analysis on the percentage of APCs expressing Ki-67. Bars are colored by spatial architecture.



Supplementary Figure 2.3: αSMA$^+$ cellular neighborhood distributions. **a**, t-SNE representation of single cells from the 'Other Immune' class colored by heatmap of mean intensity for CD163 and tryptase (mast cells) show distinct populations of cells expressing high levels of CD163 and tryptase, indicating these cell types are likely present in the 'Other Immune' class. **b**, Stacked bar chart showing the proportion (out of 1.0) of αSMA$^+$ cell neighborhood clusters present in each tumor region (n=47). **c**, Stacked bar chart showing the average proportion (out of 1.0) of αSMA$^+$ cell neighborhood clusters present in each tumor (n=18). **d**, Stacked bar chart showing the proportion (out of 1.0) of αSMA$^+$ cell neighborhood clusters present in each patient's primary and recurrent tumors averaged (n=9). **e**, Stacked bar chart showing the average proportion (out of 1.0) of αSMA$^+$ cell neighborhood clusters present in tumors collected from the three anatomic sites. **f**, Stacked bar chart showing the average proportion (out of 1.0) of αSMA$^+$ cell neighborhood clusters present in the primary tumors by their TNM stage.

Chapter 3: Comparing natural killer cell function and spatial organization across breast cancer subtypes

This chapter has been formatted for inclusion in this dissertation from the manuscript "Natural Killer cells occupy unique spatial neighborhoods in human HER2$^-$ and HER2$^+$ breast cancers" by Femke, A. I. Ehlers, Katie E. Blise, Courtney B. Betts, Shamilene Sivagnanam, Loes F. S. Kooreman, E. Shelley Hwang, Gerard M. J. Bos, Lotte Wieten, and Lisa M. Coussens, manuscript submitted (2024) [201]. The author of this dissertation is the secondary author of this manuscript and used single-cell gated data generated by LMC's laboratory to conduct the computational spatial experiments, the results of which were used to generate Main Figures 3.3, 3.4C-E, and 3.5 and Supplementary Figures 3.5 and 3.7.

3.1 Abstract

Tumor-infiltrating lymphocytes are considered clinically beneficial in breast cancer, but the significance of natural killer (NK) cells is less well characterized. As increasing evidence demonstrates that the spatial organization of immune cells in tumor microenvironments is a significant parameter for impacting disease progression as well as therapeutic responses, improved understanding of tumor-infiltrating NK cells and their location within tumor contextures is required in order to better design effective NK cell-based therapies. In this study, we developed a multiplex immunohistochemistry (mIHC) antibody panel designed to quantitatively interrogate leukocyte lineages, focusing on NK cells and their phenotypes in human breast cancer. Due to the clinical evidence supporting a significant role for NK cells in HER2$^+$ breast cancer in mediating responses to Trastuzumab, we evaluated HER2$^-$ and HER2$^+$ specimens in two independent patient cohorts (n=26 and n=30, respectively). Consistent with literature, we found CD3$^+$ T cells to be the dominant leukocyte subset across breast cancer specimens. In comparison, NK cells, identified by CD56 or NKp46, were scarce in all specimens with low granzyme B expression indicating reduced cytotoxic functionality. Whereas the immune cell context including NK cell density and phenotype did not appear influenced by

HER2 status, spatial analysis of NK cells revealed distinct phenotype profiles with regards to proximity to neoplastic tumor cells that associated with HER2 status. Spatial cellular neighborhood analysis demonstrated existence of multiple unique neighborhood compositions surrounding NK cells, where NK cells from $HER2^-$ tumors were found more frequently proximal to neoplastic tumor cells, whereas NK cells from $HER2^+$ tumors located more frequently proximal to $CD3^+$ T cells. This study establishes the utility of quantitative mIHC to evaluate NK cells on a single-cell spatial proteomics level and illustrates how spatial characteristics of NK cell neighborhoods varies within the context of $HER2^-$ and $HER2^+$ breast cancers.

3.2 Introduction

Breast cancer is the most common cancer in women, and although survival rates have improved, it remains a leading cause of death [98]. Based on their molecular classification, breast cancer subtypes are largely segregated into basal-like, luminal A, luminal B, normal-like, and $HER2^+$ breast cancer and guide therapeutic decisions [202]; however, even within subtypes, high heterogeneity is recognized. Tumor-infiltrating lymphocytes are generally accepted to be important for favorable therapeutic responses, but the significance of natural killer (NK) cells in breast cancer is less well characterized [203]. NK cells are cytotoxic leukocytes that can impact neoplastic cells through an extensive array of cell surface receptors, with one of the most potent activating receptors being CD16A (FcγRIIIA) that binds immunoglobulins and immune complexes, as well as tumor-targeting antibodies that induce tumor cell killing through antibody-dependent cellular cytotoxicity (ADCC) [204, 205]. Next to their cytotoxic capabilities, NK cells release proinflammatory cytokines and chemokines to activate and recruit other immune cell type including dendritic cells and macrophages [206, 207].

NK cell abundance and functional status are positively correlated with improved clinical outcome in several solid tumors, supporting the tenet that NK cells are critical for tumor control despite being present in rather low frequency (reviewed in [35, 208]. In experimental mouse models of $HER2^+$ breast cancer, NK cells were reported to significantly contribute to efficacy of the anti-HER2 antibody Trastuzumab by mediating ADCC [92, 209]. Indirect evidence further supports that NK cells also contribute to efficacy of Trastuzumab in patients with $HER2^+$ breast cancer, as demonstrated by increased tumor-infiltrating NK cells following Trastuzumab therapy [93, 210, 211].

Due to their cytotoxic potential and high safety profile, NK cells are promising candidates for adoptive cell therapy [212]. However, tumor-infiltrating NK cells are frequently described as dysfunctional, which might underlie limited success thus far with adoptive NK cell therapy in breast cancer [213]. NK cell effector functions could be inhibited by neoplastic tumor cells through downregulation of activating ligands and expression of inhibitory ligands such as HLA class I [94]. Moreover, immunosuppressive tumor microenvironments (TME) are known to negatively affect NK cells via various mechanisms, including immunosuppressive cytokines and hypoxic conditions [95]. Hypoxia is frequently present in solid tumors and has been reportedly identified in 50% of breast cancer patients [214]. Compared to blood, tumor-infiltrating NK cells express low levels of activating receptors and have a low cytotoxic potential, indicating that these alterations could be TME-mediated [215, 216].

To enhance NK cell anti-tumor responses, strategies have emerged, such as cytokine-activation and monoclonal antibodies (mAb), designed to be directed against inhibitory checkpoint molecules to prevent NK cell inactivation or ADCC-mediating mAb for stimulating NK cell activation [212]. Combination of anti-HER2 targeting therapy e.g., Trastuzumab, with

NK cell-based therapy could be an approach to boost NK cell activity against $HER2^+$ breast cancer [217]. In addition to enhancing NK cell potency, it will be crucial to understand the cellular spatial organizations in TMEs, including the distribution of endogenous NK cells and their interactions with neoplastic tumor cells, as well as other tumor-infiltrating leukocytes. Cellular and spatial heterogeneity has been linked to clinical outcomes; for example, spatial organization of immune cells within TMEs of triple-negative breast cancer patients, assessed by multiplexed ion beam imaging by time-of-flight (MIBI-TOF), was linked to overall patient survival [139, 218]. However, little is known regarding infiltration and distribution of NK cells within heterogeneous breast cancer TMEs. Common to the above-described studies evaluating NK cells in human breast cancer, was the use of flow cytometry or standard single-color immunohistochemistry (IHC) for quantitation, thus negating information on spatial distributions of cells or identification of cell populations being restricted. More recently, multiplexed imaging methods, such as multiplex IHC (mIHC), have become available and overcome some of these limitations, as these methods allow *in situ* phenotyping and evaluation of spatial architectures of diverse cell populations on one tissue section [219]. In this study, we developed an mIHC antibody panel designed to quantify leukocyte lineages, including deep-auditing of NK cells and their phenotypes, as well as their spatial architecture in human breast cancer. Specifically, we compared $HER2^-$ versus $HER2^+$ disease due to the clinical relevance of NK cells in combination with Trastuzumab.

3.3 Materials and methods

3.3.1 Clinical samples

Cohort 1: Breast cancer tissue was obtained by Duke University, Durham, USA, and shipped to Oregon Health & Science University (OHSU), Portland OR, USA. The study protocol

was approved by the IRB of Duke University (Durham, NC USA) ethics committee as appropriate. Informed consent was obtained from all human subjects included in this study. Cohort 2: Breast cancer tissue was obtained by Maastricht University Medical Center+, Maastricht, The Netherlands. Collection, storage, and use of tissue and patient data were performed in agreement with the "Code for Proper Secondary Use of Human Tissue in the Netherlands" and approved by the local ethics committee (METC 2019-1154). Clinical characteristics for both cohorts are shown in **Supplementary Table 3.1**.

3.3.2 Multiplex immunohistochemistry and image acquisition

Multiplex immunohistochemistry (mIHC) was performed as previously described [121, 122]. Briefly, formalin-fixed, paraffin-embedded (FFPE) tissue sections (5 μm) were baked in a 60°C heat chamber for 30 minutes, deparaffinized in xylene and rehydrated in serially graded ethanols and placed in water. After deparaffinization, the slides were fixed in 10% neutral buffered formalin for 30 minutes and subsequently washed twice in PBS to prevent tissue from falling of the slides [220].Tissue slides were counterstained with hematoxylin (Dako, S3301) for 1.0 minute at room temperature (RT), washed in water, and mounted with 1X TBST buffer (Boston Bioproducts, IBB-181R) before coverslips (Thermo Scientific, 12460S) were placed on the slides. Slides were scanned using the whole slide digital scanner Aperio ImageScope AT (Leica Biosystems) at 20X magnification. Coverslips were removed by placing slides in TBST for a few minutes with agitation. For heat-mediated antigen retrieval, slides were placed in 1X EDTA Decloaker (Biocare Medical, CB917) in a pressure cooker for 15 or 20 minutes at 115°C for staining cycle 1 and 3, and slides were placed in 1X citrate buffer, pH 6.0 (Biogenex Laboratories, HK080) for 20 minutes at 95°C for the other staining cycles (**Supplementary Table 3.2**). After antigen retrieval and washing in water and TBST, slides were incubated with

dual endogenous peroxidase block (Dako, S2003) for 10 minutes at RT. Protein blocking was performed with 5% normal goat serum and 2.5% BSA in PBS for 10 minutes at RT. Slides were incubated with the primary antibody for 30 or 60 minutes at RT or overnight at 4°C. After slides were washed, they were incubated with the secondary antibody Histofine Simple Stain MAX PO horseradish peroxidase-conjugated polymer (Nichirei Biosciences, 414134F or 414144F) for 30 minutes at RT in all staining rounds, except cycle 3 round 2 where slides were incubated for 60 minutes at RT with histofine. Information on antibodies and staining conditions are listed in **Supplementary Table 3.2**. Chromogenic signal detection was performed with AEC substrate kit, Peroxidase (Vector Laboratories, SK-4200) for all antibodies, except the anti-PD-1 antibody that was detected with AMEC Red Substrate, Peroxidase (Vector Laboratories, SK-4285). Slides were scanned using the whole slide digital scanner and coverslips were removed, as described above. AEC or AMEC were removed by placing the slides in serially graded ethanols. For cycles with two rounds of antibody development, horseradish peroxidase was inactivated by incubation with dual endogenous peroxidase block (Dako, S2003) for 10 minutes at RT. Protein blocking was repeated before each following antibody. For the following cycles, the same steps from antigen retrieval to slide scanning were performed as described above. Heat-mediated antigen retrieval stripped antibodies from the previous staining cycle. Human tonsil and spleen were used as a control in all staining cycles and a representative visualization of each antibody is shown in **Supplementary Figure 3.1**.

3.3.3 ROI selection and image processing

In cohort 1, tissue area was limited and all intact regions were captured in 1-6 regions of interest (ROIs). In cohort 2, 4-13 ROIs with lymphoid aggregates, tumor-sparse areas and tumor-dense areas were selected. When tissue size allowed, several ROIs with a minimum of 0.8 mm^2

were selected to reflect tissue heterogeneity. For each specimen, the sum of the ROIs was used for subsequent calculations. Image co-registration and processing were performed as previously described [107, 121]. Cell percentages of immune cell lineages, presented in **Figure 3.1**, were quantified by image cytometry gating strategies (**Supplementary Figure 3.2**).

3.3.4 Cellular neighborhood clustering analysis

A cellular neighborhood clustering analysis was performed, following similar steps to those as previously described [147]. Briefly, we identified all cells, excluding "other nucleated cells", within a given distance threshold for every given seed cell phenotype across the dataset. In this study, we set the distance threshold to 60 μm and the seed cell phenotype to be all NK cells. After computing the neighboring cells for each NK cell in the dataset, we then performed K-means clustering on the neighborhoods, using the normalized composition of cells comprising the neighborhoods as the features on which to cluster. The elbow method was used to determine the optimal number of clusters to generate resulting in clusters of neighborhoods, all possessing similar cellular compositions surrounding NK cells.

3.3.5 Statistical analysis

Mann-Whitney-U tests were used to determine statistically significant differences for comparisons between two groups and a Benjamini-Hochberg correction was used to account for multiple hypothesis testing. Kruskal-Wallis tests were used to determine statistically significant differences for comparisons between multiple groups and *p*-values were adjusted for multiple comparisons using Dunn's adjustments. p-values < 0.05 were considered statistically significant and reported as * $p < 0.05$ and ** $p < 0.01$. Spearman correlation coefficient was used to

determine correlations of cell densities. Statistical analysis was performed using GraphPad software (version 9.2.0).

3.3.6 Data availability

The mIHC data used for this study is available for download on Zenodo at https://doi.org/10.5281/zenodo.10632694.

3.3.7 Code availability

The code used to generate the spatial results of this study was created using Python version 3.9.4 and is available on Github at https://github.com/kblise/NKcell_mIHC_paper.

3.4 Results

3.4.1 Immune cell context in HER2$^-$ and HER2$^+$ human breast cancer

To identify NK cells and their phenotypes in human breast cancer, we used a previously validated mIHC platform and antibody panels designed to elaborate leukocyte lineages, where antibodies focusing on NK cell presence and effector status were included. A single FFPE tumor section per patient was sequentially stained with the mIHC panel, consisting of 20 well validated antibodies (**Supplementary Table 3.1**). Chromogenic detection of antibodies in their respective cycle in the mIHC panel is shown in **Supplementary Figure 3.1**. Scanned images were quantitatively analyzed by image cytometry and identified cell types, alongside their functional biomarkers, were depicted (**Figure 3.1A, Supplementary Figure 3.2**) [121, 122].

Based on the evidence for NK cells in HER2$^+$ breast cancer [92, 93, 209-211], we profiled NK cells in HER2$^-$ and HER2$^+$ tumor specimens from two independent patient cohorts, presented as cohort 1 consisting of both HER2$^-$ and HER2$^+$ specimens classified by standard IHC, and cohort 2 comprising solely HER2$^+$ therapy-naive breast cancer specimens, classified as

HER2-amplified by FISH (**Figure 3.1B**). Patient characteristics of both cohorts are provided in **Supplementary Table 3.2**. We analyzed regions of interest (ROI) on tumor tissue areas that remained after all staining cycles were completed, and due to tissue availability, more ROIs were analyzed in cohort 2 (**Figure 3.1C, Supplementary Table 3.3**). On average 12.5 mm^2 per tumor specimens were available for analysis in cohort 1, whereas 38.7 mm^2 per tumor specimens were available for analysis in cohort 2, including areas with high leukocyte density potentially reflecting lymphoid aggregates (**Figure 3.1B-C**).

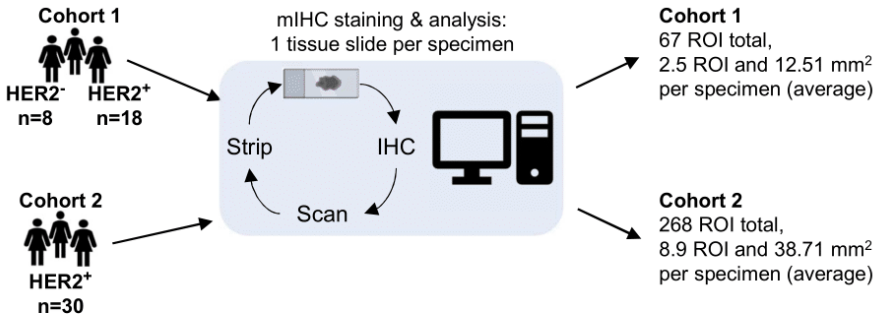
We first evaluated $CD45^+$ leukocytes, pan-cytokeratin $(PanCK)^+$ epithelial cells and other nucleated cell densities. In cohort 1, the most abundant cells in $HER2^-$ specimens reflected $PanCK^+$ epithelial cells, as compared to the median of $HER2^+$ specimens, where $PanCK^+$ epithelial cells represented comparable density to other nucleated cells (**Figure 3.2A**). Variability in cell densities was observed between specimens and ROIs, and no significant difference was found between $HER2^-$ and $HER2^+$ specimens of cohort 1 (**Figure 3.2A, Supplementary Figure 3.3A-B**). In cohort 2, comprising $HER2^+$ specimens, $CD45^+$ leukocytes were abundantly present, with a median cell density comparable to $PanCK^+$ epithelial cells (**Figure 3.2A, Supplementary Figure 3.4A-B**). Higher leukocyte density was observed in cohort 2 compared to both groups in cohort 1 likely due to differences in tissue areas available for analysis (**Figure 3.1C, 3.2A**) or due to variation in specimen acquisition at different collection sites.

A

Cell types	Identification
Epithelial cells	CD45$^-$ PanCK$^+$
CD8$^+$ T cells	CD45$^+$ CD3$^+$ CD8$^+$
CD4$^+$ T cells	CD45$^+$ CD3$^+$ CD8$^-$
NK cells	CD45$^+$ CD3$^-$ CD56$^+$ NKp46$^+$ CD45$^+$ CD3$^-$ CD56$^+$ NKp46$^-$ CD45$^+$ CD3$^-$ CD56$^-$ NKp46$^+$
Myelomonocytic cells	CD45$^+$ CD3$^-$ CD56$^-$ NKp46$^-$ CD11b$^+$ CD68$^+$ CD45$^+$ CD3$^-$ CD56$^-$ NKp46$^-$ CD11b$^-$ CD68$^+$
Dendritic cells	CD45$^+$ CD3$^-$ CD56$^-$ NKp46$^-$ CD11b$^+$ CD68$^-$ CD11c$^+$ CD45$^+$ CD3$^-$ CD56$^-$ NKp46$^-$ CD11b$^-$ CD68$^-$ CD11c$^+$
Other myeloid cells	CD45$^+$ CD3$^-$ CD56$^-$ NKp46$^-$ CD11b$^+$ CD68$^-$ CD11c$^-$
Other CD45$^+$ cells	CD45$^+$ CD3$^-$ CD56$^-$ NKp46$^-$ CD11b$^-$ CD68$^-$ CD11c$^-$

Cell function	Identification
Proliferation	Ki67
Cytotoxicity	Granzyme B (GRZB)
NK cell activation and maturation	CD16, CD57, NKG2D
Immunoregulatory	PD-1, PD-L1, TIM-3, HLA class I
Hypoxia	Carbonic anhydrase IX (CAIX)

B



C

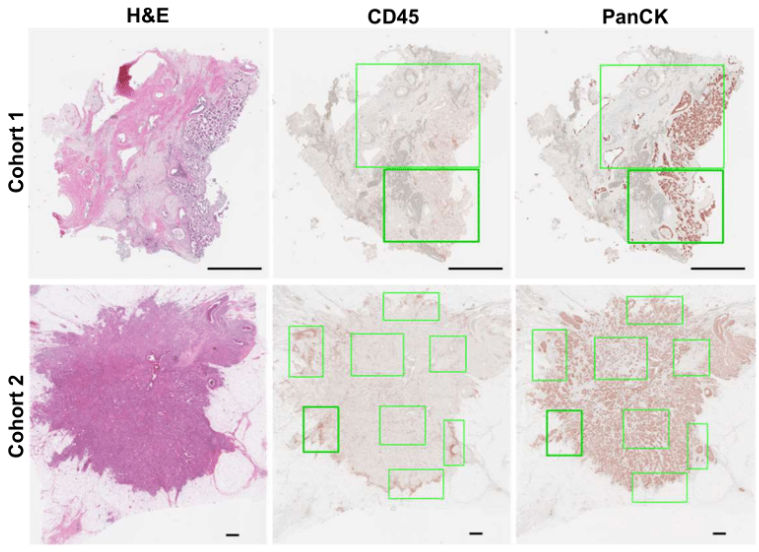


Figure 3.1: Overview of identified cell types and human breast cancer specimens. A) Cell types and corresponding markers that were used to identify the cell types. **B)** Breast cancer cohorts evaluated in this study. HER2 status was assessed by IHC (cohort 1) or FISH (cohort 2). **C)** Example images of H&E sections, CD45, and PanCK staining with the selected ROIs represented by the green boxes. Depending on tissue size, 1-6 ROI per specimen were analyzed in cohort 1 and 1-11 ROI per specimen were analyzed in cohort 2. In all images, scale bar = 1.0 mm. H&E = hematoxylin and eosin, ROI = region of interest.

3.4.2 Paucity of NK cells in breast cancer specimens independent of HER2 status

We next investigated general leukocyte subset heterogeneity by quantifying densities of T cells ($CD45^+$ $CD3^+$ cells), myeloid cells ($CD45^+$ $CD3^-$ $CD56^-$ $NKp46^-$ cells, identified by CD11b, CD68 or CD11c), other $CD45^+$ cells, as well as total NK cells ($CD3^-$ $CD56^+$ $NKp46^-$ cells, $CD3^-$ $CD56^+$ $NKp46^+$ cells, and $CD3^-$ $CD56^-$ $NKp46^+$ cells). Intra- and inter-tumoral heterogeneity was observed, and NK cells were present in low abundance (**Figure 3.2B, Supplementary Figure 3.3C-D**). $CD3^+$ T cells represented the most abundant immune subset in both in $HER2^-$ and $HER2^+$ specimens in cohort 1, followed by other $CD45^+$ cells and myeloid cells without significant differences between $HER2^-$ and $HER2^+$ disease in any of the evaluated cell populations (**Figure 3.2B, Supplementary Figure 3.3C-D**). In cohort 2, $CD3^+$ T cells remained the densest immune cell population (**Figure 3.2B, Supplementary Figure 3.4C-D**), consistent with previously published immune cell compositions in breast cancer cohorts [221, 222]. Further subset analysis within $CD3^+$ T cells revealed that $CD4^+$ and $CD8^+$ T cells were found at comparable cell densities within both the $HER2^-$ and $HER2^+$ specimens in cohort 1 (**Figure 3.2C**). As for $CD3^+$ T cells, cell densities of $CD4^+$ and $CD8^+$ T cells were higher in cohort 2 compared to cohort 1, but $CD4^+$ and $CD8^+$ T cell densities were found at comparable ratios, showing consistency with specimens in cohort 1 (**Figure 3.2C**). Within the myeloid cell fraction, the antibody panel allowed identification of $CD11b^-$ $CD11c^+$ and $CD11b^+$ $CD11b^-$ dendritic cells, $CD11b^+$ $CD68^+$ or $CD11b^-$ $CD68^+$ myelomonocytic cells, and other $CD11b^+$ cells, such as granulocytes; the myeloid subsets revealed neither significant differences between $HER2^-$ and $HER2^+$ specimens in cohort 1, nor between cohorts (**Figure 3.2D**).

NK cell density in comparison to $CD3^+$ T cells and $CD11b^+$ myeloid cells is illustrated in **Figure 3.2E**. NK cell densities ranged between 0 - 16 NK cells/mm^2 in $HER2^-$ specimens and 0 -

53 NK cells/mm^2 in HER2$^+$ specimens of cohort 1 (median of 8 NK cells/mm^2 in both groups, **Figure 3.2F, Supplementary Figure 3.3C-D**). Cell densities of NK cells in HER2$^+$ specimens were comparable in cohort 2, where we detected between 0 – 68 NK cells/mm^2 (median of 6.6 NK cells/mm^2, **Figure 2F, Supplementary Figure 3.4C-D**). NK cell presence was thus not associated with HER2 status, nor with other known tumor characteristics (tumor grade, lymph node status, hormone receptors ER and PR) or neoadjuvant treatment status, as well as CD3$^+$ or myeloid cell densities (**Supplementary Table 3.4**).

3.4.3 NK cells lack indicators of cytotoxicity

We next evaluated NK cell phenotype and functional status in HER2$^-$ versus HER2$^+$ specimens. Of note, the NK cell events were extremely low in some specimens, where the sum of total NK cell events from the three NK cell subsets and from all ROIs per specimen was below 20 events in seven of the 18 specimens in cohort 1, and in three of the 30 specimens in cohort 2 (**Figure 3.2G**). Thus, specimens with fewer than 20 events are indicated with a different color, demonstrating that they did not skew the NK cell phenotype data (**Figure 3.2H**).

Of the assessed activating receptors and functional markers, CD16 (FcγRIIIA) was prominently expressed, although a wide spread of CD16 expression was found (2-88 % of NK cells). The activating receptor NKG2D, the maturation marker CD57, as well as the functional markers Granzyme B (GRZB) and Ki67 were expressed but in a lower frequency of NK cells (average of 0-14%), thus indicating that tumor-infiltrating NK cells were not highly cytotoxic nor strongly proliferating. To test whether the low GRZB levels could be a consequence of functional exhaustion, we examined NK cell positivity for TIM-3 and PD-1. Few NK cells expressed either TIM-3 (14-20%), or PD-1 (0- 23% PD-1$^+$ NK cells), independently of HER2 status in the two cohorts (**Figure 3.2H**). Overall, these results indicate that NK cells in breast

cancer display a rather mature ($CD16^+$), but not terminally-differentiated nor exhausted phenotype combined with low GRZB and Ki67 expression, indicating a non-cytotoxic/proliferative profile. Of note, we observed no distinction between $HER2^-$ and $HER2^+$ specimens for any of the evaluated biomarkers.

3.4.4 The majority of NK cells in tumor stroma were identified as $CD56^-$ $NKp46^+$ NK cells

NK cells are frequently identified as $CD3^-$ $CD56^+$ or as $NKp46^+$ cells because the majority of NK cells in the periphery express both CD56 (NCAM1) and the activating receptor NKp46. However, neither is sufficient alone to identify NK cells as $CD56^-$ and $NKp46^-$ NK cell subsets have been reported [223-227]. To capture the majority of NK cells, we used both CD56 and NKp46 and identified $CD56^+$ $NKp46^-$, $CD56^-$ $NKp46^+$, and $CD56^+$ $NKP46^+$ NK cells. All three NK cell populations were found in the three specimen groups with $CD56^-$ $NKp46^+$ NK cells being the most abundant population (**Figure 3.2I**). Interestingly, highly similar distributions of NK cell fractions were detected in both $HER2^+$ specimen groups: A trend towards a larger fraction of $CD56^+$ $NKp46^-$ NK cells were observed in $HER2^+$ specimens of both cohorts compared to $HER2^-$ specimens of cohort 1, indicating that the $CD56^+$ $NKp46^-$ NK cell subset tended to correlate with HER2 status rather than with the cohorts. Further analysis for the three NK cell subsets was not performed due to low NK cell counts (**Figure 3.2G**).

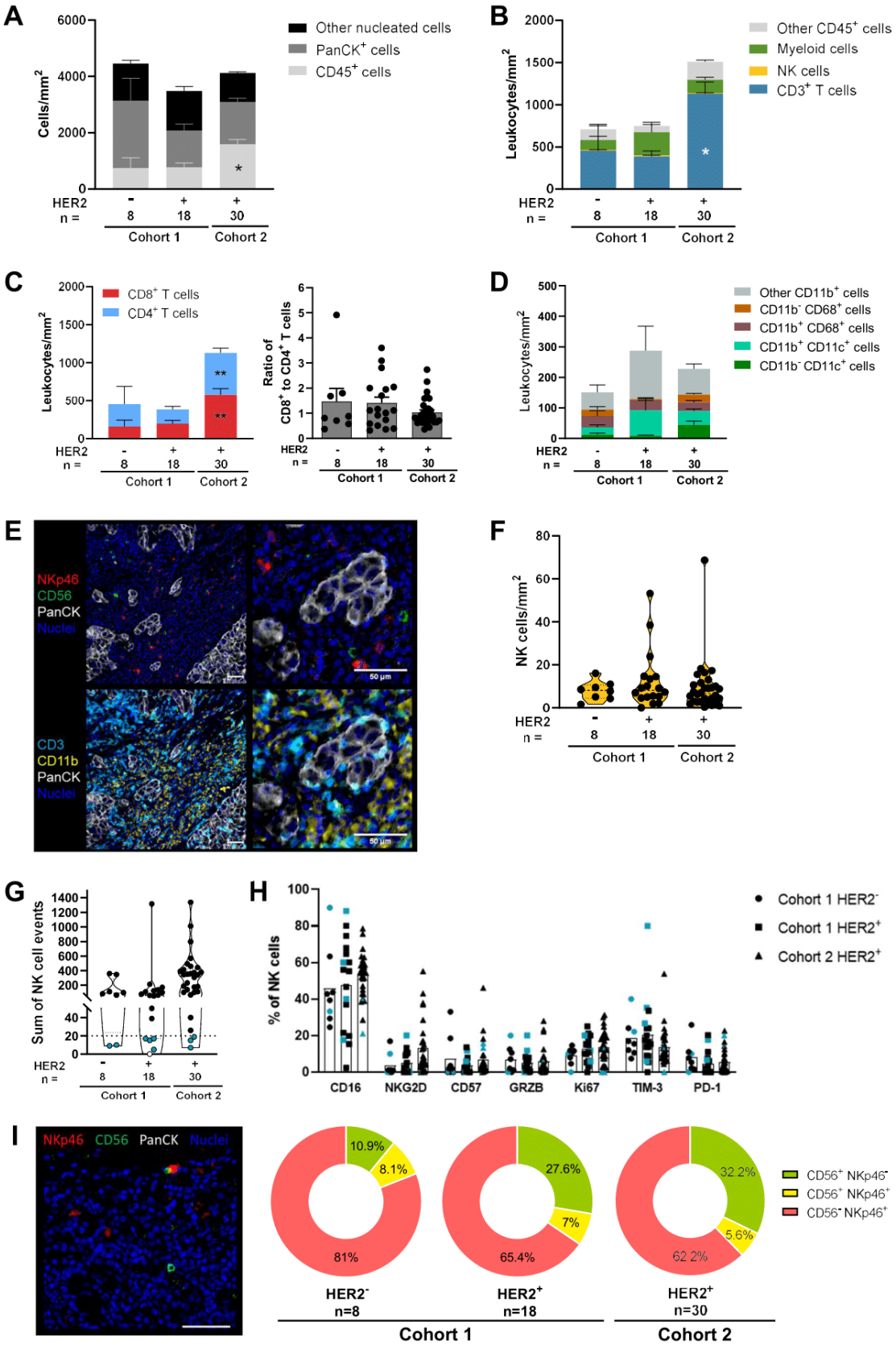

A
Cells/mm2
Other nucleated cells
PanCK+ cells
CD45+ cells
HER2
n =
Cohort 1
Cohort 2
B
Leukocytes/mm2
Other CD45+ cells
Myeloid cells
NK cells
CD3+ T cells
C
CD8+ T cells
CD4+ T cells
Ratio of CD8+ to CD4+ T cells
D
Other CD11b+ cells
CD11b- CD68+ cells
CD11b+ CD68+ cells
CD11b+ CD11c+ cells
CD11b- CD11c+ cells
E
NKp46
CD56
PanCK
Nuclei
CD3
CD11b
50 µm
F
NK cells/mm2
G
Sum of NK cell events
H
% of NK cells
Cohort 1 HER2-
Cohort 1 HER2+
Cohort 2 HER2+
CD16
NKG2D
CD57
GRZB
Ki67
TIM-3
PD-1
I
10.9%
8.1%
81%
27.6%
7%
65.4%
32.2%
5.6%
62.2%
CD56+ NKp46-
CD56+ NKp46+
CD56- NKp46+
HER2-
n=8
HER2+
n=18
HER2+
n=30

Figure 3.2: Immune cell contexture in HER2$^-$ vs. HER2$^+$ breast cancer specimens. A) Cell densities (cells/mm^2) of CD45$^+$ leukocytes, PanCK$^+$ epithelial cells, and other nucleated cells in cohort 1 (HER2$^-$ & HER2$^+$) and in cohort 2 (HER2$^+$). B) Leukocyte densities of CD3$^+$ T cells, NK cells (CD3$^-$ CD56$^+$ NKp46$^-$ cells, CD3$^-$ CD56$^+$ NKp46$^+$ cells, and CD3$^-$ CD56$^-$ NKp46$^+$ cells), myeloid cells (comprising the five cell types detailed in D) and other CD45$^+$ leukocytes. C) CD8$^+$ and CD4$^+$ T cell densities (left) and their ratio (right). D) Dendritic cells (CD11b$^+$ CD11c$^+$ cells, CD11b$^-$ CD11c$^+$ cells), myelomonocytic cells (CD11b$^+$ CD68$^+$ cells, CD11b$^-$ CD68$^+$ cells), and other CD11b$^+$ myeloid cells. Kruskal-Wallis test with Dunn correction was used to determine differences in cell population densities in the three specimen groups. Bars indicate the mean ± SEM per cell population. E) Pseudocolored image of one HER2$^+$ specimen (cohort 2) with a comparatively high density of NK cells, identified by either NKp46 or CD56 (upper panels). Density of CD3$^+$ T cells and CD11b$^+$ myeloid cells of the same region in lower panels, together with PanCK$^+$ epithelial cells and nuclear stain (Hematoxylin) in both images. Scale bars are 50 µm. F) NK cell densities with one dot representing one tumor specimen. Per specimen, the sum of 1-11 ROIs per mm^2 was calculated. G-H) One dot represents one tumor specimen. Those with <20 events are indicated by cyan color. In one specimen (open circle), no NK cells were detected. Sum of NK cell events from all analyzed ROIs (G). Expression of NK cell markers as % of total NK cells (H). I) Image with a CD56$^+$ NKp46$^-$ NK cell, CD56$^-$ NKp46$^+$ NK cells, and CD56$^+$ NKp46$^+$ double-positive NK cells, scale bar = 50 µm (left). The three identified NK cell subpopulations are depicted as % of total NK cells (right).

3.4.5 Single-cell spatial organization of NK cells revealed distinct NK cell phenotypes depending on their proximity to tumor cells and HER2 status

Although NK cells were found in low numbers in most tumor specimens, we observed that some NK cells infiltrated tumor nests, whereas other NK cells were found more distal to neoplastic tumor cells, surrounded by other cells in the TME. An example of a tumor specimen with high NK cell counts illustrates that NK cells are distributed throughout the depicted ROI (**Figure 3.3A**).

Since NK cells were found within PanCK$^+$ tumor nests, as well as distal to PanCK$^+$ cells, we further examined their spatial distribution and asked how NK cells inside or outside tumor nests differed. For this spatial analysis, total NK cells (CD3$^-$ CD56$^+$ NKp46$^-$ cells, CD3$^-$ CD56$^+$ NKp46$^+$ cells, and CD3$^-$ CD56$^-$ NKp46$^+$ cells) from all patients, and of both cohorts were analyzed on a single cell level. As PanCK$^+$ neoplastic epithelial cells can negatively affect NK cells, we first investigated whether proximity of the PanCK$^+$ cells to NK cells affected NK cell phenotypes. Given that NK cells are 5-10 µm and breast cancer cells can reach ~ 30 µm, we considered a 20 µm distance for PanCK$^+$ cells to interact with NK cells and grouped them based

on presence or absence of PanCK$^+$ cells within the set distance of 20 μm. Thus, NK cells were placed into the "proximal" (<20 μm) group when at least one PanCK$^+$ cell was located within a radius of 20 μm from its cell nucleus. NK cells were placed into the "distal" (>20 μm) group when no PanCK$^+$ cell was present within the 20 μm radius.

Both NK cell spatial groups (proximal and distal) were represented in the majority of specimens, further illustrating that not all NK cells were located outside neoplastic tumor nests but that a proportion of NK cells were located in close proximity to PanCK$^+$ neoplastic epithelial cells (**Supplementary Fig. 3.5A-C**). We found that NK cells close to PanCK$^+$ cells tended to be Ki67$^+$, reflecting proliferative status, and expressed the checkpoint molecule PD-1, but fewer expressed TIM-3 (**Figure 3.3B**).

We next examined whether differences in expression on NK cells related to HER2 status. Interestingly, increased presence of NK cells expressing both Ki67 and PD-1 proximal to PanCK$^+$ cells was observed in HER2$^+$ specimens (**Figure 3.3C-D**). Moreover, increased density of NK cells expressing TIM-3 was observed in cells distal to PanCK$^+$ cells in HER2$^-$ specimens, while the difference trended in HER2$^+$ specimens (**Figure 3.3C-D**). Although the modest sample size for the HER2$^-$ group limited conclusions, these findings indicate that NK cells localizing inside and outside tumor nests are phenotypically distinct in HER2$^-$ and HER2$^+$ disease, potentially as a result of their spatial location relative to neoplastic tumor cells and/or other cells in the TME. Subsequently, we further interrogated neoplastic tumor cell phenotype and leukocyte composition as a result of spatial proximity to NK cells.

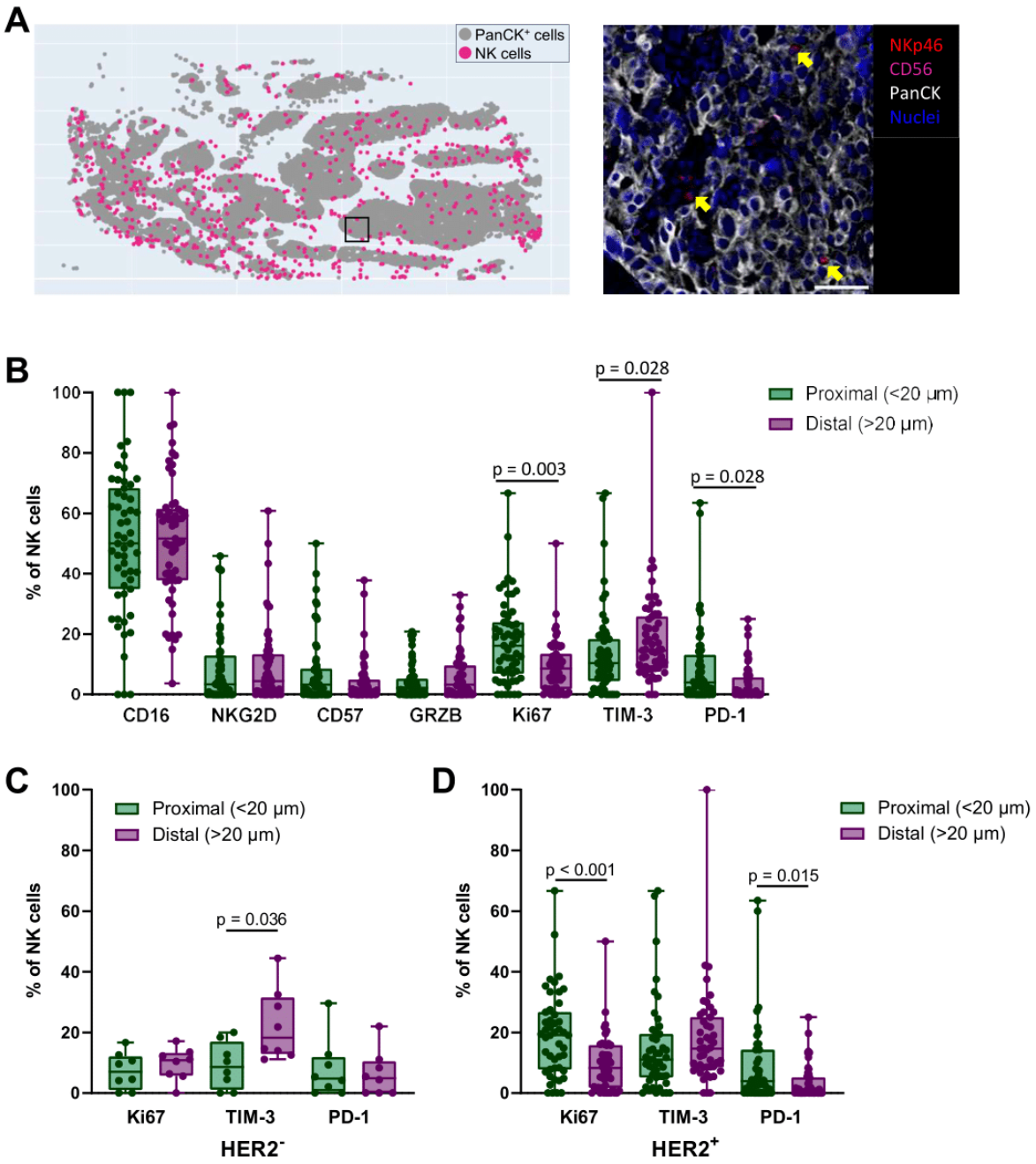


Figure 3.3: Single cell analysis of NK cells results in distinct phenotypes related to the proximity to tumor cells and HER2 status. A) Scatterplot of a specimen with, compared to other analyzed specimens, high density of NK cells, identified by either NKp46 or CD56 (left). Pseudocolored image in higher magnification with NK cells, PanCK$^+$ cells and nuclei (Hematoxylin) with yellow arrows highlighting NK cells in tumor nests; scale bar = 50 µm (right). **B-D)** Single cell distance analysis: Each NK cell from all specimens was placed into either of the two distance groups, based on the distance threshold (20 µm) of each NK cell from PanCK$^+$ cells. Expression of phenotypic and functional markers of all NK cells, grouped as close (< 20 µm) and far (> 20 µm) from PanCK$^+$ cells (**B**). Those markers that were significantly different in (**B**) were split into HER2$^-$ specimens (**C**) and HER2$^+$ specimens (**D**). One dot represents one specimen and the % is calculated of the sum of ROIs per specimen. Boxes represent Q1-Q3 with the line showing the mean and whiskers the smallest and largest data point within the quartile range.

3.4.6 PanCK$^+$ neoplastic cells exhibit extensive heterogeneity in both HER2$^-$ and HER2$^+$ specimens and retain high HLA class I expression in close proximity to NK cells

Because the phenotype of PanCK$^+$ neoplastic cells may influence susceptibility to NK cell-mediated killing, we evaluated HLA class I as a major inhibitory ligand for NK cells, cell proliferation (Ki67), the immune checkpoint ligand PD-L1 and hypoxia (carbonic anhydrase IX, CAIX) on PanCK$^+$ neoplastic cells (**Figure 3.4A**). We observed large variation in HLA-I expression between the individual tumor specimens in all three specimen groups, with slightly increased HLA class I-positive PanCK$^+$ cells in HER2$^+$ specimens (**Figure 3.4A-B**). Proliferation of PanCK$^+$ neoplastic cells was observed in a small fraction of cells in HER2$^-$ specimens (average of 6%). In comparison, higher densities of Ki67$^+$ neoplastic cells were observed in both HER2$^+$ groups (average of 14% cohort 1 and 25% cohort 2), but without significant differences between HER2$^-$ and HER2$^+$ specimens in cohort 1. The percentage of CAIX$^+$ neoplastic cells varied between specimens, and hypoxic areas were detected in a majority of specimens, independent of HER2 status. PD-L1 expression was not observed on the PanCK$^+$ cells in any of the conditions, indicating that the PD-1$^+$ NK cells are likely not inhibited through PD-L1 (**Figure 3.4A**).

Since close contact of NK cells and their target cells is crucial for NK cell-mediated killing, we subsequently asked how tumor cell phenotypes differed with regards to NK cell proximity (**Figure 3.4C**). Using the same distance threshold of 20 μm, PanCK$^+$ tumor cells were placed either in the "proximal" (<20 μm) group when at least one NK cell was located within a 20 μm radius, or they were placed in the "distal" (>20 μm) group when no NK cells were within 20 μm. Although HLA-I expression was highly heterogenous in both groups, the analysis revealed a significantly higher percentage of HLA-I expressing PanCK$^+$ cells proximal to NK

cells (**Figure 3.4C**). The expression of Ki67, PD-L1, and CAIX-positive cells was not significantly different between neoplastic tumor cells proximal or distal to NK cells.

Based on the observations of increased presence of HLA class I-expressing cells in $HER2^+$ specimens, and that these tended to be proximally located to NK cells, we evaluated the difference in HLA-I^+ cells to NK cells based on HER2 status directly. HLA-I^+ neoplastic cells proximal to NK cells trended in $HER2^-$ specimens (**Figure 3.4D**) but was significantly higher in $HER2^+$ specimens (**Figure 3.4E**). Although a larger fraction of $PanCK^+$ neoplastic cells expressed HLA-I within the $HER2^+$ cohort, neoplastic cells close to NK cells expressed HLA-I in higher fractions in both cohorts. These results indicate that HLA-I expression on neoplastic tumor cells is related to spatial proximity to NK cells as well as to HER2 status. As HLA-I molecules are major inhibitory ligands for NK cells, our finding of HLA-I^+ tumor cells close to NK cells supports that HLA-I expression may be a mechanism of breast cancer escape from NK cell-mediated killing.

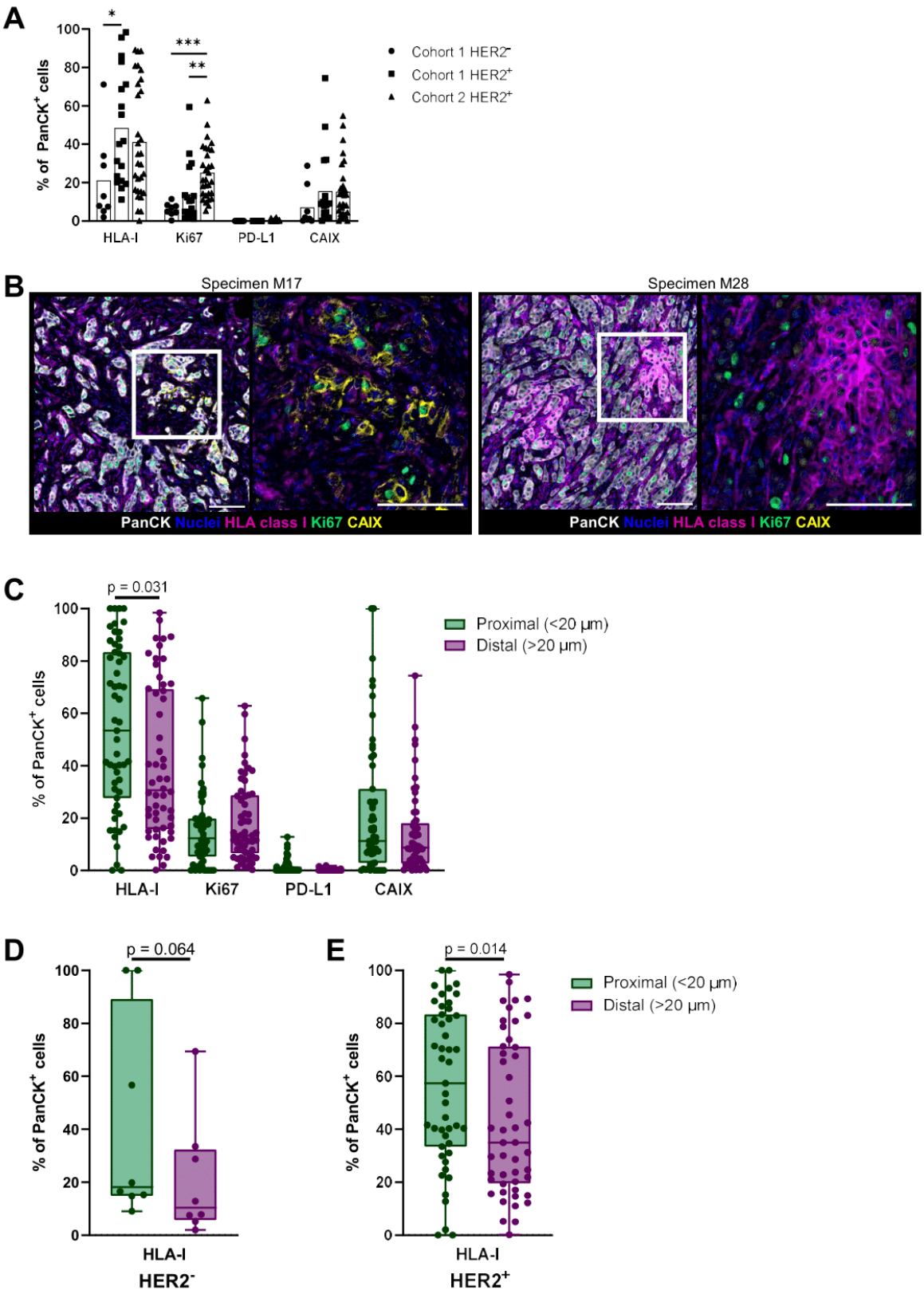


Figure 3.4: Single cell analysis of neoplastic PanCK$^+$ epithelial cells illustrate heterogeneity and high HLA class I expression in close proximity to NK cells. A) Expression of phenotypic markers on PanCK$^+$ cells. **B)** Pseudocolored images showing two examples with heterogeneous staining patterns of HLA class I,

Ki67, and CAIX in combination with nuclear staining (Hematoxylin). Higher magnification is displayed on the right without PanCK. Scale bars = 100 µm. **C-E)** Distance analysis: $PanCK^+$ cells were placed into either of the two groups based on the distance threshold (20 µm) of each $PanCK^+$ cells from NK cells. Expression of phenotypic markers of all $PanCK^+$ cells, grouped as proximal (< 20 µm) or distal (> 20 µm) from NK cells (**C**). Significantly different HLA-I expression was split into $HER2^-$ specimens (**D**) and $HER2^+$ specimens (**E**). One dot represents one specimen and the % is calculated as sum of ROIs per specimen. Boxes represent Q1-Q3 with the line showing the mean and whiskers the smallest and largest data point.

3.4.7 Distinct spatial cellular neighborhoods surrounding NK cells are associated with HER2 biology

Having analyzed the pairwise interactions between NK cells and neoplastic epithelial cells, we sought to identify cells surrounding NK cells beyond 20 µm, to include neoplastic cells as well as other leukocyte subsets. Thus, we performed a cellular neighborhood analysis and set the distance threshold to 60 µm allowing for inclusion of cells close enough to directly touch NK cells, and also accounting for neighboring cells interacting with NK cells through cytokine signaling (**Figure 3.5A**) [133, 135]. All NK cells, identified by either CD56 or NKp46, were considered for this analysis. Any NK cell that did not have a neighbor within the distance threshold was removed from downstream analyses. After calculating the cellular neighborhoods, k-means clustering was used to group the neighborhoods into five clusters, each with unique cellular compositions (**Figure 3.5B**). On average, neighborhoods assigned to cluster 1 and cluster 2 consisted mainly of T cells; cluster 1 neighborhoods were dominated by $CD8^+$ T cells (representing ~ 50% of the cellular makeup), while cluster 2 neighborhoods were characterized by a majority of $CD4^+$ T cells as neighboring cells. Cluster 3 neighborhoods were mainly defined by other $CD45^+$ leukocytes, not further characterized in this study. Neighborhoods in clusters 4 and 5 were predominantly $PanCK^+$ neoplastic cells, although differing in proportion. Cluster 4 neighborhoods represented a mixture of neoplastic tumor cells and immune cell subsets, whereas cluster 5 neighborhoods contained the largest proportion of neoplastic tumor cells (>80% of neighbors). No cluster was dominated by NK cells, and NK cells were rarely found neighboring

each other, but instead distributed across tissue regions. Despite this broad spatial localization of NK cells across a given region, NK cells localized in the T cell dominant neighborhood clusters 1 and 2 were often found outside of tumor nests, whereas NK cells in the tumor-immune cluster 4 were predominantly observed at the tumor-stroma border (**Figure 3.5C**).

Distribution of neighborhood clusters revealed that no single cluster was dominant, and that all five clusters were represented in comparable proportions (**Supplementary Figure 3.7A**). Moreover, most clusters were present in every section (**Figure 3.5D**) and in every ROI (**Supplementary Figure 3.7B**), thus indicating that no cluster was driven by the spatial organization of a single tumor or region. Interestingly, we observed a negative correlation between the presence of the neoplastic tumor dominant neighborhood cluster 5 and the T cell dominant neighborhood clusters 1 and 2 (**Figure 3.5D-E**). This finding indicates that NK cells are unlikely to have both neoplastic tumor cells and T cells as neighbors. Other cluster combinations did not correlate (data not shown). The cellular densities of NK cells were not strongly correlated with densities of other identified immune cell types (**Supplementary Fig. 3.6**), supporting the notion that the described neighborhood clusters are a result of the cellular spatial organization and that the neighborhood clusters are not driven by just the mere presence of the identified immune cell types.

Next, we determined if the proportion of NK cells assigned to each of the neighborhood clusters varied between $HER2^-$ and $HER2^+$ patients. Significant differences were observed for the $CD8^+$ T cell dominant cluster 1, the $CD4^+$ T cell dominant cluster 2, and the $PanCK^+$ tumor cell-dominant cluster 5 (**Figure 3.5F**). A larger fraction of NK cells was attributed to neighborhood cluster 5 in $HER2^-$ patients as compared to $HER2^+$ patients, indicating that NK cells in $HER2^-$ patients were more frequently near neoplastic tumor cells. In $HER2^+$ patients, we

observed a higher fraction of NK cells within the two T cell-dominant clusters 1 and 2, implying that NK cells in $HER2^+$ patients were more often surrounded by $CD8^+$ and $CD4^+$ T cells than NK cells in $HER2^-$ patients. As we detected a higher density of $CD3^+$ T cells in cohort 2 as compared to cohort 1, we repeated the comparison of HER2 status using only cohort 1 patients and found the same trends (**Supplementary Figure 3.7C**), supporting that the finding is related to HER2 biology.

Overall, we revealed that distinct cellular neighborhoods surrounding NK cells exist in human breast cancer. Each neighborhood contained unique compositions, with NK cells localizing more frequently in the tumor-dominant neighborhood cluster in the $HER2^-$ specimens, whereas NK cells localized in the two T cell-dominant neighborhood clusters more often in the $HER2^+$ specimens. These results highlight the importance for considering spatial architecture of cells surrounding NK cells when planning treatments of $HER2^-$ and $HER2^+$ breast cancer patients. Specifically, immunotherapies that leverage the coordinated anti-tumor function of both NK cells and T cells may be beneficial for $HER2^+$ patients especially, given the spatial proximity of T cells to NK cells in these patients.

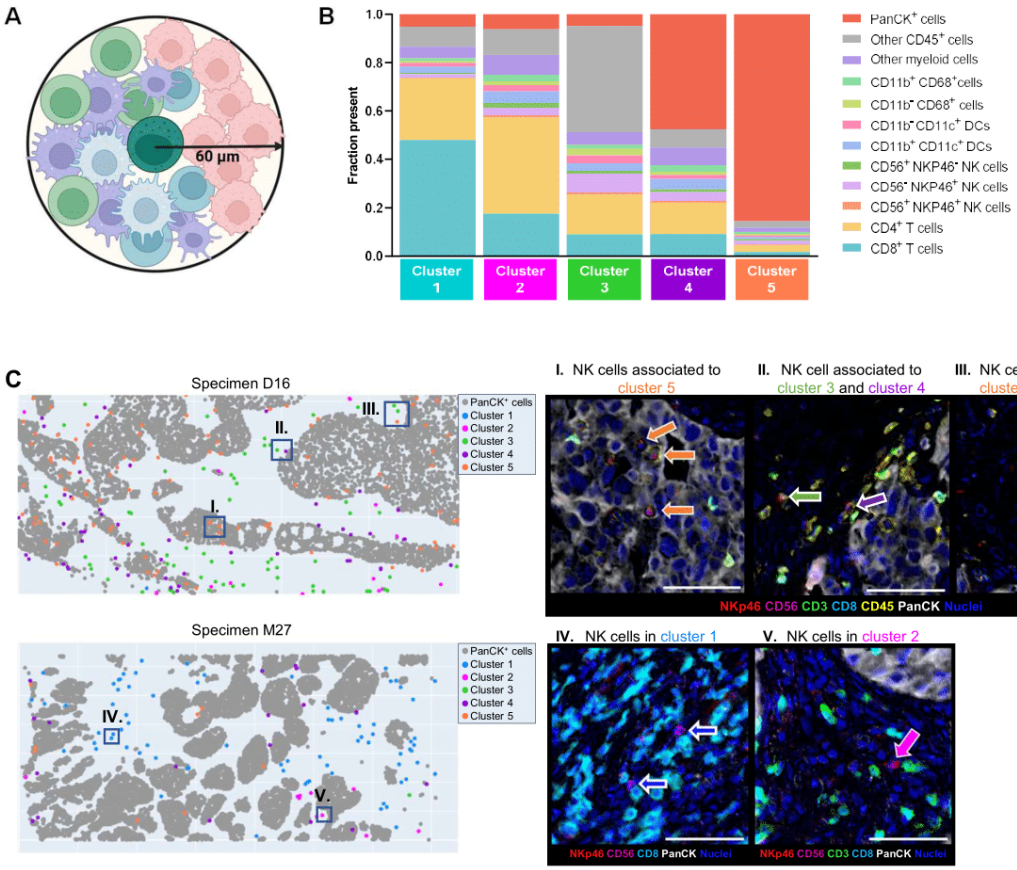

A
60 µm
B
Fraction present
1.0
0.8
0.6
0.4
0.2
0.0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
PanCK+ cells
Other CD45+ cells
Other myeloid cells
CD11b+ CD68+ cells
CD11b- CD68+ cells
CD11b- CD11c+ DCs
CD11b+ CD11c+ DCs
CD56+ NKP46- NK cells
CD56- NKP46+ NK cells
CD56+ NKP46+ NK cells
CD4+ T cells
CD8+ T cells
C
Specimen D16
PanCK+ cells
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
I.
II.
III.
I. NK cells associated to cluster 5
II. NK cell associated to cluster 3 and cluster 4
III. NK cells associated to cluster 5 and cluster 3
NKp46 CD56 CD3 CD8 CD45 PanCK Nuclei
Specimen M27
IV.
V.
IV. NK cells in cluster 1
V. NK cells in cluster 2
NKp46 CD56 CD8 PanCK Nuclei
NKp46 CD56 CD3 CD8 PanCK Nuclei
D
Fraction present
1.0
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
Specimen
E
Correlation
r = -0.8369
Cluster 5
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
T cell dominant clusters 1+2
F
Fraction present
1.0
0.8
0.6
0.4
0.2
0.0
p = 0.010
p = 0.007
p = 0.010
HER2-
HER2+
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5

Figure 3.5: Cellular neighborhood clustering of NK cells. A) Cellular neighborhood definition. A circle of 60 µm radius was drawn around each NK cell, identified by either NKp46 or CD56. The cells inside the circle were counted as neighbors of NK cells. **B)** All specimens were pooled and analyzed for NK cell neighborhood clusters. The average cellular composition of each of the five NK cell neighborhood clusters is shown. **C)** Left: Scatterplots of two specimens with comparatively high NK cell counts to illustrate the distribution of the NK cell neighborhoods clusters. PanCK$^+$ cells are shown in grey together with all NK cells, which are colored by the cluster they were assigned to. Right: Corresponding pseudocolored mIHC images of each NK cell neighborhood cluster. Scale bars = 60 µm. **D)** The average proportion of each NK cell neighborhood cluster is shown for each patient, except for one patient of cohort 1, in which no NK cells were detected. Specimens are sorted by proportions of cluster 5 and color of specimen ID represents HER2 status (HER2$^-$ in blue, HER2$^+$ in red). **E)** Negative correlation of the T cell dominant clusters 1 and 2 and the tumor dominant cluster 5. **F)** Cellular composition of the NK cell neighborhood clusters for HER2$^-$ and HER2$^+$ specimens of both cohorts.

3.5 Discussion

Improved understanding of NK cells and their spatial distribution in the TME of breast cancer can guide future design of successful NK cell-based immunotherapies. In this study, we developed an mIHC panel to analyze NK cells on a single-cell spatial proteomics level and to evaluate NK cell density, phenotype and spatial architecture in relation to tumor and tumor-infiltrating leukocytes; a summary of the key findings of this study are depicted in **Figure 3.6**. The mIHC methodology is a powerful platform, as we were able to quantify immune cell populations and phenotypes on a single cell level, while maintaining the tissue's spatial architecture. Previously, our group demonstrated that comparable results were achieved with quantitative image cytometry and quantification by flow cytometry [121]. In this study, CD3$^+$ T cells were by far the most dominant immune cells present in all analyzed groups and the identified immune cell contexture was largely comparable with the immune cell compositions reported previously in breast cancer by flow cytometry analysis [221, 222].

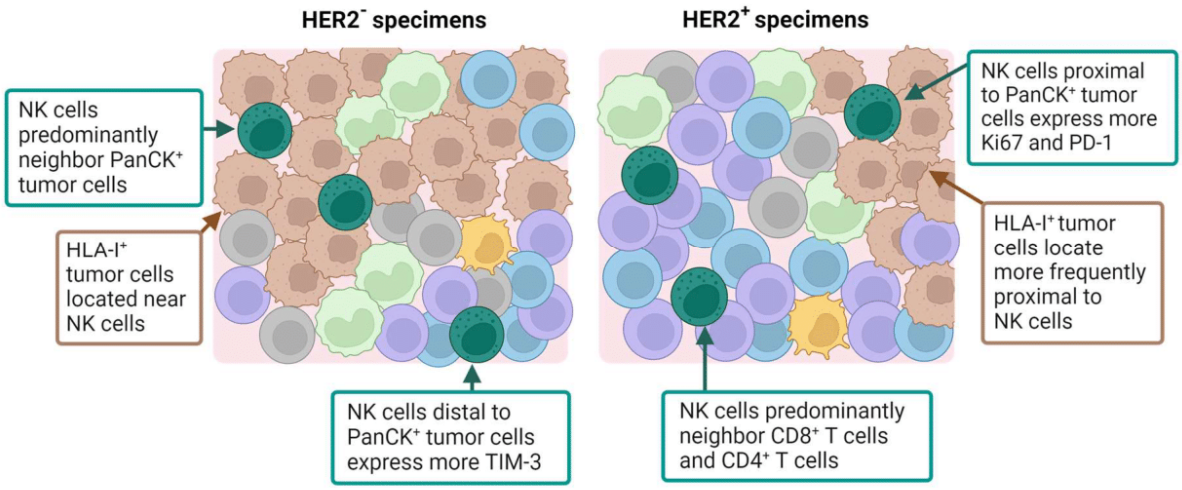


Figure 3.6: Summary of the NK cells' spatial organization in the tumor microenvironments of two breast cancer cohorts. Depicted are the significantly different findings between HER2$^-$ and HER2$^+$ specimens reported in this study.

We revealed NK cells in low frequency and without consistent differences in NK cell densities in HER2$^-$ versus HER2$^+$ specimens. Although the HER2$^-$ group contained only eight specimens, our findings indicate that the overall NK cell density is likely not influenced by HER2 status. While we focused on the role of HER2 in this study, other factors such as various molecular subtypes in cohort 1 could play a role in cellular densities of evaluated leukocytes subtypes including NK cells. Although the number of specimens per molecular subtype was low (n = 1-10 per subtype), cellular densities did not seem to differ between groups of molecular subtypes (data not shown). Moreover, the NK cell densities did not appear associated with other known tumor characteristics in the evaluated cohorts. In previous reports, higher NK cell densities were detected in ER$^-$ as compared to ER$^+$ tumors, and in lymph node-negative compared to lymph node-positive tumors [228, 229]. The differences between our study and previous reports illustrate the heterogeneity between breast cancer cohorts and could depend on

the single timepoint analyzed in each study, which does not comprehensively reflect the dynamics of the TME.

To detect NK cells, we evaluated expression of CD56 and NKp46, and found the majority of NK cells expressed the activating receptor NKp46, consistent with a report describing that NKp46 was largely retained on several solid tumors [36]. The activating receptor CD16, generally expressed on mature $CD56^{dim}$ NK cells, was expressed on all specimens except two, but with rather low frequency in some specimens. Identification of both $CD16^{+}$ and $CD16^{-}$ NK cells inside tumors has been reported previously [36, 230]. Preferred recruitment of $CD16^{-}$ $CD56^{bright}$ NK cells to breast tissues might be one possibility for the observed heterogeneous CD16 expression on tumor-infiltrating NK cells [35]. Alternatively, the receptor can be cleaved by metalloproteases following NK cell activation [231]. The heterogeneous expression of CD16 could have consequences for effectiveness of ADCC-mediating antibodies since CD16 levels were reported to positively correlate with the potency of ADCC [232]; and for therapeutic purposes, the underlying mechanism of heterogeneous CD16 expression should be unraveled. Further, NK cells were not terminally differentiated in this study based on low detection of CD57. One study reported that $CD57^{+}$ NK cells were reduced in breast cancer tissue compared to peripheral blood, indicating limited homing to tumor or decreased survival of $CD57^{+}$ tumor-infiltrating NK cells [233]. Moreover, NK cells in the cohorts evaluated herein did not abundantly express GRZB and Ki67, indicating a lack of cytotoxicity and proliferative capacity. Our image analysis approach detected GRZB expression when the signal was associated with the cell; GRZB could therefore have been released into the TME before tumor tissue was resected. While NK cells are mostly known for their anti-tumor properties, the $NKp46^{+}$ NK cells may also belong to the group of non-cytotoxic innate lymphoid cells type 1 (ILC1) [234]. The conversion

of NK cells to ILC1 has been suggested as a tumor immunoevasion strategy, although functions of ILC1 are highly variable, depending on the cytokines present in the TME [234, 235]. Additional makers would be required to distinguish NK cells from ILCs in our study.

Previous studies reported that NK cells can be confined to tumor stroma rather than infiltrating neoplastic cell nests [236, 237]. We observed both NK cells "proximal" (<20 µm) and "distal" to neoplastic tumor cells (>20 µm). Higher densities of proliferating and PD-1$^+$ NK cells, and fewer expressing TIM-3 "proximal" to PanCK$^+$ cells as compared to NK cells "distal" to PanCK$^+$ neoplastic cells indicate the likelihood that the TME shapes NK cell phenotype and spatial organization. On PanCK$^+$ neoplastic cells, we found not only HLA class I expression with large intertumoral heterogeneity, in line with previous publications [238, 239], but also increased presence of HLA class I-expressing cells in close proximity to NK cells. This finding indicates that neoplastic tumor cells may retain HLA class I expression in areas where NK cells are localized, potentially as a tumor escape mechanism. Independent of this, our observation provides a rationale for interfering with HLA class I-mediated inhibition e.g., by selecting donor NK cells lacking the corresponding KIR receptor for the HLA type expressed by breast tumors (KIR-HLA ligand mismatched NK cells). Alternatively, sufficient NK cell activating signals must be provided to overcome HLA-mediated inhibition. Similar to the heterogeneous HLA-I expression, heterogeneous CAIX expression indicated that hypoxia occurred in local areas of tumor tissues. The result that NK cells close to neoplastic tumor cells were not associated with higher CAIX expression indicates that NK cells can infiltrate independent of hypoxia. However, NK cell functions are known to be influenced negatively by the presence of hypoxia and should therefore be considered for NK cell-targeting therapies [240]. PD-L1 expression was not detected

on PanCK$^+$ cells, in agreement with published data showing that PD-L1 expression on tumor cells was uncommon in breast cancer [241, 242].

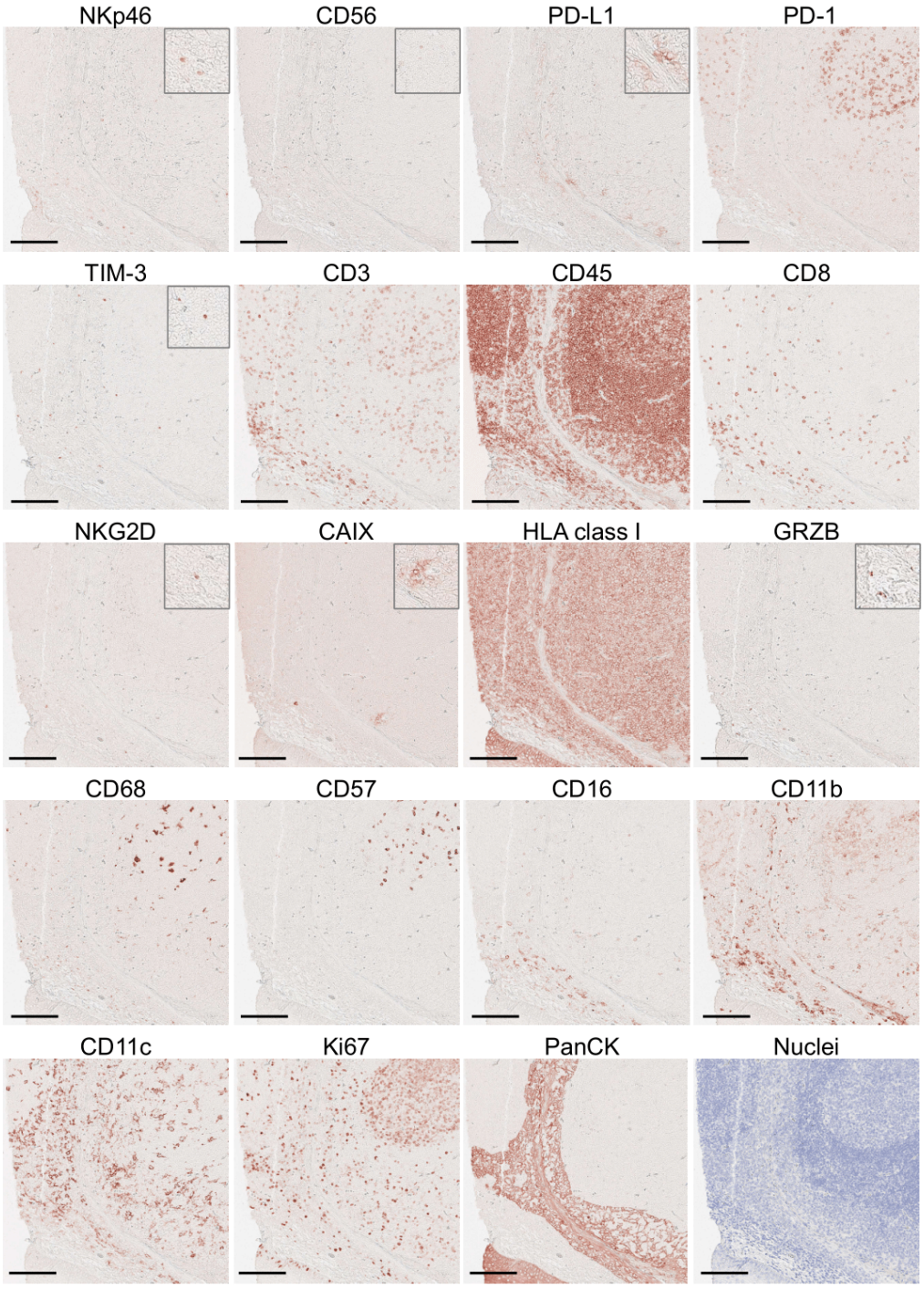
We further set out to understand the cellular spatial organization surrounding NK cells, beyond directly adjacent neoplastic tumor cells, and identified five distinct NK cell neighborhood clusters. Most strikingly was the observed negative correlation between NK cell neighborhoods consisting of predominately CD3$^+$ T cells (cluster 1 and 2) and the neoplastic tumor neighborhood (cluster 5) indicating that NK cells are predominantly surrounded by neoplastic tumor cells or T cells but not both. Our findings are consistent with a study by Namara *et al.*, that reported NK cells mostly in tumor areas, while T cells were predominantly found in the surrounding TME of treatment-naive breast cancer [243]. As we observed NK cells near T cells, a potential therapeutic approach could aim to stimulate NK cells to orchestrate an adaptive T cell-mediated immune response through cytokine secretion. A study with HER2$^+$ patients receiving Trastuzumab supports the finding that both NK cells and T cells are major contributors to observed anti-tumor responses [228]. Since NK cells and T cells are both cytotoxic effector cells, future studies may aim to interrogate the spatial distribution from the perspective of T cells, however, this is beyond of the scope of this study.

Together, this study provides detailed spatial analysis of NK cells in relation to neoplastic tumor cells as well as surrounding leukocyte subsets and describes unique NK cell neighborhoods. The fact that cellular neighborhoods of NK cells were found to be associated with HER2 biology highlights the importance of considering the spatial organization of the TME and may provide a basis to guide the design of more potent NK cell-targeted therapeutic approaches for breast cancer.

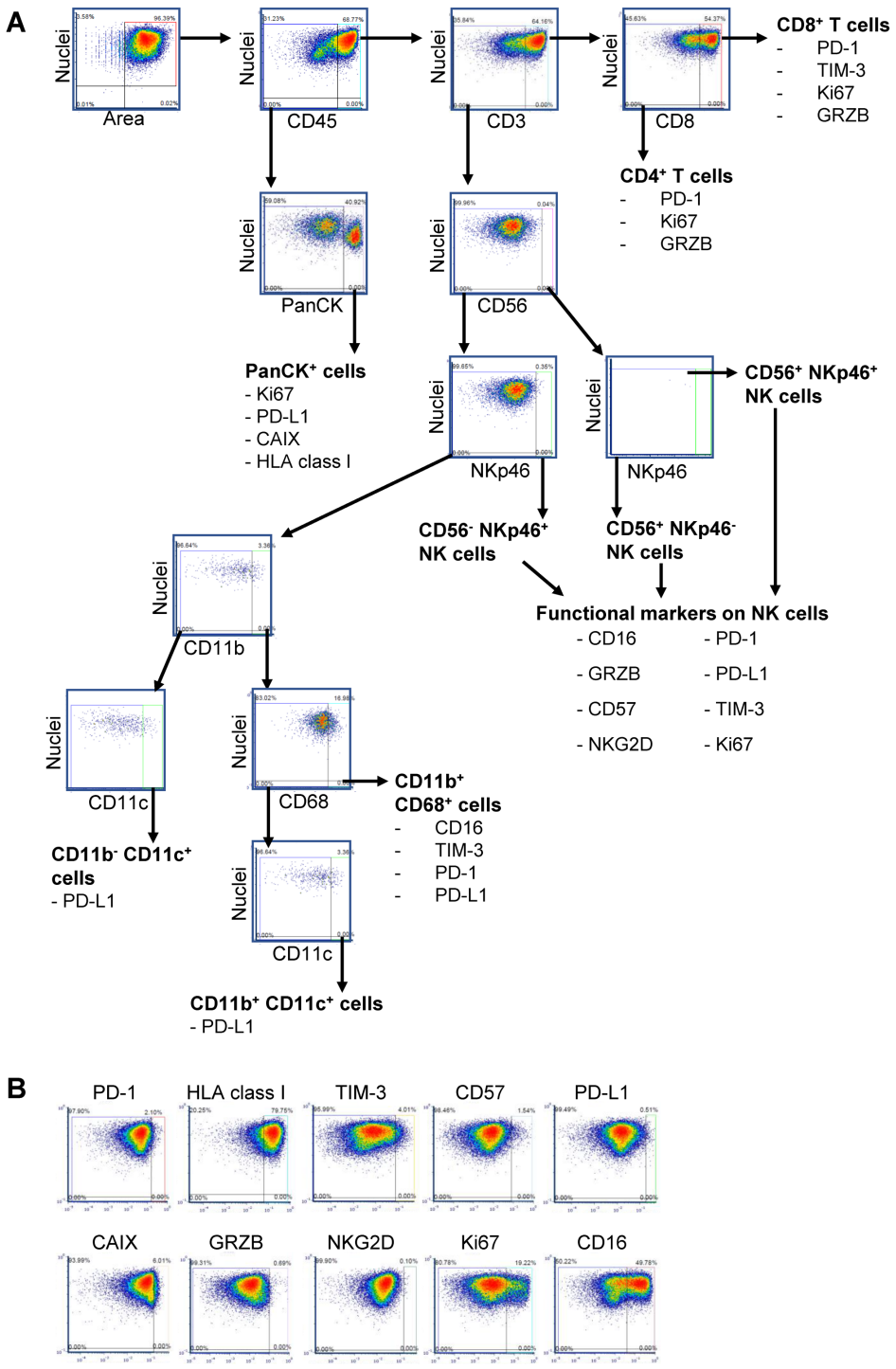
Acknowledgements

We thank the mIHC team of the Coussens laboratory for their assistance. F.A.I. Ehlers is grateful for the support by a TEFAF fellowship, awarded by GROW - School for Oncology and Reproduction, Maastricht University. G.M.J. Bos and L. Wieten received support from Kankeronderzoeksfonds Limburg (KOFL). K.E. Blise acknowledges funding from the National Cancer Institute (NCI) of the National Institutes of Health under award number T32CA254888. E. S. Hwang received funding from NIH/NCI (U2CCA233254-01, R01CA222508). L. M. Coussens and E.S. Hwang acknowledge support from the Susan G Komen Foundation, and LMC acknowledges support from the National Foundation for Cancer Research, and the Knight Cancer Institute at OHSU.

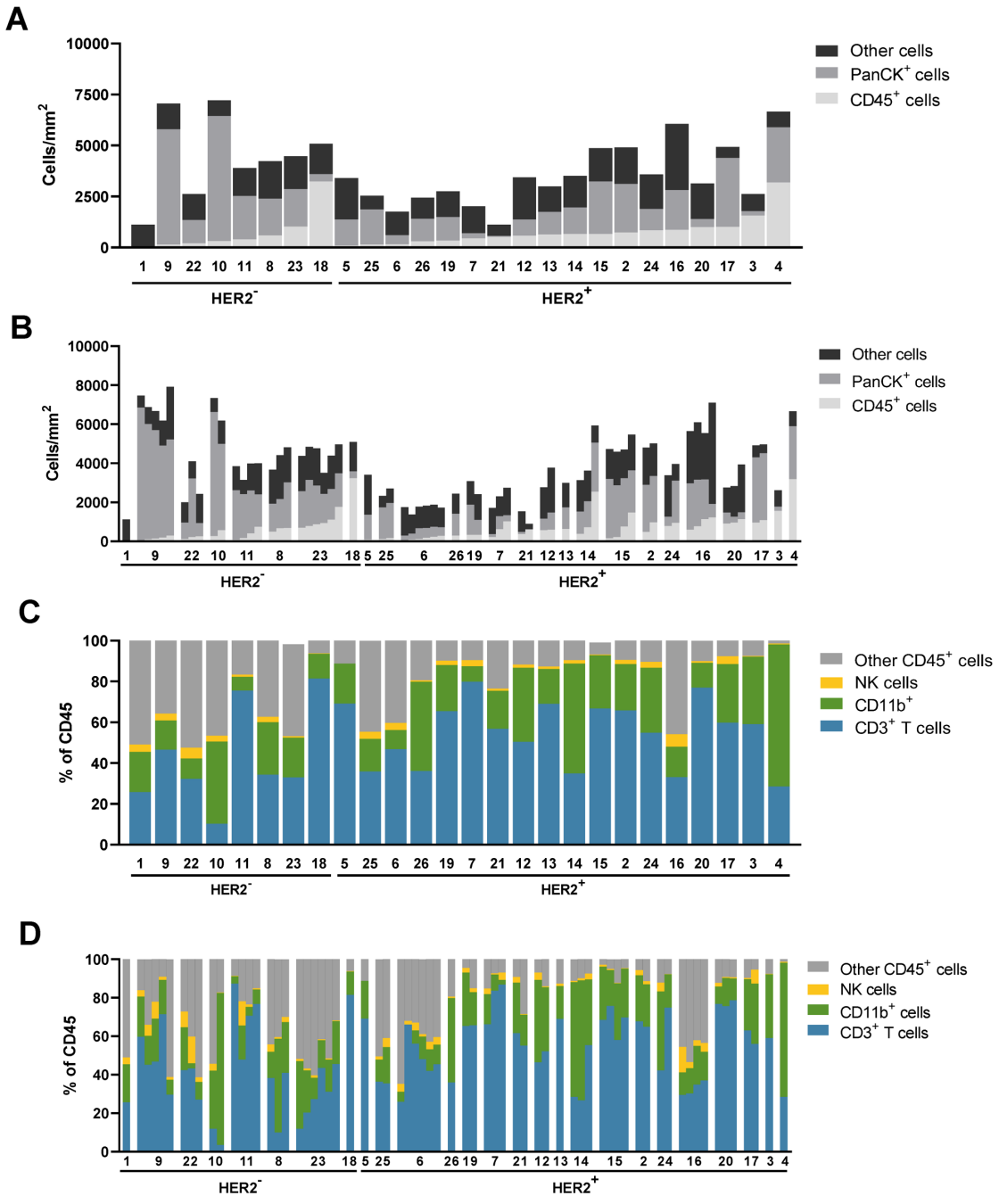
3.6 Supplementary Data



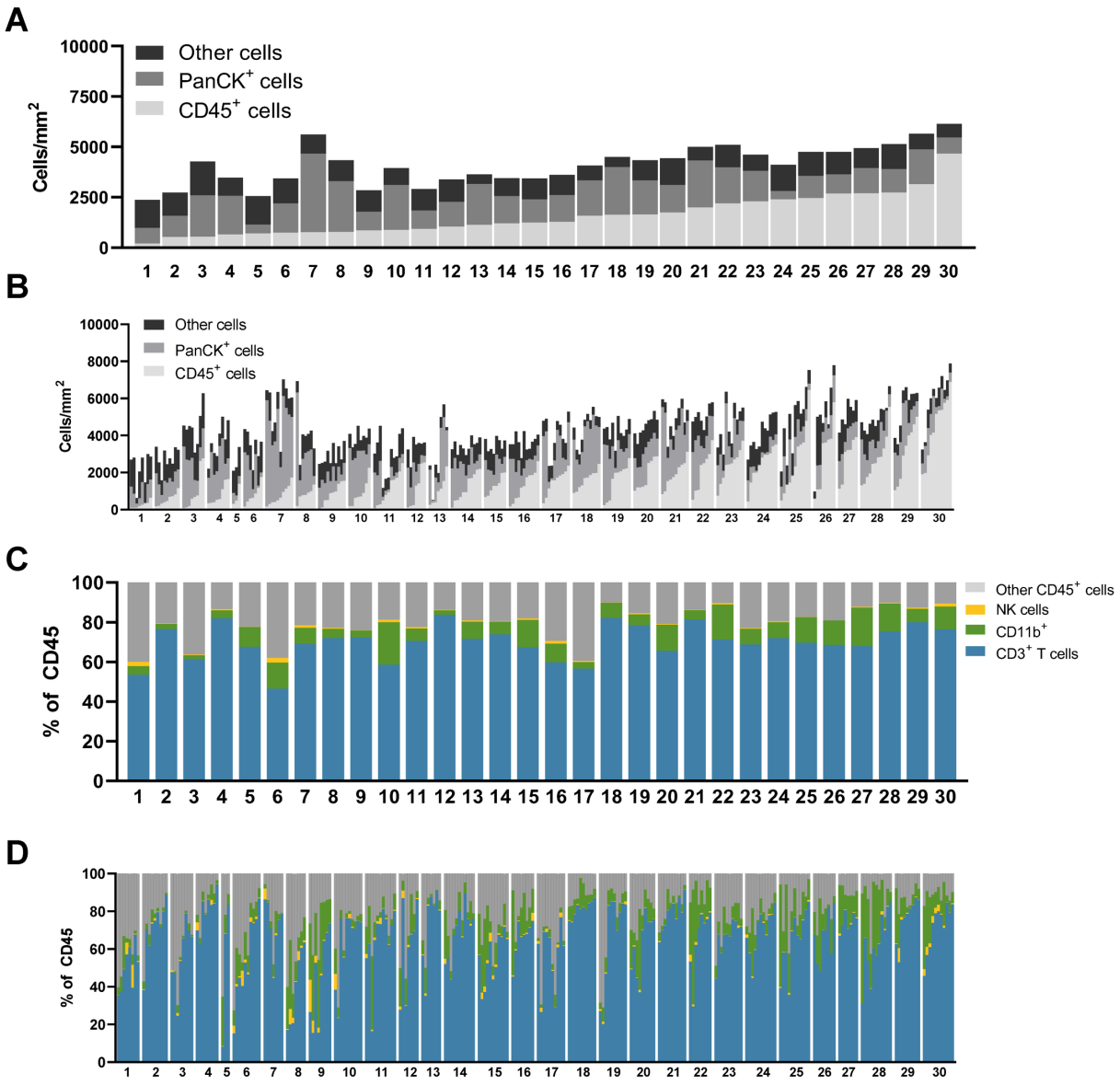
Supplementary Figure 3.1: Sequential staining with the mIHC panel. Visualization of every antibody in its respective position in the mIHC panel with AEC chromogen on human tonsil. The same ROI is shown for all antibodies. Insets with higher magnification are shown for antibodies with low expressed targets. Scale bar = 200 µm.



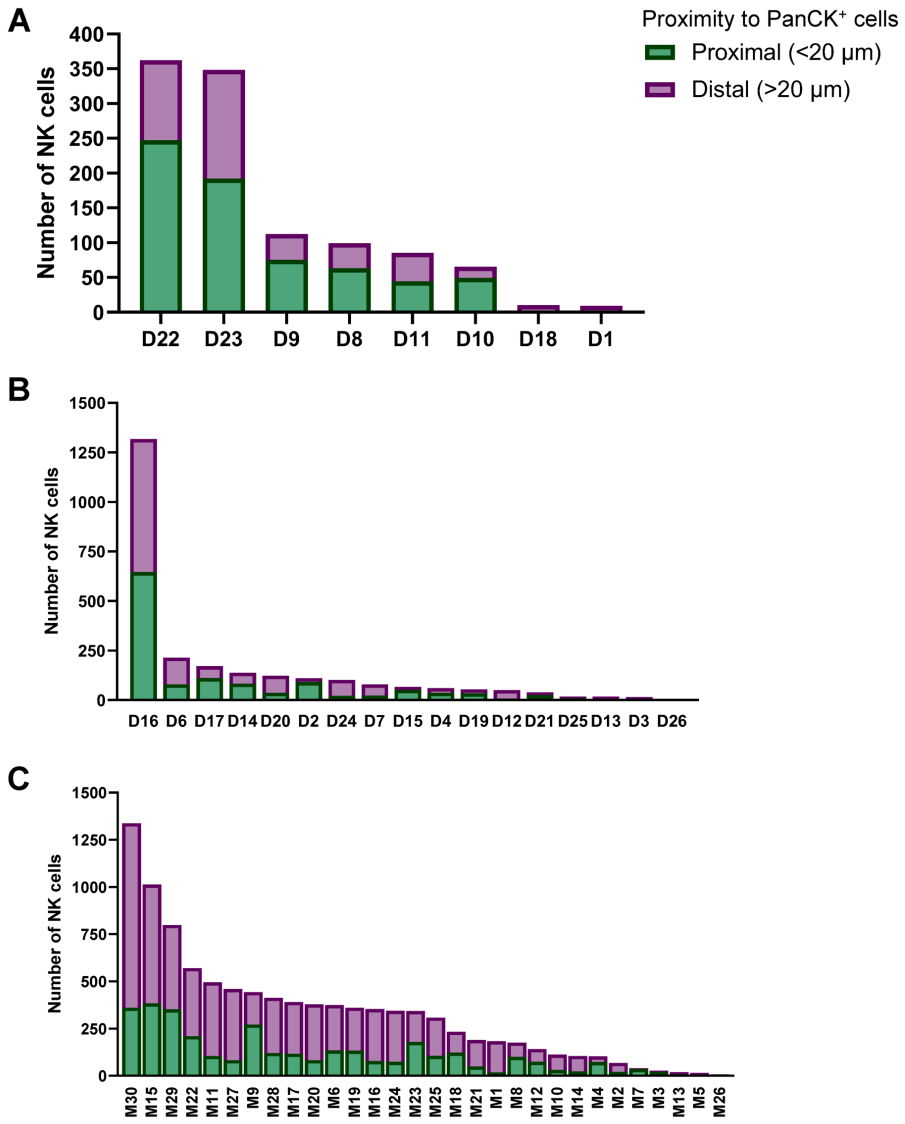
Supplementary Figure 3.2: Multiparameter cytometric image analysis to quantify the multiplex IHC. A) Image cytometry-based gating strategy in FCS Express to quantify the cell types shown in **Figure 3.1A**. Biomarker combination to identify the indicated cell types are listed in the figure with the functional biomarkers indicated below each cell type. **B)** Gating thresholds for the functional biomarkers were set on all cells and subsequently applied to the respective cell types in **A**.



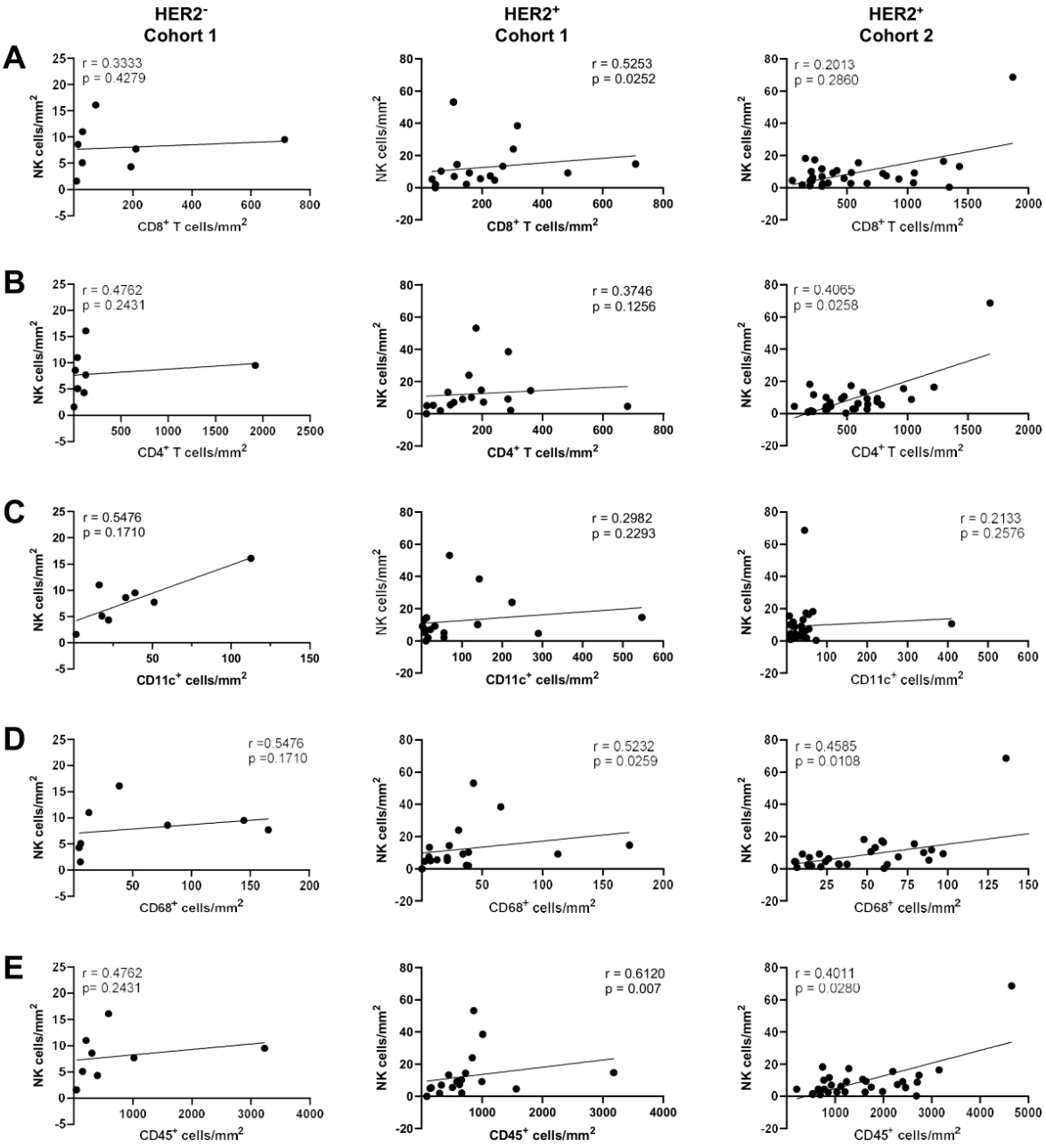
Supplementary Figure 3.3: Intra- and interpatient heterogeneity in cellular composition of HER2$^-$ and HER2$^+$ in cohort 1 (n=26). A) Cell densities (cells/mm^2) per tumor specimen. One bar depicts the sum of ROIs per specimen. Within the HER2$^-$ and HER2$^+$ group, specimens were sorted by CD45$^+$ leukocyte density. **B)** Cell densities (cells/mm^2) of individual ROIs are depicted per tumor specimen. ROIs within one specimen are sorted by CD45$^+$ cell density. **C-D)** Cell populations depicted as percentage of total leukocytes per tumor specimen **(C)** and per ROI **(D)** in the same order as in **A)** and **B)**.



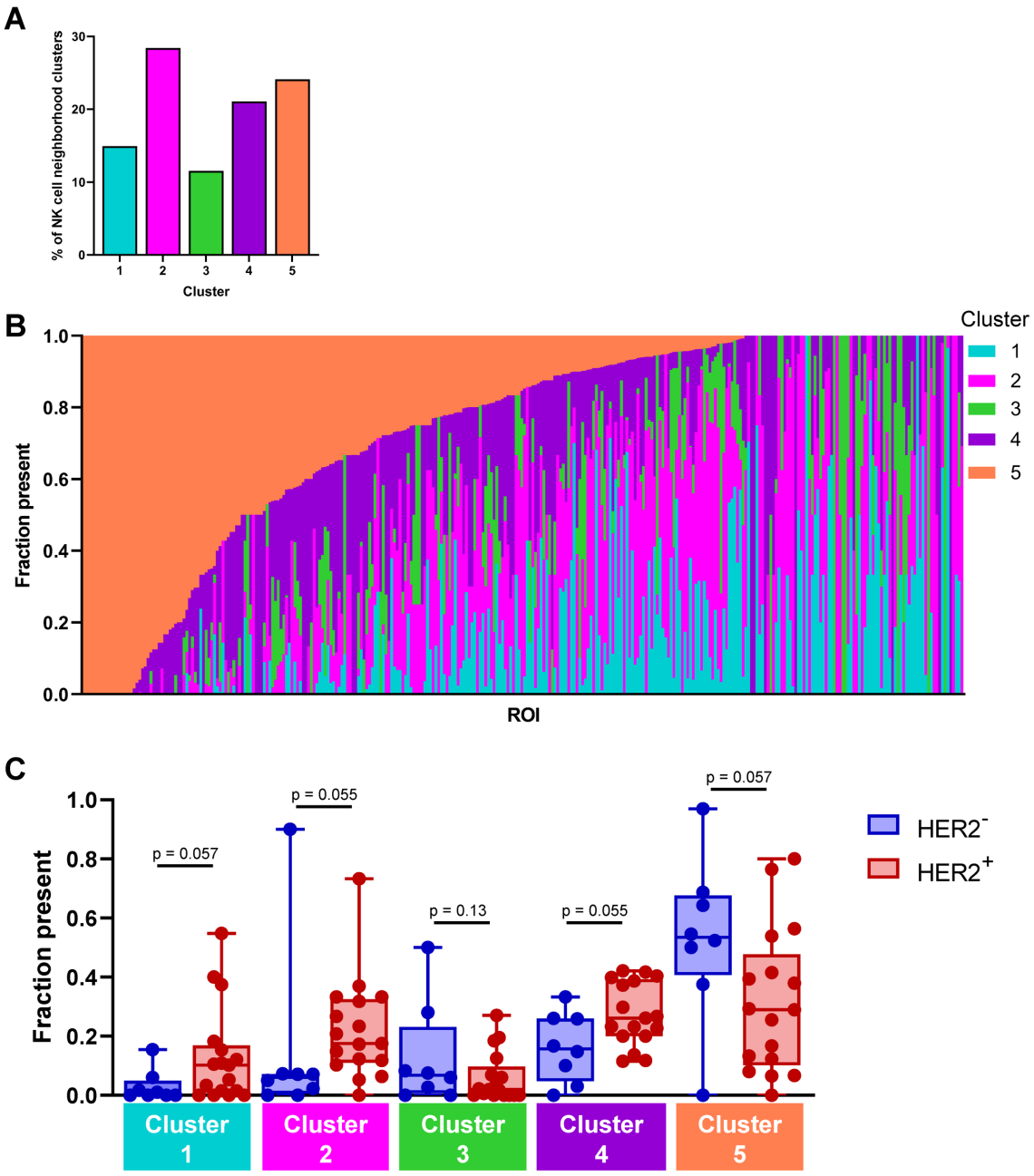
Supplementary Figure 3.4: Intra- and interpatient heterogeneity in cellular composition of HER2$^+$ specimens in cohort 2 (n=30). A) Cell densities (cells/mm^2) per tumor specimen. One bar depicts the sum of ROIs per specimen and specimens were sorted from low to high CD45$^+$ leukocyte density. **B)** Cell densities (cells/mm^2) of individual ROIs are depicted per tumor specimen. ROIs within one specimen are sorted by CD45$^+$ cell density. **C-D)** Cell populations depicted as percentage of total leukocytes per tumor specimen **(C)** and per ROI **(D)** in the same order as in **A)** and **B)**. Legend for **D** is the same as in **C**.



Supplementary Figure 3.5: Single cell distance analysis of NK cells depicted per tumor specimen. Single cell distance analysis: Each NK cell from all specimens was placed into either of the two distance groups, based on the distance threshold (20 µm) of each NK cell from PanCK$^+$ cells. Number of NK cells grouped as proximal (< 20 µm) and distal (> 20 µm) from PanCK$^+$ cells are depicted for HER2$^-$ specimens in cohort 1 **(A)**, HER2$^+$ specimens in cohort 1 **(B)**, and specimens in cohort 2 (HER2$^+$, **C**).



Supplementary Figure 3.6: NK cell density did not strongly correlate with leukocyte density in HER2$^+$ tumor specimens. A-E) Spearman correlations of NK cell densities versus CD8$^+$ T cell densities **(A)**, CD4$^+$ T cell densities **(B)**, total CD11c$^+$ cell densities (CD11b$^+$ CD11c$^+$, CD11b$^-$ CD11c$^+$) **(C)**, total CD68$^+$ cell densities (CD11b$^+$ CD68$^+$, CD11b$^-$ CD68$^+$) **(D)**, and total CD45$^+$ leukocytes **(E)** are shown with estimated regression lines in black in the HER2$^-$ specimens of cohort 1 (n = 8, left graphs), the HER2$^+$ specimens of cohort 1 (n = 18, middle graphs), and the HER2$^+$ specimens of cohort 2 (n = 30, right graphs).



Supplementary Figure 3.7: Cellular neighborhood clustering of NK cells. NK cells of all specimens in both cohorts were pooled and each NK cell that had at least one neighbor within 60 µm was analyzed, resulting in 13,329 NK cells analyzed. **A)** Distribution of NK cells assigned to each of the 5 clusters. **B)** The average proportion of each NK cell neighborhood cluster is shown for each ROI. Samples are sorted by proportions of cluster 5. **C)** The cellular composition of the NK cell neighborhood clusters is shown for the HER2$^-$ and HER2$^+$ specimens of cohort 1.

Supplementary Table 3.1: Patient and tumor characteristics of the two cohorts.

	Cohort 1	Cohort 2
Number of subjects	26	30
Age in years, median (Q1, Q3)	57 (46.3-70.3)	56.5 (47.3-67.8)
Tumor size in cm, median (Q1, Q3)	3.4 (2.1-5)	2.6 (2.2-3)
Tumor grade, n (%)		
1	4 (15%)	0
2	11 (42%)	7 (23%)
3	11 (42%)	23 (77%)
Lymph node status		
positive	13 (50%)	22 (73%)
negative	12 (46%)	8 (27%)
unknown	1 (4%)	0
Estrogen receptor status		
positive	20 (77%)	17 (57%)
negative	6 (23%)	13 (43%)
Progesterone receptor status		
positive	16 (62%)	15 (50%)
negative	10 (38%)	15 (50%)
HER2 status by IHC/FISH		
positive	18 (69%)	30 (100%)
negative	8 (31%)	0
PAM50 subtyping		
Basal	4	N/A
Luminal A	4	N/A
Luminal B	10	N/A
Multifocal	3	N/A
HER2	2	N/A
Normal-like	1	N/A
Not available	2	N/A
Neoadjuvant treatment, n (%)		
Naïve	14 (54%)	30 (100%)
Treated	12 (46%)	0
Metastasis		
yes	unknown	4 (13%)
no	unknown	26 (87%)
Vital status		
living	unknown	27 (90%)
deceased	unknown	3 (10%)

Q1 = 25th percentile, Q3 = 75th percentile, N/A = not available

IHC = immunohistochemistry, FISH = fluorescence in situ hybridization

Supplementary Table 3.2: mIHC antibody information.

	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6	Cycle 7	Cycle 8	Cycle 9	Cycle 10	Cycle 11	Cycle 12
Antigen retrieval	EDTA Pressure cooker 15 min, 115°C	Citrate Steamer 20 min	Citrate Pressure cooker 20 min, 115°C	Citrate Steamer 20 min	EDTA Steamer 20 min	Citrate Steamer 20 min	Citrate Steamer 20 min	Citrate Steamer 20 min	Citrate Steamer 20 min	Citrate Steamer 20 min	Citrate Steamer 20 min	Citrate Steamer 20 min
	Round 1	**Round 1**	**Round 1**	**Round 1**	**Round 1**	**Round 1**	**Round 1**	**Round 1**	**Round 1**	**Round 1**	**Round 1**	**Round 1**
Primary Ab	Nkp46	PD-L1	PD-1	CD3	CD8	CAIX	GRZB	CD57	CD11b	CD11c	Ki67	αSMA
Clone	195314	E1L3N	NAT105	SP7	SP16	Polyclonal	Polyclonal	HNK-1	EPR1334	AP1347Y	SP6	Polyclonal
Vendor	R&D	Cell Signaling	Abcam	Thermo Fisher	Abcam	Novus	Abcam	Thermo Fisher	Abcam	Abcam	Cell Marque	Abcam
Catalog Number	MAB1850	13684S	ab52587	MA1-90582	ab101500	NB100-417	ab4059	MA5-11605	ab133357	ab52632	275R-14	ab5694
Host species	Mouse	Rabbit	Mouse	Rabbit	Rabbit	Rabbit	Rabbit	Mouse	Rabbit	Rabbit	Rabbit	Rabbit
Conc.	1:50	100	50	150	100	200	200	100	2000	50	200	200
Reaction	O/N, 4° C	O/N, 4° C	30min, RT	1 hour, RT	30min, RT	30min, RT	30min, RT	1 hour, RT	30min, RT	30min, RT	30min, RT	30min, RT
	Round 2	**Round 2**	**Round 2**	**Round 2**	**Round 2**	**Round 2**	**Round 2**	**Round 2**	**Round 2**	**Round 2**	**Round 2**	**Round 2**
Primary Ab	-	CD56	TIM-3	CD45	NKG2D	HLA class I	CD68	CD16	-	-	-	PANCK
Clone	-	123C3	D5D5R	H130	1D11	HC10	PG-M1	SP175	-	-	-	AE1/AE3
Vendor	-	Thermo Fisher	Cell Signaling	Thermo Fisher	Novus	Nordic MU Bio	Abcam	Abcam	-	-	-	Abcam
Catalog Number	-	MA1-06801	45208S	14-0459-82	NB100-65956	MUB2037P	AB783	ab183354	-	-	-	ab27988
Host species	-	Mouse	Rabbit	Mouse	Mouse	Mouse	Mouse	Rabbit	-	-	-	Mouse
Conc.	-	1:100	1:200	1:100	1:50	1:10,000	1:50	1:100	-	-	-	1:1000
Reaction	-	30min, RT	1 hour, RT	30min, RT	30min, RT	30min, RT	30min, RT	30min, RT	-	-	-	30min, RT

Ab = antibody, Conc. = concentration, O/N = overnight, RT = room temperature, GRZB = Granzyme B, αSMA = alpha smooth muscle actin.

Supplementary Table 3.1: Overview of cell counts and cell densities analyzed per specimen.

	Cohort 1	Cohort 2
Cell counts per specimen, Median (range)	**Cohort 1**	**Cohort 2**
Epithelial cells	9974 (52 – 123,984)	54,259 (7064 – 206,749)
$CD8^+$ T cells	885 (54 – 9495)	15,007 (1495 – 64,939)
$CD4^+$ T cells	855 (11 – 9156)	18923 (1834 – 59,488)
NK cells (total)	73 (0 – 1317)	326 (7 – 1338)
$CD56^+$ $NKp46^+$ NK cells	3 (0 – 60)	11 (0 – 146)
$CD56^+$ $NKp46^-$ NK cells	6 (0 – 192)	63 (1 – 869)
$CD56^-$ $NKp46^+$ NK cells	46 (0 – 1065)	187 (1 – 826)
Myeloid cells	804 (50 – 9056)	4911 (155 – 24,674)
Myelomonocytic cells	148 (1 - 3644)	681 (24 – 6190)
Dendritic cells	90 (1 – 2236)	967 (55 – 15,101)
Other myeloid cells	316 (7 – 6151)	2105 (56 – 13,956)
Other $CD45^+$ cells	415 (26 – 16,879)	7373 (551 – 20,431)
Cell density (cells/mm^2) per specimen, Median (range)		
Epithelial cells	1215 (9 – 6134)	1339 (406 – 3886)
$CD8^+$ T cells	133 (10 – 714)	398 (46 – 1871)
$CD4^+$ T cells	126 (2 – 1919)	512 (64 – 1683)
NK cells (total)	8 (0 – 53)	7 (0 – 69)
$CD56^+$ $NKp46^+$ NK cells	0 (0 – 3)	0 (0 – 7)
$CD56^+$ $NKp46^-$ NK cells	1 (0 – 18)	2 (0 – 45)
$CD56^-$ $NKp46^+$ NK cells	4 (0 – 43)	4 (0 – 17)
Myeloid cells	126 (9 - 518)	97 (9 – 531)
Myelomonocytic cells	21 (0 – 164)	19 (1 – 132)
Dendritic cells	27 (1 – 547)	73 (14 – 447)
Other myeloid cells	65 (2 – 238)	47 (5 – 327)
Other $CD45^+$ cells	57 (7 – 374)	178 (68 – 450)

The median and range of cell counts and cell densities for each indicated cell type are depicted across analyzed regions. The median and range were calculated using the sum of ROIs per specimen. ROI = region of interest.

Supplementary Table 3.2: Immune cell densities in tumor specimens grouped by tumor characteristics in cohort 1 and cohort 2.

Cohort 1		$CD3^+$ T cells	NK cells	$CD11b^+$ cells
		Cells/mm^2	Cells/mm^2	Cells/mm^2
	n	Median (Q1-Q3)	Median (Q1-Q3)	Median (Q1-Q3)
Tumor grade				
1-2	15	214 (68-384)	7 (5-9)	107 (22-205)
3	11	356 (257-469)	10 (4-15)	131 (108-220)
Lymph node status				
negative	12	293 (169-510)	12 (5-18)	124 (26-191)
positive	14	292 (82-414)	7 (5-9)	127 (49-208)
ER status				
negative	7	460 (328-799)	6 (3-11)	165 (84-346)
positive	19	285 (71-394)	9 (5-12)	124 (25-205)
PR status				
negative	9	290 (214-478)	6 (5-14)	95 (27-165)
positive	17	294 (69-443)	9 (5-11)	131 (34-212)
HER2 status (by IHC)				
negative	8	136 (58-310)	8 (5-10)	76 (21-163)
positive	18	325 (218-473)	8 (5-14)	129 (79-254)
Neoadjuvant treatment				
naive	14	292 (205-456)	9 (6-11)	138 (79-202)
treated	12	346 (98-489)	7 (5-25)	118 (31-222)

Cohort 2		$CD3^+$ T cells	NK cells	$CD11b^+$ cells
		Cells/mm^2	Cells/mm^2	Cells/mm^2
	n	Median (Q1-Q3)	Median (Q1-Q3)	Median (Q1-Q3)
Tumor grade				
2	7	288 (215-580)	2 (1-4)	42 (13-62)
3	23	553 (324-1035)	6 (3-15)	126 (49-246)
Lymph node status				
negative	8	588 (333-737)	7 (3-10)	134 (79-236)
positive	22	464 (276-1086)	4 (2-7)	59 (20 -205)
ER status				
negative	14	736 (325-1135)	4 (1-13)	129 (71-273)
positive	16	376 (276-717)	4 (3-7)	52 (24-123)
PR status				
negative	16	666 (423-1001)	4 (2-10)	129 (64-264)
positive	14	345 (251-684)	4 (2-8)	48 (20-107)
Metastasis				
yes	3	342 (226-486)	15 (10-16)	158 (83-188)
no	27	547(282-922)	4 (2-8)	85 (43-232)
Vital status				
living	27	471 (276-876)	4 (2-8)	85 (34-232)
deceased	3	630 (486-1014)	15 (8-16)	158 (101-188)

Statistical analysis was performed for each tumor characteristic to test for differences in cell densities of $CD3^+$ T cells, NK cells and $CD11b^+$ myeloid cells; Kruskal Wallis tests were performed with Dunn's post hoc tests to correct for multiple comparisons and none of the comparisons had p values below 0.05. In cohort 2, all specimens were $HER2^+$.

Chapter 4: Interrogating the single-cell spatial landscape of T cells in pancreatic cancer

This chapter has been formatted for inclusion in this dissertation from the manuscript "Machine learning links T-cell function and spatial localization to neoadjuvant immunotherapy and clinical outcome in pancreatic cancer" by Katie E. Blise, Shamilene Sivagnanam, Courtney B. Betts, Konjit Betre, Nell Kirchberger, Benjamin J. Tate, Emma E. Furth, Andressa Dias Costa, Jonathan A. Nowak, Brian M. Wolpin, Robert H. Vonderheide, Jeremy Goecks, Lisa M. Coussens, and Katelyn T. Byrne, published in *Cancer Immunology Research* (2024) [244]. The author of this dissertation is the primary author of this manuscript and used single-cell gated data generated by LMC's laboratory to conduct computational experiments to generate all figures in this manuscript except Main Figures 4.1B and 4.1E and Supplementary Figures 4.1B, 4.1C, and 4.1D.

4.1 Abstract

Tumor molecular datasets are becoming increasingly complex, making it nearly impossible for humans alone to effectively analyze them. Here, we demonstrate the power of using machine learning (ML) to analyze a single-cell, spatial, and highly multiplexed proteomic dataset from human pancreatic cancer and reveal underlying biological mechanisms that may contribute to clinical outcome. We designed a multiplex immunohistochemistry antibody panel to compare T-cell functionality and spatial localization in resected tumors from treatment-naive patients with localized pancreatic ductal adenocarcinoma (PDAC) with resected tumors from a second cohort of patients treated with neoadjuvant agonistic CD40 (anti-CD40) monoclonal antibody therapy. In total, nearly 2.5 million cells from 306 tissue regions collected from 29 patients across both cohorts were assayed, and over 1,000 tumor microenvironment (TME) features were quantified. We then trained ML models to accurately predict anti-CD40 treatment status and disease-free survival (DFS) following anti-CD40 therapy based upon TME features. Through downstream interpretation of the ML models' predictions, we found anti-CD40 therapy reduced canonical aspects of T-cell exhaustion within the TME, as compared to treatment-naive

TMEs. Using automated clustering approaches, we found improved DFS following anti-CD40 therapy correlated with an increased presence of $CD44^{+}CD4^{+}$ Th1 cells located specifically within cellular neighborhoods characterized by increased T-cell proliferation, antigen-experience, and cytotoxicity in immune aggregates. Overall, our results demonstrate the utility of ML in molecular cancer immunology applications, highlight the impact of anti-CD40 therapy on T cells within the TME, and identify potential candidate biomarkers of DFS for anti-CD40–treated patients with PDAC.

4.2 Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the most aggressive treatment-refractory cancers with a median overall survival rate of just months [245]. Thus, there is a critical need for improved understanding of the immunobiology of PDAC to inform future treatment strategies for this disease. Recent reports reveal immunological responses in PDAC are induced via approaches that promote priming of T-cell responses against PDAC, such as occurs following agonistic CD40 monoclonal antibodies (anti-CD40) [110] and mRNA vaccination [109] strategies. We and others have previously reported that anti-CD40 binds to CD40 on dendritic cells (DCs), thereby licensing DCs to subsequently enhance T-cell activation and bolster antitumor immunity [111]. In addition, we previously described global immune contexture of the PDAC tumor microenvironment (TME) at baseline and after anti-CD40 therapy and found anti-CD40–treated TMEs contained reduced densities of M2-like tumor-associated macrophages, increased DC maturation, and increased T-cell enrichment [107, 110]. However, little is known regarding selective impact of anti-CD40 therapy on granular T-cell states within the TME, and studies interrogating how anti-CD40 stimulation sculpts the T-cell response specifically [112, 246] are needed to optimize anti-CD40 use in the clinical setting.

Upon antigen stimulation, T cells exist along a spectrum of diverse differentiation states with varying functionalities [247]. On one end, T cells possess stem-cell-like plasticity, accompanied by memory, proliferative, and cytotoxic capabilities, and are identified by expression of T-BET, and/or TCF-1 [248, 249]. On the other end, T cells are exhausted and/or dysfunctional and express TOX1 and/or EOMES [250]. Along the spectrum, expressed in varying combinations, T cells express immune checkpoint molecules such as PD-1, LAG-3, and TIM3, with increased expression of immune checkpoint molecules correlating with more exhausted T cells [22]. These partially exhausted T cells are susceptible to reinvigoration by immune checkpoint blockade (ICB) and regain the ability to proliferate and produce effector cytokines [74]. However, terminally differentiated T cells expressing TOX1 or EOMES are resistant to rescue by ICB and fail to proliferate or exert cytotoxic activity [21]. Flow cytometric analyses of T cells in preclinical tumor models or tissues following viral infection have elucidated notable T-cell states; however, characterizations of effector versus exhausted T-cell phenotypes from tumors in patients are scant. Moreover, with the advent of single-cell sequencing approaches, the diversity of T-cell subsets within tumors is seemingly endless [17, 18]. We sought to clarify T-cell characteristics within the PDAC TME and identify subsets associated with therapeutic anti-CD40 responses. Recognizing that both the cellular composition and spatial organization of cells is a critical metric associated with therapeutic response and clinical outcome [68, 140, 144, 147, 251, 252], we investigated the impact of anti-CD40 therapy on the complex spatial contexture of T cells within the PDAC TME and associated survival.

We previously developed a multiplex immunohistochemistry (mIHC) single-cell spatial proteomics imaging platform to interrogate leukocyte heterogeneity and spatial landscape within various TMEs [107, 121, 122]. Following a cyclical staining protocol, the mIHC platform

iteratively deploys up to 30 antibodies on a single tissue specimen, thus preserving TME spatial context. Resulting data provide single-cell resolution maps quantifiable by a number of metrics that can be correlated with therapeutic response or clinical outcome. However, this reductionist approach does not combine or weight TME features together to capture biological complexities of the TME. Machine learning (ML), a form of artificial intelligence, can address TME complexity by creating computational models that weigh and combine data features to predict a given output. ML models can be used to make accurate predictions for new data and analyzed to identify the feature combinations most important in the predictions. As a result of this capability, ML is becoming widely utilized in precision oncology to decipher patterns in large datasets resulting from deep interrogations [144, 146, 253, 254].

Here, we leveraged ML to elucidate the frequency of various T-cell states in PDAC and investigated the impact of anti-CD40 therapy on those states. We first designed an mIHC antibody panel to deeply audit T-cell functionality and spatial organization in patients from either treatment-naive or neoadjuvant anti-CD40–treated PDAC cohorts. Using this mIHC panel, we generated a dataset of nearly 2.5 million cells with spatially-resolved single-cell phenotypic and functional measurements. Interrogation of this dataset presented a unique opportunity to elucidate: first, the types of T cells present at baseline in a treatment-refractory disease, and second, to what degree anti-CD40 therapy sustains T-cell functionality in situ, or instead promotes T-cell dysfunction that may limit potential use of anti-CD40 or other T-cell priming therapies in the clinical setting. Given the vast amount of spatially-resolved data and complexity of T-cell function in the TME, we leveraged ML approaches to discern new biological insights regarding T cells in the pancreatic TME and their association with clinical outcome for pancreatic cancer patients.

4.3 Materials and methods

4.3.1 Tissue acquisition

Human PDAC tissue specimens from cohort 1 were obtained from patients with approval from the Oregon Pancreas Tissue Registry under Oregon Health & Science University (OHSU) IRB protocol #3609 and Dana Farber Harvard Cancer Center protocols #03-189 and #12-013. Cohort 1 consisted of 18 treatment-naive tumors, selected as a representative subset of a larger PDAC cohort (n=104 tumors) from a prior study evaluating global immune contexture [107]. These 18 samples were selected based on cellular subsets not statistically differing from the full cohort in terms of leukocyte densities or patient survival durations (**Supplementary Table 4.1**). Specimens from cohort 2 were obtained from patients treated with neoadjuvant selicrelumab with approval under the IRBs of four sites across the United States involved in an open-label phase I clinical trial (Cancer Immunotherapy Trials Network CITN11-01; NCT02588443), including 8 patients with neoadjuvant selicrelumab alone and 3 patients with neoadjuvant selicrelumab combined with gemcitabine and nab-paclitaxel. Patients from this trial were combined into a single cohort for the present study, as the anti-CD40–treated cohort was previously analyzed in two separate cohorts for total immune contexture [110]. Additionally, healthy tonsil and spleen samples were from the Knight Tissue Bank, collected under the OHSU IRB protocol #4918. All studies were conducted in accordance with the Declaration of Helsinki and written informed consent was obtained.

4.3.2 Multiplex immunohistochemistry image acquisition and analysis

Formalin-fixed, paraffin-embedded (FFPE) surgical tissue samples were sectioned and assessed using hematoxylin and eosin (H&E), as well as chromogen-based mIHC. Briefly, slides

with 5 μm tissue sections were rehydrated as described (5), followed by 2 minutes in hematoxylin, 2 minutes in tap water, 10 seconds in eosin, and 2 minutes in tap water. Slides were then dehydrated and coverslipped prior to scanning and annotation by a pathologist. Next, using the pathologist annotations overlaid from the H&E-stained slides, sections were assessed with regard to histopathologic regions of interest annotated as tumor (T), immune aggregate (IA), tumor-adjacent stroma (TAS), or normal-adjacent pancreas (NAP), as defined by the pathologist [107]. Regions were selected with the aim of maximizing both diversity of regions selected from each patient and the tissue area captured for each histopathologic region. Across all patients, a mean of 15.2% of tissue was analyzed per entire resection, ranging from 4.4% to 35.6%, and across the entire dataset, 14.6% of available tissue was analyzed (**Supplementary Fig. 4.1A**, **Supplementary Table 4.2**). Of note, our prior investigation of PDAC immune contexture found NAP regions to contain increased leukocyte density as compared to true healthy normal pancreas tissue collected from organ donors [107]. Multiplex staining was performed on 5 μm sections, as previously described, and each stained image was scanned at 20x magnification on an Aperio AT2 scanner (Leica Biosystems) [107]. The antibody panel used in the present study delineated 18 T-cell subpopulations and contained 10 functionality biomarkers to assess differentiation/exhaustion status on all T cells (TOX1, TIM3, TCF-1, CD38, PD-1, EOMES, CD39, CD44, LAG-3, and T-BET), as well as antibodies for proliferation (Ki-67) and cytotoxicity (granzyme B, GrzB) (**Supplementary Fig. 4.1B**, **Supplementary Table 4.3**). Additionally, the panel only broadly delineated epithelial cells, mesenchymal cells, B cells, and myeloid cells, given our previous efforts at describing those subsets in these same patient cohorts [107, 110, 251]. Human tonsil and spleen were included in all rounds of mIHC as staining controls. Image processing was performed using previously described methods [107]. Each

region was registered to the final hematoxylin using Matlab Computer Vision Toolbox (The Mathworks, Inc., Natick, MA), color deconvolution and watershed-based nuclei segmentation was performed using ImageJ, and single cell mean intensity for each stain was quantified using Cell Profiler [255]. Single biomarker positivity thresholds were set using FCS Express Image Cytometry RUO (De Novo Software, Glendale, CA) to visually validate protein biomarker expression overlaid on signal extracted images. Single cell classification was performed using R Statistical Software based on filtering exclusive populations in a defined hierarchy.

4.3.3 T-cell phenotyping

423,317 T cells were identified by CD3 expression and then subsequently stratified by CD8α expression. $CD8^+$ T cells were further classified as one of six cell states (T_{NAIVE}, T_{EFF}, T_{EM}, T_{EMRA}, T_{EX}, or T_{TEX}). Due to biomarker selection and positional restrictions within the cyclic multiplex panel, a CD4 antibody was not included. However, the majority (72%) of $CD3^+CD8^-$ T cells were $CD4^+$ as determined in a testing panel using a subset of the data (**Supplementary Fig. 4.1C**); therefore $CD3^+CD8^-$ T cells are referred to as $CD4^+$ T cells herein, although it is possible other minor lineages may be represented [17]. Based on this schema, $CD3^+CD8^-$ T cells were further evaluated; $CD4^+$ Th1 cells were defined by T-BET^+ expression, and further classified as one of three cell states (T_{EFF}, T_{EM}, or T_{EMRA}). Only 6% of $CD8^+$ T or $CD4^+$ Th1 cells were phenotyped as one of these six T-cell states (T_{NAIVE}, T_{EFF}, T_{EM}, T_{EMRA}, T_{EX}, or T_{TEX}) (**Supplementary Table 4.4**). The other 94% of the T cells were labeled as T_{OTHER} and were stratified based on expression of CD44, a canonical biomarker of prior cognate antigen experience [256]. A population of T-BET^-CD4^+ Th cells (non-Th1-specific T helper cells) were also stratified into two cell states based upon CD44 expression. The remaining $CD3^+CD8^-$ T cells were $FOXP3^+$ regulatory T (T_{REG}) cells and classified into three cell states (Naive T_{REG},

mT_{REG}, and T-BET^{+} T_{REG}). All 18 T-cell states were included in downstream analyses given calculations from previous single-cell studies [257], and the fact that these populations were manually gated and thus represent real phenotypes of T cells present in the PDAC TME despite low numbers of certain T-cell states (**Supplementary Table 4.4**).

4.3.4 TME feature quantifications

Three approaches were used to quantify treatment-naive and anti-CD40–treated PDAC TMEs, resulting in a total of 1,252 TME features quantified per tissue region:

(1) Cell State Densities: To identify the types and amounts of cell states present in the PDAC TME, densities of each of the 23 cell states for each tissue region assayed were quantified by dividing raw counts of cells (**Supplementary Table 4.4**) by tissue area of the region.

(2) T-cell Functionality Barcodes: To investigate T-cell functionality, we assigned all T-cell states a "Functionality Barcode," as defined by binary positive or negative expression of unique combinations of 10 T-cell functionality biomarkers (TOX1, TIM3, TCF-1, CD38, PD-1, EOMES, CD39, CD44, LAG-3, and T-BET). Among the 423,317 T cells present in the dataset, we identified expression of 961 unique barcodes (**Supplementary Table 4.5**).

(3) Cell–Cell Spatial Interactions: To address spatial organization of cells in the PDAC TME, we leveraged the fact that mIHC preserves spatial context and quantified frequency of two cell states interacting, based on their cell centers being within 20 μm from each other, as previously reported [137]. Total cell–cell interactions were normalized by dividing summed densities of cell states involved in the interactions to avoid skewing by cell states present in high abundances. We identified 268 unique pairs of cell–cell interactions present in the dataset (**Supplementary Table 4.6**).

4.3.5 Machine learning classifiers and feature importance analyses

Elastic net (EN) classifier models were built using scikit-learn's [198] LogisticRegression function to predict: 1) treatment status and 2) disease-free survival (DFS), from 1,252 TME features calculated herein. EN models perform well on datasets as generated in this present study where there are more data features than examples for learning [258]. EN models use mathematical regularization approaches to identify and upweight the most informative subset of features to make model predictions while accounting for feature collinearity [259]. Using regularization reduces model overfitting, which is important when there are a large number of data features and a limited number of examples. Further, this approach is unbiased, as no prior feature selection is performed. Instead, all 1,252 TME features were provided to the models, leveraging the EN algorithm's ability to perform aggressive feature selection within model training.

Predictions were made on an individual region basis, rather than a patient basis, to maximize sample size and model robustness, as well as to mitigate tissue region selection biases. As such, tissue regions were not averaged across patients and were instead evaluated independently, thus: 1) reducing the impact of each region on overall model performance and 2) providing the models with more examples to extract biologically-meaningful information from the dataset. Separate models were created for regions of each histopathologic site to: 1) compare performance of models derived from different histopathologic sites; 2) identify where therapy was exerting greatest impact; and 3) mitigate broad variation in average tissue area from each histopathologic site. A leave-one-patient-out cross-validation approach was used to split the train and test sets. Thus, within each cross-validation loop, a new EN model was created and trained on all regions except those from one patient, and testing was then performed on regions from the

patient withheld from training. This process was repeated until all patients were cycled through the test set. This approach prevents data leakage by ensuring regions from the same patient were not in both the train and test sets for one model, thus preventing the EN models from learning patient-specific features, which often artificially increases model accuracy.

Test set predictions were aggregated across all cross-validation loops to construct one final confusion matrix, from which performance of the models was assessed by calculating accuracy, F1 score, and area under the receiver operating characteristic curve (AUC). These metrics address both precision and recall (F1 score), in addition to the true positive rate and false positive rate (AUC) – these are often used to assess performance of classifier models. Model overfitting was mitigated by using the same model hyperparameters across cross-validation loops. The penalty term was set to "elasticnet," and the "l1_ratio" hyperparameter was set to 0.5, representing an equal balance of the lasso model and ridge model effects. All features were log10+1 normalized and scaled using a minmax [0,1] scaler to equally compare features spanning different orders of magnitude and improve model interpretability. To further prevent data leakage, in each cross-validation loop, the scaler was fit to the train set and then applied to the train set and subsequently the test set. Test feature outliers were clipped to [0,1] following this normalization. The train set was balanced within each cross-validation loop using Synthetic Minority Over-sampling Technique (SMOTE) to up-sample the minority class to equal the majority class [260]. Feature importance analyses were conducted by computing Shapley Additive exPlanations (SHAP) values for each model [261]. SHAP values enable the interpretation of which combinations of features contribute to the overall model predictions, as SHAP values denote the relative importance of a given feature in driving a model's prediction. SHAP values have been used to explain ML predictions in prior cancer studies [146, 253].

4.3.6 Recurrent cellular neighborhood analysis

Recurrent cellular neighborhoods were quantified to assess spatial organization of tissues. A neighborhood was created for every cell by counting all cells within a 60 μm radius of each seed cell's center, as inferred from previous studies [132, 133]. Using proportions of cells comprising the neighborhoods as features, neighborhoods were grouped using K-means clustering. The elbow method was used to determine the number of clusters, resulting in groupings of spatial neighborhoods that were similar in cellular composition that could be found across all regions of interest in the analysis.

4.3.7 Statistics

Mann-Whitney U tests were used to determine statistically significant differences in top TME features between treatment cohorts or DFS groups. The Benjamini-Hochberg correction was used to account for multiple hypothesis testing for each analysis. *P*-values less than 0.05 were considered statistically significant. Statistical calculations were performed with the Scipy and statsmodels packages using Python software [197, 199].

4.3.8 Data availability

mIHC data used for this study is available for download on Zenodo at https://zenodo.org/records/10476868. All other data are available in the article and its supplementary files or from the corresponding author upon reasonable request.

4.3.9 Code availability

The code used to generate all computational results of this research was created using Python version 3.9.4 and is available at https://github.com/kblise/PDAC_mIHC_paper.

4.4 Results

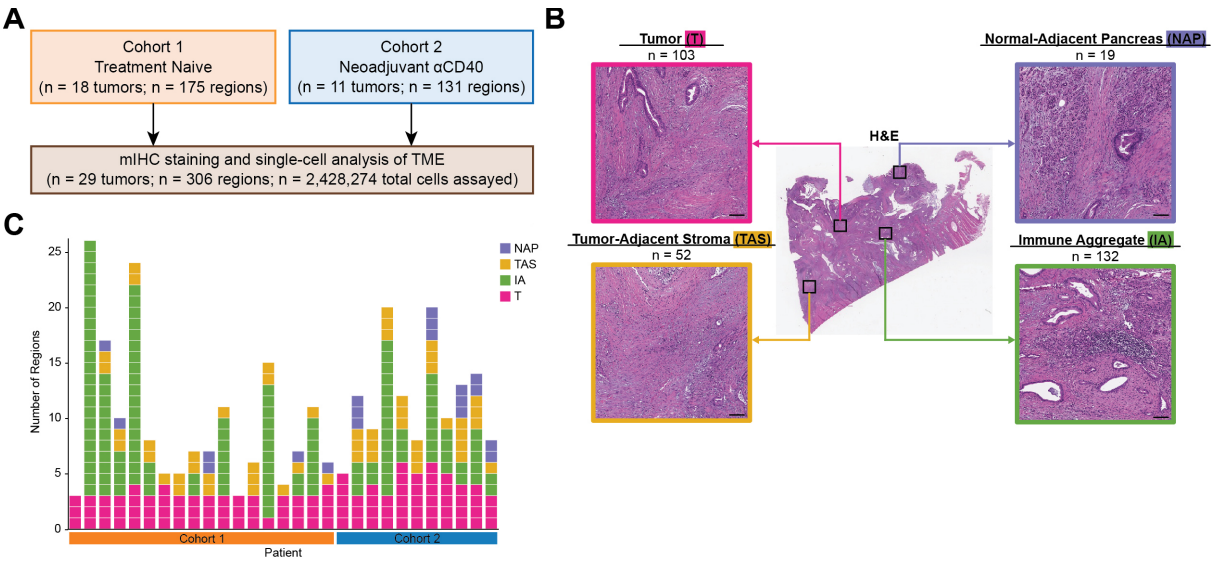
4.4.1 Deep phenotyping of T cells within the PDAC TME using mIHC

29 PDAC tumors were surgically resected from patients across two treatment cohorts (**Fig. 4.1A**). Tumors from 18 treatment-naive patients comprised cohort 1 and were previously evaluated for immune contexture in a larger study [107]. Specimens from cohort 1 served as a representative baseline comparison to the 11 specimens from cohort 2, which reflected patients who had received neoadjuvant anti-CD40 therapy alone (n=8) or in combination with gemcitabine and nab-paclitaxel (n=3) prior to resection [110], and were combined here to a single cohort to evaluate impact of CD40 stimulation on T cells given our prior study using this cohort [110]. Three to 26 tissue regions per PDAC resection were selected by a pathologist and quantitatively assayed by mIHC, with each region annotated as one of four histopathologic sites within the resected samples: tumor (T), immune aggregate (IA), tumor-adjacent stroma (TAS), or normal-adjacent pancreas (NAP) (**Fig. 4.1B**) [107]. Breakdown of region types assayed per patient are shown (**Fig. 4.1C**).

In total, nearly 2.5 million cells were assayed across 306 tissue regions by our 21-antibody mIHC panel (**Fig. 4.1D** and **E**, **Supplementary Table 4.3**, **Supplementary Fig. 4.1B**). 423,317 T cells were identified by CD3 and CD8 expression (**Supplementary Fig. 4.1C)** and subsequently classified into 18 distinct T-cell states (**Fig. 4.1F**), including eight $CD8^+$ T-cell states, five $CD4^+$ Th1-cell states, two $CD4^+$ Th-cell states (non-Th1-specific T helper cells), and three T_{REG} states. In addition, T-cell functionality was further assessed by 10 biomarkers characterizing differentiation/exhaustion, and by biomarkers of proliferation and cytotoxicity (**Fig. 4.1D**, **E**). The remaining $CD3^-$ (non-T) cells were defined by a hierarchical gating strategy and classified as B cells, myeloid cells, mesenchymal fibroblast-like cells (also referred to as

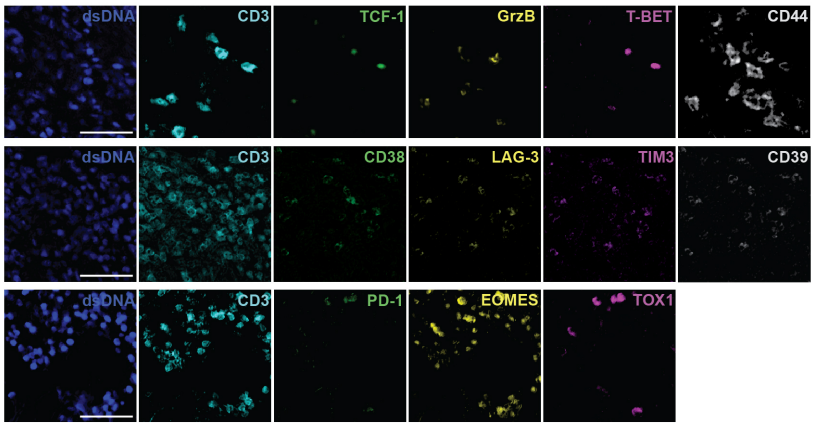
mesenchymal cells), or neoplastic epithelial cells (**Fig. 4.1F**, **Supplementary Fig. 4.1D**).

Altogether, cells were phenotyped as one of 23 different cell lineages and states (**Fig. 4.1F**).



D

mIHC Antibody Panel		
Lineage	T Cell Functionality	Proliferation/ Cytotoxicity
CD3	TOX1	Ki-67
CD20	TIM3	GrzB
CD8	TCF-1	
FOXP3	CD38	
CD45RA	PD-1	
CD45RO	EOMES	
CD11b	CD39	
αSMA	CD44	
PanCK	LAG-3	
	T-BET	



F

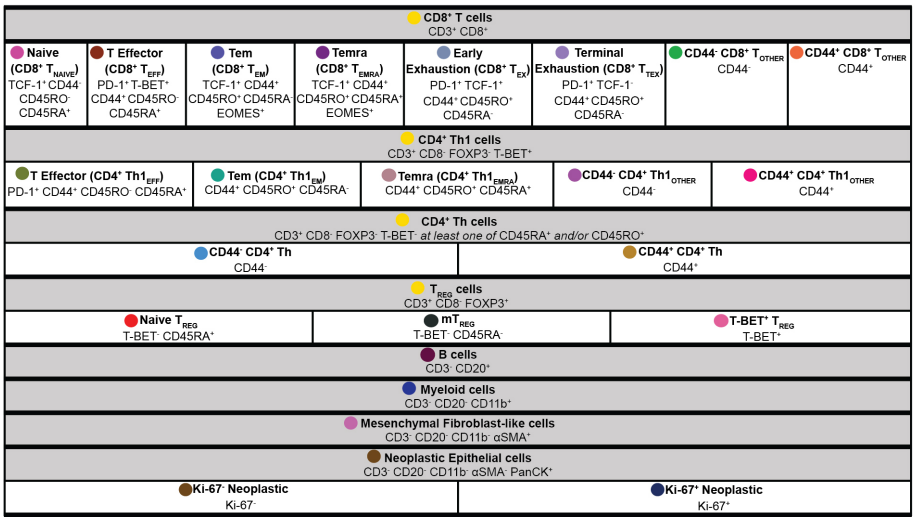


Figure 4.1: Deep phenotyping of T cells within the PDAC TME using mIHC. A) Overview of two PDAC cohorts assayed via mIHC. **B)** Representative PDAC tissue resection stained with H&E (middle) showing four histopathologic sites annotated. Total number of regions assayed per histopathologic site are listed. Scale bars = 100 μm. **C)** Number of regions assayed per patient. Each box represents one tissue region and is colored according to its histopathologic site. **D)** 21-antibody mIHC panel used to assay tissue regions. **E)** Representative pseudo-colored mIHC images showing T cell functionality biomarkers with CD3 expression. Scale bars = 50 μm. **F)** Cell phenotyping strategy by hierarchical gating of lineage and functional biomarkers. Circles indicate colors associated with each cell state in the following figures.

4.4.2 Interrogating cell states and spatial interactions within the PDAC TME

Using single-cell spatial data collected from the mIHC assay, we calculated three types of TME features to create a granular map of leukocyte infiltration, T-cell functionality status, and cellular spatial orientation in the PDAC TME.

Varying densities of leukocytes, mesenchymal fibroblast-like cells, and neoplastic epithelial cells were present in annotated histopathologic regions across treatment cohorts (**Fig. 4.2A, top**). T regions were dominated by neoplastic epithelial cells; IA regions were dominated by T and B cells, and distal NAP regions were dominated by myeloid cells. TAS regions, which encompassed tumor borders, comprised a mix of neoplastic cells, T cells, myeloid cells, and mesenchymal cells. On average, $CD4^+$ T cells were present at a density that was nearly two-fold than that of $CD8^+$ T cells (**Fig. 4.2A, middle, bottom rows**) across regions and cohorts. However, average densities of the $CD8^+$ and $CD4^+$ T-cell states often differed by histopathologic site and treatment cohort, demonstrating the importance of identifying spatial and histopathological information for interpreting how—and where—anti-CD40 therapy alters T cells in the PDAC TME, important information that could not be captured by flow cytometric methodologies.

T-cell functionality was then assessed through quantification of a T cell "Functionality Barcode," for each T cell present in the dataset. The top 15 most common barcodes by average density are shown for each histopathologic type and treatment cohort (**Fig. 4.2B**). Over half of

the most common barcoded T cells were present in both treatment cohorts, regardless of histopathologic site, as indicated by brown bars (**Fig. 4.2B**). The majority of barcodes present in the treatment-naive cohort (orange bars) contained two functionality biomarkers while the most abundant barcodes present in the anti-CD40 cohort only (**blue bars, Fig. 4.2B**) contained three or more functionality biomarkers. This result supports the hypothesis that anti-CD40 therapy shifts T-cell differentiation/functionality within the PDAC TME, as represented by an increase in the number of functionality biomarkers expressed.

Finally, cellular spatial organization was assessed by calculating cell–cell spatial interactions. On average, there were increased interactions between $CD4^+$ T cells with other $CD4^+$ T cells in the anti-CD40 cohort, regardless of histopathologic site (**Fig. 4.2C, white boxes**). Altogether, these results support the hypothesis that anti-CD40 drives an increase in $CD4^+$ T-cell density, functional capacity, and spatial proximity in the PDAC TME, as compared to treatment-naive PDAC TMEs.

Following these single-cell quantifications, all 2,428,274 cells present were phenotyped as one of 23 cell states; all 423,317 T cells were assigned one of 961 T-cell functionality barcodes, and the immediate spatial neighbors of each cell were computed and binned into one of 268 types of pairwise cell-cell interactions. This quantification resulted in 1,252 TME features computed for each of the 306 regions, each annotated as one of four histopathologic sites. Given the complexity and large amount of data, we leveraged ML and feature importance analyses to identify: 1) impact of anti-CD40 therapy on these TME metrics, and 2) the likely mechanism(s) we hypothesized underlying improved clinical outcome following anti-CD40 therapy. To do this, we trained EN classifier models to predict treatment status and DFS within the anti-CD40–treated cohort from the 1,252 TME features quantified above (**Fig. 4.2D**). Finally, SHAP values

were used to identify which combinations of TME features drove model predictions and thus interpret the cellular biology underpinning model predictions [261].

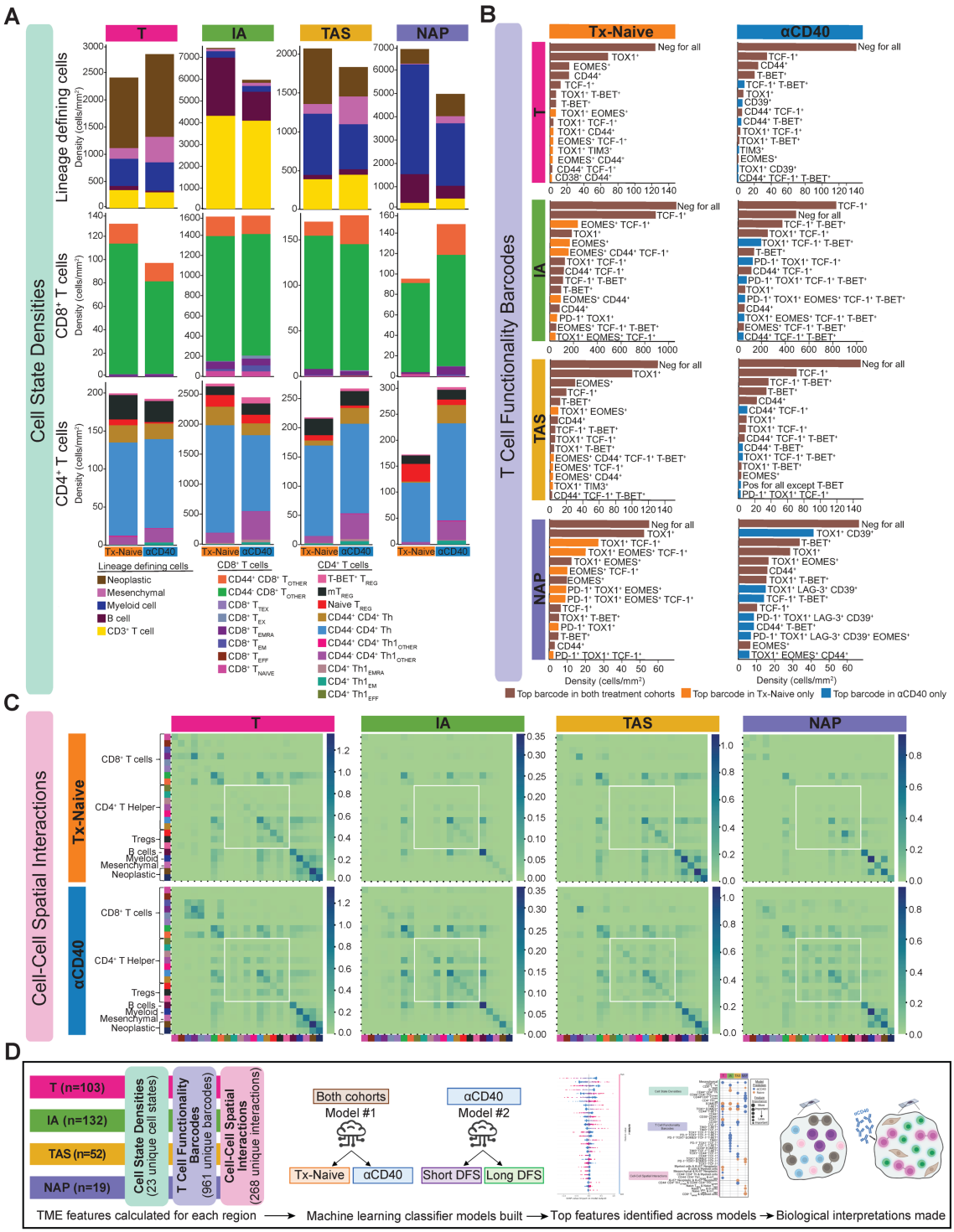


Figure 4.2: Interrogating cell states and spatial interactions within the PDAC TME. A) Stacked bar charts showing the average cell state densities for each treatment cohort and histopathologic site. Top row: lineage

defining cells including neoplastic epithelial cells, mesenchymal fibroblast-like cells, myeloid cells, B cells, and $CD3^+$ T cells; Middle row: $CD8^+$ T cell states; Bottom row: $CD4^+$ T cell states. Columns denote histopathologic site, and each plot is further broken into treatment cohort. **B)** Bar charts showing average densities of barcoded T cells for each treatment cohort and histopathologic site. Only the 15 most abundant barcodes are shown as measured by average density. Rows denote histopathologic site, and columns denote treatment cohort. Brown bars denote barcoded T cells that are in the top 15 most abundant barcodes in both cohorts. Orange bars denote barcoded T cells that are in the top 15 most abundant barcodes in the treatment-naive cohort only. Blue bars denote barcoded T cells that are in the top 15 most abundant barcodes in the anti-CD40-treated cohort only. **C)** Heatmaps showing average number of spatial interactions between two cell states for each treatment cohort and histopathologic site. Cell states are denoted by colors shown in **Figure 4.1F**. Interactions were normalized first by density of cells participating in the interaction and were then log10+1 transformed. Rows denote treatment cohort and columns denote histopathologic site. **D)** Overview schematic of analyses performed in this study. TME features were calculated for each tissue region. Two ML classifier models were built for each histopathologic site to predict treatment status and DFS. Feature importance analyses were performed to interpret biological meaning.

4.4.3 ML models classify αnti-CD40-treated TMEs as having reduced T-cell exhaustion phenotypes

To reveal the impact of anti-CD40 therapy on T cell–exhaustion phenotype, we trained four EN classifier models—one per histopathologic annotation—to predict the treatment status of the tissue. All models performed well, as measured by the accuracy, F1 score, and AUC for test sets of each of the models (**Fig. 4.3A** and **B**). Across the four models, accuracy ranged from 0.83 to 0.85, F1 score ranged from 0.73 to 0.89, and AUC ranged from 0.87 to 0.90. As models were trained to differentiate treatment-naive from anti-CD40–treated PDAC, high performance of all four models indicates that anti-CD40 modulates all types of histopathologic regions across the TME evaluated herein.

To identify the features driving model predictions, and thus reveal how anti-CD40 therapy impacted T cells in the PDAC TME, we calculated SHAP feature importance values for each of the four models (**Supplementary Fig. 4.2A**). The top 30 most important features out of 1,252 total features accounted for the majority of importance according to SHAP values (T model: 84%; IA model: 74%; TAS model: 87%; NAP model: 92%). Comparison of the top 15 features driving model predictions for each of the four histopathologic models revealed 13 of the same features were top contributors across multiple models (**Fig. 4.3C**), indicating shared T-cell

densities, differentiation states, and spatial organizations across histopathologic sites within a given treatment cohort. We then compared the normalized values for each of the top features quantified by SHAP analysis between treatment-naive samples versus anti-CD40–treated samples within each histopathologic site (**Supplementary Fig. 4.2B**, **C**, **D**, and **E**). All 15 features derived from the T model were significantly different between treatment cohorts, 14 of 15 features derived from the IA and TAS models were significantly different between treatment cohorts, and 9 of 15 features derived from the NAP models were significantly different between treatment cohorts, likely because NAP had the fewest number of regions present in the data.

Overall, the models identified anti-CD40–treated TMEs as containing increased densities of mesenchymal fibroblast–like cells and several T-cell states, including three $CD4^+$ T helper populations and two antigen-experienced $CD8^+$ T-cell populations (T_{EX} and $CD44^+$ T_{OTHER}) (**Fig. 4.3C**, **Supplementary Fig. 4.2B, C, D,** and **E**), as compared to treatment-naive TMEs, which contained increased densities of naive $CD8^+$ T cells and B cells. In addition to cell-state densities, analysis of T-cell functionality barcodes revealed that anti-CD40–treated TMEs contained increased densities of T cells expressing combinations of T-BET, CD44, CD39, TIM3, and TCF-1 (**Fig. 4.3C**, **Supplementary Fig. 4.2B, C, D,** and **E**). On the other hand, treatment-naive TMEs contained increased densities of T cells expressing combinations of TOX1 and EOMES, concordant with mIHC stained tissue images (**Fig. 4.3D**). Finally, spatial interactions involving $CD4^+$ Th1 cells were associated with the anti-CD40-treated cohort, whereas interactions involving myeloid cells, naive T_{REGs}, and Ki-67^+ neoplastic epithelial cells were associated with treatment-naive tissue (**Fig. 4.3C**, **Supplementary Fig. 4.2B, C, D,** and **E**). Altogether, these results indicate anti-CD40 TMEs contained increased presence of T cells in

close spatial proximity to one another—in particular, $CD4^+$ T helper cells—with reduced exhaustion profiles, as compared to treatment-naive TMEs.

Of the top features across all four models, the majority of features were densities of T cell–functionality barcodes. As all barcodes present on any T cell were provided to the ML models, we sought to determine whether the top barcodes identified by the models were expressed by similar or different T-cell states. Thus, barcodes were correlated based on the types and proportions of T-cell states expressing each barcode (**Fig. 4.3E**). This analysis resulted in four clusters (i. – iv.) of barcodes, each with distinct compositions of T cells and not influenced by histopathologic site. Barcodes belonging to cluster (i) were expressed by antigen-inexperienced (as defined by lack of CD44 expression) $CD8^+$ and $CD4^+$ T helper cells, naive T_{REGs}, and mT_{REGs}. However, barcodes containing $TOX1^+$ and $EOMES^+$ dominated these cluster (i) T cells when predictive of treatment-naive samples, while T cells predictive of anti-CD40–treated samples expressed barcodes containing $CD39^+$ $TIM3^+$, and $TCF\text{-}1^+$. This result supports the notion that, while the same T-cell types were present regardless of therapy exposure, their functional capacity differed following anti-CD40 treatment. Higher densities of barcodes on multiple antigen-experienced $CD4^+$ Th1 cell states and $T\text{-}BET^+$ T_{REGs}—reported to be similar to $CD4^+$ Th1 cells in their function [23]—in clusters (ii), (iii), and (iv) included combinations of $T\text{-}BET^+$, $TIM3^+$, $TCF\text{-}1^+$, and $CD44^+$ and were predictive of tissue samples from patients treated with anti-CD40 therapy.

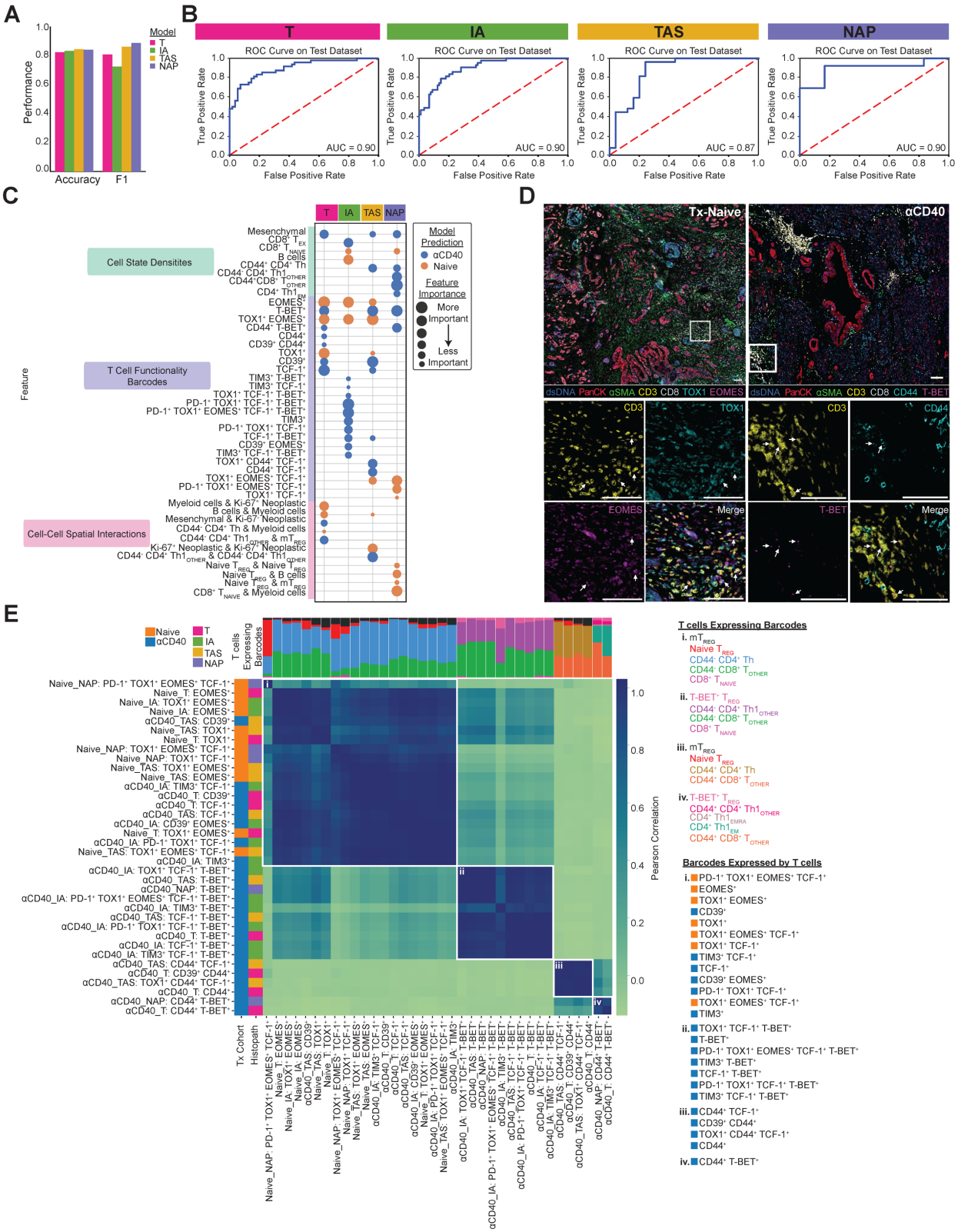


Figure 4.3: ML models classify anti-CD40-treated TMEs as having reduced T-cell exhaustion phenotypes. A) Bar chart showing accuracy and F1 score for each histopathologic model that

predicts treatment status. **B)** ROC curve with corresponding AUC for each histopathologic model. **C)** Bubble chart showing top 15 features whose increased presence drove each histopathologic model to predict treatment-naive (orange) or anti-CD40-treated (blue). Features are grouped by TME feature type (density, barcode, interaction). Bubble size denotes relative importance of the feature for a given histopathologic model. Bubbles appearing multiple times in the same row indicate TME feature is a top feature across histopathologic models. **D)** Representative pseudo-colored mIHC images showing $TOX1^+$ and/or $EOMES^+$ $CD3^+$ T cells in treatment-naive tissue (left) and $CD44^+$ and/or $T\text{-}BET^+$ $CD3^+$ T cells in anti-CD40-treated tissue (right). **E)** Matrix showing correlations between top barcodes from the models with each other based on types and proportions of T cell states expressing the barcodes. Stacked bars at the top of correlation matrixes show proportions of T cell states expressing barcodes, with T cells color coded and listed for each group to the right of the heatmap, along with corresponding barcodes in each group. Leftmost columns are color coded according to which treatment group the presence of the barcode was predicted by the model, followed by the histopathologic site the model was derived from.

4.4.4 Long disease-free survivors have enhanced T-cell effector functionality following anti-CD40 therapy

The clinical trial from which the anti-CD40–treated specimens were derived was not designed to assess correlates with survival. However, despite the small size of our cohort, we hypothesized that we could train ML models to accurately predict DFS for these patients, with the goal of identifying the combinations of TME features associated with long versus short DFS within the anti-CD40–treated cohort. The median DFS timepoint (9.8 months) across all patients in the anti-CD40–treated cohort was used to segregate long and short disease-free survivors. Separate models were built for each histopathologic site, although NAP region presence was insufficient to build a model. Only the model trained from IA regions performed well in predicting both long and short DFS, with an accuracy of 0.81, F1 score of 0.88, and AUC of 0.77 (**Fig. 4.4A** and **B**).

To identify TME features driving predictions for the IA-derived model, we followed a similar model interpretation analysis using SHAP values (**Supplementary Fig. 4.3A**, **Fig. 4.4C**). The top 30 features out of 1,252 features available to the model accounted for 78% of feature

importance. All of the top 15 ranked features identified by the SHAP analysis were significantly different between DFS groups (**Supplementary Fig. 4.3B**), demonstrating how ML can be used to reveal potential combinations of candidate biomarkers of DFS in the PDAC TME.

Of note, we observed an increased density of $CD44^+CD4^+$ Th1 cells and increased spatial interactions between $CD4^+$ Th1 T_{EM} or $CD8^+$ T_{EMRA} cells and T-BET^+ T_{REGs} in IA regions from patients with long versus short DFS (**Fig. 4.4C**, **Supplementary Fig. 4.3B**). However, increased densities of T cell–functionality barcodes were the most common feature predictive of DFS, accounting for 12 of the top 15 features. Of these, eight barcode densities were associated with short DFS and four were associated with long DFS (**Fig. 4.4C**, **Supplementary Fig. 4.3B**). TOX1 was expressed on all eight barcodes associated with short DFS, while PD-1 expression was found exclusively in five of the eight barcodes associated with short DFS. In contrast, expression of CD44, CD38, CD39, TIM3, and LAG-3 were unique to the four barcodes associated with long DFS (**Fig. 4.4C**, **D**).

To determine if the T-cell states expressing each of the predictive barcodes were similar, we again correlated the barcodes by T cell–state expression (**Fig. 4.4E**). We found four distinct clusters of barcodes, and within each cluster, the barcodes were expressed by similar types of T-cell states and in similar proportions. All but one of the barcodes in clusters (i) and (ii) were among the features whose increased densities were associated with short DFS and were expressed by antigen-inexperienced $CD8^+$ and $CD4^+$ T helper cells and T_{REGs} subsets. In contrast, increased densities of all barcodes in clusters (iii) and (iv) were associated with long DFS. Cluster (iii) consisted of one barcode that was uniquely expressed by $CD8^+$ T_{EM} and T_{EMRA} cells, $CD44^+CD4^+$ T helper cells, as well as naive T_{REG} cells. Finally, cluster (iv) barcodes were expressed by $CD44^+CD8^+$ T cells, $CD4^+$ Th1 T_{EM} and T_{EMRA} cells, $CD44^+CD4^+$ Th1 cells, and T-

BET$^+$ T_{REG}s. In summary, these findings indicate the following TME changes in IAs – an antitumor T-cell response, characterized by presence of CD44$^+$ T cells and in particular CD4$^+$ Th1 cells – are associated with prolonged DFS in our patient cohort.

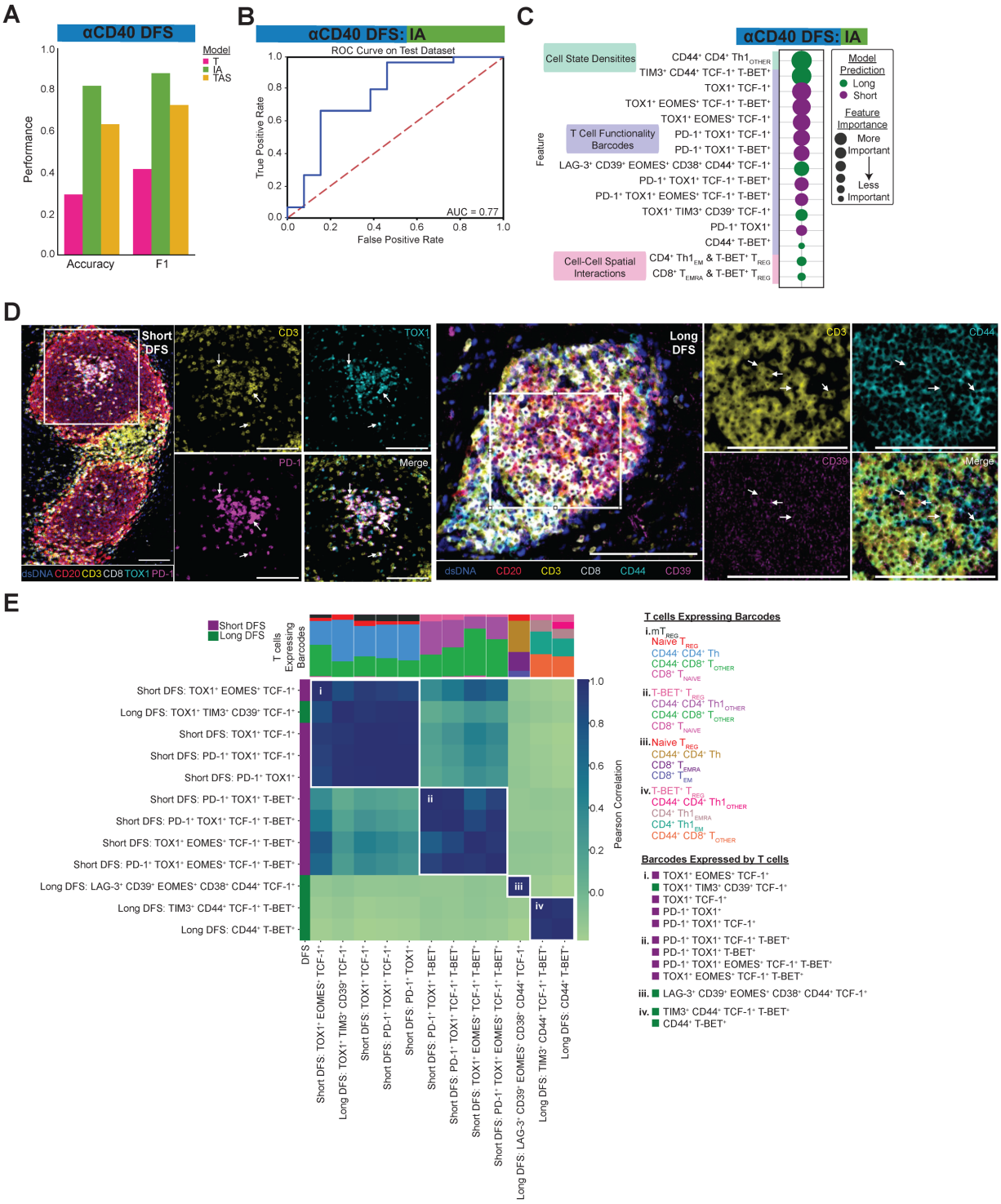


Figure 4.4: Long disease-free survivors have enhanced T-cell effector functionality following anti-CD40 therapy. A) Bar chart showing accuracy and F1 score for each histopathologic model predicting anti-CD40 DFS. **B)** ROC curve with corresponding AUC for the IA histopathologic model. **C)** Bubble chart showing the top 15 features whose increased presence drove the IA model to predict short DFS (purple) or long DFS (green). Features are grouped by TME feature type (density, barcode, interaction). Bubble size denotes relative importance of the feature. **D)** Representative pseudo-colored mIHC images showing $TOX1^+$ and/or $PD\text{-}1^+$ $CD3^+$ T cells in short DFS tissue (left) and $CD44^+$ and/or $CD39^+$ $CD3^+$ T cells in long DFS tissue (right). Scale bars = 100 μm. **E)** Matrix showing correlations between the top barcodes from the IA model with each other based on types and proportions of T cell states expressing the barcodes. Stacked bars at the top of the correlation matrix show proportions of T cell states expressing the barcodes, and T cells are color coded and listed for each group to the right of the heatmap, along with the corresponding barcodes in each group. Leftmost column is color coded according to which DFS group the presence of the barcode predicted by the model.

4.4.5 T-cell spatial organization correlates with DFS after anti-CD40 therapy

The majority of top TME feature types driving the anti-CD40 DFS IA model predictions were densities of specifically barcoded T cells. However, the spatial organization of these cells was not clear from our model predictions. TME spatial architecture is associated with clinical outcomes across cancer types [68, 140, 144, 147, 251, 252]; thus, we aimed to identify the spatial neighbors of the top barcoded T cells whose densities were associated with DFS following anti-CD40 therapy within IA regions.

To quantify the spatial organization of the tissue, we performed a recurrent cellular neighborhood (RCN) analysis across all IA regions within the anti-CD40–treated cohort [140, 147] (**Fig. 4.5A**). This resulted in seven RCNs (**Supplementary Fig. 4.4A**) – each representing spatial neighborhood of cells present across multiple IAs that were distinct in proportions and types of cell states located within the neighborhood. The average cellular composition of each RCN is shown (**Fig. 4.5B**). We confirmed that no single RCN dominated the IA regions analyzed (**Supplementary Fig. 4.4B** and **C**) and that no RCN was exclusively derived from any single IA region or patient (**Supplementary Fig. 4.4D** and **E**). Upon viewing the scatterplot reconstructions of regions, clearly defined spatial patterns within IAs were revealed (**Fig. 4.5C**, **Supplementary Fig. 4.4F**). For example, cells in RCN1, whose neighborhood consisted mostly

of B cells, were often found to be spatially clustered together, potentially representing “germinal center”-like pockets within IA regions.

Given our goal of identifying neighbors surrounding the top barcoded T cells from the anti-CD40 DFS model, we first identified which of the seven RCNs the barcoded T cells were assigned to. We then performed unsupervised clustering of barcoded T cells from long and short DFS patients together based on proportions of the seven RCNs the barcoded T cells resided in (**Fig. 4.5D**). This resulted in two distinct clusters of barcoded T cells, which also segregated according to DFS. Cluster 1 (C1) included all barcoded T cells from patients with long DFS, as well as $CD44^{+}$T-BET^{+} barcoded T cells from patients with short DFS. Cluster 2 (C2) consisted of all remaining barcoded T cells from patients with short DFS.

The most striking difference between the two clusters was proportions of barcoded T cells residing in RCN5 (**Fig. 4.5E**). Of all RCNs, RCN5 contained the greatest proportion of $CD44^{+}$ T cells, spanning both $CD8^{+}$ T cells and $CD4^{+}$ T helper lineages, and with only 28% of cells representing non-T cell lineages, including mesenchymal cells, myeloid cells, a small subset of neoplastic cells, and B cells (**Fig. 4.5B**). 45% of T cells in C1 (long DFS) resided in RCN5, whereas only 3% of T cells from C2 (short DFS) resided in RCN5 (**Fig. 4.5E**). Thus, T cells correlated with long DFS were frequently found to be surrounded by $CD44^{+}$ T cells, supporting the hypothesis that anti-CD40 therapy prolongs survival through the promotion of T-cell priming and/or recruitment of primed T cells to PDAC TMEs.

Finally, to further elucidate potential cellular mechanisms active within various RCNs, we calculated proportions of T cells expressing Ki-67 or GrzB in each RCN and compared values to the overall proportion of $Ki\text{-}67^{+}$ or $GrzB^{+}$ T cells across all anti-CD40 IA regions. Given our prior findings that the majority of T-cell barcodes whose increased density correlated

with long DFS were assigned to RCN5 (**Fig. 4.5E**), we hypothesized that T cells assigned to RCN5 would possess increased proliferative and/or cytotoxic capabilities. We found a larger proportion of T cells expressing Ki-67 residing in RCN1 and RCN5 as compared to the overall T-cell population (dashed line) (**Fig. 4.5F**). In addition, we found a larger proportion of T cells (excluding T_{REGs}) expressing GrzB residing in RCN5, RCN6, and RCN7, as compared to the overall non-T_{REG} T-cell population (dashed line) (**Fig. 4.5G**). However, raw counts of $GrzB^+$ T cells in RCN6 and RCN7 were low (RCN6: n=8 cells; RCN7: n=11 cells), whereas RCN5 contained the highest level of $GrzB^+$ T cells across all RCNs (n=168). Visualization of a representative mIHC image depicts presence of $GrzB^+$ T cells localized to RCN5 in an IA region from a patient with long DFS (**Fig. 4.5H**). Collectively, these results provide further support that T cells within RCN5 are likely activated and possess an effector phenotype capable of an antitumor cytotoxic response.

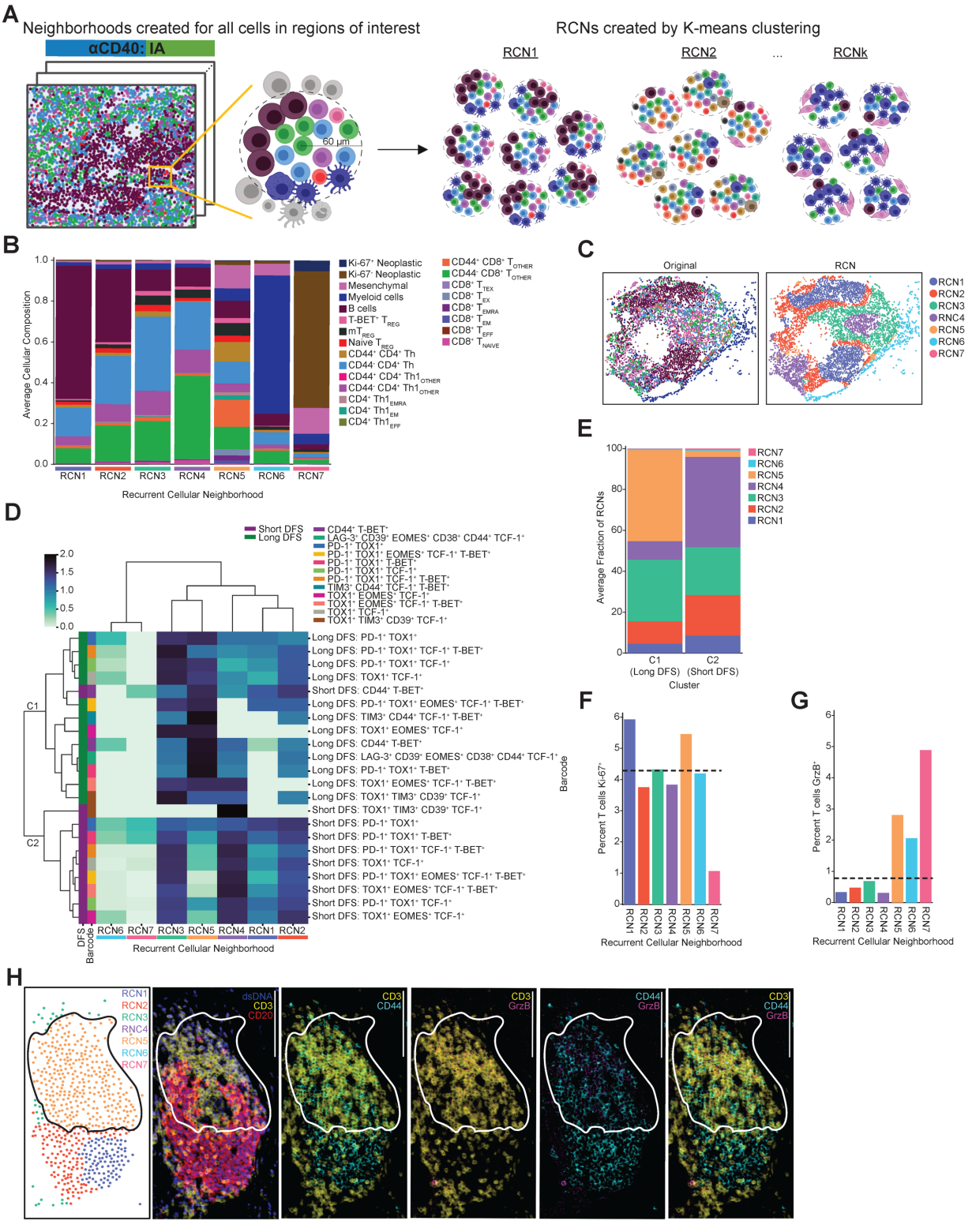


Figure 4.5: T-cell spatial organization correlates with DFS after anti-CD40 therapy. A) Schematic depicting RCN analysis. Cellular neighborhoods were defined by identifying all cells within a 60 µm radius of a given cell. Neighborhoods were calculated for all cells in anti-CD40-treated IA regions. Neighborhoods were then grouped using K-means clustering to identify RCNs. Created with BioRender.com. **B)** Stacked bar chart showing

average cellular composition of each of seven RCNs from the anti-CD40-treated IA regions. Bars are colored by cell state and represent average proportions (out of 1.0) of each cell state present in neighborhoods assigned to each RCN. **C)** Representative IA tissue region as depicted by scatterplot reconstructions. Each dot represents a cell present in the IA, and each cell is colored by its original cell state phenotype (left scatterplot) or RCN assignment (right scatterplot). **D)** Heatmap showing top T cell barcodes from the anti-CD40 IA DFS model clustered by proportion of RCNs the T cell barcodes were assigned to. Rows are barcoded T cells from IA regions from patients associated with short DFS or long DFS ordered by hierarchical clustering of their RCN assignment. Columns are RCNs used as clustering features. Proportion of RCNs was normalized using a log10+1 transformation prior to clustering. Leftmost columns are color coded by DFS group followed by barcode. **E)** Stacked bar chart showing average fraction of RCNs barcoded T cells were assigned to for each of two hierarchically clustered groups (C1 or C2). **F)** Bar chart showing percentage of T cells expressing Ki-67 residing in each of seven RCNs for anti-CD40 IA regions. Horizontal dashed line represents percentage of Ki-67$^+$ T cells across all anti-CD40 IA regions, regardless of RCN assignment. **G)** Bar chart showing percentage of T cells expressing GrzB residing in each of seven RCNs for anti-CD40 IA regions. Horizontal dashed line represents percentage of GrzB$^+$ T cells across all anti-CD40 IA regions, regardless of RCN assignment. T_{REG} populations were excluded from this analysis. **H)** Representative IA region from a patient with long DFS with cells colored by RCN assignment in the upper left scatterplot. Remaining images show mIHC staining of GrzB$^+$ CD44$^+$ CD3$^+$ T cells localized within RCN5. Scale bars = 100 μm.

4.5 Discussion

In this study, we integrated spatial proteomic imaging technology with ML approaches to understand the role of T-cell phenotypes and spatial organization in the complex TME of human pancreatic cancer. In contrast to previous single-cell spatial proteomic studies, which often group T cells together as CD8$^+$ T cells, CD4$^+$ T cells, or T_{REGs} [68, 144, 147, 251], our mIHC panel was curated to phenotype T cells as one of 18 distinct states along with functionality status from 10 different biomarkers, all while preserving the spatial orientation of each cell in the TME. In considering the full spectrum of T-cell states in addition to their spatial organization, ML approaches were necessary due to the complexity of data. ML was used to identify combinations of TME features most associated with anti-CD40 therapy exposure or prolonged DFS (**Fig. 4.6**). This study demonstrates the value of merging single-cell spatial proteomic assays with ML analyses to interrogate how immunotherapy modulates the PDAC TME and potentially drives improved survival.

Despite the multitude of unique T-cell states identified herein, our ML models identified T-cell subsets associated with antitumor characteristics and prolonged DFS. Consistent with our preclinical studies revealing that CD4$^+$ T cells are a major contributor to PDAC immunity

following anti-CD40 therapy [113], our ML models revealed that anti-CD40–treated patient tumors contain increased densities of effector memory cells specifically within the $CD4^+$ Th1 lineage, while no increase was observed in $CD8^+$ memory T-cell populations. Our models also identified antigen-experienced $CD4^+$ Th1 cells as the main cell type whose density associated with prolonged DFS following anti-CD40 therapy. This result is concordant with two independent studies, including characterization of immune cells in biopsied PDAC liver metastases following CD40 agonism [246], and the second investigating primary resected PDAC TMEs after treatment with a granulocyte-macrophage colony-stimulating factor-secreting allogenic PDAC vaccine (GVAX) [262]. Both studies reported that presence of $CD4^+$ T helper cells contribute to improved survival following immunotherapy in PDAC [246, 262]. Here, we further characterized expression features of the $CD4^+$ T helper cells as $CD44^+$ Th1 cells, which correlated with improved outcomes. Reports have also highlighted the direct role of $CD4^+$ T cells in mediating antitumor immunity, including via cytotoxicity [263] and production of effector cytokines [264]. As such, we hypothesize that future therapies designed to harness effector and memory functions of $CD4^+$ T helper cells following administration of anti-CD40 therapy may be clinically beneficial.

In addition to investigating presence of various T-cell states, our ML analyses show that localization and spatial organization of T cells within PDAC TMEs are associated with prolonged survival following anti-CD40 treatment. High performance of the IA-derived DFS prediction model indicates that IAs are a major site of anti-CD40–induced immune response contributing to prolonged DFS. We previously reported increased IAs following anti-CD40 treatment in PDAC-bearing mice [112], and in the aforementioned GVAX study, survival was linked to increased $CD4^+$ T helper pathway genes specifically within IAs (and not tumor regions)

[262]. Moreover, a recent study found enrichment of gene signatures representing mature tertiary lymphoid structures in pre-treatment PDACs associated with improved survival in patients following treatment with varying chemoimmunotherapies [265]. Our RCN analyses further revealed that the key T cells associated with prolonged DFS were often surrounded by antigen-experienced $CD8^+$ and $CD4^+$ T helper cells, as well as a higher proportion of proliferating and cytotoxic T cells, as compared to all T cells regardless of spatial neighborhood. While our mIHC panel in this study focused on T cells, we observed a minor presence of B cells (compared to other RCNs) as well as a subset of myeloid cells and mesenchymal cells in RCN5 that could also contribute locally to T-cell activation or function. Collectively, the results indicate IAs may function as sites of T-cell priming or second signal, promoting T-cell activation and function in PDAC TMEs, contributing to prolonged DFS following multiple types of immunotherapies.

Our interrogation of IAs revealed T-cell states concordant with dysfunctional tumor-infiltrating T-cell phenotypes correlated with short DFS following anti-CD40 therapy. $TOX1^+$ T cells are at the far end of the exhausted T–cell spectrum [250], and TOX1 expression correlates with PD-1 on T cells and impaired immunotherapy response in hepatocellular carcinoma [266]. Correspondingly, we observed expression of TOX1 and/or PD-1 on $CD8^+$ and $CD4^+$ T helper cells associated with shorter DFS following anti-CD40 therapy. T cells linked to short DFS expressing TOX1 and/or PD-1 were largely $CD44^-$, which may represent a population of antigen-naive T cells that aberrantly upregulated these proteins, or T cells that are terminally exhausted due to repeated T cell–receptor stimulation [267]. We also found an increased presence of $TOX1^+$ T cells within treatment-naive TMEs, indicating a baseline terminally exhausted T–cell phenotype in the PDAC TME. Despite this, and in agreement with our prior study [107], we did not identify PD-1^+ T cells as a major subset, in contrast to tumors such as

melanoma [268], where PD-1^{+} T cells are abundant. The paucity of PD-1^{+} T cells may contribute to the failure of ICBs targeting PD-1 or PD-L1 in the majority of patients with PDAC [269]. Together, our data support the conclusion that TOX1, but not PD-1, is a dominant feature of exhausted T cells in PDAC. Thus, therapies that modulate TOX1^{+} T cells in the TME—such as anti-CD40 agonism—may improve clinical outcomes for patients with PDAC.

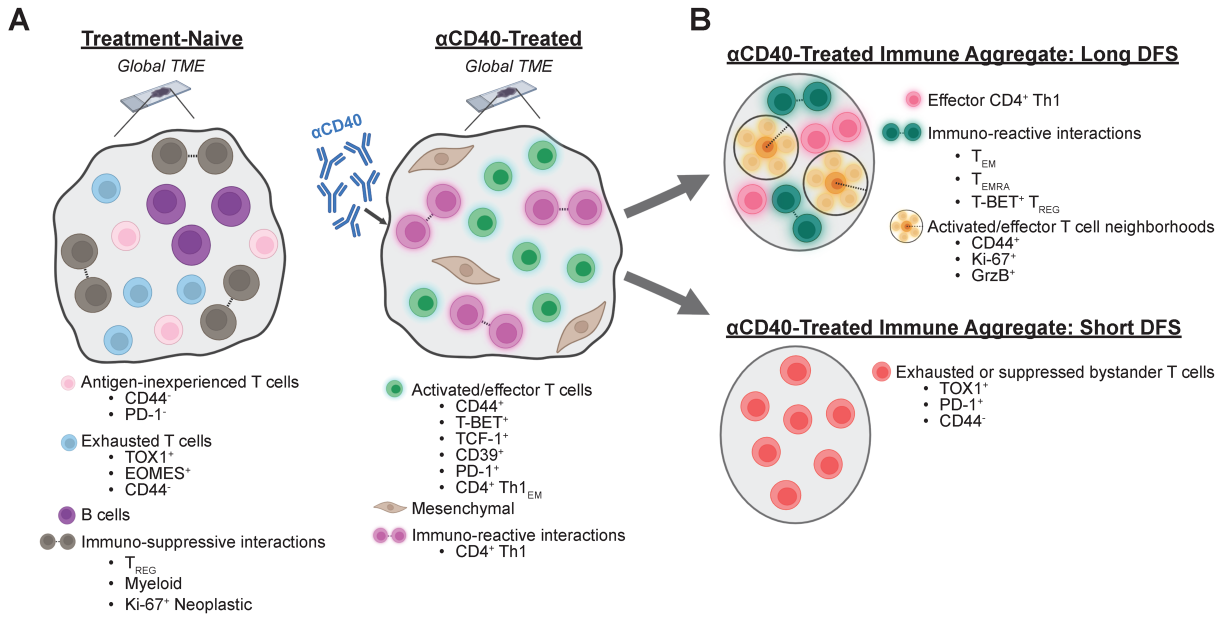


Figure 4.6: Spatial features of T cells associated with anti-CD40 therapy and prolonged DFS in the PDAC TME. A) T cell subsets that best define resected tumor samples from treatment-naïve (left) or anti-CD40-treated (right) patients. In the absence of therapy, T cells appear in an exhausted state, while T cells present with activated and effector phenotypes after CD40 agonism. Created with BioRender.com. **B)** T cell phenotypes in IAs from anti-CD40-treated patients associated with long (top) or short (bottom) DFS. IAs from patients with long DFS are characterized by the presence of spatial neighborhoods of effector T cells capable of proliferating and cytotoxicity, while IAs from patients with short DFS have a preponderance of exhausted T cell states. Created with BioRender.com.

Notably, the dataset we used to conduct this study was unique in several aspects—including therapy administered, mIHC panel deployed, and histopathologic sites assayed—making validation of our results challenging, highlighting the need for additional studies on larger cohorts to build models capable of more generalized predictions, and cautioning

integration of our results in a clinical setting prior to additional validation. Patient samples were collected from multiple institutions per treatment cohort, yet our ML models still accurately classified samples according to therapy. This indicates that any institutional or technical differences in tissue processing were not driving features in model predictions. Importantly, the ML models we trained performed comparably to or better than models derived from similar studies [144, 146, 253, 254]. Additionally, our biological conclusions on impacts of anti-CD40 are concordant with several prior studies [112, 113, 246, 250, 262, 266, 269], providing further support for methods herein and findings despite small sample sizes. Future antibody panels may incorporate additional biomarkers, such as chemokine receptors, to further characterize key T-cell subsets, or additional lineage biomarkers that denote myeloid or B-cell subsets [252], to further phenotype cell–cell interactions. It should be noted that in the present study, treatment with anti-CD40 did not prolong DFS as compared to the treatment-naive cohort, and future data analyses from clinical trials with larger cohorts powered for survival analyses will be useful to validate our findings that associate TME features with longer DFS. Finally, tumors were resected 12 days after anti-CD40 administration; thus, it is possible that T cells involved in prolonged DFS had insufficient time to transit beyond IAs and into surrounding TMEs following priming. Further analyses investigating timing of T-cell trafficking throughout PDAC TMEs are necessary to determine if analysis of T or TAS regions sampled at later timepoints following treatment could be used to assess clinical outcome for these patients.

In addition to the aforementioned caveats, this study also has several other limitations. Importantly, PDAC was the only tumor type analyzed in this study; thus, it is unclear which results may also be relevant to other cancers, or whether they are specific to PDAC. Additionally, all samples were collected from primary tumors, so the results cannot be generalized to recurrent

tumors in the pancreas or metastases to other organs. Further, due to small sample sizes, all patients treated with anti-CD40 therapy were grouped together into a single cohort despite the fact that three of the eleven patients treated with neoadjuvant anti-CD40 also received combination neoadjuvant treatment with gemcitabine and nab-paclitaxel. Thus, the biological effects of gemcitabine and nab-paclitaxel were not evaluated specifically, and the differences detected between treatment-naive tissue and anti-CD40-treated tissue may also be due to exposure to the additional therapies for the three patients treated with combination therapy. Additionally, from a data perspective, while samples were collected from multiple institutions, all data underwent mIHC imaging and processing at one institution. Therefore, it is possible the results described herein may be challenging to replicate with data assayed and processed at a different institution. Finally, while the ML models trained in this study used an unbiased approach to select for the most important TME features potentially driving biological differences, a finite number of pre-defined features were inputted into the models. Although the models performed well and top features were identified using our strategy, it is possible that additional features not computed, such as ratios of cell states and additional spatial metrics, may be relevant.

This study provides proof-of-principle for leveraging ML approaches to evaluate highly multiplexed cancer datasets and supports the use of similar analytics in future studies to identify important, and otherwise inconspicuous alterations in TMEs correlating with patient treatment or response. Future studies could utilize these findings to target pathways identified via this approach to improve treatment strategies for cancer patients.

Acknowledgements

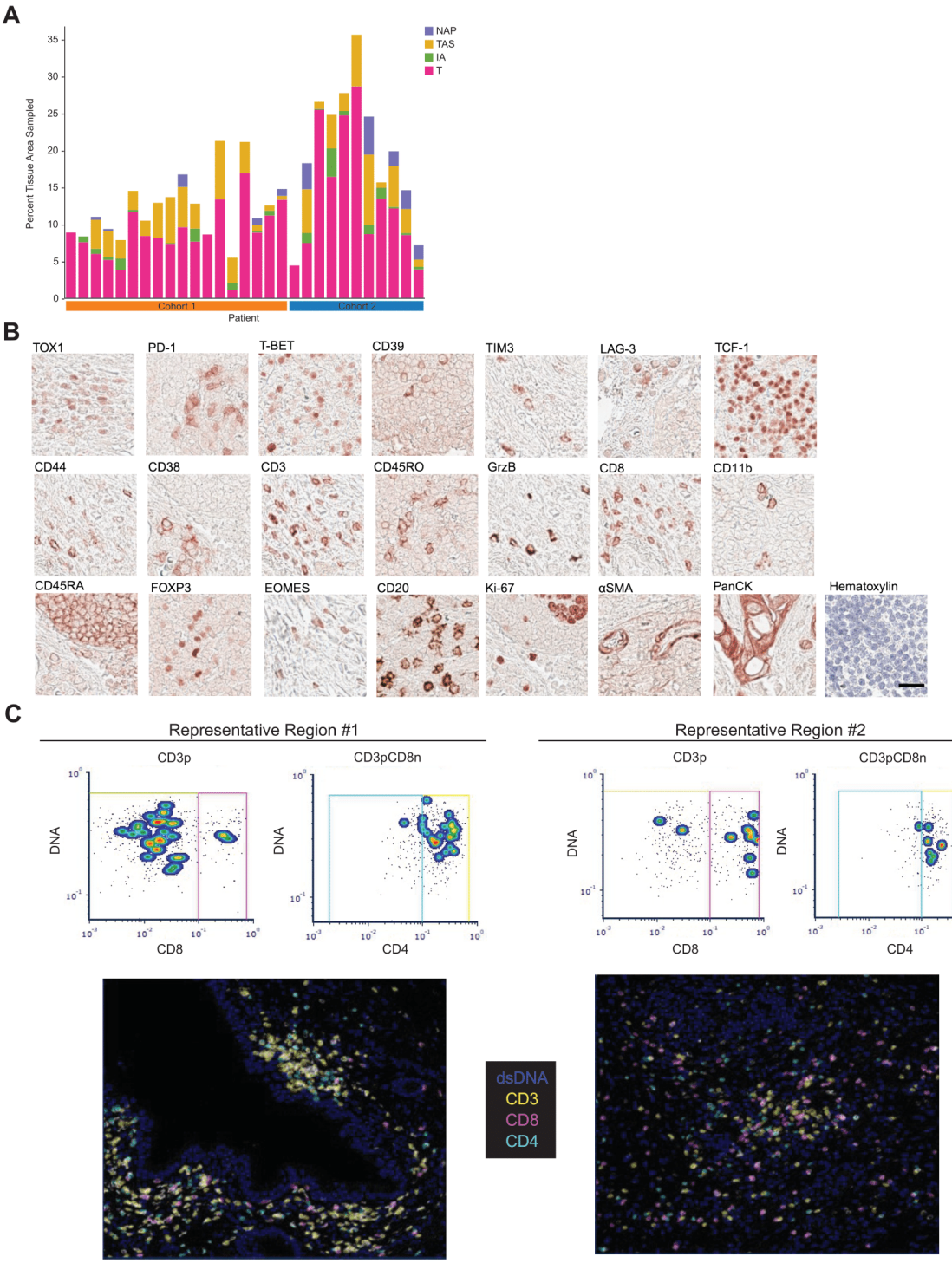
The authors thank Drs. David L. Bajor, E. Gabriela Chiorean, Daniel A. Laheru, and Mark H. O'Hara for efforts as site investigators for the neoadjuvant selicrelumab (CD40 agonist) clinical trial. We also thank John Wherry for helpful discussion related to panel design and biomarker selection. This research was supported by the National Cancer Institute (NCI) of the National Institutes of Health grants T32CA254888 (K.E.B.), P50 CA127003 (J.A.N.), R01 CA248857 (J.A.N.), R01 CA205406 (J.A.N.), R01 CA169141 (J.A.N.), R35 CA197735 (J.A.N.), U01 CA250549 (J.A.N.), U01 CA224012 (LMC), U24CA231877 (J.G.), and U2CCA233280 (J.G.), the DFCI Hale Family Center for Pancreatic Cancer Research (J.A.N. and B.M.W.), the Lustgarten Foundation Dedicated Laboratory program (J.A.N., B.M.W.), the Parker Institute for Cancer Immunotherapy (R.H.V. and K.T.B.), the Brenden-Colson Center for Pancreatic Care (L.M.C., S.S., C.B.B., K.T.B., K.B., N.K.), the Robert L. Fine Cancer Research Foundation (K.T.B.), funding from the Prospect Creek Foundation to the OHSU SMMART (Serial Measurement of Molecular and Architectural Responses to Therapy) Program (J.G. and L.M.C.), and a generous startup package from the Knight Cancer Institute and the Brenden-Colson Center for Pancreatic Care (K.T.B.).

Author contributions

K.E. Blise: Conceptualization, data curation, software, formal analysis, validation, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing. **S. Sivagnanam:** Data curation, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing. **C.B. Betts:** Data curation, investigation, methodology, writing–review and editing. **K. Betre:** Investigation, methodology, writing–review and editing. **N. Kirchberger:** Investigation, methodology, writing–review and editing. **B.J. Tate:** Data curation, methodology, writing–review and editing. **E.E.**

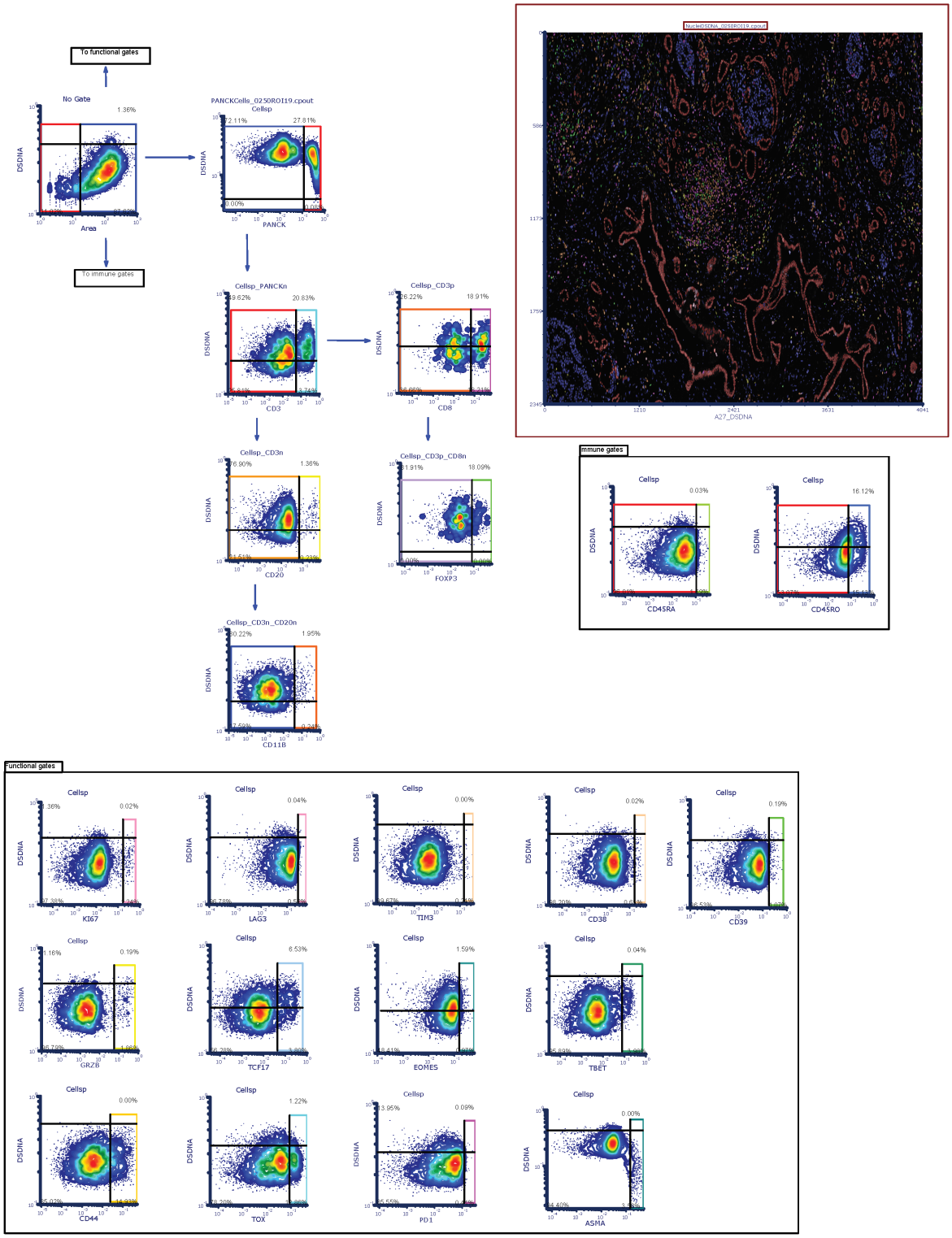
Furth: Investigation, writing–review and editing. **A. Dias Costa:** Investigation, writing–review and editing. **J.A. Nowak:** Investigation, writing–review and editing. **B.M. Wolpin:** Resources, writing–review and editing. **R.H. Vonderheide:** Resources, writing–review and editing. **J. Goecks:** Conceptualization, resources, supervision, funding acquisition, methodology, writing–original draft, writing–review and editing. **L.M. Coussens:** Conceptualization, resources, supervision, funding acquisition, methodology, writing–original draft, project administration, writing–review and editing. **K.T. Byrne:** Conceptualization, resources, supervision, funding acquisition, methodology, writing–original draft, project administration, writing–review and editing.

4.6 Supplementary Data

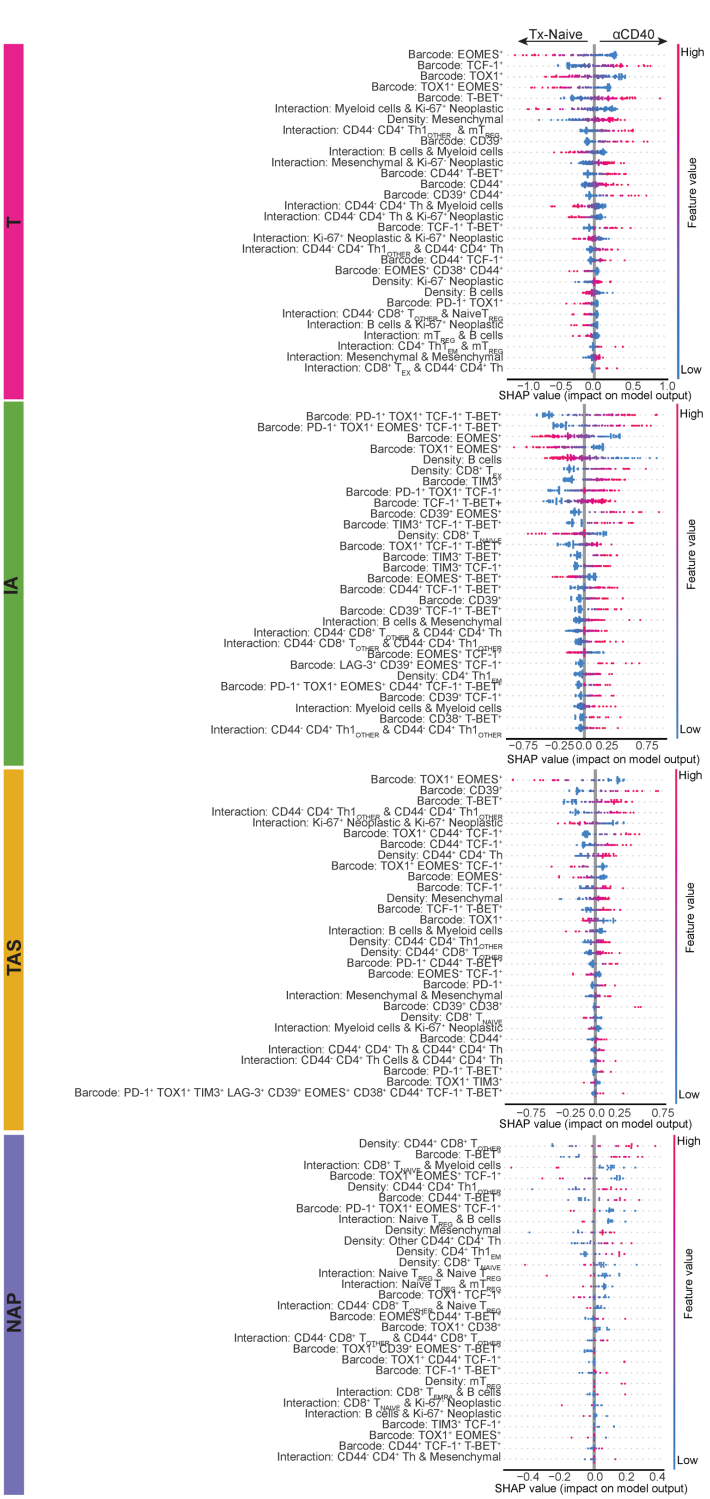


Supplementary Figure 4.1: mIHC antibody staining to single-cell phenotyping. A) Stacked bar chart showing percent tissue area (out of 100) sampled per resection from each patient. Bars are colored by histopathologic site of the regions sampled. **B)** Representative IHC staining of each antibody used in sequence in the panel. Scale bar = 50 µm. **C)** Two representative regions stained with CD3, CD8, and CD4 antibodies. For each region, top images show gates for CD8 on $CD3^+$ population (left) and CD4 on $CD3^+$ $CD8^-$ population (right), and bottom row shows pseudo-colored mIHC images.

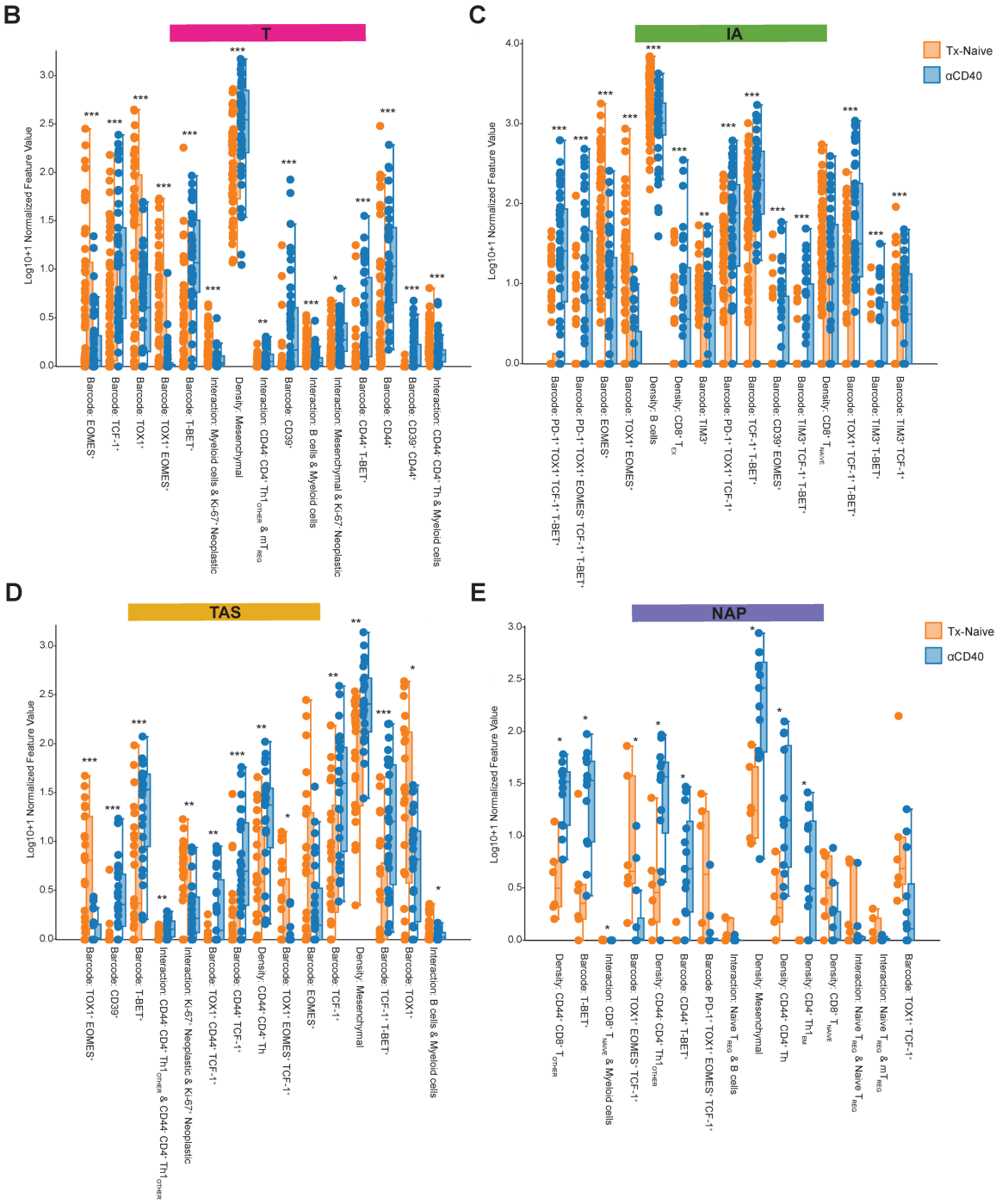
D



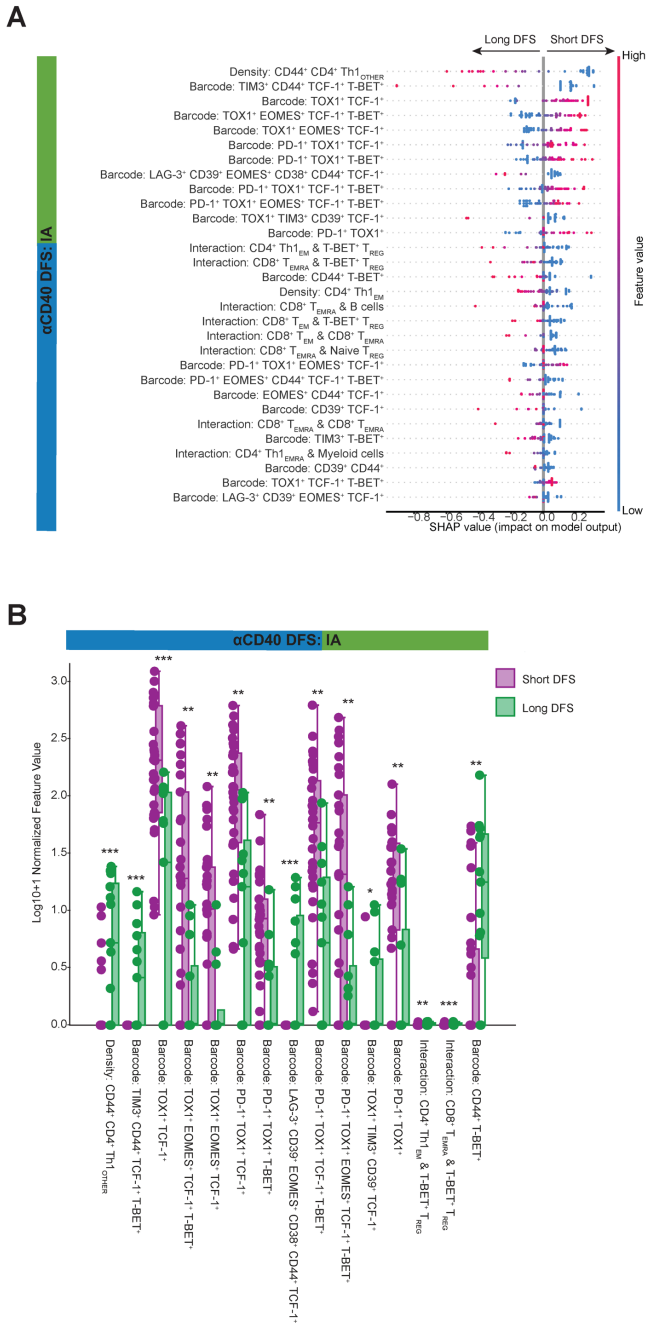
Supplementary Figure 4.1 (continued): D) Hierarchical gating template used to phenotype cells using image gating cytometry in FCS Image Cytometry RUO.



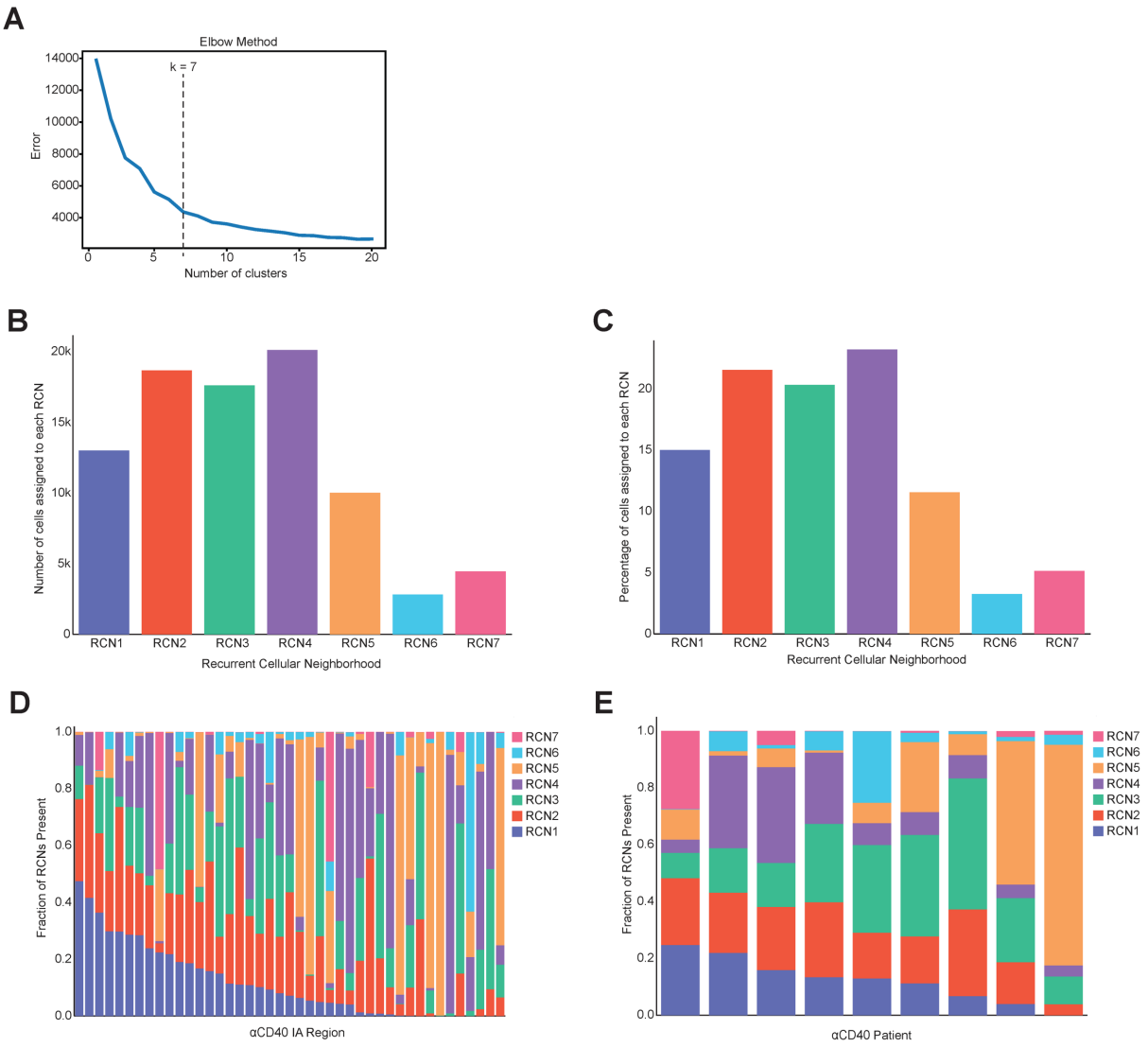
Supplementary Figure 4.2: Feature importance analyses for treatment classifier models. A) SHAP plots showing the top 30 features driving each histopathologic model. Features are ordered on the y-axis such that those with a larger impact on model's predictions appear at the top of the SHAP plots. SHAP values are shown on the x-axis, with a value of zero (center) indicating no impact on the model, and negative or positive SHAP values predicting treatment-naive or anti-CD40-treated tissues, respectively. Red or blue dots indicate presence or absence, respectively, of the corresponding feature in the tissue.



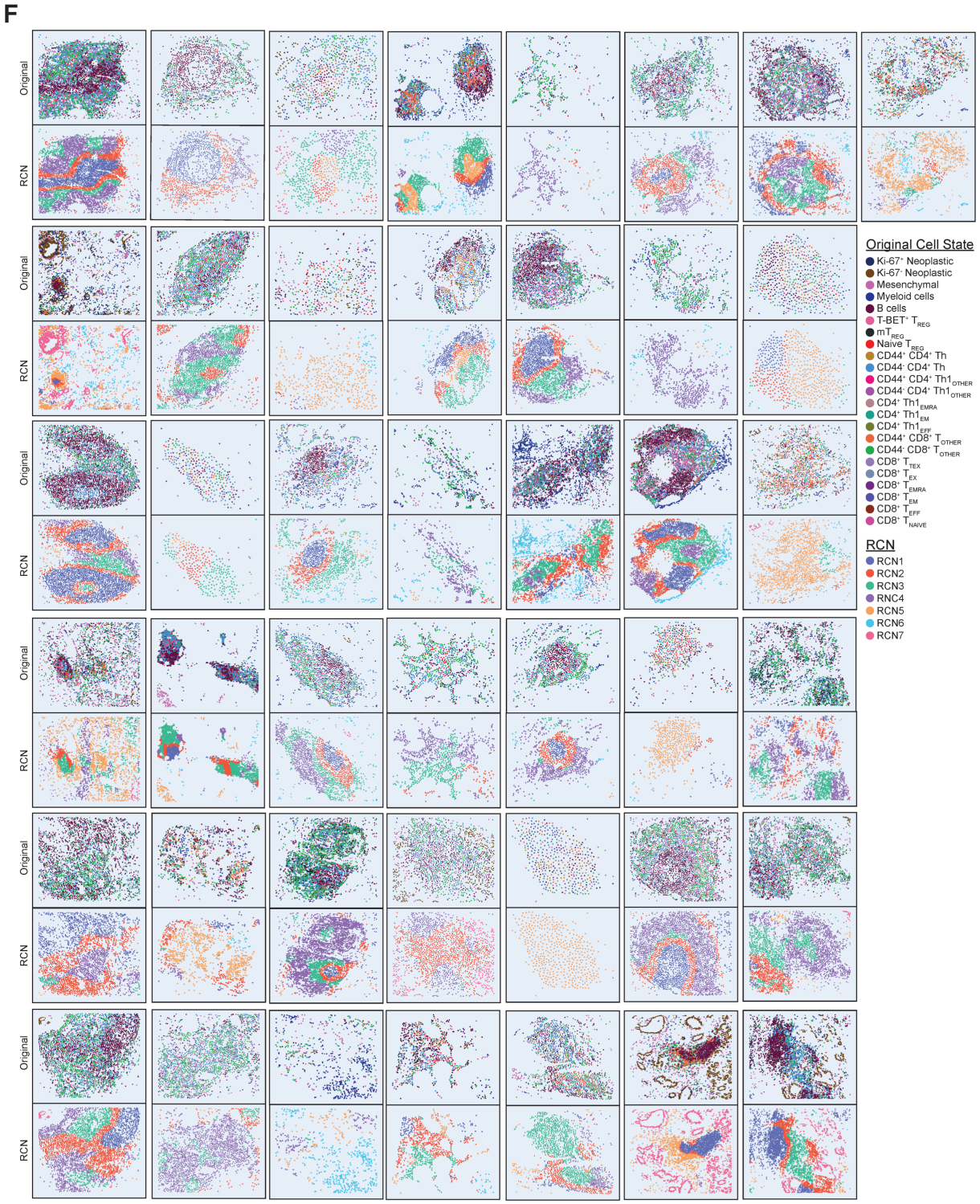
Supplementary Figure 4.2 (continued): B-E) Box plots showing feature values for each of the top 15 features for models derived from T, IA, TAS, or NAP sites, respectively, split by treatment cohort. Each dot represents the log10+1 normalized feature value for one tissue region, inputted into the classifier model. Boxes = quartile 1 (Q1) to quartile 3 (Q3); whiskers = smallest and largest datapoints within 1.5*interquartile range (IQR) +/- Q3/Q1; solid line = median. Mann–Whitney *U*-test used to determine statistical significance. *P*-values corrected using the Benjamini–Hochberg procedure. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$. **B)** T site, n= 55 treatment-naive and n = 48 anti-CD40-treated regions per feature. **C)** IA site, n= 89 treatment-naive and n = 43 anti-CD40-treated regions per feature. **D)** TAS site, n = 25 treatment-naive and n = 27 anti-CD40-treated regions per feature. **E)** NAP site, n = 6 treatment-naive and n = 13 anti-CD40-treated regions per feature.



Supplementary Figure 4.3: Feature importance analyses for DFS model derived from anti-CD40-treated IA regions. A) SHAP plot showing the top 30 features driving the IA model. Features are ordered on the y-axis such that those with a larger impact on the model's predictions appear at the top of the SHAP plot. SHAP values are shown on the x-axis, with a value of zero (center) indicating no impact on the model, and negative or positive SHAP values predicting long DFS or short DFS, respectively. Red or blue dots indicate presence or absence, respectively, of the corresponding feature in tissues. **B)** Box plot showing feature values for each of the top 15 features for the model derived from IA regions of the anti-CD40 cohort split by DFS group (n = 30 regions from short DFS patients per feature; n = 13 regions from long DFS patients per feature). Each dot represents the log10+1 normalized feature value for one tissue region, which was inputted into the classifier model. Boxes = Q1 to Q3; whiskers = smallest and largest datapoints within 1.5*IQR +/- Q3/Q1; solid line = median. Mann–Whitney *U*-test used to determine statistical significance. *P*-values corrected using the Benjamini–Hochberg procedure. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$.



Supplementary Figure 4.4: RCN distributions. A) Elbow plot showing optimal number of RCNs (k=7) for grouping cellular neighborhoods**. B)** Bar chart showing the number of cells assigned to each of the seven RCNs across all anti-CD40 IA regions. **C)** Bar chart showing the percentage (out of 100) of cells assigned to each of the seven RCNs across all anti-CD40 IA regions. **D)** Stacked bar chart showing fraction (out of 1.0) of RCNs present per anti-CD40 IA region. **E)** Stacked bar chart showing average proportion (out of 1.0) of RCNs present in IA regions for each anti-CD40-treated patient.



Supplementary Figure 4.4 (continued): F) Scatterplot reconstructions for each anti-CD40 IA region. Each dot represents a cell present in the IA, and each cell is colored by its original cell state phenotype (top scatterplot) or RCN assignment (bottom scatterplot).

Supplementary Table 4.1: Statistical comparison between the Liudahl et al. original PDAC cohort and the selected subset used as Cohort 1 in this study.

	Original Cohort	Selected Subset	*P*
OS (months) (mean ± SEM)	27.34 ± 2.25	27.27 ± 4.75	0.90
# cases OHSU/DFCI	45/59	6/12	0.61
$CD8^+$ T cell density (mean ± SEM)	144.57 ± 11.49	152.51 ± 26.35	0.63
$CD4^+$ T cell density (mean ± SEM)	185.95 ± 12.45	215.15 ± 30.41	0.33
$CD20^+$ B cell density (mean ± SEM)	98.01 ± 10.68	97.73 ± 21.11	0.85
Plasmablast density (mean ± SEM)	12.13 ± 2.13	10.78 ± 2.86	0.89
Plasma cell density (mean ± SEM)	3.31 ± 0.69	2.51 ± 1.57	0.67
Mast cell density (mean ± SEM)	30.00 ± 2.51	21.25 ± 3.32	0.34
Neutrophil/Eosinophil density (mean ± SEM)	155.74 ± 29.28	220.87 ± 94.11	0.39
Mature Dendritic cell density (mean ± SEM)	3.32 ± 0.32	3.46 ± 0.67	0.59
Immature Dendritic cell density (mean ± SEM)	118.52 ± 11.86	108.79 ± 19.72	0.89
$CD163^-$ Monocyte -Macrophage density (mean ± SEM)	51.68 ± 4.17	37.79 ± 5.50	0.30
$CD163^+$ Monocyte -Macrophage density (mean ± SEM)	73.94 ± 7.16	62.31 ± 14.26	0.59
$CD8^+/CD68^+$ Ratio (mean ± SEM)	1.86 ± 0.24	2.41 ± 0.89	0.21
Total $CD68^+$ cell density (mean ± SEM)	125.61 ± 9.84	100.09 ± 17.97	0.35

Mean value and standard error of the mean (SEM) shown for each variable. *P*-values computed using Fisher's exact test for categorical variables, Wilcoxon rank-sum test for continuous variables, and log rank test for overall survival.

Supplementary Table 4.2: Tissue area sampled per patient out of total tissue area resected per patient (in mm^2).

Tx Cohort	Total Area Sampled	Total Area Resected	% Area Sampled
Naive	21.14	237.82	8.89
Naive	37.55	448.71	8.37
Naive	39.6	359.52	11.02
Naive	26.96	287.51	9.38
Naive	25.32	321.44	7.88
Naive	26.79	184.38	14.53
Naive	15.5	147.66	10.5
Naive	33.32	258.02	12.92
Naive	41.26	301.59	13.68
Naive	52.43	312.98	16.75
Naive	12.06	94.23	12.8
Naive	24.74	286.47	8.64
Naive	40.69	191.3	21.27
Naive	16.96	308.5	5.5
Naive	18.95	89.66	21.14
Naive	24.95	230.26	10.83
Naive	31.28	249.46	12.54
Naive	31.79	215.06	14.78
αCD40	3.46	77.77	4.44
αCD40	23.87	130.7	18.27
αCD40	87.28	328.62	26.56

αCD40	50.65	204.14	24.81
αCD40	43.87	158.12	27.75
αCD40	62.69	175.92	35.63
αCD40	55	223.84	24.57
αCD40	31.69	202.15	15.68
αCD40	47.53	239.02	19.89
αCD40	20.96	143.39	14.62
αCD40	13.08	182.58	7.16

Supplementary Table 4.3: Table of antibodies used in mIHC panel.

Primary Antibody	Clone	Species	Supplier	Catalog number
TOX1 (Tox)	NAN488B	Rat	Abcam	ab237009
PD-1 (PD1)	NAT105	Mouse	Abcam	ab52597
T-BET (TBET)	D6N8B	Rabbit	Cell Signaling Technologies	13232S
CD39	A1	Mouse	Biolegend	328202
TIM3	D5D5R	Rabbit	Cell Signaling Technologies	45208S
LAG-3 (Lag3)	17B4	Mouse	Novus Biologicals	NBP1-97657
TCF-1 (TCF1/7)	C6D39	Rabbit	Cell Signaling Technologies	2203
CD44	156-3C11	Mouse	Cell Signaling Technologies	3570
CD38	38C03	Mouse	Thermofisher Scientific	MA5-14413
CD3	SP7	Rabbit	Thermofisher Scientific	RM-9107-S
CD45RO	UCHL-1	Mouse	Thermofisher Scientific	MA5-11532
GrzB (Granzyme B)	Polyclonal	Rabbit	Abcam	ab4059
CD8	C8/144B	Mouse	Thermofisher Scientific	MA5-13473
CD11b (CD11B)	EPR1344	Rabbit	Abcam	ab133357
CD45RA	F8-11-13	Mouse	Abcam	ab59168
FOXP3 (Foxp3)	236A/E7	Mouse	eBioscience	14-4777-82
EOMES	Polyclonal	Rabbit	EMD Millipore	AB2283
CD20	L26	Mouse	Abcam	ab9475
Ki-67 (KI67)	SP6	Rabbit	Sigma/Cell Marque	275R-14
αSMA (Alpha-SMA)	Polyclonal	Rabbit	Abcam	ab5694
PanCK (Pan-Cytokeratin)	AE1/AE3	Mouse	Abcam	ab27988

Supplementary Table 4.4: Raw counts of cell states defined by mIHC gating strategy present in the dataset.

Cell State	Cell Markers	Count
Ki-67- Neoplastic	CD3-CD20-CD11b-αSMA-PanCK+Ki-67-	915586
Myeloid cells	CD3-CD20-CD11b+	593571
Mesenchymal	CD3-CD20-CD11b-αSMA+	247778
CD44- CD4 Th	CD3+CD8-FOXP3-T-BET-CD44- & CD45RA+ and/or CD45RO+	163809
CD44- CD8 TOTHER	CD3+CD8+CD44-	136543

B cells	CD3-CD20+	132728
Ki-67+ Neoplastic	CD3-CD20-CD11b-αSMA-PanCK+Ki-67+	115294
mTREG	CD3+CD8-FOXP3+T-BET-CD45RA-	30836
CD44+ CD4 Th	CD3+CD8-FOXP3-T-BET-CD44+ & CD45RA+ and/or CD45RO+	24232
CD44+ CD8 TOTHER	CD3+CD8+CD44+	21711
CD44- CD4 Th1OTHER	CD3+CD8-FOXP3-T-BET+CD44-	21157
Naive TREG	CD3+CD8-FOXP3+T-BET-CD45RA+	10075
TBET+ TREG	CD3+CD8-FOXP3+T-BET+	3674
CD8 TEMRA	CD3+CD8+CD44+TCF-1+CD45RO+CD45RA+EOMES+	3550
CD8 TNAIVE	CD3+CD8+CD44-TCF-1+CD45RO-CD45RA+	2328
CD4 Th1EM	CD3+CD8-FOXP3-T-BET+CD44+CD45RA-CD45RO+	2038
CD4 Th1EMRA	CD3+CD8-FOXP3-T-BET+CD44+CD45RA+CD45RO+	1165
CD8 TEM	CD3+CD8+CD44+TCF-1+CD45RO+CD45RA-EOMES+	940
CD8 TEX	CD3+CD8+CD44+TCF-1+CD45RO+CD45RA-PD-1+	525
CD44+ CD4 Th1OTHER	CD3+CD8-FOXP3-T-BET+CD44+	449
CD8 TTEX	CD3+CD8+CD44+TCF-1-CD45RO+CD45RA-PD-1+	200
CD8 TEFF	CD3+CD8+CD44+T-BET+CD45RO-CD45RA+PD-1+	64
CD4 Th1EFF	CD3+CD8-CD44+CD45RO-CD45RA+PD-1+	21

Supplementary Table 4.5: Raw counts of T cells expressing each functionality barcode present in the dataset.

T cell Functionality Barcode	Count
Negative for all	144053
TOX1+	51606
TCF-1+	50915
CD44+	19724
T-BET+	15539
EOMES+	15405
TCF-1+ T-BET+	12775
TOX1+ TCF-1+	10565
CD44+ TCF-1+	7742
EOMES+ TCF-1+	6118
TOX1+ T-BET+	5000
TOX1+ EOMES+	4876
TOX1+ TCF-1+ T-BET+	4403
EOMES+ CD44+	3505
EOMES+ CD44+ TCF-1+	3318
TOX1+ CD44+	3002
PD-1+ TOX1+ TCF-1+	2684
PD-1+ TOX1+	2669

CD44+ T-BET+	2174
CD44+ TCF-1+ T-BET+	2103
EOMES+ TCF-1+ T-BET+	1862
CD39+	1804
TOX1+ EOMES+ TCF-1+	1629
TOX1+ CD44+ TCF-1+	1591
TOX1+ TIM3+	1574
TOX1+ CD39+	1472
CD38+ CD44+	1356
PD-1+ TOX1+ TCF-1+ T-BET+	1262
PD-1+	1177
TOX1+ EOMES+ TCF-1+ T-BET+	1153
EOMES+ T-BET+	1152
TIM3+	1084
TOX1+ CD38+	949
CD38+	936
PD-1+ TOX1+ EOMES+ TCF-1+ T-BET+	923
TOX1+ EOMES+ CD44+	851
PD-1+ TOX1+ EOMES+	840
PD-1+ TCF-1+	833
LAG-3+	819
PD-1+ TOX1+ EOMES+ TCF-1+	785
EOMES+ CD44+ TCF-1+ T-BET+	751
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	728
TOX1+ EOMES+ T-BET+	582
PD-1+ TOX1+ CD44+ TCF-1+	553
TOX1+ CD38+ CD44+	512
TOX1+ CD39+ EOMES+	495
TOX1+ CD44+ TCF-1+ T-BET+	483
CD39+ EOMES+	478
LAG-3+ TCF-1+	473
TOX11+ CD44+ T-BET+	461
PD-1+ TOX11+ T-BET+	457
CD39+ TCF-1+	448
TOX11+ EOMES+ CD44+ TCF-1+	437
TIM3+ CD44+	430
PD-1+ T-BET+	386
EOMES+ CD38+ CD44+	351
CD38+ CD44+ TCF-1+	341
EOMES+ CD38+ CD44+ TCF-1+	319
PD-1+ TOX1+ CD39+	310

TOX1+ LAG-3+ CD39+	304
TOX1+ LAG-3+	303
EOMES+ CD44+ T-BET+	292
TIM3+ TCF-1+	281
LAG-3+ T-BET+	269
PD-1+ EOMES+ TCF-1+	256
EOMES+ CD38+	252
TIM3+ T-BET+	249
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	247
PD-1+ TOX1+ CD44+	244
CD38+ TCF-1+	228
PD-1+ TCF-1+ T-BET+	216
PD-1+ EOMES+	216
TOX1+ TIM3+ T-BET+	208
LAG-3+ EOMES+ TCF-1+	198
LAG-3+ TCF-1+ T-BET+	193
CD39+ CD44+	190
TOX1+ CD38+ CD44+ TCF-1+	178
PD-1+ TOX1+ LAG-3+ CD39+	175
PD-1+ TOX1+ CD44+ TCF-1+ T-BET+	173
LAG-3+ EOMES+	167
PD-1+ CD44+ TCF-1+	166
TOX1+ TIM3+ CD39+ CD44+ TCF-1+ T-BET+	165
TOX1+ EOMES+ CD44+ TCF-1+ T-BET+	162
CD39+ T-BET+	161
TOX1+ TIM3+ CD44+	157
TOX1+ TIM3+ TCF-1+	150
PD-1+ CD44+	149
TIM3+ TCF-1+ T-BET+	144
TOX1+ EOMES+ CD44+ T-BET+	144
CD39+ EOMES+ TCF-1+	144
CD39+ TCF-1+ T-BET+	141
PD-1+ TOX1+ EOMES+ T-BET+	139
TOX1+ CD39+ EOMES+ TCF-1+	138
TOX1+ TIM3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	137
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+	135
TOX1+ EOMES+ CD38+ CD44+ TCF-1+	133
TOX1+ CD39+ TCF-1+	133
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	129
TOX1+ CD38+ T-BET+	125
PD-1+ TOX1+ CD39+ EOMES+	124

EOMES+ CD38+ TCF-1+	124
TOX1+ LAG-3+ TCF-1+	122
PD-1+ TOX1+ LAG-3+	118
PD-1+ TOX1+ EOMES+ CD44+ TCF-1+	116
TOX1+ LAG-3+ EOMES+	115
PD-1+ EOMES+ TCF-1+ T-BET+	115
TIM3+ EOMES+	113
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	112
TOX1+ EOMES+ CD38+ CD44+	111
PD-1+ TOX1+ TIM3+	109
LAG-3+ CD44+ TCF-1+	109
LAG-3+ CD39+	107
LAG-3+ CD44+	107
TOX1+ TIM3+ EOMES+	105
TIM3+ CD44+ TCF-1+	103
PD-1+ TOX1+ CD39+ EOMES+ TCF-1+	101
CD38+ T-BET+	100
TOX1+ CD39+ EOMES+ CD38+ TCF-1+	99
TIM3+ CD39+	97
TOX1+ TIM3+ TCF-1+ T-BET+	97
PD-1+ TOX1+ EOMES+ CD44+ TCF-1+ T-BET+	97
LAG-3+ EOMES+ TCF-1+ T-BET+	95
CD39+ EOMES+ TCF-1+ T-BET+	94
CD39+ CD38+	94
TOX1+ LAG-3+ CD39+ EOMES+	94
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	94
CD38+ CD44+ T-BET+	93
CD39+ EOMES+ CD38+ CD44+	93
TOX1+ CD39+ T-BET+	91
TIM3+ CD44+ T-BET+	91
CD38+ TCF-1+ T-BET+	90
CD39+ EOMES+ CD38+	90
TOX1+ EOMES+ CD38+	88
TOX1+ CD38+ TCF-1+	88
TOX1+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	88
TOX1+ LAG-3+ EOMES+ TCF-1+	86
CD38+ CD44+ TCF-1+ T-BET+	85
LAG-3+ EOMES+ CD44+ TCF-1+	83
CD39+ CD38+ CD44+	83
TOX1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	82
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	78

PD-1+ TOX1+ LAG-3+ EOMES+	75
TOX1+ EOMES+ CD38+ TCF-1+	75
PD-1+ TOX1+ CD39+ TCF-1+	73
PD-1+ CD44+ TCF-1+ T-BET+	71
PD-1+ TOX1+ CD38+	70
PD-1+ EOMES+ T-BET+	69
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	67
PD-1+ TOX1+ EOMES+ CD44+	67
PD-1+ EOMES+ CD44+ TCF-1+ T-BET+	67
TOX1+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	64
TOX1+ CD39+ TCF-1+ T-BET+	63
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	63
TIM3+ EOMES+ CD44+	62
CD39+ EOMES+ T-BET+	61
TOX1+ TIM3+ EOMES+ TCF-1+	61
TIM3+ EOMES+ T-BET+	60
TOX1+ LAG-3+ T-BET+	60
PD-1+ EOMES+ CD44+ TCF-1+	59
CD39+ EOMES+ CD44+	59
PD-1+ TOX1+ LAG-3+ EOMES+ TCF-1+	59
LAG-3+ CD39+ TCF-1+	57
PD-1+ TOX1+ CD44+ T-BET+	57
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	56
LAG-3+ CD44+ TCF-1+ T-BET+	56
TIM3+ EOMES+ TCF-1+	54
PD-1+ TOX1+ LAG-3+ T-BET+	53
PD-1+ TOX1+ CD39+ EOMES+ TCF-1+ T-BET+	53
CD39+ EOMES+ CD38+ CD44+ TCF-1+	53
PD-1+ TOX1+ TIM3+ TCF-1+	53
CD39+ CD44+ TCF-1+	52
TOX1+ TIM3+ CD44+ TCF-1+	52
PD-1+ TOX1+ LAG-3+ TCF-1+	51
TOX1+ CD39+ EOMES+ CD38+	50
CD39+ EOMES+ CD38+ TCF-1+	49
PD-1+ CD44+ T-BET+	48
TOX1+ CD39+ EOMES+ CD38+ CD44+	47
TIM3+ EOMES+ CD44+ TCF-1+	47
PD-1+ TOX1+ CD39+ TCF-1+ T-BET+	47
TOX1+ CD39+ CD44+ TCF-1+ T-BET+	47
CD39+ EOMES+ CD44+ TCF-1+	46
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	45

PD-1+ TOX1+ TIM3+ TCF-1+ T-BET+	44
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	43
TOX1+ CD39+ CD44+	43
CD39+ EOMES+ CD44+ TCF-1+ T-BET+	43
LAG-3+ CD38+ CD44+ TCF-1+	43
TOX1+ TIM3+ CD39+	43
LAG-3+ EOMES+ CD44+	42
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+	42
TOX1+ CD38+ CD44+ TCF-1+ T-BET+	42
EOMES+ CD38+ CD44+ TCF-1+ T-BET+	42
LAG-3+ EOMES+ T-BET+	41
TIM3+ CD39+ TCF-1+	41
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+	41
PD-1+ CD39+	40
LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	40
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	40
LAG-3+ CD39+ EOMES+	39
PD-1+ EOMES+ CD44+	39
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	38
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	37
PD-1+ TOX1+ LAG-3+ EOMES+ T-BET+	37
TOX1+ LAG-3+ TCF-1+ T-BET+	37
PD-1+ TOX1+ TIM3+ T-BET+	37
TOX1+ TIM3+ CD38+ CD44+	37
PD-1+ TOX1+ CD38+ CD44+	37
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ TCF-1+	36
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	36
TOX1+ CD39+ EOMES+ T-BET+	35
TOX1+ TIM3+ EOMES+ CD44+	35
PD-1+ TOX1+ TIM3+ CD38+ CD44+ TCF-1+ T-BET+	34
TIM3+ CD39+ CD44+	34
LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	34
TOX1+ TIM3+ CD38+	34
TOX1+ CD38+ TCF-1+ T-BET+	33
PD-1+ LAG-3+ EOMES+ TCF-1+	33
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	32
PD-1+ TOX1+ TIM3+ CD38+ CD44+ T-BET+	32
LAG-3+ CD44+ T-BET+	32
TOX1+ CD39+ EOMES+ CD44+ TCF-1+	32
TIM3+ CD44+ TCF-1+ T-BET+	32
TIM3+ CD39+ CD38+	32

CD39+ CD44+ TCF-1+ T-BET+	32
TIM3+ CD38+ CD44+	31
LAG-3+ EOMES+ CD38+ CD44+	31
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	31
PD-1+ LAG-3+ EOMES+	31
PD-1+ TOX1+ EOMES+ CD44+ T-BET+	31
TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	31
LAG-3+ CD38+ TCF-1+	31
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	30
TOX1+ CD39+ CD38+	30
TOX1+ TIM3+ CD44+ T-BET+	30
TIM3+ EOMES+ TCF-1+ T-BET+	30
TOX1+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	30
PD-1+ LAG-3+	30
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	30
TOX1+ CD39+ CD44+ TCF-1+	29
PD-1+ TOX1+ CD39+ T-BET+	29
TOX1+ TIM3+ CD44+ TCF-1+ T-BET+	29
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	29
LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	29
LAG-3+ CD39+ EOMES+ CD38+ CD44+	29
TIM3+ LAG-3+	29
LAG-3+ CD38+	28
TOX1+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	28
LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	28
TIM3+ LAG-3+ T-BET+	28
TOX1+ LAG-3+ EOMES+ TCF-1+ T-BET+	28
TOX1+ CD39+ EOMES+ TCF-1+ T-BET+	28
TOX1+ TIM3+ EOMES+ CD44+ TCF-1+	28
LAG-3+ CD39+ EOMES+ CD38+	28
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+ TCF-1+	27
TIM3+ EOMES+ CD44+ TCF-1+ T-BET+	27
PD-1+ TOX1+ LAG-3+ CD39+ TCF-1+	27
PD-1+ LAG-3+ TCF-1+	27
PD-1+ LAG-3+ CD38+ TCF-1+ T-BET+	27
TOX1+ LAG-3+ CD39+ TCF-1+	27
PD-1+ TOX1+ LAG-3+ CD39+ T-BET+	27
TOX1+ CD38+ CD44+ T-BET+	26
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	26
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+	26
LAG-3+ EOMES+ CD38+ TCF-1+	26

TOX1+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	26
LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	26
TOX1+ TIM3+ LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	26
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD44+ TCF-1+	26
LAG-3+ CD39+ EOMES+ TCF-1+	26
PD-1+ TIM3+	25
CD39+ CD38+ CD44+ TCF-1+	25
PD-1+ TOX1+ LAG-3+ TCF-1+ T-BET+	25
PD-1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	25
LAG-3+ CD38+ CD44+	25
TOX1+ TIM3+ LAG-3+	25
LAG-3+ CD39+ TCF-1+ T-BET+	25
TOX1+ LAG-3+ EOMES+ T-BET+	25
TOX1+ TIM3+ CD38+ CD44+ TCF-1+ T-BET+	25
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	24
PD-1+ TOX1+ CD39+ EOMES+ CD38+	24
CD39+ CD38+ TCF-1+	24
TIM3+ EOMES+ CD38+ CD44+ TCF-1+	24
TOX1+ TIM3+ EOMES+ CD38+ CD44+ TCF-1+	23
PD-1+ TOX1+ TIM3+ CD44+ TCF-1+	23
TOX1+ LAG-3+ CD39+ EOMES+ CD38+	23
PD-1+ TOX1+ CD39+ CD44+	23
PD-1+ TOX1+ CD39+ EOMES+ CD38+ TCF-1+	23
PD-1+ TOX1+ TIM3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	23
CD39+ CD44+ T-BET+	23
PD-1+ TOX1+ CD39+ EOMES+ CD44+ TCF-1+	23
TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	22
LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	22
TOX1+ TIM3+ EOMES+ T-BET+	22
TOX1+ CD39+ EOMES+ CD44+	22
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	22
EOMES+ CD38+ TCF-1+ T-BET+	22
LAG-3+ CD38+ TCF-1+ T-BET+	22
TIM3+ CD39+ T-BET+	22
TOX1+ LAG-3+ CD39+ T-BET+	22
LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	22
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	22
PD-1+ LAG-3+ EOMES+ CD44+ TCF-1+	21
PD-1+ TOX1+ CD39+ CD44+ TCF-1+	21
PD-1+ EOMES+ CD44+ T-BET+	21
TIM3+ LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	21

TIM3+ LAG-3+ EOMES+ TCF-1+	21
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+ CD44+	21
TOX1+ TIM3+ EOMES+ TCF-1+ T-BET+	21
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ T-BET+	21
TOX1+ EOMES+ CD38+ T-BET+	21
TOX1+ LAG-3+ EOMES+ CD38+ TCF-1+	20
TIM3+ LAG-3+ TCF-1+	20
TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+	20
TOX1+ TIM3+ CD38+ TCF-1+ T-BET+	20
TOX1+ TIM3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	20
PD-1+ TOX1+ TIM3+ CD44+	20
TOX1+ TIM3+ CD39+ TCF-1+	20
PD-1+ TOX1+ EOMES+ CD38+	20
EOMES+ CD38+ CD44+ T-BET+	20
PD-1+ TOX1+ CD38+ CD44+ TCF-1+	20
PD-1+ TOX1+ CD39+ EOMES+ T-BET+	19
TOX1+ CD39+ CD44+ T-BET+	19
TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	19
LAG-3+ CD39+ T-BET+	19
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+	19
LAG-3+ EOMES+ CD38+	19
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ T-BET+	19
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	19
PD-1+ CD39+ EOMES+	18
CD39+ CD38+ CD44+ TCF-1+ T-BET+	18
TIM3+ CD39+ CD38+ CD44+	18
TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	18
PD-1+ TOX1+ CD38+ TCF-1+	18
PD-1+ TOX1+ LAG-3+ EOMES+ TCF-1+ T-BET+	18
TIM3+ CD38+	18
TOX1+ LAG-3+ CD44+ TCF-1+	18
LAG-3+ CD39+ CD44+ TCF-1+	17
PD-1+ CD39+ TCF-1+	17
CD39+ CD38+ T-BET+	17
PD-1+ TOX1+ TIM3+ EOMES+ TCF-1+	17
TIM3+ CD38+ CD44+ TCF-1+ T-BET+	17
PD-1+ TOX1+ CD38+ T-BET+	17
PD-1+ TOX1+ EOMES+ CD38+ CD44+	17
PD-1+ TOX1+ CD39+ EOMES+ CD38+ CD44+	17
PD-1+ TOX1+ TIM3+ EOMES+	17
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	17

PD-1+ TOX1+ TIM3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	17
PD-1+ TOX1+ TIM3+ LAG-3+ CD44+ TCF-1+ T-BET+	17
TIM3+ EOMES+ CD44+ T-BET+	17
TOX1+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	17
LAG-3+ CD39+ CD38+	17
PD-1+ TOX1+ TIM3+ CD38+ CD44+	16
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	16
LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	16
PD-1+ LAG-3+ T-BET+	16
CD39+ CD38+ CD44+ T-BET+	16
PD-1+ LAG-3+ TCF-1+ T-BET+	16
TOX1+ LAG-3+ EOMES+ CD38+	16
TOX1+ TIM3+ EOMES+ CD38+ TCF-1+	16
TOX1+ LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	16
PD-1+ TOX1+ LAG-3+ EOMES+ CD44+	15
TOX1+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	15
PD-1+ TOX1+ LAG-3+ CD44+	15
PD-1+ LAG-3+ CD38+ TCF-1+	15
PD-1+ TOX1+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	14
PD-1+ TOX1+ EOMES+ CD38+ CD44+ TCF-1+	14
LAG-3+ CD39+ CD38+ CD44+ TCF-1+	14
LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	14
LAG-3+ CD39+ CD38+ T-BET+	14
EOMES+ CD38+ T-BET+	14
TOX1+ LAG-3+ CD44+	14
PD-1+ TOX1+ LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	14
TOX1+ TIM3+ EOMES+ CD44+ TCF-1+ T-BET+	14
PD-1+ TOX1+ TIM3+ CD39+ CD38+ CD44+ T-BET+	14
PD-1+ LAG-3+ CD39+	14
PD-1+ TOX1+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	14
PD-1+ LAG-3+ CD39+ T-BET+	14
TOX1+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	14
TOX1+ EOMES+ CD38+ TCF-1+ T-BET+	14
TIM3+ LAG-3+ EOMES+ T-BET+	14
TIM3+ CD39+ EOMES+	14
PD-1+ TIM3+ T-BET+	14
PD-1+ CD39+ TCF-1+ T-BET+	14
PD-1+ TOX1+ LAG-3+ CD39+ TCF-1+ T-BET+	13
TOX1+ LAG-3+ CD38+ TCF-1+ T-BET+	13
PD-1+ TOX1+ CD38+ TCF-1+ T-BET+	13
TIM3+ LAG-3+ CD38+ TCF-1+ T-BET+	13

LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	13
LAG-3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	13
TOX1+ LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	13
TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	13
PD-1+ TOX1+ CD39+ EOMES+ CD44+	13
PD-1+ TIM3+ LAG-3+ CD38+ TCF-1+ T-BET+	13
PD-1+ TOX1+ TIM3+ CD44+ TCF-1+ T-BET+	13
TOX1+ CD39+ CD38+ CD44+	13
TOX1+ LAG-3+ CD39+ CD38+	13
TOX1+ TIM3+ CD39+ EOMES+	13
TOX1+ CD39+ CD38+ T-BET+	13
PD-1+ TOX1+ TIM3+ CD38+ TCF-1+ T-BET+	13
PD-1+ TOX1+ TIM3+ EOMES+ TCF-1+ T-BET+	13
LAG-3+ CD39+ CD38+ CD44+	12
TIM3+ CD39+ TCF-1+ T-BET+	12
PD-1+ TOX1+ LAG-3+ EOMES+ CD44+ TCF-1+	12
PD-1+ TOX1+ LAG-3+ CD39+ CD38+ TCF-1+	12
CD39+ CD38+ TCF-1+ T-BET+	12
PD-1+ TOX1+ LAG-3+ CD44+ TCF-1+	12
PD-1+ TOX1+ LAG-3+ CD44+ TCF-1+ T-BET+	12
PD-1+ TOX1+ CD39+ CD38+	12
PD-1+ TOX1+ EOMES+ CD38+ TCF-1+	12
TOX1+ LAG-3+ EOMES+ CD44+ TCF-1+	12
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ TCF-1+	12
PD-1+ TIM3+ TCF-1+	12
TOX1+ LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	12
TOX1+ TIM3+ EOMES+ CD38+ TCF-1+ T-BET+	12
PD-1+ LAG-3+ CD39+ EOMES+ TCF-1+	12
LAG-3+ EOMES+ CD44+ T-BET+	12
PD-1+ TIM3+ CD44+	12
TOX1+ LAG-3+ CD38+	12
TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	12
TOX1+ LAG-3+ CD44+ T-BET+	12
TIM3+ LAG-3+ CD39+	12
TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	12
TOX1+ LAG-3+ EOMES+ CD38+ CD44+	12
LAG-3+ CD38+ T-BET+	12
CD39+ EOMES+ CD38+ TCF-1+ T-BET+	12
TIM3+ LAG-3+ CD38+ T-BET+	12
PD-1+ CD39+ EOMES+ TCF-1+ T-BET+	12
TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	12

PD-1+ CD39+ T-BET+	12
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	12
PD-1+ TOX1+ EOMES+ CD38+ TCF-1+ T-BET+	11
PD-1+ LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	11
PD-1+ LAG-3+ EOMES+ TCF-1+ T-BET+	11
LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	11
TOX1+ LAG-3+ CD39+ TCF-1+ T-BET+	11
TOX1+ CD39+ EOMES+ CD44+ T-BET+	11
PD-1+ LAG-3+ CD44+ TCF-1+	11
CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	11
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+ TCF-1+ T-BET+	11
TIM3+ CD38+ T-BET+	11
TOX1+ TIM3+ EOMES+ CD38+ CD44+	11
TIM3+ LAG-3+ EOMES+	11
TOX1+ TIM3+ EOMES+ CD44+ T-BET+	11
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	11
PD-1+ TOX1+ LAG-3+ CD39+ CD44+ TCF-1+	11
TIM3+ LAG-3+ CD44+	11
LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	11
LAG-3+ CD39+ EOMES+ CD44+	11
PD-1+ TOX1+ TIM3+ CD38+	11
TOX1+ TIM3+ CD38+ T-BET+	11
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	11
PD-1+ CD38+	11
TOX1+ TIM3+ CD38+ TCF-1+	11
TIM3+ CD38+ TCF-1+ T-BET+	10
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	10
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+ CD44+ T-BET+	10
PD-1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	10
TIM3+ CD38+ CD44+ T-BET+	10
TIM3+ LAG-3+ CD38+ CD44+ T-BET+	10
PD-1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ T-BET+	10
TIM3+ LAG-3+ EOMES+ TCF-1+ T-BET+	10
PD-1+ TOX1+ CD39+ CD38+ CD44+	10
PD-1+ LAG-3+ EOMES+ T-BET+	10
PD-1+ TOX1+ CD39+ EOMES+ CD44+ T-BET+	10
PD-1+ TOX1+ TIM3+ CD39+	10
PD-1+ TOX1+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	10
PD-1+ TOX1+ TIM3+ LAG-3+	10
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+ CD44+ T-BET+	10
TOX1+ LAG-3+ CD39+ EOMES+ T-BET+	10

LAG-3+ CD39+ CD44+	10
TIM3+ CD39+ CD44+ TCF-1+	10
TOX1+ TIM3+ LAG-3+ EOMES+ TCF-1+	10
TIM3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	9
PD-1+ TIM3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	9
TIM3+ CD39+ CD38+ TCF-1+	9
PD-1+ CD39+ CD44+ TCF-1+	9
PD-1+ CD38+ TCF-1+	9
PD-1+ LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	9
PD-1+ TIM3+ CD44+ TCF-1+ T-BET+	9
TOX1+ EOMES+ CD38+ CD44+ T-BET+	9
PD-1+ TOX1+ LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	9
PD-1+ TOX1+ TIM3+ CD39+ TCF-1+ T-BET+	9
LAG-3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	9
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ TCF-1+	9
TOX1+ TIM3+ CD39+ EOMES+ TCF-1+	9
TOX1+ LAG-3+ CD38+ T-BET+	9
PD-1+ TOX1+ LAG-3+ CD38+	9
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+	9
TOX1+ TIM3+ CD39+ CD38+ TCF-1+ T-BET+	9
TOX1+ TIM3+ CD38+ CD44+ TCF-1+	9
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+	9
TOX1+ LAG-3+ CD39+ CD44+ TCF-1+	9
TIM3+ CD38+ CD44+ TCF-1+	9
PD-1+ LAG-3+ EOMES+ CD38+ CD44+	8
TOX1+ LAG-3+ EOMES+ CD44+	8
PD-1+ TOX1+ TIM3+ CD39+ CD44+ TCF-1+ T-BET+	8
PD-1+ TOX1+ TIM3+ EOMES+ CD44+ TCF-1+	8
PD-1+ LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	8
PD-1+ TOX1+ TIM3+ EOMES+ CD44+	8
PD-1+ TOX1+ TIM3+ CD39+ EOMES+	8
PD-1+ LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	8
PD-1+ TOX1+ TIM3+ CD44+ T-BET+	8
LAG-3+ EOMES+ CD38+ T-BET+	8
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+ CD44+	8
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD44+ TCF-1+	8
CD39+ EOMES+ CD38+ CD44+ T-BET+	8
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+	8
CD39+ EOMES+ CD38+ T-BET+	8
CD39+ EOMES+ CD44+ T-BET+	8
PD-1+ CD39+ EOMES+ CD44+ TCF-1+	8

TOX1+ CD39+ CD38+ TCF-1+	8
PD-1+ CD39+ EOMES+ CD44+	8
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+	8
TOX1+ LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	8
PD-1+ CD39+ EOMES+ TCF-1+	8
TOX1+ TIM3+ LAG-3+ TCF-1+	8
PD-1+ TOX1+ CD39+ CD44+ T-BET+	8
PD-1+ TOX1+ CD39+ EOMES+ CD38+ CD44+ T-BET+	8
TOX1+ LAG-3+ CD39+ EOMES+ CD44+	8
PD-1+ TOX1+ LAG-3+ EOMES+ CD44+ T-BET+	8
PD-1+ TOX1+ CD39+ CD44+ TCF-1+ T-BET+	8
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	8
PD-1+ TOX1+ CD38+ CD44+ TCF-1+ T-BET+	8
PD-1+ TOX1+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	8
TIM3+ EOMES+ CD38+ CD44+	8
TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	8
PD-1+ TIM3+ EOMES+	8
TOX1+ TIM3+ LAG-3+ CD39+	8
TOX1+ TIM3+ LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	8
PD-1+ TOX1+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	8
PD-1+ LAG-3+ CD38+	7
TIM3+ EOMES+ CD38+	7
PD-1+ TOX1+ CD38+ CD44+ T-BET+	7
PD-1+ TOX1+ TIM3+ CD38+ T-BET+	7
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+	7
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+	7
PD-1+ TOX1+ LAG-3+ CD44+ T-BET+	7
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	7
PD-1+ TOX1+ LAG-3+ CD39+ CD38+	7
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	7
LAG-3+ CD39+ CD38+ TCF-1+	7
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD44+	7
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	7
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	7
PD-1+ TOX1+ EOMES+ CD38+ T-BET+	7
TOX1+ TIM3+ LAG-3+ CD39+ CD38+ T-BET+	7
TOX1+ LAG-3+ CD39+ CD38+ T-BET+	7
TIM3+ LAG-3+ CD38+ CD44+ TCF-1+	7
TOX1+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	7
TOX1+ LAG-3+ CD39+ CD44+	7
TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	7

TOX1+ TIM3+ LAG-3+ CD38+ TCF-1+ T-BET+	7
TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+	7
TOX1+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	7
PD-1+ CD38+ CD44+ TCF-1+	7
PD-1+ CD38+ CD44+	7
TOX1+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+	7
TOX1+ TIM3+ LAG-3+ EOMES+ TCF-1+ T-BET+	7
TOX1+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	7
TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	7
TIM3+ LAG-3+ TCF-1+ T-BET+	7
TOX1+ TIM3+ CD39+ TCF-1+ T-BET+	7
TOX1+ TIM3+ CD39+ T-BET+	7
TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+	7
LAG-3+ EOMES+ CD38+ CD44+ T-BET+	6
TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ T-BET+	6
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD38+ T-BET+	6
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ T-BET+	6
LAG-3+ CD39+ EOMES+ T-BET+	6
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ TCF-1+ T-BET+	6
TIM3+ LAG-3+ CD39+ EOMES+	6
PD-1+ TOX1+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	6
PD-1+ LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	6
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	6
TOX1+ TIM3+ LAG-3+ CD44+ TCF-1+ T-BET+	6
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+	6
TOX1+ LAG-3+ CD44+ TCF-1+ T-BET+	6
LAG-3+ CD39+ EOMES+ CD38+ T-BET+	6
PD-1+ TOX1+ TIM3+ LAG-3+ TCF-1+ T-BET+	6
TIM3+ LAG-3+ CD39+ CD38+	6
TOX1+ TIM3+ CD39+ CD38+ CD44+	6
PD-1+ TIM3+ CD44+ TCF-1+	6
TOX1+ LAG-3+ CD39+ CD44+ T-BET+	6
PD-1+ CD39+ CD38+ T-BET+	6
TOX1+ CD39+ CD38+ CD44+ TCF-1+	6
TIM3+ CD39+ CD44+ TCF-1+ T-BET+	6
TIM3+ EOMES+ CD38+ TCF-1+ T-BET+	6
PD-1+ TOX1+ TIM3+ CD39+ T-BET+	6
TOX1+ LAG-3+ EOMES+ CD38+ T-BET+	6
PD-1+ TOX1+ TIM3+ CD39+ TCF-1+	6
PD-1+ LAG-3+ CD44+	6
PD-1+ LAG-3+ CD44+ T-BET+	6

LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	6
PD-1+ TIM3+ EOMES+ CD44+ TCF-1+	6
TOX1+ TIM3+ CD39+ CD44+	6
TIM3+ CD39+ CD38+ T-BET+	6
TOX1+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	6
PD-1+ TOX1+ LAG-3+ CD38+ TCF-1+ T-BET+	6
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ TCF-1+ T-BET+	6
TIM3+ CD39+ CD38+ TCF-1+ T-BET+	6
PD-1+ CD38+ TCF-1+ T-BET+	6
TOX1+ CD39+ EOMES+ CD38+ T-BET+	6
PD-1+ LAG-3+ CD39+ EOMES+ T-BET+	6
TIM3+ LAG-3+ CD39+ CD38+ T-BET+	6
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+ T-BET+	5
PD-1+ TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+	5
TOX1+ TIM3+ CD39+ CD38+ CD44+ T-BET+	5
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ T-BET+	5
PD-1+ TOX1+ LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	5
TOX1+ TIM3+ CD39+ CD38+ T-BET+	5
TOX1+ TIM3+ CD39+ CD38+	5
TOX1+ TIM3+ LAG-3+ CD39+ CD38+	5
PD-1+ TIM3+ EOMES+ TCF-1+ T-BET+	5
PD-1+ TIM3+ EOMES+ CD44+ TCF-1+ T-BET+	5
PD-1+ TOX1+ TIM3+ CD38+ CD44+ TCF-1+	5
TOX1+ LAG-3+ CD39+ CD38+ TCF-1+	5
TOX1+ LAG-3+ CD39+ EOMES+ TCF-1+	5
PD-1+ TOX1+ LAG-3+ CD39+ CD44+	5
PD-1+ TOX1+ LAG-3+ CD39+ CD38+ T-BET+	5
PD-1+ TOX1+ LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	5
TOX1+ TIM3+ CD39+ CD44+ TCF-1+	5
TOX1+ LAG-3+ CD38+ CD44+	5
TOX1+ TIM3+ LAG-3+ CD38+ TCF-1+	5
PD-1+ LAG-3+ CD39+ TCF-1+ T-BET+	5
TOX1+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	5
PD-1+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	5
PD-1+ LAG-3+ EOMES+ CD44+	5
PD-1+ EOMES+ CD38+ TCF-1+	5
TIM3+ LAG-3+ CD38+ TCF-1+	5
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD44+	5
TOX1+ TIM3+ LAG-3+ CD38+ CD44+ TCF-1+	5
TOX1+ TIM3+ LAG-3+ EOMES+	5
TIM3+ CD39+ EOMES+ TCF-1+ T-BET+	5

TOX1+ TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	5
TIM3+ EOMES+ CD38+ CD44+ T-BET+	5
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+	5
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	5
PD-1+ CD39+ CD38+ TCF-1+	5
TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	5
PD-1+ LAG-3+ CD39+ TCF-1+	5
PD-1+ TOX1+ LAG-3+ CD39+ CD38+ CD44+ T-BET+	5
PD-1+ LAG-3+ CD39+ EOMES+	5
PD-1+ TOX1+ LAG-3+ CD38+ CD44+	5
TOX1+ TIM3+ CD39+ EOMES+ CD44+ TCF-1+	5
TOX1+ TIM3+ LAG-3+ TCF-1+ T-BET+	5
PD-1+ TOX1+ LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	5
PD-1+ TIM3+ LAG-3+ CD39+ CD38+ T-BET+	5
PD-1+ TOX1+ LAG-3+ CD38+ T-BET+	5
TIM3+ LAG-3+ CD38+	5
PD-1+ LAG-3+ CD39+ CD44+ TCF-1+	5
TOX1+ LAG-3+ CD38+ TCF-1+	5
TIM3+ LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	5
PD-1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	4
PD-1+ TIM3+ LAG-3+ T-BET+	4
TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+	4
TIM3+ EOMES+ CD38+ TCF-1+	4
TOX1+ TIM3+ CD39+ EOMES+ TCF-1+ T-BET+	4
TIM3+ EOMES+ CD38+ T-BET+	4
TOX1+ TIM3+ CD39+ CD38+ CD44+ TCF-1+	4
PD-1+ TIM3+ LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	4
TOX1+ TIM3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	4
PD-1+ TOX1+ CD39+ CD38+ T-BET+	4
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+	4
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	4
PD-1+ TOX1+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+	4
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	4
PD-1+ TIM3+ LAG-3+ EOMES+ CD38+ TCF-1+ T-BET+	4
PD-1+ TIM3+ EOMES+ TCF-1+	4
TIM3+ LAG-3+ EOMES+ CD44+ T-BET+	4
PD-1+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	4
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	4
PD-1+ TOX1+ TIM3+ CD39+ CD38+ CD44+	4
TOX1+ LAG-3+ CD38+ CD44+ TCF-1+	4
TOX1+ LAG-3+ CD38+ CD44+ T-BET+	4

PD-1+ CD39+ CD44+	4
PD-1+ LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	4
TIM3+ CD39+ EOMES+ CD44+ TCF-1+	4
PD-1+ LAG-3+ CD44+ TCF-1+ T-BET+	4
PD-1+ TIM3+ TCF-1+ T-BET+	4
PD-1+ EOMES+ CD38+ CD44+ TCF-1+	4
TIM3+ LAG-3+ CD39+ CD38+ CD44+ T-BET+	4
PD-1+ LAG-3+ CD38+ T-BET+	4
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD44+ T-BET+	4
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+ CD44+ TCF-1+	4
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ TCF-1+	4
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ TCF-1+ T-BET+	4
PD-1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	4
PD-1+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	4
PD-1+ TOX1+ TIM3+ EOMES+ T-BET+	4
TOX1+ LAG-3+ CD39+ CD38+ CD44+	4
PD-1+ TIM3+ CD44+ T-BET+	4
PD-1+ TOX1+ TIM3+ EOMES+ CD44+ TCF-1+ T-BET+	4
TOX1+ TIM3+ CD38+ CD44+ T-BET+	4
TOX1+ LAG-3+ EOMES+ CD38+ CD44+ T-BET+	4
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ T-BET+	4
PD-1+ TOX1+ TIM3+ EOMES+ CD44+ T-BET+	4
PD-1+ TIM3+ CD39+ TCF-1+	4
PD-1+ TOX1+ TIM3+ EOMES+ CD38+ CD44+ TCF-1+	4
PD-1+ TIM3+ EOMES+ CD44+ T-BET+	4
TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+	3
TOX1+ CD39+ CD38+ TCF-1+ T-BET+	3
TIM3+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	3
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+	3
TOX1+ TIM3+ LAG-3+ CD38+ CD44+ T-BET+	3
TOX1+ TIM3+ LAG-3+ CD39+ TCF-1+ T-BET+	3
PD-1+ TOX1+ EOMES+ CD38+ CD44+ T-BET+	3
TOX1+ TIM3+ LAG-3+ EOMES+ CD44+	3
TIM3+ CD38+ TCF-1+	3
TIM3+ LAG-3+ CD39+ EOMES+ CD38+	3
TOX1+ TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+	3
TIM3+ LAG-3+ CD39+ EOMES+ CD44+ T-BET+	3
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+ T-BET+	3
TOX1+ TIM3+ LAG-3+ CD38+ CD44+	3
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+	3
TOX1+ TIM3+ LAG-3+ T-BET+	3

PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ T-BET+	3
TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	3
TIM3+ LAG-3+ EOMES+ CD38+ T-BET+	3
TIM3+ LAG-3+ CD39+ T-BET+	3
TIM3+ LAG-3+ EOMES+ CD44+	3
TIM3+ LAG-3+ CD39+ TCF-1+	3
TIM3+ LAG-3+ CD39+ TCF-1+ T-BET+	3
TOX1+ TIM3+ CD39+ CD44+ T-BET+	3
TIM3+ LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	3
TOX1+ LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	3
PD-1+ TOX1+ TIM3+ CD39+ CD44+ TCF-1+	3
TIM3+ LAG-3+ CD39+ CD38+ CD44+	3
TOX1+ LAG-3+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ TOX1+ LAG-3+ CD38+ CD44+ TCF-1+	3
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	3
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD38+ CD44+	3
LAG-3+ CD38+ CD44+ T-BET+	3
TOX1+ TIM3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	3
TOX1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	3
TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+	3
PD-1+ TOX1+ TIM3+ LAG-3+ T-BET+	3
PD-1+ TOX1+ TIM3+ LAG-3+ CD44+ TCF-1+	3
LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	3
TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+	3
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ T-BET+	3
TOX1+ TIM3+ LAG-3+ CD39+ CD44+ T-BET+	3
PD-1+ LAG-3+ CD39+ EOMES+ CD38+ T-BET+	3
TIM3+ LAG-3+ CD44+ TCF-1+	3
PD-1+ TIM3+ EOMES+ CD38+ TCF-1+	3
TIM3+ CD39+ CD44+ T-BET+	3
PD-1+ LAG-3+ CD39+ EOMES+ CD44+	3
TIM3+ LAG-3+ CD44+ TCF-1+ T-BET+	3
PD-1+ EOMES+ CD38+	3
PD-1+ TIM3+ EOMES+ CD44+	3
TIM3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	3
PD-1+ LAG-3+ EOMES+ CD38+	3
PD-1+ LAG-3+ EOMES+ CD38+ TCF-1+	3
PD-1+ TIM3+ EOMES+ T-BET+	3
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	3
PD-1+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+	3
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+	3

PD-1+ LAG-3+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ CD38+ T-BET+	3
TIM3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ TIM3+ LAG-3+ EOMES+ TCF-1+ T-BET+	3
PD-1+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	3
PD-1+ TIM3+ LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	3
TIM3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	3
PD-1+ TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+ T-BET+	3
TIM3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ TIM3+ LAG-3+ CD38+ TCF-1+	3
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	3
PD-1+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ TIM3+ CD39+ EOMES+ CD44+ TCF-1+	3
PD-1+ CD39+ CD38+	3
PD-1+ CD39+ EOMES+ T-BET+	3
PD-1+ CD39+ EOMES+ CD38+ TCF-1+	3
PD-1+ CD39+ CD38+ TCF-1+ T-BET+	3
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+ T-BET+	3
PD-1+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	3
PD-1+ CD39+ EOMES+ CD38+ CD44+ T-BET+	3
PD-1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ CD39+ EOMES+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ TIM3+ CD39+	3
PD-1+ TIM3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	3
PD-1+ TIM3+ EOMES+ CD38+ CD44+ TCF-1+	3
PD-1+ CD39+ CD44+ T-BET+	3
PD-1+ LAG-3+ CD39+ EOMES+ CD38+	2
PD-1+ TOX1+ TIM3+ CD39+ CD38+ CD44+ TCF-1+	2
TOX1+ LAG-3+ EOMES+ CD44+ T-BET+	2
PD-1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TIM3+ EOMES+ CD38+ CD44+	2
PD-1+ TIM3+ CD39+ CD38+ T-BET+	2
TIM3+ LAG-3+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TOX1+ TIM3+ EOMES+ CD38+ CD44+	2
PD-1+ TIM3+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TIM3+ CD39+ CD44+ T-BET+	2
TOX1+ LAG-3+ CD39+ EOMES+ CD38+ T-BET+	2
TIM3+ LAG-3+ CD44+ T-BET+	2
TIM3+ LAG-3+ CD39+ CD44+ TCF-1+	2
TOX1+ TIM3+ LAG-3+ CD39+ CD44+ TCF-1+	2
TOX1+ LAG-3+ CD39+ CD38+ CD44+ T-BET+	2

PD-1+ TIM3+ CD38+ CD44+ TCF-1+ T-BET+	2
PD-1+ TOX1+ TIM3+ LAG-3+ CD44+	2
PD-1+ LAG-3+ CD39+ CD38+ CD44+ T-BET+	2
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD38+ TCF-1+	2
PD-1+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	2
TIM3+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	2
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ T-BET+	2
PD-1+ CD39+ EOMES+ CD38+ T-BET+	2
TIM3+ LAG-3+ CD39+ EOMES+ CD38+ T-BET+	2
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+	2
PD-1+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	2
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+	2
TIM3+ CD39+ EOMES+ CD44+	2
TIM3+ CD39+ EOMES+ TCF-1+	2
PD-1+ LAG-3+ CD38+ CD44+ TCF-1+	2
TIM3+ LAG-3+ CD39+ EOMES+ CD44+	2
TIM3+ CD39+ EOMES+ T-BET+	2
PD-1+ EOMES+ CD38+ TCF-1+ T-BET+	2
TOX1+ CD39+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+	2
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD44+	2
TIM3+ LAG-3+ CD39+ EOMES+ T-BET+	2
TIM3+ CD39+ EOMES+ CD38+ T-BET+	2
PD-1+ LAG-3+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+ T-BET+	2
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+	2
PD-1+ TOX1+ TIM3+ LAG-3+ CD44+ T-BET+	2
PD-1+ TIM3+ CD39+ EOMES+ TCF-1+ T-BET+	2
TIM3+ CD39+ EOMES+ CD38+ CD44+	2
PD-1+ TOX1+ TIM3+ LAG-3+ TCF-1+	2
PD-1+ LAG-3+ CD39+ CD38+	2
PD-1+ LAG-3+ CD39+ CD38+ T-BET+	2
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TOX1+ LAG-3+ CD39+ EOMES+ CD44+ T-BET+	2
PD-1+ CD39+ CD38+ CD44+	2
TIM3+ LAG-3+ EOMES+ CD38+	2
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+	2
PD-1+ CD39+ CD38+ CD44+ T-BET+	2
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+	2
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ TCF-1+	2
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+	2

PD-1+ TOX1+ CD39+ CD38+ TCF-1+	2
PD-1+ TOX1+ CD39+ CD38+ CD44+ T-BET+	2
PD-1+ TIM3+ LAG-3+ EOMES+ CD44+ T-BET+	2
PD-1+ TIM3+ LAG-3+ EOMES+ CD44+ TCF-1+	2
TOX1+ TIM3+ CD39+ EOMES+ T-BET+	2
PD-1+ TOX1+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	2
PD-1+ TIM3+ LAG-3+ CD39+ T-BET+	2
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+	2
TOX1+ TIM3+ LAG-3+ EOMES+ T-BET+	2
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ T-BET+	2
TIM3+ CD39+ CD38+ CD44+ T-BET+	2
PD-1+ TIM3+ LAG-3+ CD39+ CD44+ T-BET+	2
PD-1+ TOX1+ LAG-3+ CD38+ CD44+ T-BET+	2
PD-1+ TOX1+ LAG-3+ CD38+ TCF-1+	2
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ T-BET+	2
PD-1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+	2
TOX1+ TIM3+ LAG-3+ CD44+ TCF-1+	2
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ T-BET+	2
PD-1+ TIM3+ LAG-3+ EOMES+	2
PD-1+ TIM3+ LAG-3+ EOMES+ T-BET+	2
LAG-3+ CD39+ CD38+ CD44+ T-BET+	2
TOX1+ TIM3+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TIM3+ CD39+ EOMES+ CD44+	2
TOX1+ TIM3+ EOMES+ CD38+	2
TOX1+ TIM3+ EOMES+ CD38+ T-BET+	2
TOX1+ TIM3+ LAG-3+ CD39+ CD38+ TCF-1+	2
PD-1+ TIM3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	2
TIM3+ LAG-3+ CD38+ CD44+	2
PD-1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	2
PD-1+ TOX1+ LAG-3+ CD39+ CD38+ CD44+	2
TOX1+ TIM3+ LAG-3+ CD44+	2
PD-1+ TIM3+ LAG-3+ TCF-1+ T-BET+	2
PD-1+ TOX1+ LAG-3+ CD39+ CD44+ T-BET+	2
PD-1+ TIM3+ LAG-3+ CD38+ T-BET+	2
TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ CD44+ T-BET+	2
TOX1+ TIM3+ LAG-3+ CD39+ CD44+	1
TOX1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+	1
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+	1
PD-1+ CD39+ EOMES+ CD38+ CD44+	1
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+	1
PD-1+ CD39+ EOMES+ CD38+	1

TOX1+ TIM3+ LAG-3+ EOMES+ CD44+ T-BET+	1
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+	1
PD-1+ LAG-3+ EOMES+ CD38+ T-BET+	1
TIM3+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	1
TOX1+ TIM3+ LAG-3+ CD38+	1
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ TCF-1+ T-BET+	1
PD-1+ CD39+ CD38+ CD44+ TCF-1+	1
TOX1+ TIM3+ LAG-3+ CD44+ T-BET+	1
TOX1+ TIM3+ LAG-3+ CD38+ T-BET+	1
PD-1+ TOX1+ TIM3+ LAG-3+ EOMES+ CD38+ T-BET+	1
PD-1+ TOX1+ TIM3+ LAG-3+ CD38+ TCF-1+	1
TOX1+ TIM3+ LAG-3+ CD39+ T-BET+	1
PD-1+ TOX1+ CD39+ CD38+ TCF-1+ T-BET+	1
PD-1+ EOMES+ CD38+ CD44+ T-BET+	1
PD-1+ TOX1+ CD39+ CD38+ CD44+ TCF-1+	1
PD-1+ LAG-3+ CD38+ CD44+	1
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+	1
PD-1+ LAG-3+ CD38+ CD44+ T-BET+	1
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ T-BET+	1
PD-1+ CD39+ EOMES+ CD44+ T-BET+	1
PD-1+ LAG-3+ EOMES+ CD44+ T-BET+	1
PD-1+ TOX1+ TIM3+ LAG-3+ CD39+ CD38+ TCF-1+	1
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD44+ T-BET+	1
PD-1+ EOMES+ CD38+ CD44+	1
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ T-BET+	1
TOX1+ TIM3+ LAG-3+ CD39+ EOMES+ CD38+	1
PD-1+ EOMES+ CD38+ T-BET+	1
TOX1+ CD39+ CD38+ CD44+ T-BET+	1
TOX1+ TIM3+ LAG-3+ CD39+ TCF-1+	1
PD-1+ TIM3+ LAG-3+ CD39+ CD38+ TCF-1+	1
PD-1+ TIM3+ LAG-3+ CD39+ EOMES+	1
TIM3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	1
TOX1+ TIM3+ CD39+ CD38+ TCF-1+	1
PD-1+ TOX1+ TIM3+ EOMES+ CD38+ TCF-1+ T-BET+	1
PD-1+ TOX1+ TIM3+ EOMES+ CD38+ TCF-1+	1
PD-1+ TIM3+ CD38+	1
PD-1+ TIM3+ CD38+ T-BET+	1
PD-1+ TIM3+ CD38+ TCF-1+	1
PD-1+ TIM3+ CD38+ TCF-1+ T-BET+	1
PD-1+ TIM3+ CD38+ CD44+	1
PD-1+ TIM3+ LAG-3+ CD38+ CD44+ TCF-1+	1

PD-1+ TIM3+ LAG-3+ CD38+ CD44+ T-BET+	1
PD-1+ TIM3+ LAG-3+ CD44+ T-BET+	1
PD-1+ TIM3+ LAG-3+ CD44+	1
PD-1+ TIM3+ LAG-3+ TCF-1+	1
PD-1+ TIM3+ CD39+ T-BET+	1
TIM3+ LAG-3+ CD39+ CD44+ T-BET+	1
PD-1+ TIM3+ CD39+ TCF-1+ T-BET+	1
TIM3+ LAG-3+ CD39+ CD44+	1
TOX1+ LAG-3+ CD39+ EOMES+ CD38+ CD44+ T-BET+	1
PD-1+ TIM3+ CD39+ CD44+	1
LAG-3+ CD39+ CD44+ T-BET+	1
PD-1+ TIM3+ CD39+ CD38+	1
PD-1+ TIM3+ CD39+ CD38+ CD44+ T-BET+	1
TIM3+ LAG-3+ EOMES+ CD38+ CD44+	1
PD-1+ TOX1+ TIM3+ EOMES+ CD38+ CD44+ T-BET+	1
PD-1+ TIM3+ LAG-3+ EOMES+ TCF-1+	1
PD-1+ CD38+ CD44+ T-BET+	1
TOX1+ TIM3+ CD39+ EOMES+ CD38+	1
TOX1+ TIM3+ CD39+ EOMES+ CD38+ CD44+ TCF-1+	1
PD-1+ TIM3+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+ T-BET+	1
PD-1+ TIM3+ LAG-3+ CD39+ CD38+ TCF-1+ T-BET+	1
PD-1+ LAG-3+ CD39+ CD44+	1
PD-1+ LAG-3+ CD39+ CD44+ T-BET+	1
PD-1+ TIM3+ CD39+ EOMES+ TCF-1+	1
PD-1+ LAG-3+ CD39+ CD38+ TCF-1+	1
PD-1+ TIM3+ LAG-3+ CD39+ TCF-1+ T-BET+	1
PD-1+ TOX1+ TIM3+ CD39+ EOMES+ CD38+	1
PD-1+ LAG-3+ CD39+ CD38+ CD44+	1
PD-1+ LAG-3+ CD39+ CD38+ CD44+ TCF-1+	1
PD-1+ TIM3+ LAG-3+ EOMES+ CD44+	1
LAG-3+ CD39+ EOMES+ CD44+ T-BET+	1
TIM3+ CD39+ CD38+ CD44+ TCF-1+	1
TOX1+ TIM3+ CD39+ EOMES+ CD44+ T-BET+	1
TIM3+ LAG-3+ CD39+ CD38+ TCF-1+	1
PD-1+ LAG-3+ CD39+ EOMES+ CD44+ T-BET+	1
PD-1+ TOX1+ TIM3+ CD39+ CD38+ T-BET+	1
PD-1+ TOX1+ TIM3+ CD39+ CD38+	1
PD-1+ TOX1+ TIM3+ CD39+ CD44+ T-BET+	1
PD-1+ TOX1+ TIM3+ CD39+ CD44+	1
PD-1+ LAG-3+ CD39+ EOMES+ CD38+ TCF-1+ T-BET+	1
TOX1+ TIM3+ LAG-3+ CD39+ CD44+ TCF-1+ T-BET+	1

Supplementary Table 4.6: Raw counts of cell-cell spatial interactions present in the dataset.

Cell-Cell Spatial Interaction	Count
Ki-67- Neoplastic & Ki-67- Neoplastic	2876770
Myeloid cells & Myeloid cells	1480053
Myeloid cells & Ki-67- Neoplastic	465886
B cells & B cells	450530
Ki-67- Neoplastic & Ki-67+ Neoplastic	437638
CD44- CD4 Th & CD44- CD4 Th	226250
Mesenchymal & Mesenchymal	220227
CD44- CD4 Th & B cells	219910
CD44- CD8 T Other & CD44- CD4 Th	206383
Mesenchymal & Ki-67- Neoplastic	202107
Ki-67+ Neoplastic & Ki-67+ Neoplastic	165702
CD44- CD8 T Other & CD44- CD8 T Other	158210
CD44- CD8 T Other & B cells	111164
CD44- CD4 Th & Myeloid cells	107677
Myeloid cells & Mesenchymal	90234
CD44- CD8 T Other & Myeloid cells	89976
B cells & Myeloid cells	62552
Myeloid cells & Ki-67+ Neoplastic	58793
CD44- CD8 T Other & Ki-67- Neoplastic	49618
B cells & Ki-67- Neoplastic	49336
CD44- CD4 Th & Ki-67- Neoplastic	48406
CD44- CD4 Th1 Other & CD44- CD4 Th	42570
CD44- CD8 T Other & CD44- CD4 Th1 Other	34844
CD44- CD4 Th & mTREG	32692
CD44- CD4 Th & CD44+ CD4 Th	31233
CD44- CD4 Th & Mesenchymal	29024
CD44- CD8 T Other & mTREG	27872
CD44+ CD4 Th & B cells	27606
CD44- CD8 T Other & Mesenchymal	26494
CD44+ CD4 Th & CD44+ CD4 Th	24718
CD44+ CD4 Th & Myeloid cells	24454
CD44- CD4 Th1 Other & B cells	24437
CD44+ CD8 T Other & Myeloid cells	23784
mTREG & Myeloid cells	23570
CD44- CD8 T Other & CD44+ CD8 T Other	23249
Naive TREG & B cells	23177
CD44- CD4 Th & Naive TREG	22405

CD44+ CD8 T Other & CD44+ CD8 T Other	22351
Mesenchymal & Ki-67+ Neoplastic	20910
CD44+ CD8 T Other & CD44+ CD4 Th	20823
CD44+ CD8 T Other & CD44- CD4 Th	20030
CD44+ CD8 T Other & B cells	19305
CD44- CD8 T Other & Naive TREG	16461
B cells & Mesenchymal	15010
mTREG & B cells	14656
mTREG & Ki-67- Neoplastic	14362
CD44- CD8 T Other & CD44+ CD4 Th	12734
CD44- CD4 Th1 Other & CD44- CD4 Th1 Other	11730
mTREG & Mesenchymal	10419
CD44- CD4 Th1 Other & Myeloid cells	9616
mTREG & mTREG	9087
CD8 T NAIVE & CD44- CD8 T Other	8303
Naive TREG & Naive TREG	7839
CD8 T NAIVE & CD44- CD4 Th	7103
B cells & Ki-67+ Neoplastic	6512
CD44- CD4 Th & Ki-67+ Neoplastic	6452
CD44- CD8 T Other & Ki-67+ Neoplastic	6392
CD44- CD4 Th & T-BET+ TREG	6102
Naive TREG & Myeloid cells	6035
CD44+ CD4 Th & mTREG	5918
CD44- CD4 Th1 Other & Mesenchymal	5711
CD44- CD4 Th1 Other & Ki-67- Neoplastic	5640
CD8 TEMRA & CD44+ CD4 Th	5367
CD44- CD8 T Other & T-BET+ TREG	5300
CD44+ CD8 T Other & mTREG	5186
Naive TREG & Ki-67- Neoplastic	5185
CD8 T NAIVE & B cells	5149
T-BET+ TREG & B cells	5086
Naive TREG & mTREG	4921
CD44+ CD8 T Other & Ki-67- Neoplastic	4863
CD8 TEMRA & CD44+ CD8 T Other	4772
CD44+ CD4 Th & Mesenchymal	4593
CD44- CD4 Th1 Other & mTREG	4305
CD44+ CD4 Th & Naive TREG	4166
CD8 TEMRA & B cells	3878
CD44+ CD8 T Other & Mesenchymal	3536
CD44+ CD4 Th & Ki-67- Neoplastic	3476
CD8 TEMRA & CD44- CD8 T Other	3454

CD8 TEMRA & CD44- CD4 Th	3399
CD44- CD4 Th1 Other & T-BET+ TREG	3382
CD8 TEMRA & CD8 TEMRA	3337
CD8 TEMRA & Myeloid cells	3226
CD44- CD4 Th1 Other & CD44+ CD4 Th	3112
CD44- CD4 Th1 Other & Naive TREG	3052
CD44+ CD8 T Other & CD44- CD4 Th1 Other	2766
CD44+ CD8 T Other & Naive TREG	2300
CD4 Th1EMRA & CD44- CD4 Th	1997
CD4 Th1EM & CD44- CD4 Th	1949
mTREG & Ki-67+ Neoplastic	1924
T-BET+ TREG & Myeloid cells	1857
Naive TREG & Mesenchymal	1852
CD4 Th1EMRA & B cells	1839
Naive TREG & T-BET+ TREG	1681
CD4 Th1EM & Myeloid cells	1667
CD4 Th1EMRA & CD44+ CD4 Th	1607
CD4 Th1EM & CD44+ CD4 Th	1542
CD44- CD8 T Other & CD4 Th1EM	1486
mTREG & T-BET+ TREG	1458
CD8 TEM & CD44- CD8 T Other	1322
CD44+ CD8 T Other & CD4 Th1EM	1311
CD8 TEM & CD44+ CD8 T Other	1285
CD8 TEM & CD44+ CD4 Th	1283
CD8 T NAIVE & CD44- CD4 Th1 Other	1238
CD44- CD8 T Other & CD4 Th1EMRA	1162
CD8 TEM & CD44- CD4 Th	1140
T-BET+ TREG & T-BET+ TREG	1119
CD4 Th1EM & CD44- CD4 Th1 Other	1112
CD44+ CD4 Th & T-BET+ TREG	1092
CD8 TEX & CD44+ CD8 T Other	1071
CD44+ CD8 T Other & T-BET+ TREG	1020
T-BET+ TREG & Ki-67- Neoplastic	1013
T-BET+ TREG & Mesenchymal	969
CD44+ CD8 T Other & CD4 Th1EMRA	969
CD4 Th1EMRA & CD44- CD4 Th1 Other	934
CD4 Th1EMRA & Myeloid cells	932
CD8 TEMRA & Ki-67- Neoplastic	878
CD8 TEM & Myeloid cells	847
CD8 T NAIVE & CD44+ CD8 T Other	837
CD8 TEM & B cells	808

CD4 Th1EM & Mesenchymal	790
CD8 TEMRA & Naive TREG	769
CD8 TEX & CD44- CD8 T Other	754
CD8 T NAIVE & Myeloid cells	740
CD4 Th1EMRA & CD4 Th1EMRA	739
CD44- CD4 Th1 Other & Ki-67+ Neoplastic	727
CD44+ CD8 T Other & Ki-67+ Neoplastic	710
CD4 Th1EM & mTREG	685
CD8 T NAIVE & CD8 T NAIVE	678
CD8 TEM & CD8 TEM	628
CD8 T NAIVE & Naive TREG	618
CD4 Th1EM & B cells	609
CD44+ CD4 Th1 Other & Ki-67- Neoplastic	580
CD8 TEMRA & CD4 Th1EMRA	576
CD8 T NAIVE & mTREG	574
CD44+ CD4 Th1 Other & Myeloid cells	550
CD8 TEX & CD44+ CD4 Th	522
CD8 TEX & CD44- CD4 Th	514
CD4 Th1EM & Ki-67- Neoplastic	510
CD4 Th1EMRA & Naive TREG	507
CD44+ CD4 Th & Ki-67+ Neoplastic	495
CD8 TEM & CD8 TEMRA	478
CD44- CD4 Th1 Other & CD44+ CD4 Th1 Other	477
CD4 Th1EMRA & T-BET+ TREG	476
CD8 TEMRA & CD44- CD4 Th1 Other	467
CD8 TEMRA & mTREG	459
CD8 T NAIVE & CD44+ CD4 Th	457
Naive TREG & Ki-67+ Neoplastic	451
CD8 TEX & B cells	442
CD4 Th1EM & CD4 Th1EM	431
CD8 TEMRA & Mesenchymal	383
CD8 TEX & CD8 TEX	368
CD44+ CD4 Th1 Other & B cells	353
CD8 TEMRA & T-BET+ TREG	350
CD8 TEM & CD4 Th1EM	340
CD44+ CD4 Th1 Other & CD44+ CD4 Th1 Other	338
CD8 TEX & mTREG	316
CD8 TEX & Myeloid cells	316
CD8 TEM & mTREG	309
CD44+ CD4 Th1 Other & CD44- CD4 Th	303
CD4 Th1EMRA & CD44+ CD4 Th1 Other	290

CD8 TTEX & CD44+ CD8 T Other	285
CD4 Th1EM & T-BET+ TREG	281
CD44+ CD4 Th1 Other & Mesenchymal	278
CD8 TEMRA & CD8 TEX	250
CD4 Th1EMRA & Ki-67- Neoplastic	247
CD4 Th1EMRA & mTREG	239
CD8 TTEX & Myeloid cells	221
CD44- CD8 T Other & CD44+ CD4 Th1 Other	215
CD8 T NAIVE & T-BET+ TREG	213
CD44+ CD4 Th1 Other & CD44+ CD4 Th	206
CD8 TEM & CD44- CD4 Th1 Other	199
CD8 TEX & Ki-67- Neoplastic	199
CD8 TTEX & CD44- CD8 T Other	194
CD8 T NAIVE & Mesenchymal	193
CD44+ CD8 T Other & CD44+ CD4 Th1 Other	186
CD8 TTEX & CD44+ CD4 Th	176
CD8 TEMRA & CD4 Th1EM	174
CD4 Th1EMRA & Mesenchymal	172
CD8 TTEX & CD44- CD4 Th	168
CD8 TEM & T-BET+ TREG	159
T-BET+ TREG & Ki-67+ Neoplastic	150
CD8 T NAIVE & CD8 TEMRA	150
CD4 Th1EM & CD4 Th1EMRA	146
CD4 Th1EM & Naive TREG	137
CD8 TEM & CD4 Th1EMRA	131
CD8 T NAIVE & Ki-67- Neoplastic	116
CD8 TTEX & B cells	111
CD8 TEFF & CD44+ CD8 T Other	108
CD8 TEX & Naive TREG	101
CD8 TEX & CD44- CD4 Th1 Other	96
CD8 TEM & CD8 TEX	91
CD8 TEFF & B cells	88
CD8 TEM & Mesenchymal	85
CD8 TEX & T-BET+ TREG	83
CD8 TEX & CD8 TTEX	83
CD8 TEM & Naive TREG	82
CD8 TEMRA & Ki-67+ Neoplastic	81
CD8 TTEX & mTREG	80
CD8 TTEX & Ki-67- Neoplastic	77
CD8 TEX & Mesenchymal	75
CD8 TEX & CD4 Th1EM	72

CD8 TEM & Ki-67- Neoplastic	68
CD44+ CD4 Th1 Other & Ki-67+ Neoplastic	65
CD8 TEFF & CD44- CD8 T Other	64
CD44+ CD4 Th1 Other & T-BET+ TREG	61
CD8 TEFF & CD8 TEMRA	56
CD4 Th1EFF & B cells	56
CD44+ CD4 Th1 Other & mTREG	54
CD8 TEFF & Myeloid cells	50
CD8 TEFF & CD44- CD4 Th	49
CD4 Th1EFF & Myeloid cells	47
CD4 Th1EM & Ki-67+ Neoplastic	45
CD8 TEFF & CD8 TEFF	42
CD4 Th1EM & CD44+ CD4 Th1 Other	41
CD8 T NAIVE & CD4 Th1EMRA	35
CD8 TEMRA & CD8 TTEX	34
CD44+ CD4 Th1 Other & Naive TREG	34
CD8 TEMRA & CD44+ CD4 Th1 Other	31
CD8 TTEX & Mesenchymal	30
CD8 TEX & CD4 Th1EMRA	29
CD8 T NAIVE & CD4 Th1EM	28
CD4 Th1EFF & CD44+ CD4 Th	28
CD8 TEFF & CD44- CD4 Th1 Other	26
CD8 TEFF & Mesenchymal	25
CD8 TTEX & Naive TREG	25
CD8 TEFF & Ki-67- Neoplastic	24
CD4 Th1EMRA & Ki-67+ Neoplastic	23
CD8 T NAIVE & CD8 TEM	22
CD8 TEX & Ki-67+ Neoplastic	22
CD4 Th1EFF & CD44+ CD4 Th1 Other	21
CD8 TTEX & CD44- CD4 Th1 Other	21
CD8 T NAIVE & CD8 TEX	20
CD8 T NAIVE & CD44+ CD4 Th1 Other	20
CD4 Th1EFF & Ki-67- Neoplastic	19
CD4 Th1EFF & CD44- CD4 Th1 Other	18
CD8 TTEX & CD4 Th1EM	18
CD8 TTEX & CD8 TTEX	17
CD8 TTEX & Ki-67+ Neoplastic	15
CD4 Th1EFF & CD4 Th1EMRA	15
CD8 TEFF & CD44+ CD4 Th	13
CD8 TEM & CD8 TTEX	11
CD8 TEMRA & CD4 Th1EFF	11

CD8 TEX & CD44+ CD4 Th1 Other	10
CD44- CD8 T Other & CD4 Th1EFF	8
CD8 T NAIVE & Ki-67+ Neoplastic	8
CD4 Th1EFF & Ki-67+ Neoplastic	8
CD44+ CD8 T Other & CD4 Th1EFF	7
CD4 Th1EFF & CD44- CD4 Th	7
CD8 TTEX & CD4 Th1EMRA	7
CD8 TTEX & T-BET+ TREG	7
CD8 TEM & Ki-67+ Neoplastic	5
CD4 Th1EFF & Naive TREG	5
CD8 TEFF & CD8 TTEX	5
CD8 TEFF & T-BET+ TREG	4
CD8 TEFF & CD44+ CD4 Th1 Other	4
CD8 TEM & CD44+ CD4 Th1 Other	4
CD4 Th1EFF & CD4 Th1EM	4
CD8 TEX & CD4 Th1EFF	4
CD4 Th1EFF & Mesenchymal	4
CD8 TEFF & Naive TREG	4
CD8 T NAIVE & CD8 TEFF	3
CD4 Th1EFF & CD4 Th1EFF	3
CD4 Th1EFF & T-BET+ TREG	3
CD4 Th1EFF & mTREG	2
CD8 T NAIVE & CD8 TTEX	2
CD8 TEM & CD4 Th1EFF	2
CD8 TEFF & mTREG	1
CD8 TEFF & CD8 TEM	1

Chapter 5: Discussion

5.1 Summary

A better understanding of the complex cellular dynamics at play in the TME has the potential to improve treatment strategies and clinical outcome for cancer patients [8]. This dissertation aimed to: (1) quantitatively characterize the single-cell spatial landscape of the TME and (2) identify novel relationships between immunological features of the TME and clinical parameters for three tumor types. Chapter 2 involved a comprehensive evaluation of HNSCC TMEs—cellular composition, heterogeneity, longitudinal changes, and spatial organization were all quantified, and multiple TME spatial features associated with clinical outcome. Chapter 3 involved a targeted analysis of NK cell spatial organization in breast cancer TMEs, which resulted in an improved understanding of the differences in NK cell localization with neoplastic cells and T cells in $HER2^-$ versus $HER2^+$ breast cancer subtypes. Finally, Chapter 4 involved a deep interrogation of the complexities of T cell phenotype, function, and spatial localization in PDAC TMEs in the context of neoadjuvant immunotherapy; ML approaches were used to holistically evaluate these aspects of T cells in PDAC and detect combinations of clinically relevant T cell features in the TME.

5.2 Common themes

While the three studies presented in this dissertation analyzed different tumor types, common themes emerge across the chapters. From a computational perspective, similar methods were used to analyze the complex datasets and successfully interpret biological meaning. While cell densities sometimes correlated with various clinical parameters across the three chapters, spatial metrics were able to consistently stratify patients according to the desired clinical

parameter regardless of tumor type. These results point to the conclusion that TME spatial organization is biologically and clinically meaningful, and methods that quantify the cellular spatial landscape of the TME are valuable.

The spatial metrics utilized across chapters were selected based upon the biological hypothesis to be tested, as well as based upon the utility of the spatial metrics in prior research—including research reported in the literature, as well as research conducted as part of earlier chapters in this dissertation. Given its novelty in the single-cell spatial cancer research field at the time [139], the mixing score was adapted for use with mIHC datasets in Chapter 2 to determine whether the metric was clinically relevant beyond TNBC tumors. As it was capable of stratifying long versus short progression-free survivors in HNSCC as shown in Chapter 2, the mixing score may be a valuable method for quantifying broad neoplastic-immune cell organization in other solid tumors. While the mixing score can be adapted to study cellular organization between cells other than neoplastic and immune populations, it is limited in that it can only measure the regional mixing between two populations. On the other hand, cellular neighborhoods and cell-cell interactions can account for all cellular phenotypes present in the data, making these spatial metrics more appealing when three or more cell phenotypes must be evaluated and increased spatial resolution is needed. Cellular neighborhoods were first utilized in Chapter 2 to assess the cells surrounding αSMA cells in HNSCCs. This method enabled further understanding of the many types of cells αSMA cells were likely communicating with, as well as which neighborhoods were linked to improved clinical outcomes.

In Chapter 3, the study sought to assess how the function of NK cells was impacted by spatially proximal cells, including neoplastic cells, as well as how the function and spatial organization of NK cells differed depending on the HER2 status of the breast tumor. Because of

these goals, a highly targeted cell-cell interaction analysis was first performed to assess how NK cells and neoplastic cells were potentially communicating through close spatial proximity and impacting cell function in the context of HER2 disease. Varying functional biomarker expression was found on both NK cells and neoplastic cells depending on the spatial proximity of the cells, which also often differed depending on the HER2 status of the tumor. To further identify the differences in the types of cells NK cells were likely communicating with in $HER2^-$ versus $HER2^+$ disease, a cellular neighborhood analysis was then performed. Cellular neighborhoods were calculated for each NK cell present and differences were also identified based upon the HER2 status of the tumor.

Given the focus on T cells in the PDAC TME in Chapter 4, 18 T cell states were phenotyped by the mIHC gating strategy. To assess the spatial organization of each of these states, with every other cell state phenotyped—T cell or not—cell-cell interactions and cellular neighborhoods were utilized to quantify the spatial organization of the PDAC TME. Thus, in contrast to the use of cell-cell interaction and cellular neighborhood metrics in Chapters 2 and 3, these metrics were not limited to study specific cell phenotypes, but rather all phenotypes identified by the mIHC phenotypic gating strategy were evaluated. This meant that the number of spatial interactions were calculated for 268 phenotypically-unique pairs of proximal cells, and cellular neighborhoods were defined for every cell present, rather than just for cells of one specific phenotype—as was the case for the αSMA cells in Chapter 2 and the NK cells in Chapter 3. This unbiased approach to quantifying the spatial landscape enabled the identification of specific and clinically meaningful spatial organizations that would likely have gone undetected, if a more targeted approach, similar to that of Chapters 2 and 3, was used.

As the methods in Chapter 4 built significantly off the methods used in the prior chapters, the number of data features being evaluated for clinical relevance was also much larger. Because of this, more sophisticated statistical techniques were needed to evaluate all data features in combination with each other in this chapter. Thus, ML models were used to predict and interpret the clinical relevance of combinations of TME features, demonstrating the value of using ML approaches to interpret complex biological datasets.

In total, a common computational theme emerges from these studies: it is apparent that quantitative approaches are needed to analyze and interpret the complexities of the TME. In an ideal world, a single and simple biomarker could be used to determine treatment response and/or clinical outcome for cancer patients. However, the field is far from distilling TME biology into a single value or number that can inform medical providers of the disease trajectory of the patient, and it is quite possible that no single biomarker will ever fulfill this need. Tumors are composed of millions of cells, and each patient's tumor is unique. Additionally, single-cell spatial proteomics imaging assays are relatively new, and no algorithms have been standardized to quantify the single-cell spatial landscape of the TME from the datasets emerging from these technologies. Continuing to assess TME cellular composition, function, and, in particular, spatial organization through multiple computational algorithms will: (1) enable researchers to identify biological trends across tumors and (2) identify which computational approaches consistently detect clinically relevant TME features that may become standard clinical biomarkers in the future. This dissertation provides a quantitative framework for characterizing the single-cell spatial landscape of TMEs and interpreting their downstream clinical significance, as the computational approaches used here are appliable across tumor types and single-cell spatial proteomics datasets.

From a biological perspective, similar anti-tumor immunological features were identified across the three TMEs and tended to associate with improved outcomes. Specifically, spatial groupings of immune cells involved in antigen presentation and/or cytotoxicity were present across HNSCC, breast, and PDAC TMEs. Further, these immunoreactive cellular spatial hubs were associated with improved clinical outcome for HNSCC and PDAC patients (limited clinical outcome data prevented this type of analysis in Chapter 3). Collectively, these results support the notions that: (1) spatial proximity of immune cells contributes to increased antigen presentation and T cell activation in the TME and (2) spatial groupings of immunoreactive cells may play an important role in combating malignant cells in the TME and ultimately prolonging survival.

5.3 Clinical significance

The biological findings of this dissertation can be leveraged to generate hypotheses regarding improved treatment approaches for cancer patients. Based upon the analyses performed in Chapter 2, therapies that elicit an immune response may be valuable for patients presenting with late-stage HPV(-) HNSCC. This hypothesis is supported by the finding that later staged primary tumors tended to possess a mixed immune-neoplastic spatial organization with increased PD-1 expression on lymphocytes, which was associated with reduced PFS. Upon recurrence, these patients also experienced the greatest reduction in their $CD8^+$ T cell population. Therefore, it is plausible that treating late-staged spatially mixed primary tumors with an ICB that targets the PD-1/PD-L1 checkpoint may convert the T cells present in these TMEs into functional effector T cells capable of cytotoxicity against the neoplastic cells. In turn, a reduction in the $CD8^+$ T cell lineage may be mitigated for these patients, thus prolonging PFS. ICBs targeting the PD-1/PD-L1 checkpoint have been used with some success to treat recurrent and metastatic HNSCC [84-86]; the results from Chapter 2 indicate patients with certain primary

HPV(-) HNSCCs characterized by the single-cell spatial landscape described above may also benefit from these therapies.

The spatial results described in Chapter 3 also reveal potential opportunities to improve the treatment of breast cancer patients. The results indicate that combination therapy approaches that target both HER2 and HLA-1 expression on neoplastic cells may be beneficial for treating $HER2^+$ breast cancer patients. Anti-HER2 antibodies (e.g. Trastuzumab) are currently used to treat $HER2^+$ breast cancer patients, and NK cells have been shown to play a critical role in contributing to the efficacy of the treatment [91-93]. However, the results presented in Chapter 3 indicate that further efficacy may be reached through the disruption of HLA-1-mediated inhibition of NK cells. Despite NK cells being found in tumor nests, the spatial analyses performed in Chapter 3 indicate that effector functions of NK cells may be inhibited by nearby neoplastic cells expressing the inhibitory ligand HLA-1, which binds to killer-cell immunoglobulin-like receptors (KIRs) on NK cells. This binding suppresses NK cell activation, thus diminishing NK cell cytotoxic function [270]. Therefore, suppression of HLA-1-mediated inhibition of NK cells, in addition to treatment with anti-HER2 antibodies, may further enable NK-driven cytotoxicity in the TME and improve responses for $HER2^+$ breast cancer patients. Indeed, a recent study investigated the relationship between KIRs and their corresponding HLA ligands in breast tumors treated with Trastuzumab and found the presence of various KIR/HLA combinations to be associated with varying degrees of drug efficacy [271], highlighting the importance of also considering this inhibitory mechanism when treating breast cancer patients with Trastuzumab.

Finally, the results from Chapter 4 support several hypotheses for improving the personalized treatment of PDAC patients. First, the lack of PD-1^+ T cells found in treatment-

naive TMEs supports the increasing evidence that ICBs targeting the PD-1/PD-L1 checkpoint do not work as first-line treatment for PDAC patients [269]. Instead, the findings presented in Chapter 4 indicate therapies that modulate TOX1 expression may reduce T cell dysfunction and/or exhaustion in the PDAC TME and thus contribute to improved outcomes for these patients [272-274]. Additionally, the results from Chapter 4 support the development of a smaller antibody panel consisting of CD3, CD4, T-BET, CD44, TOX1, and PD-1, which could help determine how a PDAC patient is responding to their given immunotherapy. Following treatment with an immunotherapy, the tumor could be biopsied and assayed with the panel to determine whether the T cells present were activated or exhausted and additional combination therapies that either leverage the presence of effector $CD4^+$ Th1 cells or reinvigorate exhausted T cells could be administered. Unfortunately, as the collective results from Chapter 4 and the literature indicate immune aggregates are the key sites of T cell response [112, 262, 265], this approach is limited by the fact that biopsies sample the TME at random and are unable to target immune aggregates specifically. However, the use of this smaller antibody panel could be currently employed to assay any immune aggregates present in a full surgical resection to aid in determining therapeutic response and additional therapeutic interventions for PDAC patients.

5.4 Limitations

This dissertation provides a reliable framework for quantitatively analyzing single-cell spatial proteomics datasets. However, the resulting biological conclusions require validation through additional studies. This is due, in part, to several limitations of this work. First, the research presented herein was limited by the small sample size used in each study. Although millions of cells were analyzed in this dissertation, the number of patients from which the tissue samples were collected remained small and lacked diversity. Further, lack of knowledge of and

control over external factors, such as diet, sleep, exercise, and environment, made it challenging to identify predictive associations between TME biology and clinical outcome. Additional studies on larger cohorts that control for external variables and expand patient diversity are needed to validate the biological conclusions of this dissertation.

This research is also limited by the assumption that a two-dimensional (2D) portrayal of the TME is representative of its actual three-dimensional (3D) nature. The mIHC imaging platform assays a thin section of the 3D tissue mass, and the resulting data is a 2D image of the tissue. As TMEs are highly heterogenous ecosystems, it is possible that a 2D slice of the TME is unable to accurately portray the full heterogeneity of the ecosystem. Therefore, potentially clinically relevant TME features are lost when analyzing data generated by 2D imaging assays. Additional studies that interrogate the 3D landscape of tumors will help inform whether 2D analyses are truly representative of the entire TME [275].

Additionally, the research presented herein is limited by the antibody panels used to generate the mIHC datasets. As mIHC enables the staining of ~30 antibodies per tissue slide, a finite number of cellular phenotypes and functions were enumerated for each study. This resulted in certain studies having specific immune lineages deeply phenotyped while other lineages were lumped together into broad populations. While this was done to address biological questions specific to each chapter, an ideal study would be capable of phenotyping all immune cells deeply, while also characterizing the functional status of each lineage. Additionally, research has demonstrated that neoplastic cells also possess incredible heterogeneity [68]; thus, incorporating antibodies that interrogate this neoplastic heterogeneity in addition to leukocyte heterogeneity would further inform new candidate biomarkers of various clinical parameters.

Finally, the computational algorithms deployed in this dissertation oftentimes involved setting fixed parameters to perform the analyses. However, it was not always clear what the parameters should be set to. For example, each spatial analysis required setting a specific distance threshold. As the spatial biology field is relatively new, there is no standard distance threshold accepted by the field, in part because there are no standardized spatial metrics, and the existing metrics often assess tissues at varying spatial resolutions. For the spatial analyses performed in this dissertation, distances were chosen based upon (1) distances reported in prior literature assessing how far small molecules can travel through tissues and (2) the stability of the results when slightly adjusting the distance parameter. These rationales were also the foundation for determining the optimal parameter values for the other quantitative analyses performed in this dissertation: parameter values were selected based on prior knowledge of the algorithm, as well as for maximum stability of the results when toggling the value of the parameter. Despite these rationales, it is possible that better values could have been selected for each parameter, and future studies that report on the impact on the results of changing certain parameters, such as distance thresholds for spatial metrics, will be valuable for discerning the optimal value for various parameters moving forward.

As public data repositories are used more widely to share patient datasets and multiplex tissue imaging technologies become cheaper to use, the aforementioned limitations may be mitigated through: (1) improved access to larger patient datasets; (2) increased analysis of serial tissue sections and/or potential 3D imaging datasets; (3) deployment of bigger antibody panels capable of more deeply characterizing the phenotypes and functions of cells of all lineages in the TME; and (4) increased access to datasets to conduct parameter optimization testing.

5.5 Future directions

There are several opportunities to extend upon the work presented in this dissertation. First, studies that test new therapeutic interventions, such as those discussed above, may result in improved therapeutic responses and prolonged survival for patients with HPV(-) HNSCC, $HER2^{+}$ breast cancer, or PDAC.

Second, as the quantitative algorithms used in this dissertation are applicable across tissue types and single-cell spatial proteomics datasets, there are seemingly endless opportunities to leverage the computational methods used here to interrogate the types and locations of cells present in a diseased tissue and determine how these features contribute to clinical outcome. In the context of cancer, specific cellular structures exist within TMEs that are still poorly understood. Quantitative analysis of the single-cell spatial landscapes of these TME structures—for example, TLS—has the potential to illuminate how these structures form and contribute to tumor progression or reduction.

TLS are ectopic lymphoid organs whose presence in the TME generally associate with improved clinical outcome [276]. They are known to possess distinct cellular organizations similar to lymph nodes, however, how TLS form, the various stages of TLS maturation, and their role in anti-tumor immunity are not fully understood [276, 277]. Thus, an interesting opportunity exists to (1) assay TLS with a curated mIHC antibody panel capable of deeply phenotyping the cells known to exist within TLS (T cells, B cells, DCs, high endothelial venules) and (2) conduct a detailed quantitative analysis of the mIHC dataset to evaluate TLS cellular phenotype, function, and spatial organization by applying many of the same algorithms presented herein. Following these quantifications, downstream analyses could be performed to identify stages of TLS maturation and correlations to various clinical parameters.

As proof-of-principle for using quantitative spatial approaches to characterize TLS, a targeted cellular neighborhood analysis surrounding B cells was performed on the HNSCC mIHC dataset presented in Chapter 2. One of the resulting recurrent cellular neighborhoods was found to correspond to B cells residing specifically in the centers of the TLS present. The results were evaluated for accuracy by comparing the computationally detected TLS with the hand-annotated TLS as identified by an expert histologist on adjacent H&E tissue slides. A representative tissue region is shown (**Figure 5.1**), with the hand-annotations of the TLS overlaid to illustrate the high accuracy of the cellular neighborhood approach. Overall, these results demonstrate that single-cell spatial analyses can identify and characterize TLS. A deeper interrogation into the heterogeneity in TLS spatial neighborhoods could enable improved understanding of their role in the TME and potentially reveal new therapeutic approaches to leverage the anti-tumor effects of TLS. While TLS are one specific feature of TMEs that could be quantified through the algorithms described in this dissertation, the methods used here are applicable across cellular structures and tissues.

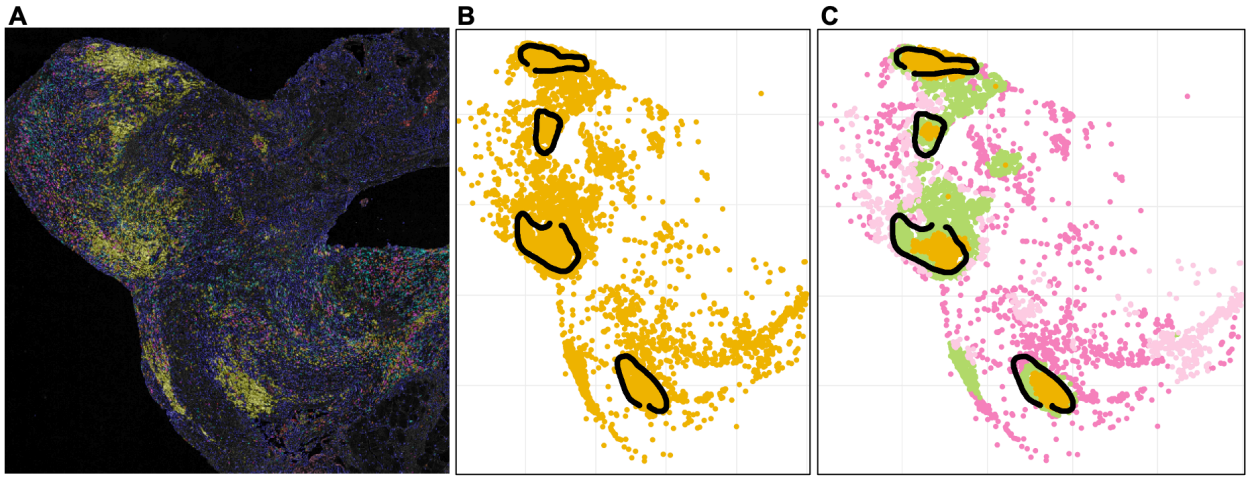


Figure 5.1: Recurrent cellular neighborhood analysis accurately identifies TLS in the HNSCC TME. A) mIHC image of all cells in TME tissue region. Yellow corresponds to B cells. **B)** Scatterplot reconstruction showing B cells only with TLS hand annotated in black. **C)** Scatterplot reconstruction showing B cells only. B cells are colored according to their neighborhood assignment. Cells belonging to the yellow neighborhood aligns with B cells participating in TLS per hand annotation.

A third extension of this work involves applying similar ML approaches to those described in Chapter 4 to answer new biological questions. For example, given the clinical importance of TLS but lack of knowledge of where they are located within the TME, it would be useful to know if TLS are present in the TME even if the tissue specimen did not contain a TLS. This is especially relevant in the case of a biopsy, where a needle randomly samples a small portion of the tumor mass, and a TLS could be easily missed in the biopsied tissue sample. In this use case, ML classifier models could be trained to predict whether or not a TLS is present in the broader TME from inputted TME features of the tissue actually sampled. As full surgical resections of tumors are not always clinically feasible, a model capable of making this prediction would enable researchers and clinicians to know if a TLS was present in a given TME without the need for a full resection. Additionally, downstream feature importance analyses could then identify which combinations of TME features were associated with presence versus absence of TLS in the surrounding tissue. This would help researchers better understand how and why a TLS forms in a tumor.

A final extension of this work involves developing an interactive visualization tool for future researchers tasked with analyzing single-cell spatial proteomics datasets. All of the quantitative analyses used in this dissertation were performed using Python software, and all of the analyses are applicable across single-cell spatial proteomics datasets. Thus, generation of a user-friendly dashboard for researchers with no programming experience would enable greater use of these quantitative methods. Streamlit is an open-source web application for running Python-based data analytics, and it is an appealing option for building an interactive dashboard. Upon creation of the dashboard, an end-user would simply upload their dataset to the dashboard, specify various metadata, and select which single-cell spatial algorithms to run. The results

would appear in interactive figures with adjustable parameters, such as distance thresholds or desired cell phenotypes to analyze. As proof-of-principle, a Streamlit dashboard was created to calculate various TME single-cell spatial metrics, including the immune-neoplastic mixing score described in Chapter 2, for a given tissue region (**Figure 5.2**). Expanding access to sophisticated computational methods through the design and publication of an interactive visualization tool, such as a Streamlit dashboard, would enable more researchers to easily conduct single-cell spatial analyses of tissues.

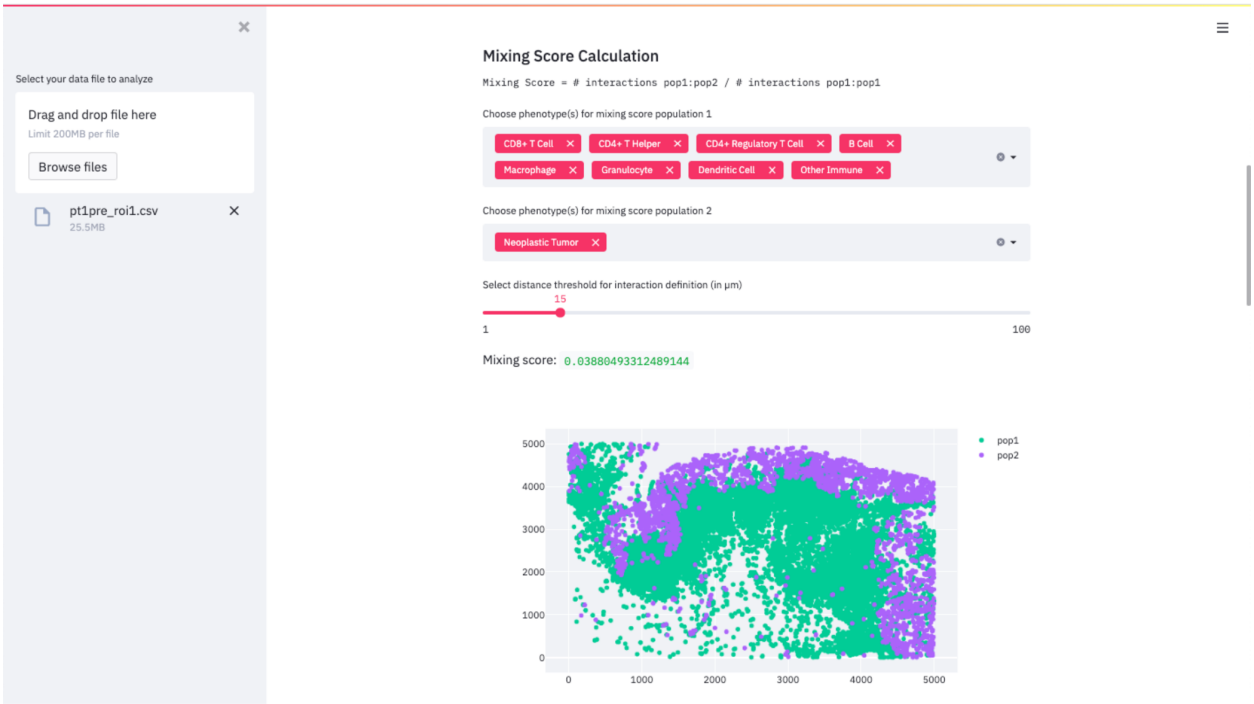


Figure 5.2: Streamlit dashboard enables interactive visualization of TME single-cell spatial analyses. The mixing score is calculated for a specified HNSCC tissue region, and the results are shown in the dashboard. Desired cell phenotypes for the calculation and a distance threshold are adjustable by the user. The analysis requires no programming experience.

5.6 Concluding remarks

It is clear that the cellular composition and spatial architecture of the TME play a significant role in tumor progression, response to therapy, and clinical outcome. Improving our understanding of the TME through quantitative analysis of single-cell spatial proteomics datasets

has the potential to improve patient stratification and identify increasingly personalized therapeutic interventions for cancer patients. The work in this dissertation advances cancer research by: (1) laying a foundation for future quantitative assessments of TME cellular composition, heterogeneity, and spatial organization, and (2) identifying potential candidate biomarkers of various clinical parameters, which may result in improved treatment strategies for future cancer patients. More broadly, this work can be used to guide the quantitative analysis of any single-cell spatial proteomics dataset, as the computational methods used here are applicable across diseases and data resulting from any multiplex tissue imaging assay.

References

[1] D. Hanahan and R. A. Weinberg, "The hallmarks of cancer," *Cell,* vol. 100, no. 1, pp. 57-70, Jan 7 2000, doi: 10.1016/s0092-8674(00)81683-9.

[2] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *Cell,* vol. 144, no. 5, pp. 646-74, Mar 4 2011, doi: 10.1016/j.cell.2011.02.013.

[3] K. E. de Visser and J. A. Joyce, "The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth," *Cancer Cell,* vol. 41, no. 3, pp. 374-403, Mar 13 2023, doi: 10.1016/j.ccell.2023.02.016.

[4] D. Hanahan and L. M. Coussens, "Accessories to the crime: functions of cells recruited to the tumor microenvironment," *Cancer Cell,* vol. 21, no. 3, pp. 309-22, Mar 20 2012, doi: 10.1016/j.ccr.2012.02.022.

[5] B. L. Brucher and I. S. Jamall, "Cell-cell communication in the tumor microenvironment, carcinogenesis, and anticancer treatment," *Cell Physiol Biochem,* vol. 34, no. 2, pp. 213-43, 2014, doi: 10.1159/000362978.

[6] M. R. Junttila and F. J. de Sauvage, "Influence of tumour micro-environment heterogeneity on therapeutic response," *Nature,* vol. 501, no. 7467, pp. 346-54, Sep 19 2013, doi: 10.1038/nature12626.

[7] G. P. Dunn, A. T. Bruce, H. Ikeda, L. J. Old, and R. D. Schreiber, "Cancer immunoediting: from immunosurveillance to tumor escape," *Nat Immunol,* vol. 3, no. 11, pp. 991-8, Nov 2002, doi: 10.1038/ni1102-991.

[8] M. Binnewies *et al.*, "Understanding the tumor immune microenvironment (TIME) for effective therapy," *Nat Med,* vol. 24, no. 5, pp. 541-550, May 2018, doi: 10.1038/s41591-018-0014-x.

[9] A. Labani-Motlagh, M. Ashja-Mahdavi, and A. Loskog, "The Tumor Microenvironment: A Milieu Hindering and Obstructing Antitumor Immune Responses," *Front Immunol,* vol. 11, p. 940, 2020, doi: 10.3389/fimmu.2020.00940.

[10] N. Zhang and M. J. Bevan, "CD8(+) T cells: foot soldiers of the immune system," *Immunity,* vol. 35, no. 2, pp. 161-8, Aug 26 2011, doi: 10.1016/j.immuni.2011.07.010.

[11] V. Shankaran *et al.*, "IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity," *Nature,* vol. 410, no. 6832, pp. 1107-11, Apr 26 2001, doi: 10.1038/35074122.

[12] L. Poncette, J. Bluhm, and T. Blankenstein, "The role of CD4 T cells in rejection of solid tumors," *Curr Opin Immunol,* vol. 74, pp. 18-24, Feb 2022, doi: 10.1016/j.coi.2021.09.005.

[13] Y. Togashi, K. Shitara, and H. Nishikawa, "Regulatory T cells in cancer immunosuppression - implications for anticancer therapy," *Nat Rev Clin Oncol,* vol. 16, no. 6, pp. 356-371, Jun 2019, doi: 10.1038/s41571-019-0175-7.

[14] D. Ostroumov, N. Fekete-Drimusz, M. Saborowski, F. Kuhnel, and N. Woller, "CD4 and CD8 T lymphocyte interplay in controlling tumor growth," *Cell Mol Life Sci,* vol. 75, no. 4, pp. 689-713, Feb 2018, doi: 10.1007/s00018-017-2686-7.

[15] J. R. Giles, A. M. Globig, S. M. Kaech, and E. J. Wherry, "CD8(+) T cells in the cancer-immunity cycle," *Immunity,* vol. 56, no. 10, pp. 2231-2253, Oct 10 2023, doi: 10.1016/j.immuni.2023.09.005.

[16] D. E. Speiser, O. Chijioke, K. Schaeuble, and C. Munz, "CD4(+) T cells in cancer," *Nat Cancer,* vol. 4, no. 3, pp. 317-329, Mar 2023, doi: 10.1038/s43018-023-00521-2.
[17] Y. Chu *et al.*, "Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance," *Nat Med,* vol. 29, no. 6, pp. 1550-1562, Jun 2023, doi: 10.1038/s41591-023-02371-y.
[18] L. Zheng *et al.*, "Pan-cancer single-cell landscape of tumor-infiltrating T cells," *Science,* vol. 374, no. 6574, p. abe6474, Dec 17 2021, doi: 10.1126/science.abe6474.
[19] E. J. Wherry *et al.*, "Molecular signature of CD8+ T cell exhaustion during chronic viral infection," *Immunity,* vol. 27, no. 4, pp. 670-84, Oct 2007, doi: 10.1016/j.immuni.2007.09.006.
[20] J. M. Angelosanto, S. D. Blackburn, A. Crawford, and E. J. Wherry, "Progressive loss of memory T cell potential and commitment to exhaustion during chronic viral infection," *J Virol,* vol. 86, no. 15, pp. 8161-70, Aug 2012, doi: 10.1128/JVI.00889-12.
[21] A. Schietinger *et al.*, "Tumor-Specific T Cell Dysfunction Is a Dynamic Antigen-Driven Differentiation Program Initiated Early during Tumorigenesis," *Immunity,* vol. 45, no. 2, pp. 389-401, Aug 16 2016, doi: 10.1016/j.immuni.2016.07.011.
[22] E. J. Wherry, "T cell exhaustion," *Nat Immunol,* vol. 12, no. 6, pp. 492-9, Jun 2011, doi: 10.1038/ni.2035.
[23] T. Duhen, R. Duhen, A. Lanzavecchia, F. Sallusto, and D. J. Campbell, "Functionally distinct subsets of human FOXP3+ Treg cells that phenotypically mirror effector Th cells," *Blood,* vol. 119, no. 19, pp. 4430-40, May 10 2012, doi: 10.1182/blood-2011-11-392324.
[24] E. Lanitis, D. Dangaj, M. Irving, and G. Coukos, "Mechanisms regulating T-cell infiltration and activity in solid tumors," *Ann Oncol,* vol. 28, no. suppl_12, pp. xii18-xii32, Dec 1 2017, doi: 10.1093/annonc/mdx238.
[25] M. M. Melssen, N. D. Sheybani, K. M. Leick, and C. L. Slingluff, Jr., "Barriers to immune cell infiltration in tumors," *J Immunother Cancer,* vol. 11, no. 4, Apr 2023, doi: 10.1136/jitc-2022-006401.
[26] S. M. Downs-Canner, J. Meier, B. G. Vincent, and J. S. Serody, "B Cell Function in the Tumor Microenvironment," *Annu Rev Immunol,* vol. 40, pp. 169-193, Apr 26 2022, doi: 10.1146/annurev-immunol-101220-015603.
[27] Z. Qin, G. Richter, T. Schuler, S. Ibe, X. Cao, and T. Blankenstein, "B cells inhibit induction of T cell-dependent tumor immunity," *Nat Med,* vol. 4, no. 5, pp. 627-30, May 1998, doi: 10.1038/nm0598-627.
[28] M. A. Perricone *et al.*, "Enhanced efficacy of melanoma vaccines in the absence of B lymphocytes," *J Immunother,* vol. 27, no. 4, pp. 273-81, Jul-Aug 2004, doi: 10.1097/00002371-200407000-00003.
[29] P. B. Olkhanud *et al.*, "Tumor-evoked regulatory B cells promote breast cancer metastasis by converting resting CD4(+) T cells to T-regulatory cells," *Cancer Res,* vol. 71, no. 10, pp. 3505-15, May 15 2011, doi: 10.1158/0008-5472.CAN-10-4316.
[30] D. Schrama *et al.*, "Targeting of lymphotoxin-alpha to the tumor elicits an efficient immune response associated with induction of peripheral lymphoid-like tissue," *Immunity,* vol. 14, no. 2, pp. 111-21, Feb 2001, doi: 10.1016/s1074-7613(01)00094-2.
[31] B. A. Helmink *et al.*, "B cells and tertiary lymphoid structures promote immunotherapy response," *Nature,* vol. 577, no. 7791, pp. 549-555, Jan 2020, doi: 10.1038/s41586-019-1922-8.

[32] R. Cabrita *et al.*, "Tertiary lymphoid structures improve immunotherapy and survival in melanoma," *Nature,* vol. 577, no. 7791, pp. 561-565, Jan 2020, doi: 10.1038/s41586-019-1914-8.
[33] F. Petitprez *et al.*, "B cells are associated with survival and immunotherapy response in sarcoma," *Nature,* vol. 577, no. 7791, pp. 556-560, Jan 2020, doi: 10.1038/s41586-019-1906-8.
[34] E. Vivier *et al.*, "Innate or adaptive immunity? The example of natural killer cells," *Science,* vol. 331, no. 6013, pp. 44-9, Jan 7 2011, doi: 10.1126/science.1198687.
[35] T. Bald, M. F. Krummel, M. J. Smyth, and K. C. Barry, "The NK cell-cancer cycle: advances and new challenges in NK cell-based immunotherapies," *Nat Immunol,* vol. 21, no. 8, pp. 835-847, Aug 2020, doi: 10.1038/s41590-020-0728-z.
[36] L. Gauthier *et al.*, "Multifunctional Natural Killer Cell Engagers Targeting NKp46 Trigger Protective Tumor Immunity," *Cell,* vol. 177, no. 7, pp. 1701-1713 e16, Jun 13 2019, doi: 10.1016/j.cell.2019.04.041.
[37] L. M. Coussens and Z. Werb, "Inflammation and cancer," *Nature,* vol. 420, no. 6917, pp. 860-7, Dec 19-26 2002, doi: 10.1038/nature01322.
[38] M. L. Broz *et al.*, "Dissecting the tumor myeloid compartment reveals rare activating antigen-presenting cells critical for T cell immunity," *Cancer Cell,* vol. 26, no. 5, pp. 638-52, Nov 10 2014, doi: 10.1016/j.ccell.2014.09.007.
[39] J. P. Bottcher and C. Reis e Sousa, "The Role of Type 1 Conventional Dendritic Cells in Cancer Immunity," *Trends Cancer,* vol. 4, no. 11, pp. 784-792, Nov 2018, doi: 10.1016/j.trecan.2018.09.001.
[40] F. A. Verreck *et al.*, "Human IL-23-producing type 1 macrophages promote but IL-10-producing type 2 macrophages subvert immunity to (myco)bacteria," *Proc Natl Acad Sci U S A,* vol. 101, no. 13, pp. 4560-5, Mar 30 2004, doi: 10.1073/pnas.0400983101.
[41] A. J. Boutilier and S. F. Elsawa, "Macrophage Polarization States in the Tumor Microenvironment," *Int J Mol Sci,* vol. 22, no. 13, Jun 29 2021, doi: 10.3390/ijms22136995.
[42] A. Sica, T. Schioppa, A. Mantovani, and P. Allavena, "Tumour-associated macrophages are a distinct M2 polarised population promoting tumour progression: potential targets of anti-cancer therapy," *Eur J Cancer,* vol. 42, no. 6, pp. 717-27, Apr 2006, doi: 10.1016/j.ejca.2006.01.003.
[43] S. Ghaffari and N. Rezaei, "Eosinophils in the tumor microenvironment: implications for cancer immunotherapy," *J Transl Med,* vol. 21, no. 1, p. 551, Aug 16 2023, doi: 10.1186/s12967-023-04418-7.
[44] G. Marone, A. R. Gambardella, F. Mattei, J. Mancini, G. Schiavoni, and G. Varricchi, "Basophils in Tumor Microenvironment and Surroundings," *Adv Exp Med Biol,* vol. 1224, pp. 21-34, 2020, doi: 10.1007/978-3-030-35723-8_2.
[45] A. Dutta, S. Bhagat, S. Paul, J. P. Katz, D. Sengupta, and D. Bhargava, "Neutrophils in Cancer and Potential Therapeutic Strategies Using Neutrophil-Derived Exosomes," *Vaccines (Basel),* vol. 11, no. 6, May 26 2023, doi: 10.3390/vaccines11061028.
[46] M. Yan *et al.*, "Roles of tumor-associated neutrophils in tumor metastasis and its clinical applications," *Front Cell Dev Biol,* vol. 10, p. 938289, 2022, doi: 10.3389/fcell.2022.938289.

[47] S. Raftopoulou, P. Valadez-Cosmes, Z. N. Mihalic, R. Schicho, and J. Kargl, "Tumor-Mediated Neutrophil Polarization and Therapeutic Implications," *Int J Mol Sci,* vol. 23, no. 6, Mar 16 2022, doi: 10.3390/ijms23063218.
[48] G. Marone *et al.*, "Is There a Role for Basophils in Cancer?," *Front Immunol,* vol. 11, p. 2103, 2020, doi: 10.3389/fimmu.2020.02103.
[49] G. Varricchi *et al.*, "Eosinophils: The unsung heroes in cancer?," *Oncoimmunology,* vol. 7, no. 2, p. e1393134, 2018, doi: 10.1080/2162402X.2017.1393134.
[50] J. Chauhan *et al.*, "Clinical and Translational Significance of Basophils in Patients with Cancer," *Cells,* vol. 11, no. 3, Jan 27 2022, doi: 10.3390/cells11030438.
[51] D. E. A. Komi and F. A. Redegeld, "Role of Mast Cells in Shaping the Tumor Microenvironment," *Clin Rev Allergy Immunol,* vol. 58, no. 3, pp. 313-325, Jun 2020, doi: 10.1007/s12016-019-08753-w.
[52] T. Fu *et al.*, "Spatial architecture of the immune microenvironment orchestrates tumor immunity and therapeutic response," *J Hematol Oncol,* vol. 14, no. 1, p. 98, Jun 25 2021, doi: 10.1186/s13045-021-01103-4.
[53] N. A. Giraldo *et al.*, "The clinical role of the TME in solid cancer," *Br J Cancer,* vol. 120, no. 1, pp. 45-53, Jan 2019, doi: 10.1038/s41416-018-0327-z.
[54] O. Elhanani, R. Ben-Uri, and L. Keren, "Spatial profiling technologies illuminate the tumor microenvironment," *Cancer Cell,* vol. 41, no. 3, pp. 404-420, Mar 13 2023, doi: 10.1016/j.ccell.2023.01.010.
[55] X. Q. Wang *et al.*, "Spatial predictors of immunotherapy response in triple-negative breast cancer," *Nature,* vol. 621, no. 7980, pp. 868-876, Sep 2023, doi: 10.1038/s41586-023-06498-3.
[56] H. R. Ali *et al.*, "Association between CD8+ T-cell infiltration and breast cancer survival in 12,439 patients," *Ann Oncol,* vol. 25, no. 8, pp. 1536-43, Aug 2014, doi: 10.1093/annonc/mdu191.
[57] Y. Naito *et al.*, "CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer," *Cancer Res,* vol. 58, no. 16, pp. 3491-4, Aug 15 1998. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/9721846.
[58] N. Guo *et al.*, "CD8 + T cell infiltration is associated with improved survival and negatively correlates with hypoxia in clear cell ovarian cancer," *Sci Rep,* vol. 13, no. 1, p. 6530, Apr 21 2023, doi: 10.1038/s41598-023-30655-3.
[59] Y. Yang *et al.*, "High intratumoral CD8(+) T-cell infiltration is associated with improved survival in prostate cancer patients undergoing radical prostatectomy," *Prostate,* vol. 81, no. 1, pp. 20-28, Jan 2021, doi: 10.1002/pros.24068.
[60] S. K. Bisheshar *et al.*, "The prognostic role of tumor associated macrophages in squamous cell carcinoma of the head and neck: A systematic review and meta-analysis," *Oral Oncol,* vol. 135, p. 106227, Dec 2022, doi: 10.1016/j.oraloncology.2022.106227.
[61] Y. Xue *et al.*, "Tumor-infiltrating M2 macrophages driven by specific genomic alterations are associated with prognosis in bladder cancer," *Oncol Rep,* vol. 42, no. 2, pp. 581-594, Aug 2019, doi: 10.3892/or.2019.7196.
[62] E. Allison, S. Edirimanne, J. Matthews, and S. J. Fuller, "Breast Cancer Survival Outcomes and Tumor-Associated Macrophage Markers: A Systematic Review and Meta-Analysis," *Oncol Ther,* vol. 11, no. 1, pp. 27-48, Mar 2023, doi: 10.1007/s40487-022-00214-3.

[63] J. Galon *et al.*, "Cancer classification using the Immunoscore: a worldwide task force," *J Transl Med,* vol. 10, p. 205, Oct 3 2012, doi: 10.1186/1479-5876-10-205.
[64] F. Pages *et al.*, "In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer," *J Clin Oncol,* vol. 27, no. 35, pp. 5944-51, Dec 10 2009, doi: 10.1200/JCO.2008.19.6147.
[65] N. A. Giraldo *et al.*, "Multidimensional, quantitative assessment of PD-1/PD-L1 expression in patients with Merkel cell carcinoma and association with response to pembrolizumab," *J Immunother Cancer,* vol. 6, no. 1, p. 99, Oct 1 2018, doi: 10.1186/s40425-018-0404-0.
[66] T. N. Gide *et al.*, "Close proximity of immune and tumor cells underlies response to anti-PD-1 based therapies in metastatic melanoma patients," *Oncoimmunology,* vol. 9, no. 1, p. 1659093, 2020, doi: 10.1080/2162402X.2019.1659093.
[67] D. B. Johnson *et al.*, "Quantitative Spatial Profiling of PD-1/PD-L1 Interaction and HLA-DR/IDO-1 Predicts Improved Outcomes of Anti-PD-1 Therapies in Metastatic Melanoma," *Clin Cancer Res,* vol. 24, no. 21, pp. 5250-5260, Nov 1 2018, doi: 10.1158/1078-0432.CCR-18-0309.
[68] H. W. Jackson *et al.*, "The single-cell pathology landscape of breast cancer," *Nature,* vol. 578, no. 7796, pp. 615-620, Feb 2020, doi: 10.1038/s41586-019-1876-x.
[69] E. I. Buchbinder and A. Desai, "CTLA-4 and PD-1 Pathways: Similarities, Differences, and Implications of Their Inhibition," *Am J Clin Oncol,* vol. 39, no. 1, pp. 98-106, Feb 2016, doi: 10.1097/COC.0000000000000239.
[70] G. J. Freeman *et al.*, "Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation," *J Exp Med,* vol. 192, no. 7, pp. 1027-34, Oct 2 2000, doi: 10.1084/jem.192.7.1027.
[71] M. F. Krummel and J. P. Allison, "CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation," *J Exp Med,* vol. 182, no. 2, pp. 459-65, Aug 1 1995, doi: 10.1084/jem.182.2.459.
[72] K. E. Pauken *et al.*, "Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade," *Science,* vol. 354, no. 6316, pp. 1160-1165, Dec 2 2016, doi: 10.1126/science.aaf2807.
[73] S. C. Wei *et al.*, "Distinct Cellular Mechanisms Underlie Anti-CTLA-4 and Anti-PD-1 Checkpoint Blockade," *Cell,* vol. 170, no. 6, pp. 1120-1133 e17, Sep 7 2017, doi: 10.1016/j.cell.2017.07.024.
[74] A. C. Huang *et al.*, "A single dose of neoadjuvant PD-1 blockade predicts clinical outcomes in resectable melanoma," *Nat Med,* vol. 25, no. 3, pp. 454-461, Mar 2019, doi: 10.1038/s41591-019-0357-y.
[75] D. Schadendorf *et al.*, "Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma," *J Clin Oncol,* vol. 33, no. 17, pp. 1889-94, Jun 10 2015, doi: 10.1200/JCO.2014.56.2736.
[76] A. Cercek *et al.*, "PD-1 Blockade in Mismatch Repair-Deficient, Locally Advanced Rectal Cancer," *N Engl J Med,* vol. 386, no. 25, pp. 2363-2376, Jun 23 2022, doi: 10.1056/NEJMoa2201445.
[77] G. Morad, B. A. Helmink, P. Sharma, and J. A. Wargo, "Hallmarks of response, resistance, and toxicity to immune checkpoint blockade," *Cell,* vol. 184, no. 21, pp. 5309-5337, Oct 14 2021, doi: 10.1016/j.cell.2021.09.020.

[78] M. Labrie *et al.*, "Multiomics analysis of serial PARP inhibitor treated metastatic TNBC inform on rational combination therapies," *NPJ Precis Oncol,* vol. 5, no. 1, p. 92, Oct 19 2021, doi: 10.1038/s41698-021-00232-w.
[79] D. E. Johnson, B. Burtness, C. R. Leemans, V. W. Y. Lui, J. E. Bauman, and J. R. Grandis, "Head and neck squamous cell carcinoma," *Nat Rev Dis Primers,* vol. 6, no. 1, p. 92, Nov 26 2020, doi: 10.1038/s41572-020-00224-3.
[80] A. Barsouk, J. S. Aluru, P. Rawla, K. Saginala, and A. Barsouk, "Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma," *Med Sci (Basel),* vol. 11, no. 2, Jun 13 2023, doi: 10.3390/medsci11020042.
[81] A. A. Bhat *et al.*, "Tumor microenvironment: an evil nexus promoting aggressive head and neck squamous cell carcinoma and avenue for targeted therapy," *Signal Transduct Target Ther,* vol. 6, no. 1, p. 12, Jan 12 2021, doi: 10.1038/s41392-020-00419-w.
[82] P. Zhou, Y. F. Yu, C. L. Lian, J. Wang, R. G. Zhuo, and S. G. Wu, "Survival Outcomes and Treatment Decision by Human Papillomavirus Status Among Patients With Stage IVC Head and Neck Squamous Cell Carcinoma," *Front Oncol,* vol. 11, p. 668066, 2021, doi: 10.3389/fonc.2021.668066.
[83] Q. Li, Y. Tie, A. Alu, X. Ma, and H. Shi, "Targeted therapy for head and neck cancer: signaling pathways and clinical studies," *Signal Transduct Target Ther,* vol. 8, no. 1, p. 31, Jan 16 2023, doi: 10.1038/s41392-022-01297-0.
[84] R. L. Ferris *et al.*, "Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck," *N Engl J Med,* vol. 375, no. 19, pp. 1856-1867, Nov 10 2016, doi: 10.1056/NEJMoa1602252.
[85] T. Y. Seiwert *et al.*, "Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial," *Lancet Oncol,* vol. 17, no. 7, pp. 956-965, Jul 2016, doi: 10.1016/S1470-2045(16)30066-3.
[86] B. Burtness *et al.*, "Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study," *Lancet,* vol. 394, no. 10212, pp. 1915-1928, Nov 23 2019, doi: 10.1016/S0140-6736(19)32591-7.
[87] H. Zeng, X. Song, J. Ji, L. Chen, Q. Liao, and X. Ma, "HPV infection related immune infiltration gene associated therapeutic strategy and clinical outcome in HNSCC," *BMC Cancer,* vol. 20, no. 1, p. 796, Aug 24 2020, doi: 10.1186/s12885-020-07298-y.
[88] Y. Xu, M. Gong, Y. Wang, Y. Yang, S. Liu, and Q. Zeng, "Global trends and forecasts of breast cancer incidence and deaths," *Sci Data,* vol. 10, no. 1, p. 334, May 27 2023, doi: 10.1038/s41597-023-02253-5.
[89] N. Cancer Genome Atlas, "Comprehensive molecular portraits of human breast tumours," *Nature,* vol. 490, no. 7418, pp. 61-70, Oct 4 2012, doi: 10.1038/nature11412.
[90] H. Tower, M. Ruppert, and K. Britt, "The Immune Microenvironment of Breast Cancer Progression," *Cancers (Basel),* vol. 11, no. 9, Sep 16 2019, doi: 10.3390/cancers11091375.
[91] A. Muntasell *et al.*, "Interplay between Natural Killer Cells and Anti-HER2 Antibodies: Perspectives for Breast Cancer Immunotherapy," *Front Immunol,* vol. 8, p. 1544, 2017, doi: 10.3389/fimmu.2017.01544.

[92] R. A. Clynes, T. L. Towers, L. G. Presta, and J. V. Ravetch, "Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets," *Nat Med,* vol. 6, no. 4, pp. 443-6, Apr 2000, doi: 10.1038/74704.
[93] L. Arnould *et al.*, "Trastuzumab-based treatment of HER2-positive breast cancer: an antibody-dependent cellular cytotoxicity mechanism?," *Br J Cancer,* vol. 94, no. 2, pp. 259-67, Jan 30 2006, doi: 10.1038/sj.bjc.6602930.
[94] E. O. Long, H. S. Kim, D. Liu, M. E. Peterson, and S. Rajagopalan, "Controlling natural killer cell responses: integration of signals for activation and inhibition," *Annu Rev Immunol,* vol. 31, pp. 227-58, 2013, doi: 10.1146/annurev-immunol-020711-075005.
[95] B. Cozar, M. Greppi, S. Carpentier, E. Narni-Mancinelli, L. Chiossone, and E. Vivier, "Tumor-Infiltrating Natural Killer Cells," *Cancer Discov,* vol. 11, no. 1, pp. 34-44, Jan 2021, doi: 10.1158/2159-8290.CD-20-0655.
[96] A. Ignatov, H. Eggemann, E. Burger, and T. Ignatov, "Patterns of breast cancer relapse in accordance to biological subtype," *J Cancer Res Clin Oncol,* vol. 144, no. 7, pp. 1347-1355, Jul 2018, doi: 10.1007/s00432-018-2644-2.
[97] C. J. Halbrook, C. A. Lyssiotis, M. Pasca di Magliano, and A. Maitra, "Pancreatic cancer: Advances and challenges," *Cell,* vol. 186, no. 8, pp. 1729-1754, Apr 13 2023, doi: 10.1016/j.cell.2023.02.014.
[98] H. Sung *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA Cancer J Clin,* vol. 71, no. 3, pp. 209-249, May 2021, doi: 10.3322/caac.21660.
[99] H. Oettle *et al.*, "Adjuvant chemotherapy with gemcitabine and long-term outcomes among patients with resected pancreatic cancer: the CONKO-001 randomized trial," *JAMA,* vol. 310, no. 14, pp. 1473-81, Oct 9 2013, doi: 10.1001/jama.2013.279201.
[100] A. D. Rhim *et al.*, "Stromal elements act to restrain, rather than support, pancreatic ductal adenocarcinoma," *Cancer Cell,* vol. 25, no. 6, pp. 735-47, Jun 16 2014, doi: 10.1016/j.ccr.2014.04.021.
[101] B. C. Ozdemir *et al.*, "Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival," *Cancer Cell,* vol. 25, no. 6, pp. 719-34, Jun 16 2014, doi: 10.1016/j.ccr.2014.04.005.
[102] J. J. Lee *et al.*, "Stromal response to Hedgehog signaling restrains pancreatic cancer progression," *Proc Natl Acad Sci U S A,* vol. 111, no. 30, pp. E3091-100, Jul 29 2014, doi: 10.1073/pnas.1411679111.
[103] H. Jiang *et al.*, "Pancreatic ductal adenocarcinoma progression is restrained by stromal matrix," *J Clin Invest,* vol. 130, no. 9, pp. 4704-4709, Sep 1 2020, doi: 10.1172/JCI136760.
[104] R. E. Royal *et al.*, "Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma," *J Immunother,* vol. 33, no. 8, pp. 828-33, Oct 2010, doi: 10.1097/CJI.0b013e3181eec14c.
[105] Y. Zhu *et al.*, "Tissue-Resident Macrophages in Pancreatic Ductal Adenocarcinoma Originate from Embryonic Hematopoiesis and Promote Tumor Progression," *Immunity,* vol. 47, no. 2, pp. 323-338 e6, Aug 15 2017, doi: 10.1016/j.immuni.2017.07.014.
[106] S. A. Vayrynen *et al.*, "Composition, Spatial Characteristics, and Prognostic Significance of Myeloid Cell Infiltration in Pancreatic Cancer," *Clin Cancer Res,* vol. 27, no. 4, pp. 1069-1081, Feb 15 2021, doi: 10.1158/1078-0432.CCR-20-3141.

[107] S. M. Liudahl *et al.*, "Leukocyte Heterogeneity in Pancreatic Ductal Adenocarcinoma: Phenotypic and Spatial Features Associated with Clinical Outcome," *Cancer Discov,* vol. 11, no. 8, pp. 2014-2031, Aug 2021, doi: 10.1158/2159-8290.CD-20-0841.
[108] M. Yarchoan, A. Hopkins, and E. M. Jaffee, "Tumor Mutational Burden and Response Rate to PD-1 Inhibition," *N Engl J Med,* vol. 377, no. 25, pp. 2500-2501, Dec 21 2017, doi: 10.1056/NEJMc1713444.
[109] L. A. Rojas *et al.*, "Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer," *Nature,* vol. 618, no. 7963, pp. 144-150, Jun 2023, doi: 10.1038/s41586-023-06063-y.
[110] K. T. Byrne *et al.*, "Neoadjuvant Selicrelumab, an Agonist CD40 Antibody, Induces Changes in the Tumor Microenvironment in Patients with Resectable Pancreatic Cancer," *Clin Cancer Res,* vol. 27, no. 16, pp. 4574-4586, Aug 15 2021, doi: 10.1158/1078-0432.CCR-21-1047.
[111] R. H. Vonderheide, "CD40 Agonist Antibodies in Cancer Immunotherapy," *Annu Rev Med,* vol. 71, pp. 47-58, Jan 27 2020, doi: 10.1146/annurev-med-062518-045435.
[112] K. T. Byrne and R. H. Vonderheide, "CD40 Stimulation Obviates Innate Sensors and Drives T Cell Immunity in Cancer," *Cell Rep,* vol. 15, no. 12, pp. 2719-32, Jun 21 2016, doi: 10.1016/j.celrep.2016.05.058.
[113] A. H. Morrison, M. S. Diamond, C. A. Hay, K. T. Byrne, and R. H. Vonderheide, "Sufficiency of CD40 activation and immune checkpoint blockade for T cell priming and tumor immunity," *Proc Natl Acad Sci U S A,* vol. 117, no. 14, pp. 8022-8031, Apr 7 2020, doi: 10.1073/pnas.1918971117.
[114] X. Yang *et al.*, "High-Throughput Transcriptome Profiling in Drug and Biomarker Discovery," *Front Genet,* vol. 11, p. 19, 2020, doi: 10.3389/fgene.2020.00019.
[115] R. M. Barnard, "Flow cytometry: a flexible tool for biomarker research," *Bioanalysis,* vol. 4, no. 20, pp. 2471-83, Oct 2012, doi: 10.4155/bio.12.225.
[116] S. R. Shi *et al.*, "Evaluation of the value of frozen tissue section used as "gold standard" for immunohistochemistry," *Am J Clin Pathol,* vol. 129, no. 3, pp. 358-66, Mar 2008, doi: 10.1309/7CXUYXT23E5AL8KQ.
[117] J. R. Lin, M. Fallahi-Sichani, and P. K. Sorger, "Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method," *Nat Commun,* vol. 6, p. 8390, Sep 24 2015, doi: 10.1038/ncomms9390.
[118] C. Giesen *et al.*, "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry," *Nat Methods,* vol. 11, no. 4, pp. 417-22, Apr 2014, doi: 10.1038/nmeth.2869.
[119] Y. Goltsev *et al.*, "Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging," *Cell,* vol. 174, no. 4, pp. 968-981 e15, Aug 9 2018, doi: 10.1016/j.cell.2018.07.010.
[120] M. Angelo *et al.*, "Multiplexed ion beam imaging of human breast tumors," *Nat Med,* vol. 20, no. 4, pp. 436-42, Apr 2014, doi: 10.1038/nm.3488.
[121] T. Tsujikawa *et al.*, "Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis," *Cell Rep,* vol. 19, no. 1, pp. 203-217, Apr 4 2017, doi: 10.1016/j.celrep.2017.03.037.
[122] G. Banik *et al.*, "High-dimensional multiplexed immunohistochemical characterization of immune contexture in human cancers," *Methods Enzymol.,* vol. 635, pp. 1-20, 2020, doi: 10.1016/bs.mie.2019.05.039.

[123] E. B. Garon *et al.*, "Pembrolizumab for the treatment of non-small-cell lung cancer," *N Engl J Med,* vol. 372, no. 21, pp. 2018-28, May 21 2015, doi: 10.1056/NEJMoa1501824.
[124] A. A. Davis and V. G. Patel, "The role of PD-L1 expression as a predictive biomarker: an analysis of all US Food and Drug Administration (FDA) approvals of immune checkpoint inhibitors," *J Immunother Cancer,* vol. 7, no. 1, p. 278, Oct 26 2019, doi: 10.1186/s40425-019-0768-9.
[125] S. L. Topalian, J. M. Taube, R. A. Anders, and D. M. Pardoll, "Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy," *Nat Rev Cancer,* vol. 16, no. 5, pp. 275-87, May 2016, doi: 10.1038/nrc.2016.36.
[126] R. Bai, Z. Lv, D. Xu, and J. Cui, "Predictive biomarkers for cancer immunotherapy with immune checkpoint inhibitors," *Biomark Res,* vol. 8, p. 34, 2020, doi: 10.1186/s40364-020-00209-0.
[127] S. Nawaz and Y. Yuan, "Computational pathology: Exploring the spatial dimension of tumor ecology," *Cancer Lett,* vol. 380, no. 1, pp. 296-303, Sep 28 2016, doi: 10.1016/j.canlet.2015.11.018.
[128] C. C. Maley, K. Koelble, R. Natrajan, A. Aktipis, and Y. Yuan, "An ecological measure of immune-cancer colocalization as a prognostic factor for breast cancer," *Breast Cancer Res,* vol. 17, no. 1, p. 131, Sep 22 2015, doi: 10.1186/s13058-015-0638-4.
[129] S. Nawaz, A. Heindl, K. Koelble, and Y. Yuan, "Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer," *Mod Pathol,* vol. 28, no. 6, pp. 766-77, Jun 2015, doi: 10.1038/modpathol.2015.37.
[130] L. C. Borish and J. W. Steinke, "2. Cytokines and chemokines," *J Allergy Clin Immunol,* vol. 111, no. 2 Suppl, pp. S460-75, Feb 2003, doi: 10.1067/mai.2003.108.
[131] M. E. Hoekstra *et al.*, "Long-distance modulation of bystander tumor cells by CD8(+) T cell-secreted IFNgamma," *Nat Cancer,* vol. 1, no. 3, pp. 291-301, Mar 2020, doi: 10.1038/s43018-020-0036-4.
[132] A. Oyler-Yaniv *et al.*, "A Tunable Diffusion-Consumption Mechanism of Cytokine Propagation Enables Plasticity in Cell-to-Cell Communication in the Immune System," *Immunity,* vol. 46, no. 4, pp. 609-620, Apr 18 2017, doi: 10.1016/j.immuni.2017.03.011.
[133] D. Busse *et al.*, "Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments," *Proc Natl Acad Sci U S A,* vol. 107, no. 7, pp. 3058-63, Feb 16 2010, doi: 10.1073/pnas.0812851107.
[134] A. J. Muller, O. Filipe-Santos, G. Eberl, T. Aebischer, G. F. Spath, and P. Bousso, "CD4+ T cells rely on a cytokine gradient to control intracellular pathogens beyond sites of antigen presentation," *Immunity,* vol. 37, no. 1, pp. 147-57, Jul 27 2012, doi: 10.1016/j.immuni.2012.05.015.
[135] R. Thibaut *et al.*, "Bystander IFN-gamma activity promotes widespread and sustained cytokine signaling altering the tumor microenvironment," *Nat Cancer,* vol. 1, no. 3, pp. 302-314, Mar 2020, doi: 10.1038/s43018-020-0038-2.
[136] K. Francis and B. O. Palsson, "Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion," *Proc Natl Acad Sci U S A,* vol. 94, no. 23, pp. 12258-62, Nov 11 1997, doi: 10.1073/pnas.94.23.12258.
[137] J. L. Carstens *et al.*, "Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer," *Nat Commun,* vol. 8, p. 15095, Apr 27 2017, doi: 10.1038/ncomms15095.

[138] C. Siret *et al.*, "Deciphering the Crosstalk Between Myeloid-Derived Suppressor Cells and Regulatory T Cells in Pancreatic Ductal Adenocarcinoma," *Front Immunol,* vol. 10, p. 3070, 2019, doi: 10.3389/fimmu.2019.03070.
[139] L. Keren *et al.*, "A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging," *Cell,* vol. 174, no. 6, pp. 1373-1387 e19, Sep 6 2018, doi: 10.1016/j.cell.2018.08.039.
[140] C. M. Schurch *et al.*, "Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front," *Cell,* Jul 31 2020, doi: 10.1016/j.cell.2020.07.005.
[141] A. Farkkila *et al.*, "Immunogenomic profiling determines responses to combined PARP and PD-1 inhibition in ovarian cancer," *Nat Commun,* vol. 11, no. 1, p. 1459, Mar 19 2020, doi: 10.1038/s41467-020-15315-8.
[142] D. Bzdok, N. Altman, and M. Krzywinski, "Statistics versus machine learning," *Nat Methods,* vol. 15, no. 4, pp. 233-234, Apr 2018, doi: 10.1038/nmeth.4642.
[143] J. Goecks, V. Jalili, L. M. Heiser, and J. W. Gray, "How Machine Learning Will Transform Biomedicine," *Cell,* vol. 181, no. 1, pp. 92-101, Apr 2 2020, doi: 10.1016/j.cell.2020.03.022.
[144] T. Risom *et al.*, "Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma," *Cell,* vol. 185, no. 2, pp. 299-310 e18, Jan 20 2022, doi: 10.1016/j.cell.2021.12.023.
[145] M. Sorin *et al.*, "Single-cell spatial landscapes of the lung tumour immune microenvironment," *Nature,* vol. 614, no. 7948, pp. 548-554, Feb 2023, doi: 10.1038/s41586-022-05672-3.
[146] L. Buk Cardoso *et al.*, "Machine learning for predicting survival of colorectal cancer patients," *Sci Rep,* vol. 13, no. 1, p. 8874, Jun 1 2023, doi: 10.1038/s41598-023-35649-9.
[147] K. E. Blise, S. Sivagnanam, G. L. Banik, L. M. Coussens, and J. Goecks, "Single-cell spatial architectures associated with clinical outcome in head and neck squamous cell carcinoma," *NPJ Precis Oncol,* vol. 6, no. 1, p. 10, Feb 25 2022, doi: 10.1038/s41698-022-00253-z.
[148] H. R. Ali *et al.*, "Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer," *Nature Cancer,* vol. 1, no. 2, pp. 163-175, 2020, doi: 10.1038/s43018-020-0026-6.
[149] S. Barua *et al.*, "Spatial interaction of tumor cells and regulatory T cells correlates with survival in non-small cell lung cancer," *Lung Cancer,* vol. 117, pp. 73-79, Mar 2018, doi: 10.1016/j.lungcan.2018.01.022.
[150] K. S. S. Enfield *et al.*, "Hyperspectral cell sociology reveals spatial tumor-immune cell interactions associated with lung cancer recurrence," *J Immunother Cancer,* vol. 7, no. 1, p. 13, Jan 16 2019, doi: 10.1186/s40425-018-0488-6.
[151] J. Galon *et al.*, "Type, density, and location of immune cells within human colorectal tumors predict clinical outcome," *Science,* vol. 313, no. 5795, pp. 1960-4, Sep 29 2006, doi: 10.1126/science.1129139.
[152] J. Eng, G. Thibault, S. W. Luoh, J. W. Gray, Y. H. Chang, and K. Chin, "Cyclic Multiplexed-Immunofluorescence (cmIF), a Highly Multiplexed Method for Single-Cell Analysis," *Methods Mol. Biol.,* vol. 2055, pp. 521-562, 2020, doi: 10.1007/978-1-4939-9773-2_24.

[153] E. Alsahafi *et al.*, "Clinical update on head and neck cancer: molecular biology and ongoing challenges," *Cell Death Dis,* vol. 10, no. 8, p. 540, Jul 15 2019, doi: 10.1038/s41419-019-1769-9.
[154] E. E. W. Cohen *et al.*, "The Society for Immunotherapy of Cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC)," *J Immunother Cancer,* vol. 7, no. 1, p. 184, Jul 15 2019, doi: 10.1186/s40425-019-0662-5.
[155] M. Canning *et al.*, "Heterogeneity of the Head and Neck Squamous Cell Carcinoma Immune Landscape and Its Impact on Immunotherapy," *Front Cell Dev Biol,* vol. 7, p. 52, 2019, doi: 10.3389/fcell.2019.00052.
[156] C. Fakhry *et al.*, "Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial," *J Natl Cancer Inst,* vol. 100, no. 4, pp. 261-9, Feb 20 2008, doi: 10.1093/jnci/djn011.
[157] K. K. Ang *et al.*, "Human papillomavirus and survival of patients with oropharyngeal cancer," *N Engl J Med,* vol. 363, no. 1, pp. 24-35, Jul 1 2010, doi: 10.1056/NEJMoa0912217.
[158] J. Predina *et al.*, "Changes in the local tumor microenvironment in recurrent cancers may explain the failure of vaccines after surgery," *Proc Natl Acad Sci U S A,* vol. 110, no. 5, pp. E415-24, Jan 29 2013, doi: 10.1073/pnas.1211850110.
[159] B. Zheng *et al.*, "Changes in the tumor immune microenvironment in resected recurrent soft tissue sarcomas," *Ann Transl Med,* vol. 7, no. 16, p. 387, Aug 2019, doi: 10.21037/atm.2019.07.43.
[160] A. Walens *et al.*, "Adaptation and selection shape clonal evolution of tumors during residual disease and recurrence," *Nat Commun,* vol. 11, no. 1, p. 5017, Oct 6 2020, doi: 10.1038/s41467-020-18730-z.
[161] C. Watermann *et al.*, "Recurrent HNSCC Harbor an Immunosuppressive Tumor Immune Microenvironment Suggesting Successful Tumor Immune Evasion," *Clin Cancer Res,* vol. 27, no. 2, pp. 632-644, Jan 15 2021, doi: 10.1158/1078-0432.CCR-20-0197.
[162] D. J. Hartman, F. Ahmad, R. L. Ferris, D. L. Rimm, and L. Pantanowitz, "Utility of CD8 score by automated quantitative image analysis in head and neck squamous cell carcinoma," *Oral Oncol,* vol. 86, pp. 278-287, Nov 2018, doi: 10.1016/j.oraloncology.2018.10.005.
[163] P. Balermpas *et al.*, "CD8+ tumour-infiltrating lymphocytes in relation to HPV status and clinical outcome in patients with head and neck cancer after postoperative chemoradiotherapy: A multicentre study of the German cancer consortium radiation oncology group (DKTK-ROG)," *Int J Cancer,* vol. 138, no. 1, pp. 171-81, Jan 1 2016, doi: 10.1002/ijc.29683.
[164] A. Nasman *et al.*, "Tumor infiltrating CD8+ and Foxp3+ lymphocytes correlate to clinical outcome and human papillomavirus (HPV) status in tonsillar cancer," *PLoS One,* vol. 7, no. 6, p. e38711, 2012, doi: 10.1371/journal.pone.0038711.
[165] C. Nordfors *et al.*, "CD8+ and CD4+ tumour infiltrating lymphocytes in relation to human papillomavirus status and clinical outcome in tonsillar and base of tongue squamous cell carcinoma," *Eur J Cancer,* vol. 49, no. 11, pp. 2522-30, Jul 2013, doi: 10.1016/j.ejca.2013.03.019.

[166] P. Balermpas *et al.*, "Tumour-infiltrating lymphocytes predict response to definitive chemoradiotherapy in head and neck cancer," *Br J Cancer,* vol. 110, no. 2, pp. 501-9, Jan 21 2014, doi: 10.1038/bjc.2013.640.
[167] Y. K. So *et al.*, "An increase of CD8(+) T cell infiltration following recurrence is a good prognosticator in HNSCC," *Sci Rep,* vol. 10, no. 1, p. 20059, Nov 18 2020, doi: 10.1038/s41598-020-77036-8.
[168] C. R. Schutt *et al.*, "Genomic and neoantigen evolution from primary tumor to first metastases in head and neck squamous cell carcinoma," *Oncotarget,* vol. 12, no. 6, pp. 534-548, Mar 16 2021, doi: 10.18632/oncotarget.27907.
[169] S. Russell *et al.*, "Immune cell infiltration patterns and survival in head and neck squamous cell carcinoma," *Head Neck Oncol,* vol. 5, no. 3, p. 24, Feb 27 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24723971.
[170] G. T. Wolf *et al.*, "Tumor infiltrating lymphocytes (TIL) and prognosis in oral cavity squamous carcinoma: a preliminary study," *Oral Oncol,* vol. 51, no. 1, pp. 90-5, Jan 2015, doi: 10.1016/j.oraloncology.2014.09.006.
[171] X. Wang *et al.*, "PD-1-expressing B cells suppress CD4(+) and CD8(+) T cells via PD-1/PD-L1-dependent pathway," *Mol Immunol,* vol. 109, pp. 20-26, May 2019, doi: 10.1016/j.molimm.2019.02.009.
[172] M. L. Thibult *et al.*, "PD-1 is a novel regulator of human B-cell activation," *Int Immunol,* vol. 25, no. 2, pp. 129-37, Feb 2013, doi: 10.1093/intimm/dxs098.
[173] W. H. Fridman, F. Pages, C. Sautes-Fridman, and J. Galon, "The immune contexture in human tumours: impact on clinical outcome," *Nat Rev Cancer,* vol. 12, no. 4, pp. 298-306, Mar 15 2012, doi: 10.1038/nrc3245.
[174] O. Wood *et al.*, "Head and Neck Squamous Cell Carcinomas Are Characterized by a Stable Immune Signature Within the Primary Tumor Over Time and Space," *Clin Cancer Res,* vol. 23, no. 24, pp. 7641-7649, Dec 15 2017, doi: 10.1158/1078-0432.CCR-17-0373.
[175] A. Reuben *et al.*, "Genomic and immune heterogeneity are associated with differential responses to therapy in melanoma," *NPJ Genom Med,* vol. 2, 2017, doi: 10.1038/s41525-017-0013-8.
[176] Q. Zhang *et al.*, "Integrated multiomic analysis reveals comprehensive tumour heterogeneity and novel immunophenotypic classification in hepatocellular carcinomas," *Gut,* vol. 68, no. 11, pp. 2019-2031, Nov 2019, doi: 10.1136/gutjnl-2019-318912.
[177] W. Lin *et al.*, "Single-cell transcriptome analysis of tumor and stromal compartments of pancreatic ductal adenocarcinoma primary tumors and metastatic lesions," *Genome Med,* vol. 12, no. 1, p. 80, Sep 29 2020, doi: 10.1186/s13073-020-00776-9.
[178] R. Gyanchandani *et al.*, "Intratumor Heterogeneity Affects Gene Expression Profile Test Prognostic Risk Stratification in Early Breast Cancer," *Clin Cancer Res,* vol. 22, no. 21, pp. 5362-5369, Nov 1 2016, doi: 10.1158/1078-0432.CCR-15-2889.
[179] C. H. Chung *et al.*, "Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression," *Cancer Cell,* vol. 5, no. 5, pp. 489-500, May 2004, doi: 10.1016/s1535-6108(04)00112-6.
[180] V. Walter *et al.*, "Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes," *PLoS One,* vol. 8, no. 2, p. e56823, 2013, doi: 10.1371/journal.pone.0056823.
[181] H. Parajuli *et al.*, "Integrin alpha11 is overexpressed by tumour stroma of head and neck squamous cell carcinoma and correlates positively with alpha smooth muscle actin

expression," *J Oral Pathol Med,* vol. 46, no. 4, pp. 267-275, Apr 2017, doi: 10.1111/jop.12493.
[182] Q. Zhang *et al.*, "Integrated analysis of single-cell RNA-seq and bulk RNA-seq reveals distinct cancer-associated fibroblasts in head and neck squamous cell carcinoma," *Ann Transl Med,* vol. 9, no. 12, p. 1017, Jun 2021, doi: 10.21037/atm-21-2767.
[183] K. P. Lim *et al.*, "Fibroblast gene expression profile reflects the stage of tumour progression in oral squamous cell carcinoma," *J Pathol,* vol. 223, no. 4, pp. 459-69, Mar 2011, doi: 10.1002/path.2841.
[184] K. Shiga, M. Hara, T. Nagasaki, T. Sato, H. Takahashi, and H. Takeyama, "Cancer-Associated Fibroblasts: Their Characteristics and Their Roles in Tumor Growth," *Cancers (Basel),* vol. 7, no. 4, pp. 2443-58, Dec 11 2015, doi: 10.3390/cancers7040902.
[185] L. Tao, G. Huang, H. Song, Y. Chen, and L. Chen, "Cancer associated fibroblasts: An essential role in the tumor microenvironment," *Oncol Lett,* vol. 14, no. 3, pp. 2611-2620, Sep 2017, doi: 10.3892/ol.2017.6497.
[186] E. Sahai *et al.*, "A framework for advancing our understanding of cancer-associated fibroblasts," *Nat Rev Cancer,* vol. 20, no. 3, pp. 174-186, Mar 2020, doi: 10.1038/s41568-019-0238-1.
[187] A. C. Johansson *et al.*, "Cancer-associated fibroblasts induce matrix metalloproteinase-mediated cetuximab resistance in head and neck squamous cell carcinoma cells," *Mol Cancer Res,* vol. 10, no. 9, pp. 1158-68, Sep 2012, doi: 10.1158/1541-7786.MCR-12-0030.
[188] A. B. Rodriguez *et al.*, "Immune mechanisms orchestrate tertiary lymphoid structures in tumors via cancer-associated fibroblasts," *Cell Rep,* vol. 36, no. 3, p. 109422, Jul 20 2021, doi: 10.1016/j.celrep.2021.109422.
[189] N. Fujii *et al.*, "Cancer-associated fibroblasts and CD163-positive macrophages in oral squamous cell carcinoma: their clinicopathological and prognostic significance," *J Oral Pathol Med,* vol. 41, no. 6, pp. 444-51, Jul 2012, doi: 10.1111/j.1600-0714.2012.01127.x.
[190] B. Gok Yavuz *et al.*, "Cancer associated fibroblasts sculpt tumour microenvironment by recruiting monocytes and inducing immunosuppressive PD-1(+) TAMs," *Sci Rep,* vol. 9, no. 1, p. 3172, Feb 28 2019, doi: 10.1038/s41598-019-39553-z.
[191] S. M. Markwell and S. A. Weed, "Tumor and stromal-based contributions to head and neck squamous cell carcinoma invasion," *Cancers (Basel),* vol. 7, no. 1, pp. 382-406, Feb 27 2015, doi: 10.3390/cancers7010382.
[192] H. Munir *et al.*, "Stromal-driven and Amyloid beta-dependent induction of neutrophil extracellular traps modulates tumor growth," *Nat Commun,* vol. 12, no. 1, p. 683, Jan 29 2021, doi: 10.1038/s41467-021-20982-2.
[193] X. Cong *et al.*, "CD66b(+) neutrophils and alpha-SMA(+) fibroblasts predict clinical outcomes and benefits from postoperative chemotherapy in gastric adenocarcinoma," *Cancer Med,* vol. 9, no. 8, pp. 2761-2773, Apr 2020, doi: 10.1002/cam4.2939.
[194] C. Gaggioli *et al.*, "Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells," *Nat Cell Biol,* vol. 9, no. 12, pp. 1392-400, Dec 2007, doi: 10.1038/ncb1658.
[195] J. G. Goetz *et al.*, "Biomechanical remodeling of the microenvironment by stromal caveolin-1 favors tumor invasion and metastasis," *Cell,* vol. 146, no. 1, pp. 148-63, Jul 8 2011, doi: 10.1016/j.cell.2011.05.040.

[196] J. R. Lin *et al.*, "Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes," *Elife,* vol. 7, Jul 11 2018, doi: 10.7554/eLife.31657.
[197] P. Virtanen *et al.*, "SciPy 1.0: fundamental algorithms for scientific computing in Python," *Nat. Methods,* vol. 17, no. 3, pp. 261-272, Mar 2020, doi: 10.1038/s41592-019-0686-2.
[198] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *JMLR,* vol. 12, pp. 2825-2830, 2011.
[199] S. Seabold and J. Perktold, "statsmodels: Econometric and statistical modeling with python," *Proceedings of the 9th Python in Science Conference,* 2010.
[200] C. Davidson-Pilon, "lifelines: survival analysis in Python," *Journal of Open Source Software,* vol. 4, no. 40, p. 1317, 2019, doi: 10.21105/joss.01317.
[201] F. A. I. Ehlers *et al.*, "Natural killer cells occupy unique spatial neighborhoods in human HER2- and HER2+ breast cancers," *Manuscript in preparation,* 2023.
[202] A. G. Rivenbark, S. M. O'Connor, and W. B. Coleman, "Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine," *The American journal of pathology,* vol. 183, no. 4, pp. 1113-1124, 2013.
[203] P. Savas *et al.*, "Clinical relevance of host immunity in breast cancer: from TILs to the clinic," *Nature reviews Clinical oncology,* vol. 13, no. 4, pp. 228-241, 2016.
[204] W. Wang, A. K. Erbe, J. A. Hank, Z. S. Morris, and P. M. Sondel, "NK cell-mediated antibody-dependent cellular cytotoxicity in cancer immunotherapy," *Front. Immunol.,* vol. 6, p. 368, 2015.
[205] Y. T. Bryceson, M. E. March, H.-G. Ljunggren, and E. O. Long, "Synergy among receptors on resting NK cells for the activation of natural cytotoxicity and cytokine secretion," *Blood,* vol. 107, no. 1, pp. 159-166, 2006.
[206] E. Vivier, E. Tomasello, M. Baratin, T. Walzer, and S. Ugolini, "Functions of natural killer cells," *Nat. Immunol.,* vol. 9, no. 5, pp. 503-510, 2008.
[207] K. C. Barry *et al.*, "A natural killer–dendritic cell axis defines checkpoint therapy–responsive tumor microenvironments," *Nat. Med.,* vol. 24, no. 8, pp. 1178-1191, 2018.
[208] O. Melaiu, V. Lucarini, L. Cifaldi, and D. Fruci, "Influence of the Tumor Microenvironment on NK Cell Function in Solid Tumors," *Front Immunol,* vol. 10, p. 3038, 2019, doi: 10.3389/fimmu.2019.03038.
[209] S. Park *et al.*, "The therapeutic effect of anti-HER2/neu antibody depends on both innate and adaptive immunity," *Cancer Cell,* vol. 18, no. 2, pp. 160-70, Aug 9 2010, doi: 10.1016/j.ccr.2010.06.014.
[210] A. Beano *et al.*, "Correlation between NK function and response to trastuzumab in metastatic breast cancer patients," *J Transl Med,* vol. 6, p. 25, May 16 2008, doi: 10.1186/1479-5876-6-25.
[211] A. Muntasell *et al.*, "NK Cell Infiltrates and HLA Class I Expression in Primary HER2(+) Breast Cancer Predict and Uncouple Pathological Response and Disease-free Survival," *Clin Cancer Res,* vol. 25, no. 5, pp. 1535-1545, Mar 1 2019, doi: 10.1158/1078-0432.CCR-18-2365.
[212] J. A. Myers and J. S. Miller, "Exploring the NK cell platform for cancer immunotherapy," *Nature Reviews Clinical Oncology,* vol. 18, no. 2, pp. 85-100, 2021.
[213] J. Fuentes-Antras *et al.*, "Adoptive cell therapy in breast cancer: a current perspective of next-generation medicine," *Front. Oncol.,* p. 2376, 2020.

[214] E. Tan *et al.*, "The key hypoxia regulated gene CAIX is upregulated in basal-like breast tumours and is associated with resistance to chemotherapy," *Br. J. Cancer,* vol. 100, no. 2, pp. 405-411, 2009.
[215] E. Mamessier *et al.*, "Human breast tumor cells induce self-tolerance mechanisms to avoid NKG2D-mediated and DNAM-mediated NK cell recognition," *Cancer Res.,* vol. 71, no. 21, pp. 6621-6632, 2011.
[216] E. Mamessier *et al.*, "Human breast cancer cells enhance self tolerance by promoting evasion from NK cell antitumor immunity," *The Journal of clinical investigation,* vol. 121, no. 9, pp. 3609-3622, 2011.
[217] F. A. Ehlers *et al.*, "ADCC-Inducing Antibody Trastuzumab and Selection of KIR-HLA Ligand Mismatched Donors Enhance the NK Cell Anti-Breast Cancer Response," *Cancers (Basel),* vol. 13, no. 13, p. 3232, 2021.
[218] B. E. Wilson, C. Gorrini, and D. W. Cescon, "Breast cancer immune microenvironment: from pre-clinical models to clinical therapies," *Breast Cancer Res Treat,* vol. 191, no. 2, pp. 257-267, Jan 2022, doi: 10.1007/s10549-021-06431-0.
[219] T. Tsujikawa *et al.*, "Prognostic significance of spatial immune profiles in human solid cancers," *Cancer Sci.,* vol. 111, no. 10, pp. 3426-3434, 2020.
[220] N. Sorrelle *et al.*, "Improved multiplex immunohistochemistry for immune microenvironment evaluation of mouse formalin-fixed, paraffin-embedded tissues," *The Journal of Immunology,* vol. 202, no. 1, pp. 292-299, 2019.
[221] B. Ruffell, A. Au, H. S. Rugo, L. J. Esserman, E. S. Hwang, and L. M. Coussens, "Leukocyte composition of human breast cancer," *Proceedings of the National Academy of Sciences,* vol. 109, no. 8, pp. 2796-2801, 2012.
[222] C. R. G. Del Alcazar *et al.*, "Immune escape in breast cancer during in situ to invasive carcinoma transition," *Cancer Discov.,* vol. 7, no. 10, pp. 1098-1115, 2017.
[223] E. Mamessier *et al.*, "Peripheral blood NK cells from breast cancer patients are tumor-induced composite subsets," *The Journal of Immunology,* vol. 190, no. 5, pp. 2424-2436, 2013.
[224] S. Mahgoub, H. Abosalem, M. Emara, N. Kotb, A. Maged, and S. Soror, "Restoring NK cells functionality via cytokine activation enhances cetuximab-mediated NK-cell ADCC: A promising therapeutic tool for HCC patients," *Mol. Immunol.,* vol. 137, pp. 221-227, 2021.
[225] A.-S. Chretien *et al.*, "High-dimensional mass cytometry analysis of NK cell alterations in AML identifies a subgroup with adverse clinical outcome," *Proceedings of the National Academy of Sciences,* vol. 118, no. 22, 2021.
[226] T. Garcia-Iglesias *et al.*, "Low NKp30, NKp46 and NKG2D expression and reduced cytotoxic activity on NK cells in cervical cancer and precursor lesions," *BMC Cancer,* vol. 9, no. 1, pp. 1-8, 2009.
[227] B. Han *et al.*, "Altered NKp30, NKp46, NKG2D, and DNAM-1 expression on circulating NK cells is associated with tumor progression in human gastric cancer," *Journal of immunology research,* vol. 2018, 2018.
[228] A. Muntasell *et al.*, "NK cell infiltrates and HLA class I expression in primary HER2+ breast cancer predict and uncouple pathological response and disease-free survival," *Clin. Cancer Res.,* vol. 25, no. 5, pp. 1535-1545, 2019.

[229] S. Rezaeifard, A. Talei, M. Shariat, and N. Erfani, "Tumor infiltrating NK cell (TINK) subsets and functional molecules in patients with breast cancer," *Mol. Immunol.,* vol. 136, pp. 161-167, 2021.
[230] B. Stankovic *et al.*, "Immune Cell Composition in Human Non-small Cell Lung Cancer," *Front Immunol,* vol. 9, p. 3101, 2018, doi: 10.3389/fimmu.2018.03101.
[231] J. Wu, H. K. Mishra, and B. Walcheck, "Role of ADAM17 as a regulatory checkpoint of CD16A in NK cells and as a potential target for cancer immunotherapy," *J. Leukoc. Biol.,* vol. 105, no. 6, pp. 1297-1303, 2019.
[232] E. Hatjiharissi *et al.*, "Increased natural killer cell expression of CD16, augmented binding and ADCC activity to rituximab among individuals expressing the FcγRIIIa-158 V/V and V/F polymorphism," *Blood, The Journal of the American Society of Hematology,* vol. 110, no. 7, pp. 2561-2564, 2007.
[233] A. Muntasell *et al.*, "High numbers of circulating CD57+ NK cells associate with resistance to HER2-specific therapeutic antibodies in HER2+ primary breast cancer," *Cancer immunology research,* 2019.
[234] L. Chiossone, P.-Y. Dumas, M. Vienne, and E. Vivier, "Natural killer cells and other innate lymphoid cells in cancer," *Nature Reviews Immunology,* vol. 18, no. 11, pp. 671-688, 2018.
[235] Y. Gao *et al.*, "Tumor immunoevasion by the conversion of effector NK cells into type 1 innate lymphoid cells," *Nat. Immunol.,* vol. 18, no. 9, pp. 1004-1015, 2017.
[236] P. Carrega *et al.*, "Natural killer cells infiltrating human nonsmall-cell lung cancer are enriched in CD56brightCD16− cells and display an impaired capability to kill tumor cells," *Cancer,* vol. 112, no. 4, pp. 863-875, 2008.
[237] L. Schmidt *et al.*, "Enhanced adaptive immune responses in lung adenocarcinoma through natural killer cell stimulation," *Proceedings of the National Academy of Sciences,* vol. 116, no. 35, pp. 17460-17469, 2019.
[238] M. A. Garrido *et al.*, "HLA class I alterations in breast carcinoma are associated with a high frequency of the loss of heterozygosity at chromosomes 6 and 15," *Immunogenetics,* vol. 70, no. 10, pp. 647-659, 2018.
[239] G. B. R. F. d. Silva *et al.*, "Expression of the classical and nonclassical HLA molecules in breast cancer," *Int. J. Breast Cancer,* vol. 2013, 2013.
[240] S. Sarkar *et al.*, "Hypoxia induced impairment of NK cell cytotoxicity against multiple myeloma can be overcome by IL-2 activation of the NK cells," *PLoS One,* vol. 8, no. 5, p. e64835, 2013, doi: 10.1371/journal.pone.0064835.
[241] C. Van Berckelaer *et al.*, "Infiltrating stromal immune cells in inflammatory breast cancer are associated with an improved outcome and increased PD-L1 expression," *Breast Cancer Res.,* vol. 21, no. 1, pp. 1-12, 2019.
[242] V. Pelekanou *et al.*, "Tumor-infiltrating lymphocytes and PD-L1 expression in pre-and posttreatment breast cancers in the SWOG S0800 phase II neoadjuvant chemotherapy trial," *Mol. Cancer Ther.,* vol. 17, no. 6, pp. 1324-1331, 2018.
[243] K. L. McNamara *et al.*, "Spatial proteomic characterization of HER2-positive breast tumors through neoadjuvant therapy predicts response," *Nature cancer,* vol. 2, no. 4, pp. 400-413, 2021.
[244] K. E. Blise *et al.*, "Machine learning links T-cell function and spatial localization to neoadjuvant immunotherapy and clinical outcome in pancreatic cancer," *Cancer Immunol Res,* Feb 21 2024, doi: 10.1158/2326-6066.CIR-23-0873.

[245] M. J. Pishvaian *et al.*, "Overall survival in patients with pancreatic cancer receiving matched therapies following molecular profiling: a retrospective analysis of the Know Your Tumor registry trial," *Lancet Oncol,* vol. 21, no. 4, pp. 508-518, Apr 2020, doi: 10.1016/S1470-2045(20)30074-7.
[246] L. J. Padron *et al.*, "Sotigalimab and/or nivolumab with chemotherapy in first-line metastatic pancreatic cancer: clinical and immunologic analyses from the randomized phase 2 PRINCE trial," *Nat Med,* vol. 28, no. 6, pp. 1167-1177, Jun 2022, doi: 10.1038/s41591-022-01829-9.
[247] G. Oliveira and C. J. Wu, "Dynamics and specificities of T cells in cancer immunotherapy," *Nat Rev Cancer,* vol. 23, no. 5, pp. 295-316, May 2023, doi: 10.1038/s41568-023-00560-y.
[248] C. Kao *et al.*, "Transcription factor T-bet represses expression of the inhibitory receptor PD-1 and sustains virus-specific CD8+ T cell responses during chronic infection," *Nat Immunol,* vol. 12, no. 7, pp. 663-71, May 29 2011, doi: 10.1038/ni.2046.
[249] Y. Wang *et al.*, "The Transcription Factor TCF1 Preserves the Effector Function of Exhausted CD8 T Cells During Chronic Viral Infection," *Front Immunol,* vol. 10, p. 169, 2019, doi: 10.3389/fimmu.2019.00169.
[250] O. Khan *et al.*, "TOX transcriptionally and epigenetically programs CD8(+) T cell exhaustion," *Nature,* vol. 571, no. 7764, pp. 211-218, Jul 2019, doi: 10.1038/s41586-019-1325-x.
[251] H. Mi *et al.*, "Quantitative Spatial Profiling of Immune Populations in Pancreatic Ductal Adenocarcinoma Reveals Tumor Microenvironment Heterogeneity and Prognostic Biomarkers," *Cancer Res,* vol. 82, no. 23, pp. 4359-4372, Dec 2 2022, doi: 10.1158/0008-5472.CAN-22-1190.
[252] A. Dias Costa *et al.*, "Neoadjuvant Chemotherapy Is Associated with Altered Immune Cell Infiltration and an Anti-Tumorigenic Microenvironment in Resected Pancreatic Cancer," *Clin Cancer Res,* vol. 28, no. 23, pp. 5167-5179, Dec 1 2022, doi: 10.1158/1078-0432.CCR-22-1125.
[253] C. Jiang, Y. Xiu, K. Qiao, X. Yu, S. Zhang, and Y. Huang, "Prediction of lymph node metastasis in patients with breast invasive micropapillary carcinoma based on machine learning and SHapley Additive exPlanations framework," *Front Oncol,* vol. 12, p. 981059, 2022, doi: 10.3389/fonc.2022.981059.
[254] S. J. Sammut *et al.*, "Multi-omic machine learning predictor of breast cancer therapy response," *Nature,* vol. 601, no. 7894, pp. 623-629, Jan 2022, doi: 10.1038/s41586-021-04278-5.
[255] A. E. Carpenter *et al.*, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes," *Genome Biol,* vol. 7, no. 10, p. R100, 2006, doi: 10.1186/gb-2006-7-10-r100.
[256] B. J. Baaten, R. Tinoco, A. T. Chen, and L. M. Bradley, "Regulation of Antigen-Experienced T Cells: Lessons from the Quintessential Memory Marker CD44," *Front Immunol,* vol. 3, p. 23, 2012, doi: 10.3389/fimmu.2012.00023.
[257] A. Davis, R. Gao, and N. E. Navin, "SCOPIT: sample size calculations for single-cell sequencing experiments," *BMC Bioinformatics,* vol. 20, no. 1, p. 566, Nov 12 2019, doi: 10.1186/s12859-019-3167-9.

[258] I. S. Jang, E. C. Neto, J. Guinney, S. H. Friend, and A. A. Margolin, "Systematic assessment of analytical methods for drug sensitivity prediction from cancer cell line data," *Pac Symp Biocomput,* pp. 63-74, 2014.
[259] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* vol. 67, no. 2, pp. 301-320, 2005, doi: 10.1111/j.1467-9868.2005.00503.x.
[260] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *JAIR,* vol. 16, pp. 321-357, 2002.
[261] L. S.M. and L. S., "A Unified Approach to Interpreting Model Predictions," *NIPS,* pp. 4765-4774, 2017.
[262] E. R. Lutz *et al.*, "Immunotherapy converts nonimmunogenic pancreatic tumors into immunogenic foci of immune regulation," *Cancer Immunol Res,* vol. 2, no. 7, pp. 616-31, Jul 2014, doi: 10.1158/2326-6066.CIR-14-0027.
[263] A. Sledzinska *et al.*, "Regulatory T Cells Restrain Interleukin-2- and Blimp-1-Dependent Acquisition of Cytotoxic Function by CD4(+) T Cells," *Immunity,* vol. 52, no. 1, pp. 151-166 e6, Jan 14 2020, doi: 10.1016/j.immuni.2019.12.007.
[264] D. Chen *et al.*, "CTLA-4 blockade induces a microglia-Th1 cell partnership that stimulates microglia phagocytosis and anti-tumor function in glioblastoma," *Immunity,* vol. 56, no. 9, pp. 2086-2104 e8, Sep 12 2023, doi: 10.1016/j.immuni.2023.07.015.
[265] G. S. Kinker *et al.*, "Mature tertiary lymphoid structures are key niches of tumour-specific immune responses in pancreatic ductal adenocarcinomas," *Gut,* vol. 72, no. 10, pp. 1927-1941, Oct 2023, doi: 10.1136/gutjnl-2022-328697.
[266] X. Wang *et al.*, "TOX promotes the exhaustion of antitumor CD8(+) T cells by preventing PD1 degradation in hepatocellular carcinoma," *J Hepatol,* vol. 71, no. 4, pp. 731-741, Oct 2019, doi: 10.1016/j.jhep.2019.05.015.
[267] R. Abu Eid *et al.*, "Enhanced Therapeutic Efficacy and Memory of Tumor-Specific CD8 T Cells by Ex Vivo PI3K-delta Inhibition," *Cancer Res,* vol. 77, no. 15, pp. 4135-4145, Aug 1 2017, doi: 10.1158/0008-5472.CAN-16-1925.
[268] M. Ahmadzadeh *et al.*, "Tumor antigen–specific CD8 T cells infiltrating the tumor express high levels of PD-1 and are functionally impaired," *Blood,* vol. 114, no. 8, pp. 1537-1544, 2009, doi: 10.1182/blood-2008-12-195792.
[269] A. H. Morrison, K. T. Byrne, and R. H. Vonderheide, "Immunotherapy and Prevention of Pancreatic Cancer," *Trends Cancer,* vol. 4, no. 6, pp. 418-428, Jun 2018, doi: 10.1016/j.trecan.2018.04.001.
[270] W. Zhang, Z. Zhao, and F. Li, "Natural killer cell dysfunction in cancer and new strategies to utilize NK cell potential for cancer immunotherapy," *Mol Immunol,* vol. 144, pp. 58-70, Apr 2022, doi: 10.1016/j.molimm.2022.02.015.
[271] E. Muraro *et al.*, "KIR-HLA Functional Repertoire Influences Trastuzumab Efficiency in Patients With HER2-Positive Breast Cancer," *Front Immunol,* vol. 12, p. 791958, 2021, doi: 10.3389/fimmu.2021.791958.
[272] C. Liang, S. Huang, Y. Zhao, S. Chen, and Y. Li, "TOX as a potential target for immunotherapy in lymphocytic malignancies," *Biomark Res,* vol. 9, no. 1, p. 20, Mar 20 2021, doi: 10.1186/s40364-021-00275-y.
[273] J. E. Darnell, Jr., "Transcription factors as targets for cancer therapy," *Nat Rev Cancer,* vol. 2, no. 10, pp. 740-9, Oct 2002, doi: 10.1038/nrc906.

[274] Z. Tao and X. Wu, "Targeting Transcription Factors in Cancer: From "Undruggable" to "Druggable"," *Methods Mol Biol,* vol. 2594, pp. 107-131, 2023, doi: 10.1007/978-1-0716-2815-7_9.
[275] L. Kuett *et al.*, "Three-dimensional imaging mass cytometry for highly multiplexed molecular and cellular mapping of tissues and the tumor microenvironment," *Nat Cancer,* vol. 3, no. 1, pp. 122-133, Jan 2022, doi: 10.1038/s43018-021-00301-w.
[276] T. N. Schumacher and D. S. Thommen, "Tertiary lymphoid structures in cancer," *Science,* vol. 375, no. 6576, p. eabf9419, Jan 7 2022, doi: 10.1126/science.abf9419.
[277] Q. Zhang and S. Wu, "Tertiary lymphoid structures are critical for cancer prognosis and therapeutic response," *Front Immunol,* vol. 13, p. 1063711, 2022, doi: 10.3389/fimmu.2022.1063711.