The Texas Medical Center Library

DigitalCommons@TMC

Dissertations and Theses (Open Access)

MD Anderson UTHealth Houston Graduate School

6-2021

Impact Of Intratumor Heterogeneity And The Tumor Microenvironment In Shaping Tumor Evolution And Response To Therapy

Akash Mitra

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations

Part of the Bioinformatics Commons, Computational Biology Commons, Genomics Commons, Immunology and Infectious Disease Commons, Oncology Commons, Systems Biology Commons, and the Translational Medical Research Commons

Recommended Citation

Mitra, Akash, "Impact Of Intratumor Heterogeneity And The Tumor Microenvironment In Shaping Tumor Evolution And Response To Therapy" (2021). *Dissertations and Theses (Open Access)*. 1111.
https://digitalcommons.library.tmc.edu/utgsbs_dissertations/1111

This Dissertation (PhD) is brought to you for free and open access by the MD Anderson UTHealth Houston Graduate School at DigitalCommons@TMC. It has been accepted for inclusion in Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digcommons@library.tmc.edu.

Impact of Intratumor Heterogeneity and the Tumor Microenvironment in Shaping Tumor Evolution and Response to Therapy

by

Akash Mitra, BS

APPROVED:

P. Andrew Futreal

P. Andrew Futreal, Ph.D.
Advisory Professor

Patrick Hwu, M.D.

Alexander Lazar MD/PhD

Alexander Lazar, M.D., Ph.D.

Han Liang, Ph.D.

Nicholas Navin, Ph.D.

Hussein Tawbi, M.D., Ph.D.

Linghua Wang

Linghua Wang, Ph.D.

APPROVED:

Dean, The University of Texas

MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

Impact of Intratumor Heterogeneity and the Tumor Microenvironment in Shaping Tumor Evolution and Response to Therapy

A

DISSERTATION

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Akash Mitra, BS

Houston, Texas

June 2021

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Andrew Futreal for being an incredible scientist, mentor and downright good human being. Regardless of his schedule, he spent tremendous time mentoring me and discussing our projects from inception to completion. During these meetings, I've learned the important of being inquisitive and openly seeking guidance from those with expert insights. He also allowed me to work independently and laid the groundwork that has given me the tools and skillset to explore my future career path. He provided me with all the resources I could ask in order to undertake my research and helped build and foster collaborations with other incredible physicians and scientists. Above all, Andy is committed to academic integrity, and his passion and drive in science is evident to anyone who comes in contact with him. I am grateful to him for providing me an excellent example of a scholar and mentor.

I would also like to extend my deepest gratitude to all of my committee members. Dr. Patrick Hwu for providing me with training opportunities in the field of melanoma and immunology along with his excellent background in cancer immunology and immunotherapy. Dr. Alexander Lazar for always engaging with my scientific curiosity and establishing an incredible resource in terms of tissue collection to further out scientific understanding. Dr. Han Liang for his expertise in bioinformatics and for building translational skillsets that interface between the computational and clinical world. Dr. Nicholas Navin for his expertise in single cell sequencing technologies and for providing helpful comments during committee meetings and encouraging me to build upon my scientific hypothesis. Dr Hussein Tawbi for his dual expertise in both the melanoma and sarcoma realms and for conducting ground-breaking clinical trials that are changing the paradigm for patients with late stage disease. I thank Dr. Linghua Wang for her robust computational advice and for showcasing an incredible example of what it means to be productive as early career faculty.

I am incredibly grateful to all the non-formal relationships and mentorships I have built along the way. First and foremost, Dr. Miles Andrews for being the epitome of an amazing scientist, clinician and mentor. Through chance collaborations, Miles encouraged me to delve deeper into the immunology realm and always entertained my questions, regardless of how basic they might have been. He is incredibly patient, and his work speaks for itself. I would also like to thank Dr. Won-Chul Lee and Dr. Whijae Roh, for their time and computational guidance. They always made time for me and provided me with intellectually stimulating discussions. I would like to thank Dr. Hannah Beird for her helpful comments and feedback on manuscripts, research proposals and presentations. I would also like to extend a thank you to Latasha Little, Curtis Gumbs and the entire Cancer Genomic Laboratory team for their astounding scientific rigor and work ethic. I thank Dr. Jianhua Zhang and his entire team for always being available for computational discussions and guidance. I would also like to thank the entirety of the Futreal lab and the incredible departmental support we have through Genomic Medicine for making this work possible.

Finally, I would like to thank my parents, Jyoti Mitra and Srirupa Mitra for supporting me through all my career trajectories. I would also like to thank Penny, for always being a bundle of joy whenever I would be home after work.

Impact of Intratumor Heterogeneity and the Tumor Microenvironment in Shaping Tumor Evolution and Response to Therapy

Akash Mitra, B.S.

Advisory Professor: P. Andrew Futreal, Ph.D.

Intratumor heterogeneity (ITH) is a crucial challenge in cancer treatment. The genotypic and phenotypic heterogeneity underlying diverse cancer types leads to subclonal variation, which may result in mixed or failed response to therapy. The heterogeneity at the tumor level, along with the tumor microenvironment (TME), often shapes tumor evolution and ultimately clinical outcome. Given that modern treatment paradigms increasingly expose patients with metastatic disease to multiple treatment modalities through the course of their disease, there exists a need to characterize robust and predictive biomarkers of response to therapy. In order to accurately characterize tumor evolution, we need to account for both intra-tumoral genomic factors as well as how these factors interact with the components of the tumor microenvironment to allow for response and resistance to various forms of therapy.

Through multidimensional profiling of a tumor and its cellular ecosystem we assessed spatial heterogeneity in response to ICB in metastatic melanoma and highlighted the effects of copy number alterations being associated to varied immune phenotypes. We characterized localized regions of immune activation

by both T and B cells however, sub-regions with chromosome 7 gain displayed a distinct lack of immunocytic infiltrate but evidence of neutrophil activation that was recapitulated in TCGA samples and was associated with lack of response to ICB across three separate cohorts.

Through a longitudinal sampling strategy paired to a clinical trial, we characterized the impact of heterogeneity on therapy outcomes after combination ICB agents in multiple sarcoma histologies. Amongst the genomic, transcriptomic and immune-based molecular correlates analyzed, elevated intratumoral levels of B cells were most significantly correlated with response at both time-points, as measured by both gene expression and immunostaining. Responsive tumors were also associated with higher diversity and richness of the intratumoral T cell repertoire at baseline.

In order to characterize the effects of the immune-tumor microenvironment, we used multiplex immunofluorescence to quantitatively analyze T-cell subsets and sequenced the T-cell receptors of matched Barrett's Esophagus (BE) and Esophageal adenocarcinoma (EAC) samples from our patient cohort. Through our work we identified a more infiltrated and diverse immune microenvironment in EAC as compared to BE, however this was primarily characterized by an immunosuppressive T-cell infiltrate, the clonal expansion of which may have been limited with public antigenicity.

In summary, our work frames the spatial and temporal molecular features of heterogeneity characterized across tumors and their surrounding microenvironment, and how their interplay in turn may influence disease progression, tumor evolution and response to therapy.

Table of Contents

Approval i

Title ii

Acknowledgements iii

Abstract vi

Table of Contents ix

CHAPTER ONE – Introduction 1

1.1 Intratumor Heterogeneity 1

1.2 Tumor Microenvironment 2

1.3 The immune system and cancer 3

1.4 Cancer Immunology 3

1.5 Immune Checkpoint Blockade 5

1.6 Cancer Immunogenomics – Neoantigen Prediction 6

1.7 Cancer Immunogenomics – T-cell Receptor Profiling 7

1.8Cancer Immunogenomics – Immune Deconvolution 9

1.9 Co-evolution of tumor genomics with immune landscape of ICB 10

1.10 Hypothesis and Specific Aims and Rationale 13

CHAPTER TWO – Materials and Methods 16

Aim 1 – Biospecimen collection 17

Aim 1 – Sample processing: spatial intratumoral analysis. 17

Aim 1 – Sample processing: longitudinal time points 19

Aim 1 – Nucleic acid extraction 19

Aim 1 – Cancer gene panel DNA sequencing 20

Aim 1 – Whole exome sequencing 20

Aim 1 – Phylogenetic tree construction 21

Aim 1 – RNA Sequencing 22

Aim 1 – Methylation analysis 23

Aim 1 – TCR sequencing 24

Aim 1 – Stimulation of ZDNNC17 p.H507Y neoantigen-specific T cells 25

Aim 1 – Single-cell sequencing 26

Aim 1 – Neoantigen prediction 26

Aim 1 – Copy number alteration analysis 27

Aim 1 – Reverse-phase protein array 27

Aim 1 – Ba/F3 transformation assay 27

Aim 1 – Immunohistochemistry 28

Aim 1 – Data analysis and statistical considerations 29

Aim 2 – Biospecimen collection 29

Aim 2 – Response Criteria 30

Aim 2 – Nucleic acid extraction 30

Aim 2 – Whole exome sequencing 30

Aim 2 – Mutational Signatures 31

Aim 2 – Copy Number Alterations 31

Aim 2 – RNA sequencing 32

Aim 2 – Nanostring Analysis 34

Aim 2 – B-cell Receptor repertoire inference 34

Aim 2 – Immunohistochemistry staining for CD20 and CD56 35

Aim 2 – Multiplex immunofluorescence staining 36

Aim 2 – TCR Sequencing 37

Aim 3 – Cohort Selection 37

Aim 3 - Multiplex immunofluorescence (mIF) staining 38

Aim 3 - TCR sequencing 39

Aim 3 – Statistical Analysis 39

CHAPTER THREE – Spatial Intratumor Heterogeneity 41

3.1 Introduction 42

3.2 Results 43

3.2 Results - Longitudinal tumor sampling 43

3.2 Results - Mutational ITH is highly prevalent and spatially restricted 44

3.2 Results - Genomic ITH is dominated by copy number alterations 44

3.2 Results - Immune cell content is highly and spatially diverse 45

3.2 Results - Sites of similar immune composition may be spatially remote 46

3.2 Results - ITH implicates simultaneous methods of immune exclusion 47
3.2 Results - Integrative analyses of multimodal molecular phenotypic data 49
3.2 Results - Tumor-specific and agnostic T-cell recruitment occurs on a regional basis 50
3.2 Results - T-cell clone persistence reveals recurrent priming and functional diversity 52
3.2 Results - Chromosome 7 gain is associated with an unfavorable immune outcome 53
3.3 Discussion 55
3.4 Figures and Tables 60
CHAPTER FOUR – Predictors of Response and Resistance to CICB in sarcoma 104
4.1 Introduction 105
4.2 Results 106
4.2 Results - Patient Profile 107
4.2 Results - Copy number alterations acquired during ICB exposure may lead to therapy resistance 107
4.2 Results - Co-expression networks modulating response to C-ICB in sarcoma 109
4.2 Results - Distinct immune microenvironments present in sarcomas 111

4.2 Results - B-cells predict response to C-ICB 112

4.2 Results - B cell enrichment significantly associated with response across multiple cohorts 113

4.2 Results - Baseline and on-treatment TCR diversity correlates with response to C-ICB 114

4.3 Discussion 116

4.4 Figures and Tables. 118

CHAPTER FIVE – Immune Spatial Microenvironment 135

5.1 Introduction 136

5.2 Results 137

5.2 Results – Patient Characteristics 137

5.2 Results - Multiplex immunofluorescence 138

5.2 Results- TCR sequencing analysis 140

5.3 Discussion 142

5.4 Figures and Tables 146

CHAPTER SIX – Discussion, Conclusions and Future Directions 157

6.1 Discussion and Conclusions 158

6.2 Future Directions 163

References 168

VITA 191

Chapter 1 INTRODUCTION

1.1 Intratumor Heterogeneity (ITH)

Cancer originates from histologically normal cells that progress towards neoplasia through the accumulation of various genetic and epigenetic reprogramming events. In particular, normal cells progress towards hyperplasia and dysplasia with an increase in cell numbers and size, and loss of cellular orientation(Douglas Hanahan, 2000). The majority of these changes are reversible until the first pre-invasive step of carcinoma in-situ which may then lead to invasive carcinoma. Various molecular changes occur through these steps and beyond in the context of metastasis and while different tumor types may exhibit different alterations for transformation events, most share a common feature of variability present within the cells of a single neoplastic lesion which may downstream lead to tumor progression(Fidler, 1978, Heppner, 1984). This variability or heterogeneity may result in changes in antigen expression, cellular behavior, metastatic potential and therapeutic response(Axelson et al., 2005, Brown et al., 2014).

While cells acquire somatic alterations in an evolutionary process across time, the majority of these aberrations are relatively benign passenger events that confer no selective advantages to these cells(Kandoth et al., 2013, Lawrence et al., 2014). Cancer cells, however, may eventually acquire driver mutations though these mutations may not uniformly be seen within a tumor(Lawrence et al., 2014, Kandoth et al., 2013). Additionally, increasing evidence has shown that low frequency sub-clones can assist and aid in the growth of dominant clones. These types and distributions of mutations present in tumors may be viewed through the lens of space, wherein tumor cells may exhibit differing mutational events in a spatially distributed manner or through the context of time and exposure to therapeutic interventions. Multiple facets may govern the shaping of the cancer genome including interactions of tumor cells with accompanying cells present in the tumor microenvironment.

1.2 Tumor Microenvironment (TME)

Bidirectional communication between cells and their microenvironment is essential for healthy tissue and cellular homeostasis. Tumors do not exist in silos, but rather in conjunction with normal cells found in the surrounding stroma. Infiltrating immune cells which form the hallmark of an effector immune response in tumors have been implicated as early as 1863 with the work of Virchow demonstrating infiltrating leukocytes as a hallmark of cancer and thereby establishing the role of inflammation and cancer(Balkwill, 2001). While ITH leads to aberrant genomic changes in cancerous cells, tumors are also shaped by microenvironmental populations and their stromal contributions or activation states. Moreover, this reprogramming occurs in a cyclic manner with environmental and oncogenic signals changing over the course of cancer progression and metastasis.

Fibroblasts and macrophages contribute towards a growth suppressive state within the TME at early stages, however through tumor education, these cells can be re-programmed to a pro-tumorigenic state(Qian and Pollard, 2010). Tumor associated macrophages (TAMs) can support primary tumor growth, angiogenesis and invasion, and can secrete growth signaling factors such as EGF which participates in a paracrine signaling loop through tumor-induced CSF-1(Pyonteck et al., 2013). Myeloid derived stem cells (MDSCs) and regulatory T-cells also create an anti-inflammatory phenotype in response to tumor activated cytokine axes including TGF-B and CXCL5-CXCR2(Almand et al., 2001). These immune-regulatory cells can inhibit antigen presentation by dendritic cells

(DCs), inhibition of T and B-cell proliferation and activation and inhibition of NK-cell derived killing(Gabrilovich et al., 2001, Gabrilovich et al., 2012, Sinha et al., 2005, Liu et al., 2007). Cancer associated fibroblasts (CAFs), activated through tumor induced TGF-B signaling, can secrete VEGF, thereby supporting angiogenesis through tumor growth(Kalluri and Zeisberg, 2006).

1.3 The immune system and cancer

Dysregulated microenvironments through the context of tissues with chronic inflammation have been linked to higher levels of cancer incidence. In particular, through a retrospective study of 417 cancer-free patients with liver cirrhosis, 27% of individuals went on to develop hepatocellular carcinoma over a follow up time of 12 years(Sangiovanni et al., 2004). Additional work in the context of inflammatory bowel disease and colitis has also shown a causal link from inflammation to the incidence of tumorigenesis through colorectal cancer(Beaugerie et al., 2013). An unresolved inflammatory response can lead to the accumulation of stromal cell-types that eventually become dysregulated and lead to the formation of a pro-tumorigenic niche.

However, the immune system can act in paradoxical formats in the context of cancer. In an analysis of chronically immunocompromised organ transplant recipients, Stewart and colleagues observed an increase incidence of multiple tumor types including lung, GI, skin and genital cancers(Stewart, 1995). The balance of a suppressive immune system and inflammation have both been shown to increase pro-tumorigenic functions, thereby underscoring the need of better understanding the opposing functions of immune cells in cancer.

1.4 Cancer immunology

Tumor cells traditionally function through driver genes or through the activation of oncogenes and inactivation of various tumor suppressor genes. However, the yin and yang of the immune system functions in tandem with the tumor ecosystem to induce both pro-tumorigenic and pro-immunogenic factors. Moreover, while the majority of cancer immunotherapies involve shifting the balance in favor of immune-activation and T/NK-cell killing, only a subset of patients benefit from such immunotherapies, emphasizing the need for better molecular stratification of patients and therapeutics.

Cancer cells may typically be recognized as foreign cells by the immune system as a result of mutated peptides, neoantigens, that are present on the surface of cancer cells(Yarchoan et al., 2017). Briefly, the cancer immunity cycle implies release of cancer cell antigens which are then presented by DCs for the priming and activation by T-cells which leads to trafficking of cytotoxic lymphocytes (CTLs) from secondary lymphoid structures to tumor periphery.(Chen and Mellman, 2013) Once in the tumor ecosystem, CTLs recognize cancer cells and secrete interferon-gamma to initiate tumor killing.

However, tumors may employ multiple strategies to evade immune attack which includes losing the ability to present neoantigens through the loss of the human leukocyte antigen (HLA)(Douglas Hanahan, 2000, Schumacher and Schreiber, 2015). Since neoantigen presentation and an effector T-cell driven immune response is driven through Class 1 HLA presentation on the cell surface for recognition by T-cell receptors, certain tumors evolve mechanisms to downregulate HLA genes thereby facilitating immune evasion. Additionally, tumors may employ immune checkpoints at the priming and activation stage including CTLA4:B7.1, PD-L1:PD-1, PD-L1:B7.1 as well as prevent infiltration of CTLs into the tumor bed through the secretion of VEGF and endothelin B receptors(Lippitz, 2013, Franciszkiewicz et al., 2012, Riella et al., 2012). Recognition of tumor cells by T-cells may also be impaired through reduced MHC expression on cancer cells and tumor-killing evaded through PD-L1:PD-1, PD-L1:B7.1, TIM-3, BTLA, VISTA, LAG-3, IDO and other factors(Chen et al., 2012, Mellman et al., 2011, Greaves and Gribben, 2013). Given the dawn of next generation sequencing and a better understanding of the genomic and molecular factors that govern tumors and their immune evasion, their lies a need to better understand predictors of response and resistance for cancer immunotherapies in order to allow for better patient stratification for improving overall outcomes in patients, and reducing immune related toxicities.

1.5 Immune checkpoint blockade (ICB)

The goal of cancer immunotherapies is to initiate or re-invigorate a self-sustaining cycle of cancer immunity and to potentially overcome negative feedback loops that may be employed by cancer cells. In the current age of personalized therapy, Topalian and colleagues have discussed the concept of the "common denominator" approach in the context of ICB(Topalian et al., 2015). A large number of mutational diversity and causal driver mutations have been identified through the vast profiling of tumors over the past 20 years. ICB takes advantage of the mutational diversity by exposing the plethora of antigens created for potential immune recognition. However, through the implication of immune checkpoints, the innate and adaptive immune systems are often held in check with their effector and activated functions being suppressed.

Seminal work by Leach and colleagues laid the groundwork for blocking the cytotoxic T lymphocyte antigen 4 (CTLA-4), that led to spontaneous tumor regression in murine models and eventually the clinical development and approval of the first in man anti-

CTLA-4 agent for patients with metastatic melanoma(DR Leach, 1996, Hodi et al., 2010). Subsequent work highlighted the role of blocking Programmed Death-1 (PD-1) and it's ligand, PD-L1 in the treatment of diverse cancer types. Anti CTLA-4 agents are used to enhance the priming of in lymphoid organs while the anti-PD-1/PD-L1 axis reverses inhibition of intratumoral T-cells and improves antigenic response(Ribas, 2012, DR Leach, 1996). Moreover, Wei and colleagues have shown distinct mechanisms through combination therapy that is implicated in the re-activation of previously phenotypically exhausted T-cells(Wei et al., 2019). The 5 year survival rate for diseases such as metastatic melanoma have dramatically improved with the advent of ICB however, this finding is extended only to a subset of patients afflicted with melanoma and an even smaller subset of cancer patients overall(Larkin et al., 2019). Several groups have recently undertaken the task of elucidating pre- and early on-treatment predictors of various ICB agents through advances in tumor and host-immune factor sequencing(Chen et al., 2016a, Reuben et al., 2017, Roh et al., 2017, Snyder et al., 2014a, Rizvi et al., 2016, Hellmann et al., 2018, Hugo et al., 2016, Zaretsky et al., 2016, Cloughesy et al., 2019, Ock et al., 2017, Riaz et al., 2017b, Samstein et al., 2019, Mandal et al., 2019).

1.6 Cancer Immunogenomics – Neoantigen Prediction

The advent of next generation sequencing has led to many advances in discovering and characterizing normal as well as malignant genomes. Early efforts by the Cancer Genome Project and The Cancer Genome Atlas program, have allowed us to understand drivers and dependencies in cancer genomes and how they manifest and evolve through disease and therapy. The work initially started with genomic next generation sequencing (NGS) technologies that allowed for the detection of single nucleotide variants, insertions and deletions, copy number alterations and structural variants present in the cancer exome/genome. This was eventually expanded upon to include other modalities including evaluating the transcriptome through RNA-sequencing and the epigenetic landscape through bisulfite/methylation-based sequencing. Through a multidimensional lens, we can characterize multiple molecular markers to understand the mechanisms through which cancer cells establish and expand their niche.

While cancer cells may possess several alterations, neoantigens represent somatic mutations that can be presented by DCs and recognized by the immune system as foreign(Chen and Mellman, 2013). Over the course of the past decade, large amounts of data have been used to generate and predict candidate neo-antigens in silico to allow for more efficient therapeutic targeting. Briefly, somatic mutations are called using variant callers with sequences spanning the somatic point mutation being translated into mutated peptides with k-mer amino acids varying in length from 8-11 amino acids(Lundegaard et al., 2008, Karosiene et al., 2013). Following which, we can perform HLA-typing and binding affinities of the predicted mutated peptides can be computed against patient specific HLA molecules through the use of neural networks and large databases(Lundegaard et al., 2008, Karosiene et al., 2013). A candidate short list of predicted neoantigens can then be computed by sub-setting the IC50 binding affinities under a particular nanomolar threshold.

1.7 Cancer Immunogenomics – T-cell Receptor Profiling

An active effector immune response is employed through the engagement of T-lymphocytes against neoantigens. T-cells recognize and bind to these foreign molecules through the use of antigen binding receptors. The vast amounts of diversity present in the immune repertoire is caused due to V(D)J recombination which results in the somatic recombination of the gene segments present in the variable (V), diversity (D), and joining (J) region of the T-cell receptor (TCR) loci(Market and Papavasiliou, 2003, Schatz, 1992). Following recombination, a sub-sample of nucleotide are removed from the junction sites and the alpha and beta subunits combine to form one TCR(Market and Papavasiliou, 2003). The uniqueness of the TCR rearrangement is determined by the complementary-determining regions 3B (CDR3B), which is the junctional site of VDJ recombination, accounting for the most variability in the TCR repertoire(Janeway CA Jr, 2001).

Through recent advances in targeted repertoire sequencing, we can now sequence the CDR3 region of the TCRα and TCRβ subunits to investigate TCRs present in the TME or in the peripheral blood of patients. Immuno-sequencing of the TCR allows for the comprehensive assessment of the repertoire and enables us to correlate immune responses pertaining to tumor growth or shrinkage present within a tumor and its ecosystem, or through the patient as a whole. However, given that genomic material for TCR-sequencing may not always be available, several groups have implemented algorithms to probe RNA-sequencing data for TCR rearrangements(Bolotin et al., 2015, Li et al., 2016a). Following the sequencing or de-convoluting the immune repertoire, multiple ecology-based characteristics can be employed to detect the clonality and diversity of the repertoire. The richness of the repertoire can be obtained through the total number of unique nucleotide rearrangements present in a sample. Simpsons clonality allows us to determine the shape of the repertoire, wherein values approaching 1 indicate a highly monoclonal and values approaching 0 indicating a highly polyclonal repertoire(Simpson, 1949). Simpsons clonality is defined as the square root of the sum of all observed rearrangements of the square fractional abundance of each rearrangement(Simpson, 1949). Additionally, it can also be computed as the square root of Simpsons Dominance (D)(Simpson, 1949). Values for Simpsons Dominance are also categorized in the same 0 (highly polyclonal) to 1 (highly monoclonal) range. Simpsons dominance can be transformed to Simpsons diversity through either the complement (1-D) or the reciprocal (1/D)(Simpson, 1949). The former ranges from 0 (highly monoclonal) to 1 (highly polyclonal) while the latter is characterized between 1 and the richness of the repertoire. Additionally, the Shannon entropy can also be used to represent diversity with high entropy samples being highly diverse and low entropy samples having lower diversity(Schneider-Hohendorf et al., 2016, CE, 1963). Mathematically, its defined as the sum of frequency of each clone multiplied by the log2 base of the same frequency over all productive rearrangements in a sample.

1.8 Cancer Immunogenomics – Immune Deconvolution

Immunohistochemistry (IHC) or multiplexed immunofluorescence (mIF) profiling allows us to quantitate the characterize the presence of various immune and non-immune markers present within tumor bed and it’s periphery. However, sampling all immune markers may not be feasible with a certain block of tissue. In order to quantitate the various immune components, present within a sample, recent algorithms have taken advantage of RNA-sequencing technologies to deconvolute various immune and stromal compartments present within a sample. Newman and colleagues released the first version of their tool, CIBSERSORT, which was built on microarray data in 2015(Newman et al., 2015). Following which, several groups have built algorithms and classifiers to characterize cell composition from complex tissue(Aran et al., 2017, Finotello et al., 2019, Sturm et al., 2020). Li and colleagues developed TIMER to probe tumor infiltration lymphocytes from the TCGA and their interactions with cancer cells(Li et al., 2016b). Becht et al developed MCP-counter to allow for the robust identification of multiple immune and stromal cell-based populations(Becht et al., 2016). Additionally, Newman and colleagues re-built CIBERSORT absolute to allow for an absolute quantification of immune and stromal cells as well as CIBERSORTx for implementation on single cell RNA-seq data(Chen et al., 2018a, Newman et al., 2019). In addition to characterizing specific cell populations, an absolute quantification of the cell populations allows us to understand relationships and interactions between different immune and stromal cell types and could be further correlated with tumor shrinkage and growth. These analyses could be performed in addition/in conjunction with typical transcriptomic workflows involving differential expression and gene set enrichment analysis.

1.9 Co-evolution of tumor genomics with immune landscape

Significant work through various groups have extended our understanding of genetic factors influencing tumor evolution(Nowell, 1976, Greaves and Maley, 2012). More recently, non-genomic features have been implicated in the evolution of cancer cells.

Epigenetic dysregulation through means such as promoter hypermethylation, altered enhancer activity changes in chromatin configuration can influence transcription and gene expression(Flavahan et al., 2017). Alterations in the cancer genome can induce transient changes in gene expression and regulate highly plastic gene expression networks. Dysregulation of the cancer epigenome has been implicated wherein increased global enhancer expression was depicted across multiple cancer types with enhancer activity being correlated with the fraction of the genome being affected by CNAs(Chen et al., 2018b). While preliminary work has been undertaken in the potential consequences of the

CNAs and its impact on chromatin state, the evolutionary context of epigenetic dysregulation still requires further work.

The cellular transcriptome has also been shown to be implicated through alternative splicing, alternative promoter usage, gene fusions and aberrant oncogenic signaling, to ultimately affect the genomic determinants of cancer(Kahles et al., 2018, Demircioglu et al., 2019, Group et al., 2020, Shiraishi et al., 2018). Through a variance component analysis, the PCAWG group found that CNAs was the most determinant genomic event influencing gene expression(Group et al., 2020). Additional work has revealed that RNA variants through editing enzymes also result in diversity within tumors that may ultimately impact protein function(Chen et al., 2013, Baysal et al., 2017).

Transcriptomic variation has been leveraged due to its ability to implement expression based biomarkers to predict clinical and therapeutic outcome (Carter et al., 2006). Multi-region sequencing studies help us further elucidate the ITH that may be caused due to gene expression through paired genomic and transcriptomic sequencing. Biswas and colleagues identified a dominant role of CNAs in influencing gene expression and the ITH described as a result of copy number events was strongly correlated to the ITH observed through gene expression(Biswas et al., 2019). Additionally, through further granular modalities like single cell sequencing, we can capture epigenetic and transcriptomic changes that may govern tumor evolution. Through work conducted in model organism, Lafave et al described the epigenetic and landscape of lung adenocarcinoma cells en-route to metastasis revealing important changes in chromatin state, loss of cellular identity and alterations in transcription factors(LaFave et al., 2020). Through this work, the authors were able to not only determine the stemness of cells undergoing metastasis, but also the diverse transcriptional and epigenetic landscapes that may influence the evolutionary trajectory of a tumor combined with their ability clinically correlate these findings to survival in human patients.

Given that a genomic silo of cancer evolution may not always be representative, remarkable breakthroughs in cancer immunotherapies have occurred as a result of disrupting signaling between cancer and immune cells. In the context of therapeutic intervention, particularly with the advent of ICB, it is now well known that various non-genetic factors can influence response and resistance to therapy. Given that a directed adaptive immune response is directed through a neoantigen specific genetic route, tumor cells undergo negative selection pressures to evade immune detection and elimination. Mutations in the B2-microglobulin gene (B2M), a component of the MHC, can disrupt antigen presentation in response to immune attack(Rooney et al., 2015, Zaretsky et al., 2016). In a seminal study conducted by McGranahan et al, HLA Loss of Heterozygosity (LOH) was detected in 40% of lung adenocarcinomas with a subclonal alteration in 65% of cases(McGranahan et al., 2017). CNA of non-HLA loci can also promote immune evasion through copy number loss of neoantigens capable of stimulation an effector T-cell response(Rosenthal et al., 2019). Multi-region sampling also went on to show immune-excluded tumor regions to be characterized through copy number amplifications of MYC target genes and WNT signaling in ovarian cancer(Jiménez-Sánchez et al., 2017).

Given that not all neoantigens may stimulate a functional T-cell response, substantial work by the TRACERx consortia characterized clonal neoantigens that must be targeted for effective response to ICB(McGranahan et al., 2016). In a multi-region study, Joshi et al described the spatial heterogeneity present in the TCR repertoire and correlated the number of T-cells clones found in tumor regions to the clonal mutations thus emphasizing the importance oof neoantigens in stimulating an early response in tumor evolution(Joshi et al., 2019). Moreover, the clonality of T-cells within a tumor has also been associated with improved response to anti PD-1 therapy in the context of metastatic melanoma(Tumeh et al., 2014, Roh et al., 2017). Important to note that clonal diversity of neoantigens can also influence the antitumor immune response. In particular, a mouse model of UV induced melanoma illustrated increased clonal diversity of a developing tumor is associated with evasion of an anticancer immune response(Wolf et al., 2019). This clonal diversity can provide negative selection pressures to a growing tumor and shape it's clonal composition.

1.10 Hypothesis and Specific Aims

Hypothesis: Integrative spatiotemporal approaches for biomarker discovery and immune-tumor evolution

While past studies have focused and succeeded on identifying genetic biomarkers and non-genetic/immune biomarkers to therapy response and tumor evolution, there is a lack of information about mechanistic links between genetic and immune biomarkers and how they work in combination with the TME to shape tumor evolution, and response and resistance to therapy. To this end, our studies focus on a multiplatform approach incorporating genomic data through whole exome sequencing or deep targeted panel sequencing, transcriptomic data through RNA-sequencing or targeted Nanostring gene expression profiling, epigenetic profiling through the EPIC 450K methylation array, immunological data through 12 marker IHC or mIF combined with TCR-sequencing along with proteomic approaches through the reverse phase protein assay. The genomic data allows us to unravel genetic somatic mutational and copy number events at play along with understating the mutational signatures that may govern response and tumor evolution in these contexts. The transcriptome allows us to additionally deconvolute the various immune and stromal components present along with identifying gene fusions that may also be candidate neoantigens.

I hypothesize that through this multimodality approach, we are able to paint an integrative picture that allows us to discern tumor biology and evolution, along with the characterization of the tumor-immune microenvironment and how they work in tandem to shape response and resistance to ICB. This work is performed in three sub-projects that aim to understand spatial heterogeneity in a single patient, genetic and immunological biomarkers for response and resistance to ICB, and to understand the co-evolution of tumors along with their immune contexture.

Specific Aims

Aim 1: To characterize spatial intratumor heterogeneity within metastatic melanoma and investigate cellular interaction within a tumor ecosystem by integrative analysis of multidimensional profiling

Aim 2: To delineate genetic and immunobiological biomarkers of response and resistance to ICB through longitudinal sampling in the context of metastatic sarcoma

Aim 3: To characterize the landscape of immune microenvironment of Barrett`s esophagus and high-grade dysplasia by multiplex immunofluorescence and TCR sequencing

Rationale

ITH has been recognized as a crucial challenge in cancer treatment. Sequencing of tumors in spatially distinct regions showed considerable amount of heterogeneity within tumors. Recently, single cell sequencing has been utilized to decipher genotypic and phenotypic states that exist within bulk tumors. However, previous approaches of multi-region sequencing and single cell sequencing required the process of tissue dissociation or micro-dissection, which led to loss of three-dimensional spatial information of a tumor. Through multidimensional profiling of a three-dimensional dissection of a single tumor, we can obtain rich information about the cellular ecosystem within a single tumor at an intermediate resolution between a bulk tumor and a single cell. Additionally, ITH has been implicated in lack of response to multiple therapeutic interventions. Given that response rates to ICB remain modest, there exists an unmet need to find biomarkers predictive of response to ICB and to better understand resistance mechanisms. Through longitudinal pre-treatment and early on-treatment sampling, we aim to discover biomarkers of response and resistance by identifying predictive signatures from the pre-treatment biopsies and better understand the dynamic changes that tumors undergo through therapy. Given that cancers exploit local environments composed of immune cells and stroma, we need to incorporate non-genetic/immune markers to characterize evolutionary changes within tumors. Through comprehensive immunophenotyping, we can uncover the immunologic compositions present in pre-tumorigenic stages and how this may play out in invasive carcinoma.

Chapter Two: Materials and Methods

Aim 1: To characterize spatial intratumor heterogeneity within metastatic melanoma and investigate cellular interaction within a tumor ecosystem by integrative analysis of multidimensional profiling

Biospecimen collection

Patient data, tumor samples, and matched peripheral blood leukocyte samples were obtained and used in accordance with research protocols approved by the local Institutional Review Board of the University of Texas MD Anderson Cancer Center. Biospecimens were retrieved, collected and analysed under UT MD Anderson Cancer Center Institutional Review Board approved protocols in accordance with the Declaration of Helsinki.

Sample processing – spatial intratumoral analysis

We developed a 3-dimensional model for processing a whole resected metastatic lesion. The lesion measured 2.5cm x 2.4cm x 1.5cm and was obtained from abdominal wall soft tissue.

Processing consisted of the following steps:

1) Following resection, the specimen was measured and oriented according to its largest diameter. Lateral (short axis, “left”/”right”) orientation was preserved by differential inking of the outside surface with red or blue ink.

2) The specimen was serially sectioned perpendicularly to its largest axis resulting in 8 slices of 2mm to 3 mm thickness. The cut surface was then painted prior to each cut to preserve the true orientation.

3) Alternate slices were submitted for FFPE (4 slices; odd numbered slices) or frozen (4 slices; even numbered slices) processing. FFPE slices were used for pathological assessment and immunohistochemistry analysis. Frozen slices were embedded in optimal cutting temperature compound (OCT) and used for DNA, RNA and protein extraction and downstream analyses.

4) For frozen sections, H&E staining was performed for histological quality control (QC). Frozen sections were further squared into a 0.2-0.4cm grid as shown in Fig. 1B and Extended Data Fig.3, generating a total of 67 sub-regions of tumor. Due to the variation in tumor shape throughout three-dimensional space, each tumor slice presented a distinct cross-sectional area and thus a unique grid was applied to each frozen tumor slice (slice *n*) and the immediately adjacent FFPE slices (slices *n-1* face B and *n+1* face A). Thus, whilst sub-region numbering generally proceeded bottom-to-top and right-to-left, specific slice sub-region numbering is not directly comparable between slices. Each piece was labeled and numbered.

5) Histopathological review for assessment of viable tumor, inflammatory infiltrate, necrosis and connective tissue of each sub-region piece was performed by a pathologist.

> 5.1) Designation of sub-regions as located at the tumor core or margin was performed by inspection of all regions annotated on SOX10-stained IHC slides to infer a volumetric estimate of tumor content and location. Three-dimensional variation throughout the frozen slices was accounted for by considering the immediately adjacent FFPE slices (i.e.: above AND below, when both were available) in order to arrive at a consensus call to categorize tumors as core or margin.
>
> 5.2) Immune infiltrates were evaluated using CD45-LCA positive cell density measured by digital image quantification using the Aperio ImageScope software. CD45+ density was then categorized as low, medium or high by binning into lower, middle, or upper tertiles considering all regions analyzed. Several additional factors required consideration before arriving at a final semi-quantitative categorization as having focal, low, moderate or high immune infiltration: between-slide variation in staining efficiency, spatial distribution of immune cell content (focal, broad, intra/peri-/extra-tumoral), and a consensus estimate for genomic frozen slices by considering the adjacent FFPE sections on both sides for which the higher immune content was assigned precedence.

6) Frozen section squares were submitted for dual DNA and RNA extraction, and for protein extraction.

7) DNA samples were submitted for T200 targeted sequencing (n=38), methylation (n=38), whole exome sequencing (n=6) and TCR sequencing (n=46). RNA samples were submitted for gene expression profiling by RNA sequencing (n=48) and TCR sequencing (n=46).

Sample processing – longitudinal time points

Archival formalin-fixed paraffin-embedded specimens from the lung metastasis (pre-treatment) and right gluteal mass (post-PD-1 inhibitor) were obtained from the institutional pathology department and utilized for tumor evaluation by a Pathologist, DNA/RNA extraction, and immunohistochemistry as described below.

A single cell suspension was generated from the post-PD-1 inhibitor time point tumor by gentle mechanical digestion of fresh tumor material, followed by enzymatic digestion with 2mg/mL collagenase A (Roche, Cat. No. 11 088 793 001) and DNase I (Roche, Cat. No. 11 284 932 001) in serum-free RPMI-1640 (Gibco, Cat. No. 11875119) for 1 hour at 37°C with agitation. The cryopreserved single cell suspension was thawed and purified for viable cells by negative selection using the MACS Dead Cell Removal Kit (Cat. No. 130-090-101) and a LS Column (Cat. No. 130-042-401) on the QuadroMACS Separator (Cat. No. 130-090-976, all Miltenyi Biotec). The purified single cell suspension was stained with SYTOX blue dead cell stain (Cat. No. S34857, Thermofisher Scientific), anti-human CD45 PerCP-Cy5.5 (clone HI30, Cat. No. 564105, BD Biosciences), anti-human CD3 FITC (clone SK7, Cat. No. 340542, BD Biosciences) and anti-human melanoma (MCSP) APC (Cat. No. 130-091-252, Miltenyi Biotec) prior to cell sorting on a BD FACSAria III flow cytometer (BD Biosciences) to enrich for a live T cell fraction (CD45+CD3+) and live tumor fraction (CD45-MCSP+)(Campoli et al., 2004).

Nucleic acid extraction

DNA and RNA isolation were performed using the AllPrep DNA/RNA/miRNA Universal kit (Cat. No. 80224, Qiagen) for fresh frozen samples and the AllPrep DNA/RNA FFPE kit (Cat. No. 80234, Qiagen) for FFPE samples according to the manufacturer's instructions. Tumor viability of 80% estimated from corresponding IHC samples was set as a minimum threshold for genomic analyses. Samples with DNA integrity numbers (DIN) greater than 7 were used for targeted panel sequencing and EPIC array methylation profiling. RNAseq was performed on samples with a minimum RNA integrity number (RIN) of 5.5 except for two cases (6A10 and 8A3) with RINs greater than 3. A minimum of 700ng of RNA were required for all samples undergoing RNAseq.

Cancer gene panel DNA sequencing

Samples with cancer cell purity greater than 80% based on pathologic assessment were used for cancer gene panel DNA sequencing. Mean sequencing coverage was 861x in tumors and 1,314x in germline samples. Paired-end reads in FASTQ format were generated by the Illumina pipeline and aligned to the reference human genome hg19 build using the Burrows-Wheeler Alignment Tool (BWA, v0.7.5) with default settings(Li and Durbin, 2009). Aligned reads were further processed using GATK with best practices for removing duplicates, indel removal and recalibration(Van der Auwera et al., 2013).

To detect potential single nucleotide variants, MuTect (v1.1.4) was used with default parameters including a variant allele frequency of >10% in tumor DNA, variants present on both strands, a high read count of tumor DNA and the removal of positions listed in dbSNP 129(Cibulskis et al., 2013). Pindel (v0.2.4) was used to identify small insertions and deletions(Ye et al., 2009). Copy number was called using Sequenza (v2.1.2)(Favero et al., 2015). Tumor purities and ploidies were calculated from Sequenza calls using the sequencing data with default parameters. The content of the cancer gene panel is given in Supplementary Data 2.

Whole exome sequencing

Exome sequencing data was generated using methods as previously described, including library preparation using the Agilent SureSelect XT Target Enrichment protocol (#5190-8646) prior to sequencing on an Illumina HiSeq 2000/2500 v3 system using 76bp paired-end reads[3]. Raw sequencing data was then processed using Saturn V, the next generation sequencing data processing and analysis pipeline developed by the Department of Genomic Medicine at the UT MD Anderson Cancer Center. BCL files were pre-processed using CASAVA (Consensus Assessment of Sequence and Variation, v1.8.2) for demultiplexing and converting to FASTQ. The files were aligned using the Burrows-Wheeler Alignment Tool (BWA, v0.7.5) using the hg19 reference genome build(Li and Durbin, 2009). Picard (v2.5.0) was used to convert SAM files to BAM files and remove duplicates. BAM files were realigned and recalibrated using GATK. Mean coverage was 181X for tumors and 81X for matched germline DNA. MuTect and Pindel were used to identify somatic point mutations and small insertions and deletions respectively(Cibulskis et al., 2013, Ye et al., 2009). Somatic mutations in HLA genes were called using POLYSOLVER (v1.0)(Shukla et al., 2015). Data was annotated by ANNOVAR (v20180118) using the NCBI Reference Sequence Database(Wang et al., 2010).

Phylogenetic tree construction

Mutations that passed our WES Mutect filtering criteria were considered for the purpose of constructing phylogenetic trees. A tumor power of 0.8 was used to filter mutations that reflected the power to detect a mutation at 0.30 allelic fraction. Trees were built using binary presence/absence matrices built from the distribution of mutations within the tumor samples. A representative sub-region sample was chosen from the four frozen slices of the on-PD-1 inhibitor tumor (**Figure 1B**), thus producing a total set of 6 samples being compared (pre-treatment x1, on-PD-1 inhibitor x4, post-PD-1 inhibitor x1). Since the most-inferior section (section 8) of the on-PD-1i tumor was largely necrotic, we sampled two regions from the preceding frozen section (final samples: 2A2, 4A11, 6A3, 6A16). In order to compare the three timepoints at bulk tumor level, we combined the multiple on-PD-1 inhibitor tumor sub-samples. The R Bioconductor package phangorn (v2.5.5) was utilized to compute the hamming distance under the neighbor joining tree method, and generated unrooted trees(Schliep, 2011). The distance was computed after 100 bootstrap iterations with the bootstrap value reflected on the branch. We identified somatic mutations using MuTect and both DNA copy-number changes and tumor purity using Sequenza(Cibulskis et al., 2013, Favero et al., 2015). We estimated the cancer cell fractions identified with a particular mutation, accounting for tumor purity, using PyClone (v0.13.0), which was used to infer cancer cell fractions and assign clonal clusters(Roth et al., 2014). Further clonal evolution was evaluated through ClonEvol (v0.1)(Dang et al., 2017).

RNA sequencing

Paired-end transcriptome reads were aligned using TopHat2, to the UCSC hg19 reference genome(Kim et al., 2013). Gene read counts were generated using Htseq-count(Anders et al., 2015). Bioconductor R package DESeq2 (v1.24.0) was used to normalize the read counts and for downstream analysis, vsn (v3.52.0) was used for variance stabilization(Love et al., 2014, Lin et al., 2008). Differential gene expression analysis between the heavy and low infiltrated samples were performed after adjusting for variation in tumor content due to core or margin location by excluding samples with very low tumor purity (i.e.: sampling predominantly surrounding stroma). Genes with significant changes in expression were assessed by including the top 1,000 most variant genes after performing median absolute deviation. The genes were clustered based on Euclidean distance and the samples based on Pearson correlation with complete linkage. DAVID (v6.8) online functional annotation tools, showed immune-regulated pathways from the most variable genes with an FDR cutoff of 1%(Huang et al., 2007). Pathway analysis was performed on the most differentially expressed genes. ssGSEA was run through GSVA (v1.32.0). Pathway level enrichment was run on the output of DESeq2 for each condition through DOSE (v3.10.2) and ClusterProfler (v3.12.0)(Hänzelmann et al., 2013, Yu et al., 2015, Yu et al., 2012, Yu, 2018). Cell type-specific gene expression was evaluated using immune and melanoma-specific markers. All heatmaps were constructed using ComplexHeatmap (v2.0.0)(Gu et al., 2016). ESTIMATE (v1.0.13) was used to detect tumor purity and the presence of infiltrating stromal/immune cells in tumor tissues using gene expression data(Yoshihara et al., 2013). Sample distances were visualized using Circlize (v0.4.8) and Plotly (v4.9.0)(Gu et al., 2014, Inc). Consensus clustering was performed using the Consensus Cluster Plus (v1.48.0)(Wilkerson and Hayes, 2010). Hierarchical clustering was used to group the 48 transcriptomic samples with a maximum cluster count of 20. The delta area plot was used to determine the relative increase in consensus clustering of samples within a given cluster, and to determine a value of k beyond which no appreciable increase was achieved. The tracking plot was used to depicts which samples were allocated to which cluster, and lack of correlation with geographic location of the sample. For validation across public datasets, we used RPKM/FPKM values to build linear models of expression, given the design matrix of binary responders and non-responders using response classifications based on RECIST v1.1 criteria provided with each paper. In order to compare the most representative samples across all datasets, we restricted our analysis from the Riaz dataset to include only the on-treatment time point samples. Ribosomal L and S (RPL and RPS respectively) genes were excluded from downstream analysis, following which the expression values were log-transformed. Linear modelling was performed to fit a model of expression values for each gene, given the binary response status(Ritchie et al., 2015). Empirical Bayes moderation was then carried out by utilizing information across all genes to obtain precise estimates of gene-wise variability(Smyth, 2004). The differentially expressed genes were characterized by FDR-adjusted p-values of less than 0.05 and log fold change in the positive direction for the non-responders. Downstream GO and KEGG pathway enrichment was performed using ClusterProfiler based on a logFC greater than 7 and an adjusted P-value of less than 1e-5(Yu et al., 2012).

Methylation analysis

We studied methylation levels through the pipeline integrated into the Bioconductor package, ChAMP (v2.14.0) using R(Morris et al., 2014). The data was imported as raw idat files and a variety of quality control plots were evaluated. The data was imported and filtered based on detection p-value (<0.01) and probes with less than 3 beads in 5% of samples per probes. Non-CPG probes and SNP-related probes were then removed. Finally, multi-hit probes and probes located on the X chromosome were filtered out.

Type II probe normalization was performed using BMIQ (v1.5)(Teschendorff et al., 2013). We also assessed the number and nature of significant components of variation by using singular value decomposition to look at batch effects and COMBAT (v3.32.1) for batch correction(Johnson et al., 2007). For the identification of Differentially Methylated Regions (DMRs), the BumpHunter (v1.26.0) method was used to identify extended segments of the genome that show quantitative alteration in DNA methylation levels(Jaffe et al., 2012).

We used the mean normalized value of beta to collapse probe level data to gene-wise data for integrative analysis. Genes on chromosome 7 were used to compare the correlation of expression and methylation. The gene level beta values for the most variable genes from expression were used to compare with copy number and expression data.

TCR sequencing

DNA-based: DNA sequencing of the variable region of the beta chain of the T cell receptor was performed by ImmunoSeq (Adaptive Biotechnologies, Seattle, WA)(Robins et al., 2009a, Carlson et al., 2013). RNA-based: RNA sequencing of the variable chain of the T cell receptor was performed using Immunoverse TCR (ArcherDX, Boulder, CO)(Chasseriau et al., 2004). The TCR clonality was used for linear regression. The top 5 clones with respect to DNA and RNA clonal fractions were calculated for each sample, and their residuals from the line of best fit were used as a measure of activation of TCR clones. Spearman’s rank correlation was performed between the residuals and the clonality measured from DNA-based TCR sequencing. The top clones were analyzed using both platforms independently and then concurrently. Plot3D was used to map the potential trajectory of the most dominant clones across all regions(Soetart, 2013). TCR statistics were computed using the R package tcR (v2.2.4)(Nazarov et al., 2015).

Stimulation of ZDNNC17 p.H507Y neoantigen specific T cells

To evaluate the potential *in vitro* immunogenicity of the ZDNNC17 p.H507Y neoantigen, we synthesized 12 overlapping candidate 9-mer peptides (ELIM Biopharmaceuticals, Inc, Hayward, CA), spanning the neoantigenic point mutation, and used these peptides to elicit T cell responses from HLA-A*0301 donor PBMC. Peptides were dissolved in 1× PBS at a concentration of 10 mg/mL. HLA-A*0301-transfected K562 (A3-K562) cells pulsed with 2 μg/mL of each ZDNNC17 p.H507Y peptide were used as antigen presenting cells for stimulating CD8+ T cells from each of two HLA-A*0301 donors using methods previously established in our lab[81]. For each peptide stimulation, irradiated (8000 rads) peptide pulsed K562 cells were co-cultured in a 48 well plate with 1 million PBMC from each donor at a ratio of 1:20 in RPMI-1640 containing 25 mM HEPES, 2 mM L-glutamine, 10% human AB serum (CTL medium) and β2-microglobulin (3 μg/mL)(Pollack et al., 2014). Three rounds of PBMC stimulation with peptide pulsed A3-K562 were performed at 7-day intervals. During the first stimulation, IL-21 (30 ng/mL; Peprotech, Rocky Hill, NJ, USA) was included in the cell culture medium and during the second and third stimulation cycle, IL-21 (30 ng/mL), IL-2 (10 ng/mL; Bayer, Terrytown, NY, USA) and IL-7 (5 ng/mL; R&D Systems, Minneapolis, MN, USA) were added to the growth medium as previously described(Pollack et al., 2014, Li et al., 2005). Controls included PBMC co-cultured with non-pulsed A3-K562 cells and A3-K562 pulsed with a pool of all 12 peptides.

After three rounds of stimulation, an aliquot of 100,000 cells from each well was co-cultured overnight with peptide pulsed K562 cells at a ratio of 10:1 to assay for antigen specific T cells using flow cytometry-based intracellular IFN-γ production assay(Fan et al., 2018). The cells were cultured in the presence of the intracellular protein transport inhibitor Brefeldin A (Thermo Fisher, USA). After overnight culture, cells were washed and then stained with CD8-APC (Clone K1, BioLegend, San Diego, CA) for 20 min. Intracellular staining for IFN-γ-PE (Clone B27, BioLegend, San Diego, CA) was performed according to the manufacturer's protocol. After staining, cells were resuspended in 100μL of FACS buffer and data was acquired using a NovoCyte Flow Cytometer (ACEA Biosciences, San Diego, CA). Data were analyzed using FlowJo™ software (Tree Star, Ashland, OR, USA). The percentage of background IFN-γ positive cells was determined by the response of PBMC co-cultured with non-pulsed K562 cells and peptide-specific responses were registered as positive if the proportion of T cells producing IFN-γ in response to stimulation with ZDNNC17 p.H507Y derived peptide was ≥2-fold higher than the background proportion of IFN-γ+ CD8 T cells(Doubrovina et al., 2012). PMA + ionomycin treatment of PBMC served as a positive control for the IFN- γ production assay.

Single cell sequencing

Three technical replicates of FACS-sorted T cells (CD45+CD3+) and one replicate of FACS-sorted tumor cells (MCSP+) were loaded to a targeted 10,000 cells per lane on the 10X Genomics Chromium Controler with the single cell 5' Immune Repertoire and Gene Expression profiling kit. In total, we loaded ~30,000 individual tumor infiltrating lymphocytes (TILs) and ~10,000 melanoma cells on the 10X platform (10X Genomics, CA, USA). Reverse transcription, TCR enrichment, and library preparations were performed according to the 10X Genomics 5' V(D)J protocol revision C. Transcriptome libraries were pooled and sequenced on the Illumina NovaSeq 6000 S2 flow cell with 26 R1, 8 i7, and 91 R2 cycles respectively. The TCR libraries were pooled and sequenced on the Illumina MiSeq V2 150 cycles paired-end. Single cell transcriptomic and TCR data was processed with the 10X Genomics Cell Ranger Pipeline version 2.2.0 with the software-provided GRCh38 reference transcriptomes(Zheng et al., 2017). After quality control, there was RNAseq profile data available from 6267 immune and 4303 melanoma cells. Downstream processing and visualization was encompassed through Seurat and tSNE plots(Satija et al., 2015, Van Der Maaten and Hinton, 2008).

Neoantigen prediction

HLA Class I neoepitopes were predicted for each sample and affinity was predicted for the predicted peptides using NetMHCpan (v2.8)(Karosiene et al., 2013). Patient HLA-A, HLA-B and HLA-C variants were identified using ATHLATES (v2014_04_26)(Liu et al., 2013). All possible 9- to 11-mer peptides flanking a nonsynonymous exonic mutation were generated and binding affinity was predicted based on patient HLA and compared to the wild-type normal peptide counterpart from NetMHCpan(Karosiene et al., 2013). MuTect calls were filtered using tumor count >30, normal count >10, tumor VAF >0.05, normal VAF <0.01 and tumor power >0.8(Cibulskis et al., 2013). Additionally, a FPKM count >1 and an alternate allele count >4 was leveraged from the RNA-seq data. Mutated peptides with predicted IC_{50} <500nM were considered to be predicted neoantigens.

Copy number alteration (CNA) analysis

Sequenza was used to obtain copy number segments of $\log_2$ copy ratios for tumor samples(Favero et al., 2015). CNTools (v1.24.0) was used to identify copy number gain/loss events at $\log_2$ thresholds of 0.6(Jianhua, 2018). The burden of copy number gain or loss was extrapolated from the total number of genes with copy number events in each sample. ExomeCNV (v1.4) was used to calculate the $\log_2$ copy ratios(Sathirapongsasuti et al., 2011). For the TCGA dataset, processed segmented values were used. Whole chromosome 7 events were characterized as log segmented mean values greater 0.3 and covering greater than 70% of the length of the chromosome.

Reverse-phase protein array

Frozen tumor sub-region samples were processed for RPPA analysis in the UT MD Anderson Cancer Center RPPA Core Facility using previously described methods (https://www.mdanderson.org/research/research-resources/core-facilities/functional-proteomics-rppa-core/rppa-process.html). Briefly, tumor lysates were prepared in RPPA lysis buffer, serially diluted and printed onto nitrocellulose-coated slides prior to being probed with ~300 validated primary antibodies and detection with biotinylated secondary antibodies specific for the primary antibody species. Signal amplification and visualization by a DAB colorimetric reaction was performed prior to slide scanning and quantification using the Array-Pro Analyzer software (MediaCybernetics), relative protein level estimation using SuperCurve GUI, correction for spatial bias and QC check of each slide.

Ba/F3 transformation assay

Transforming potential of JAK1 wild-type (WT) and variants were assayed in IL-3-dependent Ba/F3 cell model as described previously(Ng et al., 2018). Briefly, lentivirus vector of JAK1 WT and variants were generated with pHAGE-PURO backbone by High-Throughput Mutagenesis and Molecular Barcoding (HiTMMoB) technique as described previously(Tsang et al., 2016). All clones were full-length validated by Sanger sequencing. Virus were produced by transfecting LentiX-293T cells (Clontech) with pHAGE-PURO backbone and two packaging plasmids (psPAX2 and pMD2.G) and were harvested by filtration through 0.45μm PVDF filter 3 days post-transfection. Six hundred thousand Ba/F3 cells were transduced by spinoculation at 1000g for 3 hours in the presence of polybrene (EMD Millipore; final concentration: 8 μg/ml), and then incubated in the assay medium without IL-3 (Advanced RPMI 1640 with 1x GlutaMAX and 5% FBS; Thermo Fisher Scientific) for 2 weeks. Cell viability was measured by CellTiter-Glo luminescent cell viability assay (Promega).

Immunohistochemistry

Immunohistochemistry (IHC) was performed on each of the 4 FFPE sections using an automated stainer (Leica Bond Max, Leica Biosystems) using primary antibodies against SOX10 (polyclonal, 1:50, Cell Marque, Cat. No. 383A-7), CD45-LCA (clones 2B11 + PD7/26, 1:300, Dako, Cat. No. M0701), CD45-RO (clone UCHL1, undiluted, Leica Biosystems, Cat. No. PA0146), CD4 (clone 4B12, 1:80, Leica Biosystems, Cat. No. NCL-L-CD4-368), CD8α (clone C8/144B, 1:25, ThermoScientific, Cat. No. MA5-13473), Granzyme B (clone GrB-7, 1:25, ThermoScientific, Cat. No. MA1-35461), FoxP3 (clone 206D, 1:50, BioLegend, Cat. No. 320102), LAG-3 (clone D2G4O, 1:100, Cell Signaling Technology, Cat. No. 15372), PD-1 (clone EPR4877(2), 1:250, Abcam, Cat. No. ab137132), PD-L1 (clone E1L3N, 1:100, Cell Signaling Technology, Cat. No. 13684), PAX5 (clone 1EW, undiluted, Leica Biosystems, Cat. No. PA0552), CD68 (clone PG-M1, 1:450, Dako, Cat. No. M0876), CD57 (clone HNK1/Leu-7, 1:250, Abcam, Cat. No. ab187274), and phospho-p44/42(Erk1/2)(Thr202/Tyr204) (clone D13.14.4E, 1:300, Cell Signaling Technology, Cat. No. 4370). Slides were counter-stained with hematoxylin, scanned using an Aperio slide scanner (Aperio AT Turbo, Leica Biosystems) and digitized images analyzed using the Aperio ImageScope software (Aperio - Leica Biosystems). Three-dimensional reconstruction re-connecting the frozen and FFPE slices in a sequential order was performed based on the documented inter-slice relationships and histological findings. The IHC slices were gridded into smaller pieces in the ImageScope software to match the gridding of frozen sections, and the results were obtained for each sub-region. IHC-derived cell subset results were quantified as the number of positive-staining cells for each antibody per mm^2, using custom-tuned algorithms based on nuclear v9, membrane v9, or cytoplasmic v1 algorithms as appropriate for the staining pattern of each antibody.

Data analysis and statistical considerations

Statistical analyses were performed using R v3.5.0(R Core Team, 2019). Analysis packages and tools used are described in the relevant methods sections. Statistical tests included Welch's two sample t-test and Spearman's rank correlation with the Benjamini-Hochberg correction for an adjusted p-value threshold of 0.05. The R package plot3D and Plotly were used to map sequencing-derived data to spatial locations(Soetart, 2013, Inc). Data were parsed and organized through R packages tidyr, reshape2 and dplyr(Wickham et al., 2017b, Wickham, 2017, Wickham et al., 2017a). Dendograms and tanglegrams were constructed using dendextend(Galili, 2015). Plotting was done through ggplot2 and ggrepel(Wickham, 2011a, Slowikowski et al., 2018).

Aim 2: To delineate genetic and immunobiological biomarkers of response and resistance to ICB through longitudinal sampling in the context of metastatic sarcoma

Biospecimen collection

Patient data, tumor samples, and matched peripheral blood leukocyte samples were obtained and used in accordance with research protocols approved by the local Institutional Review Board of the University of Texas MD Anderson Cancer Center. Biospecimens were retrieved, collected, and analyzed under UT MD Anderson Cancer Center Institutional Review Board-approved protocols in accordance with the Declaration of Helsinki.

Response Criteria

Patient responses were coded using irRC and irRECIST with patients deriving clinical benefit based on a complete response of all lesions with no new lesions (CR), partial response as decrease in tumor burden > 50% from baseline (PR) or stable disease, with no progression beyond 25% relative to baseline (SD), and patients who did not receive clinical benefit with progressive disease > 25% relative to nadir (PD). All molecular analysis was conducted based on best response criteria at the 12-week time point based on irRC criteria.

Nucleic acid extraction

DNA and RNA isolation were performed using the AllPrep DNA/RNA/miRNA Universal kit (catalog number 80224, Qiagen) for fresh-frozen samples according to the manufacturer's instructions.

Whole exome sequencing

Exome sequencing data were generated using methods as previously described, including library preparation using the Agilent SureSelect XT Target Enrichment protocol (#5190-8646) prior to sequencing on an Illumina HiSeq 2000/2500 v3 system using 76 bp paired-end reads and the Illumina NovaSeq 6000(Mitra et al., 2020). Sequencing data was processed through the Saturn V pipeline maintained by the department of Genomic Medicine at the UT MD Anderson Cancer Center. Briefly, BCL files were processed through the implementation of CASAVA (Consensus Assessment of Sequence and Variation, v1.8.2) for demultiplexing and converting to FASTQ. This was followed by alignment using BWA (v0.75) with the hg19 reference genome build(Li and Durbin, 2009). Duplicates were removed and BAM files were constructed from SAM files using Picard. BAM files were then realigned and calibrated through the implementation of GATK. Mean target coverage for tumor samples was at 200x while target coverage for PBMC at 100x was achieved. MuTect and Pindel were used to identify somatic point mutations and small insertions and deletions, respectively(Cibulskis et al., 2013, Ye et al., 2009). Somatic mutations in HLA genes were called using POLYSOLVER (v1.0)(Shukla et al., 2015). Data were annotated by ANNOVAR (v20180118) using the NCBI Reference Sequence Database(Wang et al., 2010). Copy number events were called using Sequenza and the cghMCR package implemented in R(Favero et al., 2015, Zhang and Feng). Tumor purities and ploidies were calculated from the genomic data using the default parameters from Sequenza(Favero et al., 2015). GISTIC was used to identify recurrent copy number events present in responders and non-responders to ICB at the pre-treatment timepoint. Arm level amplifications and deletions were computed using default parameters at an FDR threshold of 0.05.

Mutational Signatures

De novo mutational signature discovery resulted in an optimal solution of three signatures that most contributed towards the mutational processes in the cohort. We then attributed all the mutational processes based on established COSMIC signature profiles and found the highest contributions from SBS1 (26%), SBS5 (10.5%), SBS15 (7.7%) and SBS10b (5.3%). While SBS1 and SBS5 represent spontaneous deamination of 5-methylcytosine or a clock-like signature known to occur over time, SBS15 represents defective DNA mismatch repair and SBS10b occurs as a result of a POLE mutations and have not been characterized in sarcoma before.

Copy Number Alterations

Sequenza was used to obtain copy number segments of log2 copy ratios for tumor samples(Favero et al., 2015). CNTools (v1.24.0) was used to identify copy number gain/loss events at log2 thresholds of 0.6(Jianhua, 2018). The burden of copy number gain or loss was extrapolated from the total number of genes with copy number events in each sample. ExomeCNV (v1.4) was used to calculate the log2 copy ratios(Sathirapongsasuti et al., 2011). Recurrent copy number events were computed using two modalities: GISTIC and the cghMCR package implemented in R(Favero et al., 2015, Zhang and Feng, Mermel et al., 2011). Tumor purities and ploidies were calculated from the genomic data using the default parameters from Sequenza(Favero et al., 2015). GISTIC was used to identify recurrent copy number events present in responders and non-responders to ICB at the pre-treatment timepoint. Arm level amplifications and deletions were computed using default parameters at an FDR threshold of 0.05.

RNA sequencing

Paired-end transcriptome reads were aligned using TopHat2, to the UCSC hg19 reference genome(Kim et al., 2013). Gene read counts were generated using Htseq-count(Anders et al., 2015). Bioconductor R package DESeq2 (v1.24.0) was used to normalize the read counts and for downstream analysis, vsn (v3.52.0) was used for variance stabilization(Love et al., 2014, Lin et al., 2008). Differential gene expression analysis between responders and non-responders as defined as patients with partial responses and those with progressive disease was performed at the pre-treatment time point.

Consensus clustering was performed on the pre-treatment samples. A max iteration of K was defined at 10, using a 1000 fold repetition with the Manhattan distance being utilized using the pam clustering algorithm(Wilkerson and Hayes, 2010). The delta area plot was used to determine a value of k (=3) beyond which no appreciable increase in heterogeneity was observed. All heatmaps were constructed using ComplexHeatmap (v2.0.0)(Gu et al., 2016). ESTIMATE (v1.0.13) was used to detect tumor purity and the presence of infiltrating stromal/immune cells in tumor tissues using gene expression data(Yoshihara et al., 2013). Consensus clustering was performed using the Consensus Cluster Plus (v1.48.0)(Wilkerson and Hayes, 2010). Hierarchical clustering was used to group the pre and on-treatment samples with a maximum cluster count of 10. The delta area plot was used to determine the relative increase in consensus clustering of samples within a given cluster and to determine a value of k beyond which no appreciable increase was achieved. A suite of immune deconvolution tools were employed in the pre-treatment setting to call consensus based calls of immune infiltrate proportions including CIBERSORT, MCP-counter, EPIC, x-Cell, quantiSeq, TIMER and immundeconv(Newman et al., 2015, Becht et al., 2016, Aran et al., 2017, Li et al., 2016b, Racle et al., 2017, Finotello et al., 2019, Sturm et al., 2020). A consensus of these calls was then used to estimate proportions of monocytes, neutrophils, B-cells, T-cells, NK-cells, dendritic cells, endothelial cells and fibroblasts. The consensus of these immune-cell compositions was then consensus clustered based on the optimal k of the transcriptomic data to result in immune cold, immune intermediate and immune hot clusters. Similar analysis for immune deconvolution was performed at the on-treatment time point. Sankey plots were constructed to depict change in immune clusters between the pre-treatment and on-treatment time point for patients using ggalluvial(Brunson, 2020).

LIONESS or the linear interpolation to obtain network estimates for single samples was used to reconstruct single-sample networks(Kuijjer et al., 2019). Briefly the workflow included subsetting the top 1000 most variable genes present in the transcriptome at the pre-treatment time-point. Following which a design model matrix was built to model response in the cohort. We then calculated differences in the condition specific matrices between responders and non-responders and converted these adjacency matrices to edge lists. The single sample networks were then computed based on co-expression using LIONESS. Limma analysis was then used to select the top 100 most differentially expressed gene edges(Ritchie et al., 2015). Coefficient edges were then colored based on the weights of these edges between responders and non-responders while nodes were colored using based on the t-statistic from differential expression analysis.

For validation across melanoma public datasets, we used reads/fragments per kilobase of transcript per million mapped reads (FPKM) values to build linear models of expression, given the design matrix of binary responders and non-responders using response classifications based on RECIST v1.1 criteria provided with each paper. For each of the publicly available melanoma datasets, MCP-counter was used to ascertain immune cell proportions using default parameters(Becht et al., 2016). Paired histogram, density and scatter plots were constructed using GGally(Barret Schloerke, 2021).

Nanostring Analysis

RNA from the SARC028 cohort was analyzed using the nCounter Technology (Nanostring Technologies) as per the manufacturer's protocol. Filtering was performed utilizing quality assurance checking, normalization and batch effects adjustments of Nanostring data using Nanostringr(Talhouk et al., 2016). NanostringNorm was used for data normalization and for filtering out samples with failed flagged values of quality check(Waggott et al., 2012). Downstream computation of immune cell infiltrate was performed using a modified version of MCP-counter(Becht et al., 2016).

B-cell Receptor repertoire inference

A modified version of the TRUST algorithm was executed to quantify the B-cell immunoglobulin hypervariable regions from the RNA-seq data(Hu et al., 2019). CDR3 sequences of the B-cell Immunoglobulin Heavy (IGH) chain and Immunoglobulin Light (IGL) chain were assessed to infer the BCR clonotypes. In order to use the most robust data, in frame productive sequences were used for downstream analysis along with filtering out samples with low template counts. Following filtering, the number of BCR clonotypes detected in each sample was normalized based on sequencing depth. The immunarch package was used to quantify Simpson clonality and Simpson diversity(Team, 2019). Largely dominant IGH and IGL BCR clonotypes were computed using templates that occupied greater than 1% of the entire repertoire for any given sample.

Immunohistochemistry staining for CD20 and CD56

We performed singlet chromogenic immunohistochemistry to detect CD20+ and CD56+ positive cells. We used Leica Bond Max (Leica Bio systems Nussloch GmbH) automated strainer, the protocol is briefly described: four-µm formalin-fixed paraffin embedded tumor tissue sections were deparaffinized and rehydrated following the Leica Bond protocol. Antigen retrieval was performed with Bond ER Solution #2 ((Leica Bio systems, equivalent to EDTA based buffer, pH 9.0). The primary antibody (CD20, clone L-26, Agilent Dako, Cat#M0755, dilution 1:1400; CD56, clone 123C3, DAKO, Cat#M7304) was incubated for 15 minutes at room temperature and detected using the Bond Polymer Refine Detection kit with DAB as chromogen. The slides were counterstained with hematoxylin, and then they were dehydrated and cover-slipped for image analysis.

Immunohistochemistry stained slides were scanned using Aperio AT2 scanner (Leica Biosystem) at 20x, the images were visualized and evaluated using image analysis software (Halo software, Indicia Labs). A pathologist trained in image analysis selected the viable tumor tissue area (areas of necrosis or artifacts were excluded for analysis), then the cytonuclear algorithm v2.0.5 was used to identify CD20+ and CD56+ positive cells. We reported the results as cell densities (n/mm2) using the following formula: number of positive cells for each biomarker divided by the tissue area (mm2) analyzed.

Multiplex immunofluorescence staining

Multiplex immunofluorescence (mIF) was conducted on pre-treatment and on-treatment samples. Briefly, four micrometer-thick formalin fixed, paraffin embedded sections were stained using an automated staining system (BOND-RX; Leica Biosystems, Buffalo Grove, IL) on two panels containing the following markers: CD3, CD8, PD-1, PD-L1, CD68, Granzyme B, CD45 RO and FOXP3. Pre and on-treatment samples were then immunophenotyped based on the co-localization of these markers including: T lymphocytes [(CD3+CD8+)+(CD3+PD-1+)+(CD3+CD8+PD-1+)+(CD3+)], Cytotoxic T cells [(CD3+CD8+)+(CD3+CD8+PD-1+)], T cells antigen-experienced [(CD3+PD-1+)+(CD3+CD8+PD-1+)], Cytotoxic T cells antigen-experienced (CD3+CD8+PD-1+), Macrophages [(CD68+PD-L1-)+(CD68+PD-L1+)] and Percentage of macrophages PD-L1+ (CD68+PD-L1+) for Panel 1. Panel 2 was characterized using: T lymphocytes[(CD3+CD8)+(CD3+CD8+G&B+)+(CD3+CD8+CD45RO+)+(CD3+FOXP3+) +(CD3+CD8+FOXP3+)+(CD3+)], Cytotoxic T cells activated (CD3+CD8+G&B+), Effector/memory cytotoxic T cells (CD3+CD8+CD45RO+) and Regulatory T cells [(CD3+FOXP3+)-(CD3+CD8+FOXP3+)]. Immunofluorescence was performed simultaneously for each fluorochrome to make a spectral library and stained slides were scanned using the multispectral microscope, Vectra 3.0.3 imaging system (Akoya Biosciences/PerkinElmer PerkinElmer), under fluorescence conditions at 40x magnification . Analysis was performed based on regions of interest determined by the pathologist. Final immune cell densities were reported as the average density of cells per mm2. Samples without tumor areas or severe necrosis accompanied by lack of viable tumor cells were excluded from analysis.

TCR Sequencing

DNA sequencing of the variable region of the Beta chain was performed by Immunoseq (Adaptive Biotechnologies, WA)(Robins et al., 2009a, Carlson et al., 2013). Immunarch was used to compute the clonality and diversity present within our cohort(Team, 2019). The top clonal proportions for each sample was calculated following the clonal proportions defined as a highly expanded clone (>0.1 of repertoire), largely expanded clones ($0.001 < x < 0.1$), medium ($0.0001 < x < 0.001$) and rare (< 0.00001) clones. Using the ImmunoSeq analyzer, we estimated the levels of Simpson clonality, maximum productive frequency (frequency of the most productively forming clone), Simpson diversity, Simpson Evenness and Shannon diversity. Simpson diversity was defined as the complement of the Simpsons' Dominance (1-Dominance). Simpsons' Evenness is calculated from Simpson's D by dividing by the richness of the repertoire ((1-Dominance/Richness). Maximum productive frequency is computed using the highest frequency of the most dominant clone in a sample accounting for only in frame sequences that would produce a functional template.

Aim 3: To characterize the landscape of immune microenvironment of Barrett`s esophagus and high-grade dysplasia by multiplex immunofluorescence and TCR sequencing

Cohort Selection

We searched our institutional databases and identified 10 patients of EAC who underwent esophagogastrectomy or endoscopy-based resection without preoperative therapy. All patients had BE on esophagogastroduodenoscopy and on pathology examination.

Multiplex immunofluorescence (mIF) staining

Multiplex immunofluorescence staining was performed using similar methods that have been previously described and validated(Parra et al., 2017). Briefly, four micrometer-thick formalin fixed, paraffin embedded sections were stained using an automated staining system (BOND-RX; Leica Biosystems, Buffalo Grove, IL) and a panel of antibodies against: pan cytokeratin, CD3, CD8, CD45RO, Granzyme B, and FOXP3 (Supplemental table-1). All the markers were stained in sequence using their respective fluorophore in the Opal 7-color fIHC kit (catalogue #NEL797001KT; Akoya Biosciences/PerkinElmer, Waltham, MA) (Parra et al., 2019). Uniplex IF was performed simultaneously for each fluorochrome to make a spectral library in human tonsil FFPE tissues as control and was used in the multispectral analysis. The stained slides were scanned using the multispectral microscope, Vectra 3.0.3 imaging system (Akoya Biosciences/PerkinElmer PerkinElmer), under fluorescence conditions at 10x magnification(Parra et al., 2019). Entire EAC or BE region on a slide was selected as regions of interest (each ROI, 0.335 mm^2) by a pathologist using the phenochart 1.0.9 viewer (Akoya Biosciences/PerkinElmer PerkinElmer). Multispectral Analysis: InForm 2.4.0 Image Analysis software (Akoya Biosciences/PerkinElmer) was used to analyze the scanned images. Using the tissue segmentation tool, each ROI was divided into intraepithelial and stromal compartment. The phenotype was defined based on co-localization of antibodies and a training algorithm was created as described earlier (Parra et al., 2017, Parra et al., 2019). The algorithm identifies each cell separately for one of the phenotypes mentioned. The final report of the various cell phenotypes is created by the InForm software expressing the results as number of cells of each phenotype in the individual compartment per mm^2. All the data was consolidated using the R studio 3.5.3 (Phenopter 0.2.2 packet, Akoya Biosciences/PerkinElmer) and SAS 7.1 Enterprise.

TCR sequencing

DNA sequencing of the variable region of the beta chain of the TCR (TCR-beta) was performed with immunoSEQ (Adaptive Biotechnologies, Seattle, WA) (Robins et al., 2009b). TCR clonality was computed using the Simpson Clonality Index (Simpson, 1949). The observed richness was computed as the number of unique nucleotide rearrangements per sample. The top 10 clones for each sample were calculated using the read proportions derived from the immunoSEQ analysis. We used the Shannon Equitability Index, computed as normalized Shannon entropy, to determine each repertoire's degree of uniform distribution; values approaching 0 indicated a very skewed distribution of frequencies (i.e., more variation in abundance), whereas values approaching 1 indicated nearly identical frequencies of every rearrangement (i.e., less variation in abundance). The inverse Chao, Efron thisted and Daley Smith estimator metrics are non-parametric models of determining repertoire richness and abundance based that were computed based on sample richness, extrapolation of the rarefaction curve and accounting for sampling depth (Chiu et al., 2014, Efron and Thisted, 1976, Daley and Smith, 2013).

Statistical Analysis

Statistical analyses were performed using R v3.6.1 (R Core Team, 2019). Analysis packages and tools used are described in the relevant methods sections. Statistical tests included basing the distribution of the samples based on a Shapiro test for normality. Mean analysis for groups with normal distribution were carried out using a two-sample t-test (for two groups) or ANOVA (for multiple groups) while non-parametric tests were carried out using the Wilcoxon

Rank sum test (for two groups) and the Kruskal-Wallis test (for multiple groups). TCR statistics were computed using the R packages tcR (v2.2.4) and immunarch(Team, 2019). Data were parsed and organized through R packages tidyr, reshape2, and dplyr (Wickham et al., 2017b, Wickham, 2007, Wickham et al., 2017a). All TCR-seq based Plotting was done through ggplot2 and ggrepel (Wickham, 2011b, Slowikowski et al., 2018).TCR statistics were computed using the R packages tcR (v2.2.4) and immunarch(Team, 2019).

Chapter Three: Spatial Intratumor Heterogeneity

Contents of this chapter is based on:

Akash Mitra[*], Miles C. Andrews[*], Whijae Roh, Marianna Petaccia De Macedo, Courtney W. Hudgens, Fernando Carapeto, Shailbala Singh, Alexandre Reuben, Feng Wang, Xizeng Mao, Xingzhi Song, Khalida Wani, Samantha Tippen, Kwok-Shing Ng, Aislyn Schalck, Donald A. Sakellariou-Thompson, Eveline Chen, Sangeetha M. Reddy, Christine N. Spencer, Diana Wiesnoski, Latasha D. Little, Curtis Gumbs, Zachary A. Cooper, Elizabeth M. Burton, Patrick Hwu, Michael A. Davies, Jianhua Zhang, Chantale Bernatchez, Nicholas Navin, Padmanee Sharma, James P. Allison, Jennifer A. Wargo, Cassian Yee, Michael T. Tetzlaff, Wen-Jen Hwu, Alexander J. Lazar, P. Andrew Futreal. Spatially resolved analyses link genomic and immune diversity and reveal unfavorable neutrophil activation in melanoma. Nat Commun 11, 1839 (2020). https://doi.org/10.1038/s41467-020-15538-9

Copyright permissions are not required since *Nature Communications* state that articles are published open access under a CC BY license (Creative Commons Attribution 4.0 International License). The CC BY license allows for maximum dissemination and re-use of open access materials and is preferred by many research funding bodies. Under this license users are free to share (copy, distribute and transmit) and remix (adapt) the contribution including for commercial purposes, providing they attribute the contribution in the manner specified by the author or licensor.

3.1 Introduction

Modern treatment paradigms increasingly expose patients with metastatic melanoma to multiple treatment modalities through the course of their disease(Coit et al., 2016). Immune checkpoint blockade in particular has revolutionized the therapeutic landscape, yet durable clinical benefit remains limited to a subset of patients(Chen et al., 2016a, Reuben et al., 2017, Roh et al., 2017). Numerous biomarker studies aiming to elucidate why the majority of patients fail to respond have revealed both immune and genomic contributors to therapeutic activity but incorporation of such factors into clinical practice is not yet routine(Chen et al., 2016a, Roh et al., 2017, Snyder et al., 2014a, Van Allen et al., 2014, Van Allen et al., 2015a).

Intra- and inter-tumoral heterogeneity can influence lesion-specific and overall patient response to therapy and may contribute significantly to tumor immune evasion(Reuben et al., 2017, Joyce and Pollard, 2009, Akbani et al., 2015). Studying the influence of *intra*-tumoral heterogeneity (ITH) using standard approaches such as bulk tumor sequencing or single cell sequencing generally loses spatial information. Thus, here we perform spatially detailed immune and genomic analysis of three metastatic lesions, including 67 sub-regions of one tumor sampled throughout its entire mass, from a heavily treated but long-term surviving melanoma patient. Through molecular analyses coupled with strict retention of spatial detail, we reconstruct the striking relationship between genomic and immune heterogeneity. We identify a remarkable link between copy number gain of chromosome 7 and an unfavorable immune composition driven by neutrophil activation recapitulated within TCGA melanoma samples and dominating non-responders to checkpoint blockade immunotherapy across multiple published cohorts. We also identify a long-term persistent T cell clonotype having potential relevance to vaccine and cellular immunotherapy.

3.2 Results

Longitudinal tumor sampling

Tumor and blood biospecimens were obtained from a Caucasian female diagnosed with de novo stage IV M1b melanoma of unknown primary metastatic to the left lung at the age of 77. Following initial curative intent wedge resection of the solitary *NRAS*^{Q61R} mutated lung metastasis (**Fig. 1A**, lesion 1), her clinical course was remarkable for long-term survival despite multiple lines of therapy for widely distributed soft tissue metastases with limited to no objective response over the following 8 years (**Fig. 1A**). To explore the relevance of *intra*-tumoral heterogeneity (ITH) to the setting of long-term survival with metastatic disease, we studied a ventral abdominal wall metastasis resected due to isolated progression during therapy with the PD-1 inhibitor pembrolizumab. This mass (**Fig. 1A**, lesion 2) was subjected to extensive multi-dimensional spatial and immunogenomic profiling by serial sectioning and use of alternate tumor sections for region matched immunohistochemistry analyses (odd numbered slices) and genomic and proteomic analyses (even numbered slices; **Fig. 1B**). Individual sections were further sub-divided into up to 20 regions (**Fig. 1B**, **Supplementary Fig. 1**), producing a total of 67 regions assessed by at least one analytical platform (**Supplementary Data 1**).

Mutational ITH is highly prevalent and spatially restricted

To characterize genomic ITH within the tumor specimen progressing during PD-1 inhibitor treatment ("on-PD-1 inhibitor" tumor), we performed deep targeted DNA sequencing for a panel of 265 cancer-related genes (**Supplementary Data 2**) of DNA from 41 tumor sub-regions. Of 53 identified somatic mutations, 28% (15 of 53) were shared in all 41 regions whereas 30% (16 of 53) were restricted to a single region, consistent with a degree of mutational ITH not previously described at this resolution. Somatic mutations in putative melanoma driver genes including $NRAS^{Q61R}$, $BRAF^{G421R}$ and $MAP2K1^{P124S}$, all key components of the MAPK pathway, were ubiquitously detected in all 41 regions, supporting the notion that somatic mutational heterogeneity is predominantly attributable to passenger mutations. A $JAK1^{P1044S}$ mutation affecting the activation loop of JAK1 detected in all 41 regions that conferred signaling hypomorphism by Ba/F3 mutant transformation assay (**Fig. 1C**), potentially contributed to the immunotherapy resistance displayed by this tumor clinically(Zaretsky et al., 2016).

Genomic ITH is dominated by copy number alterations

Analysis of copy number alterations (CNAs) detected across all 41 deeply sequenced samples identified gains of chromosome 6p and 20q, and losses of chromosome 6q and 9p, each of which has previously been identified in melanoma clinical samples (**Fig. 1D**)(Akbani et al., 2015). Subclonal alterations were also seen, including chromosome 7 gain in 4 samples, whole chromosome 10 loss in 5 samples, 10p loss in 1 sample, and chromosome 13 gain in 4 samples. Samples with subclonal loss of chromosome

10 were localized in adjacent tumor slices, but subclonal gains of chromosome 7 and 13 were found at non-contiguous sites (**Fig. 1D**). While previous studies have shown metastatic potential being associated with the loss of chromosome 10, we found evidence of regional losses of chromosome 10, most extensively along the tumor margin, suggesting this may be selected for in the context of stromal interactions at advancing tumor margins(Kabbarah et al., 2010). Nearly half (17/39, 44%) of the differentially expressed genes associated with chromosome 10 copy number losses were located on chromosome 10 itself, characterized by relatively high expression but low fold-change. Additional, more pronounced changes (at fold-change level) were observed in differentially expressed genes located on other chromosomes, such as *MT1B* (chr16), *TNNT3* (chr11) and *MUC12* (chr7) and *RPS6KA6* (chrX) (**Supplementary Data 3**). Additionally, unique chromosomal CNAs were found in nearly all (12 of 14) regions, demonstrating that CNAs may develop along spatially distinct trajectories even within a single metastasis. Comparing CNAs across longitudinal metastases of this patient, we also identified progressive stepwise regional loss of chromosome 10 in relation to therapy (pre-, on-, and post-PD-1 inhibitor therapy), thus implicating this CNA in both tumor margin dynamics and overall disease progression (**Supplementary Fig. 2**).

Immune cell content is highly and spatially diverse

We next characterized the ITH of gene expression patterns in the tumor, to gain insight into the nature of local tumor-immune microenvironments. Unexpectedly, unsupervised hierarchical clustering based on transcriptomic profiling revealed limited association between regional gene expression and histologic features such as intra-tumoral site (e.g.: “core” surrounded only by tumor mass versus “margin” spanning the tumor edge and including surrounding tissue) (**Fig. 2A**, **Supplementary Fig. 3A**). We then used several immune deconvolution tools to enumerate separate immune, stromal, and tumor cell populations as well as melanoma-, AXL- and MITF-related gene expression programs(Tirosh et al., 2016, Yoshihara et al., 2013, Rooney et al., 2015, Becht et al., 2016, Li et al., 2016b, Newman et al., 2015). Sub-regions with high content of one immune cell subset generally displayed an enrichment for multiple cell subsets, indicative of a broadly diverse infiltrating immune population, and consistent with the observed high correlations between immune cell marker stains by immunohistochemistry (**Fig. 2B-C**, **Supplementary Fig. 3B-C**). Samples with higher immune activity were over-represented at tumor margin sites ($p<0.001$, Fisher's exact test), reflecting the spatially excluded (i.e.: peri-tumoral) leukocytic accumulation observed on IHC (**Fig. 2A, Supplementary Fig. 1**). A notable exception was particularly high T and B cell signatures in multiple samples of section 8 (8A6, 8A7, 8A8, and 8A13), which was highly necrotic and displayed heavy neutrophil infiltration on matched FFPE slices (CD15 stain; **Supplementary Fig. 3D**) although this could not be determined as the cause or consequence of necrosis. In many cases, high immune cell signatures were accompanied by high expression of interferon-stimulated genes. Relatively low levels of variability were observed in expression of genes linked to melanoma cell phenotype (i.e.: melanocytic, MITF-related and AXL-related gene sets), however the same samples that displayed prominent and uniform immune signature upregulation also displayed a non-melanocytic phenotype that is known to be associated with mesenchymal-like and pro-invasive cellular behaviors, evidenced by high expression of the AXL-related gene set, and correspondingly low expression of the melanocytic and MITF-related gene sets (**Supplementary Fig. 3A**)(Konieczkowski et al., 2014, Garraway et al., 2005).

Mirroring transcriptional ITH, a high degree of ITH was observed between samples at the protein level, measured by a 296-target reverse-phase protein array (**Supplementary Fig. 3E**). Within the most spatially variable proteins, two notable modules of co-expressed proteins emerged; one including AIM1, ARID1A, MTOR, STAT5A, DUSP4 and SOX2, resembling a melanocytic cellular origin, and an anti-correlated set comprising AXL, PDGFR, JAK2, STAT, PDCD1 (PD-1) and PREX1, suggesting a mesenchymal-shifted and/or immune-infiltrated set (**Fig. 2D**). Several proteins were significantly associated with either locally low (VIM, SOX2) or high (MYH11, LCK, PTK2) immune infiltrate ($FDR<0.10$, Spearman's rho rank correlation, **Fig. 2D**).

Sites of similar immune composition may be spatially remote

Using sample-wide Euclidean distance metrics to connect samples with highly similar immune composition based on immune deconvolution rather than reductive immune scores or overall immune cell densities, we found that similar immunophenotypes were unrestricted by location at core or margin sites (e.g.: core 4A7 vs margin 6A3), or by spatial proximity (**Fig. 2E**). Three-dimensional mapping across all sampled regions of the tumor revealed clear but disconnected pockets of immune activation and suppression as typified by signatures derived for cytolytic activity, type I IFN activity and an anti-inflammatory signature (**Fig. 2F**)(Rooney et al., 2015), indicative of a degree of immunophenotype convergence. To address the implications of regional immune phenotype variation for clinical biomarker assessment, we performed consensus clustering of samples based on gene expression data and identified an optimal four cluster solution, being the minimum number of distinct regional "phenotypes" that would need to be sampled in order to approximately represent the transcriptional heterogeneity present across the entire tumor mass (**Supplementary Fig. 3F**). Importantly, we found that each of these clusters contained non-contiguous samples, indicating a low chance of serendipitously sampling all microenvironmental types with any single biopsy of the lesion (**Supplementary Fig. 3F**).

ITH implicates simultaneous methods of immune exclusion

Given progression of this tumor through anti-PD-1 immunotherapy, and previous findings suggesting a predictive significance of the immune status at the invasive tumor margin, we next compared tumor regions having either a high or low immune cell content as measured by a pan-leukocyte stain (CD45LCA) on IHC (**Supplementary Fig. 1**)(Tumeh et al., 2014). The most differentially expressed genes enriched in heavily infiltrated sub-regions included *FCRL1*, *CADM3*, *CR2* and *PAX5* (**Fig. 3A**, **Table 1**, **Supplementary Data 4 Top 100**), as well as genes involved in T cell function including *CD3D*, *CD28*, *ZAP70*, and *CD40LG*, in agreement with extensive CD8 and CD4 staining of mononuclear cells within these highly immune infiltrated regions by IHC (**Fig. 3B-D**). At the gene ontology pathway level, these differentially expressed genes contributed to mixed T and B cell enrichments and a substantial degree of functional gene connectivity (**Fig. 3C-D**), that was maintained even when specifically comparing samples located at the tumor margin (**Supplementary Fig. 4A-B**). We also identified a clear B cell gene enrichment in highly immune infiltrated samples, driven by *PAX5*, *BLK*, *CD19*, *CLECL1*, *CD180*, *CD22*, *CD79A* (**Fig. 3C-D**, **Table 2**). Parallel PAX5 immunostaining of tumor sections confirmed B cell lineage presence within these immune infiltrated samples localized to intra- and peri-tumoral leukocytic infiltrates, or within dense para-tumoral clusters associated with blood vessels and other immune cell types, suggestive of tertiary lymphoid structures (**Fig. 3E**). Additionally, a pro-tumorigenic M2 macrophage signature was evident throughout most regions of the tumor (**Supplementary Data 5**)(Newman et al., 2015, Coates et al., 2008). Tumor associated macrophages at the tumor periphery are known to be associated with tumor progression thus these data implicate active participation of immunosuppressive macrophages in the observed clinical progression of this tumor despite anti-PD-1 therapy(Qian and Pollard, 2010).

Reasoning that grouped analyses may obscure the true extent of variability in gene expression between individual samples, we performed single sample GSEA (ssGSEA) to gain a finer resolution of the functional transcriptomic activity(Barbie et al., 2009). Strikingly, unsupervised hierarchical clustering of the samples based on ssGSEA of Hallmark gene sets again revealed little similarity in terms of physical location within the tumor or the extent of peri-/intra-tumoral immune infiltrate (**Fig. 3F**). Samples with prominent enrichment of WNT/β-catenin signaling (2A10, 2A13, 2A16, 4A11, 8A4), which is a known tumor cell-intrinsic mechanism of immune cell exclusion(Spranger et al., 2015), were typically located at the tumor margin but did not show consistent association with immune cell content, although when immune cells were present, they were largely peri- or extra-tumoral in distribution. These data suggest that WNT/β-catenin signaling may contribute to exclusion of an immune infiltrate when one is present, but additional factors are necessary to explain the complete absence of an immune infiltrate from some regions. Intriguingly, despite known presence of activating *NRAS* and *MAP2K1* mutations, phospho-ERK1/2 (pERK) expression (by IHC) was largely absent from tumor cells except when located at or immediately beneath the tumor margin (**Fig. 3G**), suggesting MAPK activation in response to factors originating near the tumor surface. Areas of strong tumor cell pERK staining were frequently observed in association with overlaid peri-tumoral immune infiltrates, thus tumor cell ERK activation may be actively involved in the maintenance of immune cell exclusion, and at a scale that is significantly more spatially localized than previously thought based on pre-clinical models and broad assessments of patient samples(Khalili et al., 2012, Frederick et al., 2013).

Integrative analyses of multimodal molecular phenotypic data

Having identified clear links between immune and genomic heterogeneity throughout sub-regions of this tumor, we sought to identify genomic features underlying this through an integrative analysis of CNA and mRNA data, with the addition of methylation as a potential modulator of transcriptional activity. Examining the 560 most variably expressed genes for which all genomic data were available, unlike CNA, underlying methylation patterns appeared more variable between samples rather than between genes, implying that a genome-wide methylation state interacts with more localized genomic and post-transcriptional influences to affect gene expression in this context (**Fig. 4A**). From an immune standpoint, transcriptome-derived ESTIMATE immune scores trended lower in regions with chromosome 10 losses (p=0.088, two sided t-test), and were significantly lower in regions with sub-clonal gain of chromosome 7 (p=0.018, two sided t-test)[26]. Similarly, differentially expressed genes were enriched within CNA-affected regions of chromosome 6 and 7 (p=3.53e-7 and 1.22e-5, respectively, Benjamini-Hochberg corrected p-value). As expected, the clear majority of genes on chromosome 7 showed positive correlations between copy number and transcript abundance, consistent with CNA representing a dominant method of regulation of these genes (**Fig. 4B**, upper panel). However, copy number versus transcript correlations were negligible for four genes (*CALD1*, *CCT6A*, *CHCHD2*, and *ESYT2*) and negative for six genes (*ACTB*, *AEBP1*, *COL1A2*, *GIMAP4*, *GIMAP7*, and *SFRP4*), suggesting that additional mechanisms regulate transcript abundance of these genes, such as methylation. Gene methylation was inversely correlated with transcript abundance and thus consistent with negative regulation of transcription for most chromosome 7 genes (**Fig. 4B**, lower panel), including all but three of the copy number discordant genes (*GIMAP4*, *GIMAP7* and *SFRP4*). Notably, *SFRP4* is a soluble modulator of Wnt signaling that may antagonize Wnt-driven immune exclusion when highly expressed. *GIMAP4* is known to be involved in the regulation of Th1 versus Th2 T cell phenotypes. Using the unique genome-phenotype associations of chromosome 7 to model overall regulatory complexity and ITH, we found strikingly little similarity in the unsupervised clustering patterns of sub-regions based on copy number, methylation or transcript abundances, evidenced by a high degree of cross-cluster entanglement indicating the action of additional (unmeasured) factors in regulation of these genes (**Fig. 4C**). Furthermore, sub-regions of similar immune cell content (measured by IHC) were generally dispersed throughout the clusters, thus demonstrating unequivocally the presence of profound ITH underlying broadly similar appearing immune microenvironments at the cellular level.

Tumor-specific and agnostic T cell recruitment occurs on a regional basis

T cell function is central to current immunotherapy efficacy, hence, we performed sequencing of the variable region of the T cell receptor beta chain using both DNA and RNA approaches to study T cell ITH. T cell repertoire clonality was substantially variable between samples (**Fig. 4D**), suggesting highly localized patterns of clonal expansion and contraction that result in variable repertoire composition throughout the volume of an individual tumor. Only 0.02% of all TCR rearrangements were detectable in all regions of the tumor and the vast majority (74.6%) were restricted to a single region. We identified the top five highly transcriptionally active T cell clones per sample by plotting complementary productive frequencies generated from DNA- and RNA-based approaches (**Supplementary Fig. 5**): Three dominant clonotypes were present at high abundances across multiple regions of the metastasis (**Fig. 4E**). One dominant clonotype, at the amino acid level, present as a top five clone in all samples (CSVPTSGSRDNEQFF), was most prevalent in the upper sections (2 and 4) and least prevalent in the lowest section 8, which also had the lowest proportion of viable tumor. The next two most prevalent clones present in 72% (CASSSLQGARREETQYF) and 69% (CASSLHGDQPQHF) of all samples were particularly enriched in sections 6 and 8.

We examined repertoire overlap between samples to infer intra-tumor trajectories of T cell clones and found a high level of T cell clonal overlap between samples *within* section 8 (**Fig. 4F**). Conversely, there was greater sharing of clonotypes *between* sections 4 and 6, and more sparsely with section 2, paralleled by evidence of greater immune activation in key regions of these sections. The greater restriction of T cell clones within section 8 may reflect a functionally distinct T cell repertoire reacting to the prevailing necrotic conditions seen histologically throughout much of this section. Overall, this spatial variation suggests underlying differences in regional immunogenicity and antigenicity driving local accumulation of different T cell clones.

The observation of marked T cell repertoire ITH (**Fig. 4D**) and apparently distinct T cell repertoires between distinct regions of the tumor was surprising given the comparatively similar mutational landscape between tumor sub-regions. To explore the relationship between T cell clonal composition and tumor mutations we compared the truncal set of 15 mutations found in every subsample of the metastasis with the most highly expanded TCR-Vβ sequences found simultaneously in high proportions across all regions. In general, the productive frequencies of these expanded T cell clones correlated positively with the mutation variant allele frequencies (VAF), suggesting a surrogate relationship between VAF, tumor content and T cell clones reactive to tumor (but not necessarily these exact mutations). However, some expanded T cell clones showed inverse or mixed patterns of correlation with this set of truncal mutations, (**Supplementary Fig. 6**), including several clones negatively correlated with all the shared mutations, such as one (CASSLHGDQPQHF) that was predominantly found expanded in the necrotic slices. Together, these data indicate that while certain expanded T cell clones correlate positively with a set of truncal tumor mutations and are likely tumor specific, a distinct population of expanded T cell clones generally anti-correlate with truncal tumor mutations and although present within some regions of the tumor, are likely not specific for tumor antigens and may be non-specifically recruited into regions of inflamed/necrotic TME.

T cell clone persistence reveals recurrent priming and functional diversity

Leveraging the availability of peripheral blood samples and tumor samples obtained from distinct metastatic deposits spanning treatment-naïve, progression on-PD-1 inhibitor and progression post-PD-1 inhibitor time points (**Fig. 1A**), the temporal dynamics of the T cell repertoire were evaluated. Strikingly, the dominant clone present within the progressing abdominal wall tumor during PD-1 inhibitor therapy was not only present over time but was the most hyperexpanded clone within the treatment naive lung tumor sampled 7 years earlier (**Fig. 4E**). Evaluation of predicted neoantigens revealed one (*ZDHHC17* p.H507Y; IC50=77.17nM) that was shared amongst all tumor specimens. These data are at least consistent with a common neoepitope driving a persistent T cell response over time. In order to validate the potential *in vitro* immunogenicity of the ZDHHC17 p.H507Y neoantigen, we synthesized 12 overlapping candidate 9-mer peptides spanning the point mutation and used these peptides to elicit CD8 T cell responses from HLA-A*0301 donor peripheral blood mononuclear cells (PBMC) in peptide stimulation assays *in vitro* (see Methods). Compared with donor PBMC co-cultured with non-peptide pulsed A3-K562 cells, we observed elevated IFN-γ production by CD8 T cells of two HLA-A*0301 donors with several peptides (4, 6, 7, 9, 11 and 12) but most particularly from peptides 4 and 12 which induced the most robust responses at an average of 3-5 fold greater than background levels (i.e.: unpulsed cells), thus representing immunogenic epitope candidates (**Supplementary Fig. 7A-B**).

To evaluate the functional characteristics of this remarkably persistent T cell clonotype, we harnessed matched single cell *TCRα*, *TCRβ* and transcriptome sequencing of sorted CD45+CD3+ T cells within the post-PD-1 inhibitor tumor. The T cells clustered broadly into a population of activated cytolytic T cells (49%, CTL) expressing *CD8A*, *GZMA*, and *PRF1* and checkpoint-regulated T cells (20%) expressing multiple immune checkpoints including *ICOS*, *CTLA-4* and *TNFRSF18* but which were also dominantly *CD4*-expressing (**Fig. 4G-H, Supplementary Fig. 7C**). We recovered 11 counts of the persistent TCR-Vβ rearrangement in a population of 6267 T-cells (0.17%), and identified multiple *TCRα* (*TRAV35*-*TRAJ23* and *TRAV2*-*TRAJ2*) and *TCRβ* (*TRBV29*-*TRBJ2*, *TRBV7*-*TRBJ2* and *TRBV5*-*TRBJ1*) partners to the TCR-Vβ sequence of interest comprising this T cell population, including T cells with dual TCR-Vβ rearrangements. Based on VDJ combinatorics, a minimum of two distinct T cell clones contributed to this recurrent TCR rearrangement at the amino acid level. Interestingly, when immunoprofiling these cells using matched RNA-seq data, we found eight cells within the cluster expressing multiple immune checkpoint molecules, and five cells in the cluster resembling activated CTLs. The detection of multiple clones at nucleotide level expressing a synonymous CDR3 amino acid sequence, their persistence over nearly a decade, and simultaneous presence of both activated and exhausted phenotypes suggests that this T cell population arose from multiple independent T-cell priming events rather than functional divergence following a single more recent priming/activation event.

Chromosome 7 gain is associated with an unfavorable immune outcome

To further explore the link between genomic copy number alterations and immune ITH we focused on the observation of decreased ESTIMATE immune scores in regions with subclonal gain of chromosome 7 (p=0.018, two-sided t-test). Immune deconvolution revealed low counts of multiple immune cell subsets including T cells (p=0.00096), CD8+ T cells (p=0.084), cytotoxic lymphocytes (p=0.036), NK cells (p=0.0013), B cells (p=0.015), monocytic lineage (p=7.6e-5), myeloid-derived dendritic cells (p=0.0039) and most significantly neutrophils (p=3.4e-5; all two-sided t-test comparison of means) in these sub-regions (**Supplementary Fig. 8A**; neutrophil signature genesets)(Becht et al., 2016). However, whilst overall neutrophil counts (derived from transcriptome data) were low and this was consistent with generally low neutrophil densities identified by CD15 immunostaining of corresponding FFPE sections, interrogation of neutrophil activation gene sets to assess putative neutrophil functional status revealed a net neutrophil activation (slightly higher levels of 'positive neutrophil activation', p=0.51; significantly lower levels of 'negative neutrophil activation', p=0.00087, and 'negative regulation of neutrophil degranulation', p=0.0068; all Benjamini-Hochberg corrected p-values) in the sub-regions with gain of chromosome 7 compared with chromosome 7 stable regions (**Fig. 5A-B**). We sought to validate this relationship using TCGA SKCM samples and identified samples with both copy number and mRNA expression data (n=470), within which 50 samples harbored whole chromosome gains of chromosome 7. Differential expression analysis comparing samples with chromosome 7 gain versus non-gain revealed marked enrichment for neutrophil-related genes and associated pathway level enrichment (**Fig. 5C-D**) despite marginally lower neutrophil enumeration by CIBERSORT (**Supplementary Fig. 8B**): The top four enriched gene ontology (GO) terms were neutrophil degranulation, neutrophil involved in immune response, neutrophil activation and neutrophil mediated immunity (all p=1e-6, two-sided t-test comparison of means with Benjamini-Hochberg correction). In parallel, KEGG pathways enriched in the chromosome 7 gain samples included response to bacterial infections, phagosome and lysosome formation, and antigen processing, consistent with the observed strength of gene enrichments in neutrophil-related GO terms (**Supplementary Fig. 8C**).

As immune infiltrate correlates with OS and has been shown to correlate with responsiveness to anti-PD-1 and anti-CTLA-4 immunotherapy, we then investigated the significance of such neutrophil signatures in three publicly available immunotherapy-treated melanoma cohorts (n=119)(Van Allen et al., 2015b, Tumeh et al., 2014, Hugo et al., 2016, Riaz et al., 2017b). Within the anti-CTLA-4 cohort (Van Allen, n=36) and two anti-PD-1 cohorts (Hugo, n=27; Riaz, n=56), overall neutrophil estimation was again largely similar (**Supplementary Fig. 8D-F**) however the genes significantly enriched in non-responders compared with responders to therapy showed pathway level enrichments dominated by the same neutrophil signatures observed in chromosome 7 gain TCGA SKCM samples, namely neutrophil degranulation (p=6e-5, 3e-10, 5e-6, respectively), neutrophil involved in immune response (p=6e-5, 3e-10, 5e-6, respectively), neutrophil activation (p=6e-5, 3e-10, 5e-6, respectively) and neutrophil mediated immunity (p=6e-5, 3e-10, 5e-6 respectively; all Benjamini-Hochberg corrected p-values) (**Fig. 5E-G**). A core group of differentially expressed genes (*FTH1, FTL, HSPA8, HSP90AA1 and HSP90B1*) was recurrently identified within significantly enriched pathways across TCGA SKCM samples and clinical cohorts (**Fig. 5H**). Together, these data suggest a recurrent immunosuppressive role of chromosome 7 gain, potentially mediated by neutrophil accumulation and/or activation, although it is unclear whether neutrophil density or activation status are acting as a surrogate for the typical co-localisation of tumor necrosis observed at such sites. However, these associations appear active both locally within tumors and at the bulk tumor level where it has clinical implications for immune checkpoint blockade.

3.3 Discussion

In this study, we performed matched genomic and immune analysis of 67 distinct regions of a melanoma metastasis coupled to longitudinal analyses in a patient treated with multiple therapies, including prolonged exposure to (and progression on) anti-PD-1 immunotherapy. Consistent with previous studies, we observed minimal ITH in oncogenic mutations in canonical melanoma driver genes, but reveal striking genomic ITH in CNAs, including distinct copy number loss in chromosome 10, and gains of chromosomes 7 and 13, which may contribute to differences in the immune landscape. The loss of chromosome 10, and thus *PTEN*, has been implicated in resistance to PD-1 blockade previously, and in the context of this immunotherapy treated patient was observed to be lost in a stepwise fashion between tumors sampled prior to, during, and after anti-PD-1 therapy(Peng et al., 2016, Roh et al., 2017).

The most immediately apparent implication of the extent of heterogeneity observed, and its diverse representation across space even within a single metastatic deposit, is how inherently limited the prediction of clinical outcomes can be when based on limited physical sampling of tumor material, especially if only one metastatic deposit is sampled. Indeed, based on transcriptional heterogeneity alone, a complete understanding of the immunogenomic TME of the extensively profiled lesion in this study would likely require a minimum of four separate passes if subjected to needle biopsy. Whilst the degree of immunogenomic spatial heterogeneity in any given tumor mass cannot yet be predicted non-invasively, the spontaneous nature of immune-tumor interactions implies that relevant spatial heterogeneity will be found, irrespective of prior therapeutic exposures. To the extent that additional non-mutational features begin to emerge as clinically meaningful biomarkers for treatment response/resistance, these facets of multidimensional heterogeneity will need to be considered when planning biopsy-derived, biomarker driven trials.

Tumor heterogeneity has been linked to the emergence of treatment-resistant tumor cell sub-populations which expand under the selective pressure of therapy. Conceptually, heterogeneity encompasses multiple domains (e.g.: spatial, temporal, clonal) and can be applied to any measurable feature of a tumor, thus it remains unclear exactly which molecular constituents of heterogeneity are most consequential to clinical outcomes. Previous studies of heterogeneity in other tumor types (e.g.: renal, prostate, lung) have focused primarily or exclusively on phylogenetic mutational analyses to characterize clonal heterogeneity of tumor cell content(Linch et al., 2017, Morris et al., 2016, Jamal-Hanjani et al., 2017a, Gerlinger et al., 2012b). When performed, multi-region sequencing either for tumor cell mutations or TCR profiling has surveyed minimal numbers of regions (e.g.: 3-5 per tumor), often in relatively small numbers of samples, whilst truly multi-platform analyses have effectively evaluated *inter*-tumoral rather than *intra*-tumoral heterogeneity(Jamal-Hanjani et al., 2017a, Zhang et al., 2013, Gerlinger et al., 2013, Roper et al., 2019, Gerlinger et al., 2014). At the extreme of cellular resolution, studies employing single cell techniques, whilst informative of the multi-dimensional cellular heterogeneity within bulk tumor cell populations, necessarily destroy spatial information during sample processing and arguably do not *comprehensively* survey the transcriptome within any individual cell(Tirosh et al., 2016). Thrane and colleagues performed a proof-of-principle high resolution spatial transcriptomics analysis of four lymph node metastases obtained from patients with stage III melanoma, finding evidence of variably distinct gene expression profiles between regions of tumor, lymphoid tissue, and an apparent transition zone that may have represented functional interaction between tumor, stroma and lymphoid cells(Thrane et al., 2018). Relative intratumoral transcriptomic homogeneity in one sample was associated with long-term overall survival, however other domains of heterogeneity were not evaluable with this technique.

We found chromosome 7 gain to be significantly associated with features of an unfavorable immune microenvironment, including a paucity of effector cell populations and signatures of neutrophil activation. This relationship was confirmed amongst melanoma samples of The Cancer Genome Atlas. Furthermore, a strikingly consistent set of neutrophil enrichments was observed in melanoma tumors failing to respond to either anti-CTLA-4 or anti-PD-1 therapy across three independent published cohorts. This reveals two key messages, with the important caveat that additional studies are required to clarify the nature of neutrophil recruitment and activation in anti-tumor immune responses and whether their presence is largely as a consequence of cellular destruction by other mechanisms. Firstly, the insights from regional immunogenomic differences within a single tumor metastasis can directly translate to the bulk tumor level, and secondly, chromosome 7 gain may drive an immunologically adverse phenotype associated directly or indirectly with neutrophil activation. Several prominent oncogenes (*BRAF*, *EGFR*, *MET*) are located on chromosome 7 and may thus be subject to amplification in the setting of copy number gains. In our patient's spatially profiled tumor, *BRAF* amplification compounded by an activating $BRAF^{G421R}$ mutation may have contributed to immunosuppressive MAPK signaling, potentially in conjunction with other chromosome 7 oncogenes(Khalili et al., 2012, Boni et al., 2010). We also observed recurrent dysregulation of ferritin and HSP90-related genes, suggestive of an enhanced acute-phase protein reaction, iron loading and molecular stress in the context of chromosome 7 gain and immunotherapy failure. Iron availability is known to influence tumor cell survival and the function of numerous immune cell types including T cells, however these competing outcomes have been poorly studied in solid tumors such as melanoma(Pfeifhofer-Obermair et al., 2018, Wang et al., 2019). Nevertheless, a potential role for immunosuppressive neutrophil phenotypes and iron trafficking within the TME warrants further evaluation.

We identified a persistent, high frequency T cell clonotype prevalent in multiple tumor deposits over many years in this patient, with evidence of both activated and checkpoint molecule regulated (likely *previously* activated) cells present simultaneously. The time frame, and multiplicity of independent genomic rearrangements leading to this clonotype indicates repeated priming events, potentially in response to a highly persistent tumor antigen robust to multiple lines of treatment. Identification of such a persistent T cell population, or its persistent antigen, might be specifically useful for the development of defined antigen therapies such as vaccines (definitive or adjunctive therapy) or engineered T cell therapies based upon these targets, and warrants wider sampling of multiple tumors in patients – including use of archival tissues – in order to identify persistent tumor features that may be exploited for therapeutic advantage. Furthermore, our integrative immunogenomic analysis strongly suggests that high frequency T cell clonotypes may be recruited to the tumor microenvironment not only due to tumor cell reactivity, but as passengers in the inflammatory milieu. Further work will be required to determine to what extent such "passenger" T cell clonotypes contribute usefully to the anti-cancer immune response.

Our findings of extensive immunogenomic heterogeneity at the intra-tumoral level are inherently limited by detailed multi-platform profiling of a single lesion, thus it is difficult to determine how typical the observed extent of heterogeneity is to broader patient populations, particularly those having differing clinical scenarios and treatment outcomes. Nonetheless, considering that sub-clonal variation has now been described in numerous tumor types, these findings serve to highlight the potential sensitivity of the immune microenvironment to local factors, including tumor genomic features that appear to have functional impact on local tumor immunity. Through analyses of several clinical datasets we found certain immunogenomic features from our deeply profiled tumor to have meaningful correlates in additional cohorts of patient samples, but additional studies are clearly required to refine these inferences towards therapeutically manipulable strategies. Whilst overall objective clinical responses were not achieved in this patient over a period approaching 10 years and 7 lines of therapy, substantial clinical benefit was derived. Although this patient represents only a subset of long-term survivors with metastatic melanoma, considering the increasing availability of disease modifying therapies, it is likely that this group will become increasingly prevalent. Further molecular characterization will ultimately aid in understanding long term survivors of metastatic disease, providing therapeutic insights transferrable to the greater majority of patients.

3.4 Figures and Tables

Figure 1.

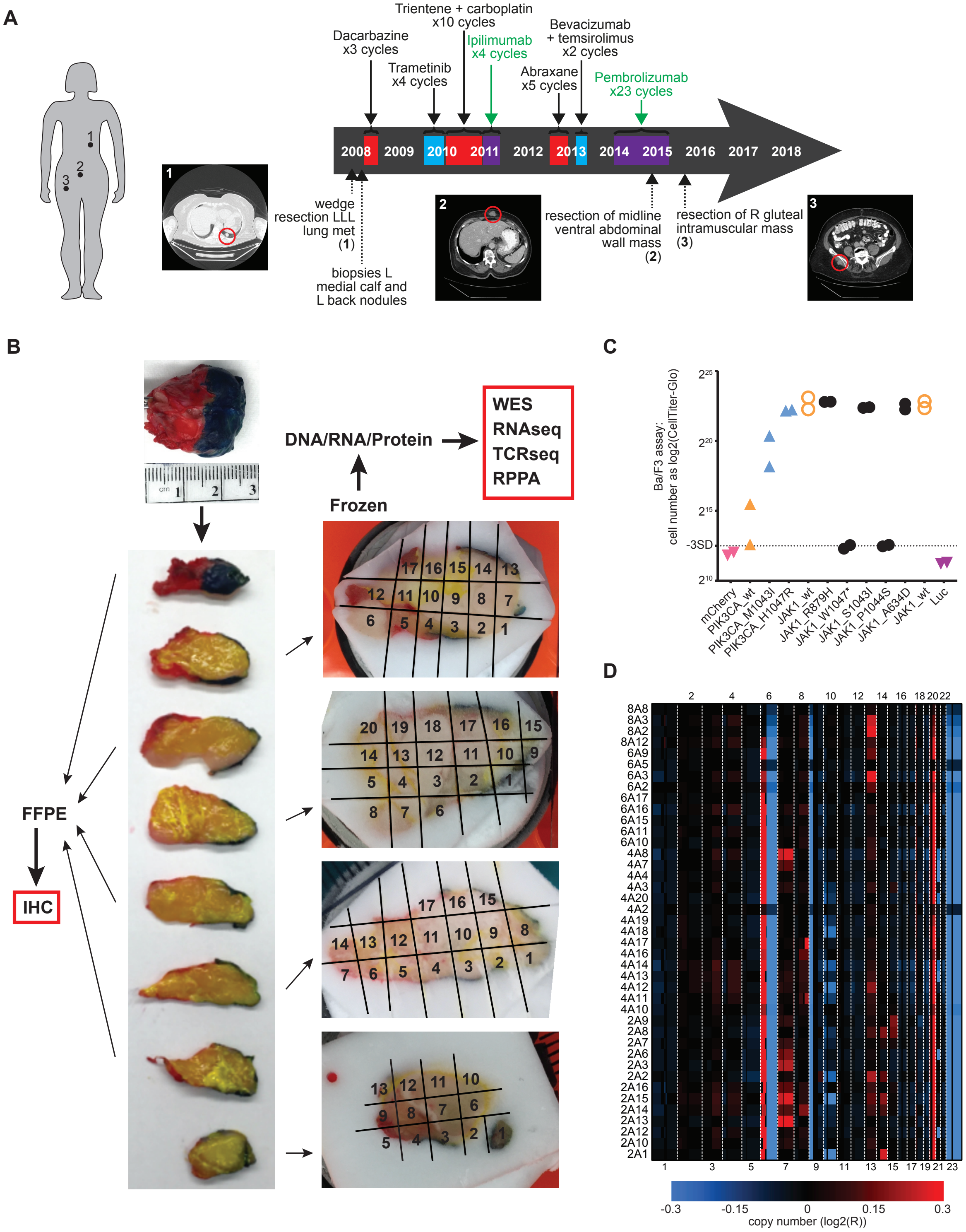


Figure 1: Genomic inter- and intra-tumoral heterogeneity in a heavily treated melanoma patient are driven by copy number alterations. (**A**) Timeline of treatments and surgical sampling of three distinct melanoma tumors from a long-term surviving patient with largely treatment unresponsive metastatic melanoma. Treatment modality is indicated by color (red - chemotherapy, blue - targeted therapy, purple – immunotherapy). Molecularly profiled lesions are indicated: index left lower lobe (LLL) lung metastasis (lesion 1), progressing ventral abdominal wall mass (lesion 2) and slowly progressing right gluteal mass (lesion 3). (**B**) Sectioning and use of the on-PD-1 inhibitor abdominal wall lesion. The tumor was oriented by lateral inking (red – left, blue – right), sliced and laid on a grid. The odd numbered slices were processed for FFPE and used for immunohistochemistry whilst the even numbered slices were fresh-frozen and used for genomic and proteomic analyses (whole exome sequencing - WES, RNA sequencing, TCR sequencing, reverse-phase protein array - RPPA). (**C**) Functional hypomorphism of the identified *JAK1* mutation (*JAK1*^{P1044S}) was identified by Ba/F3 transformation assay. Also shown are known oncogenic *JAK1* variants (*JAK1*^{R879H}, *JAK1*^{S1043I}, *JAK1*^{A634D}), wild-type *JAK1*, a truncating *JAK1* hypomorph (*JAK1*^{W1047*}), and oncogenic *PIK3CA* variants. (**D**) Copy number alterations in each region of the tumor are shown in the chromosome coordinate as log2-transformed copy number probe intensities R (observed intensity/reference intensity); copy number gains are shown as red and copy number losses as blue.

Figure 2.

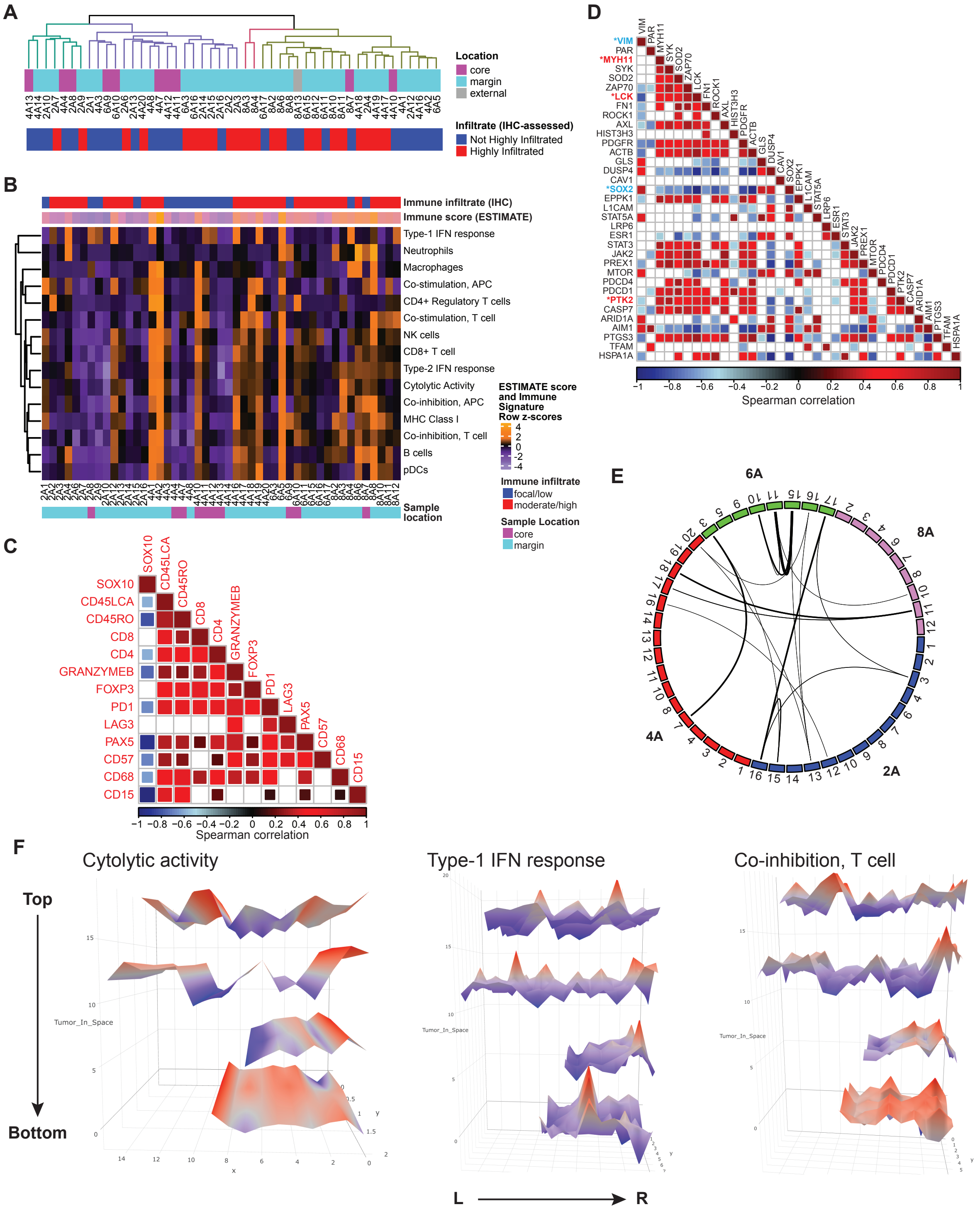

A
Location
core
margin
external
Infiltrate (IHC-assessed)
Not Highly Infiltrated
Highly Infiltrated
B
Immune infiltrate (IHC)
Immune score (ESTIMATE)
Type-1 IFN response
Neutrophils
Macrophages
Co-stimulation, APC
CD4+ Regulatory T cells
Co-stimulation, T cell
NK cells
CD8+ T cell
Type-2 IFN response
Cytolytic Activity
Co-inhibition, APC
MHC Class I
Co-inhibition, T cell
B cells
pDCs
ESTIMATE score and Immune Signature Row z-scores
4
2
0
-2
-4
Sample location
Immune infiltrate
focal/low
moderate/high
Sample Location
core
margin
C
SOX10
CD45LCA
CD45RO
CD8
CD4
GRANZYMEB
FOXP3
PD1
LAG3
PAX5
CD57
CD68
CD15
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
Spearman correlation
D
*VIM
PAR
*MYH11
SYK
SOD2
ZAP70
*LCK
FN1
ROCK1
AXL
HIST3H3
PDGFR
ACTB
GLS
DUSP4
CAV1
*SOX2
EPPK1
L1CAM
STAT5A
LRP6
ESR1
STAT3
JAK2
PREX1
MTOR
PDCD4
PDCD1
*PTK2
CASP7
ARID1A
AIM1
PTGS3
TFAM
HSPA1A
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
Spearman correlation
E
6A
8A
4A
2A
F
Cytolytic activity
Type-1 IFN response
Co-inhibition, T cell
Top
Bottom
Tumor_In_Space
x
y
L
R

Figure 2: Immune-driven transcriptional heterogeneity implicates diverse immune cell populations and highly localized immune activation or suppression. (**A**) Unsupervised hierarchical clustering based on the top 1000 most variant (mean absolute deviation) genes across all samples of the on-PD-1 inhibitor tumor, demonstrating limited associations between clustering of samples and tumor location based on the general transcriptome, but apparent association between high immune infiltrate and location at the tumor margin. (**B**) Heatmap of immune signature gene sets (from Rooney, *et al*[13]) across tumor sub-regions demonstrate dispersed pockets of immune activation or suppression throughout the tumor, wherein immune-high samples (e.g.: 4A1/4A2, 6A5 and 8A6) are physically distant from each other within the tumor mass. IHC-based immune infiltrate and ESTIMATE immune scores (top) and IHC-based tumor sample location (bottom) are indicated. (**C**) Immunohistochemical marker inter-correlations demonstrating generally diverse representation of immune cell types when infiltrates are present. Data are Spearman correlation values (with Benjamini-Hochberg correction; only $p<0.05$ are shown) indicated according to the color scale shown. (**D**) Correlation of most variably abundant proteins measured by reverse-phase protein array, revealing two main modules of highly correlated molecules. Proteins displaying statistically significant (FDR $p<0.10$) correlation with immune infiltrate are indicated by * and color (blue = anti-correlated with immune infiltrate, red = directly correlated with immune infiltrate). Data are Spearman correlation values (with Benjamini-Hochberg correction; only $p<0.05$ are shown) indicated according to the color scale shown. (**E**) Sample-wide similarity of immune activity was estimated by calculation of the distance matrix between samples using the immune activation signature expression data; lines connect samples in the top quartile of similarity scores, demonstrating global immune signature similarities that are not restricted by intra-tumoral location. (**F**) Three-dimensional spatial mapping of subregion Cytolytic activity signature, Type-I IFN response signature and Co-inhibition, T cell signature scores derived from transcriptomic data, in the manner of Rooney *et al*[13]. Data map the geometric mean of genes included in each gene set onto three-dimensional space representing the tumor slices, with color and height indicating expression value (higher expression = red peaks, lower expression = blue troughs).

Figure 3.

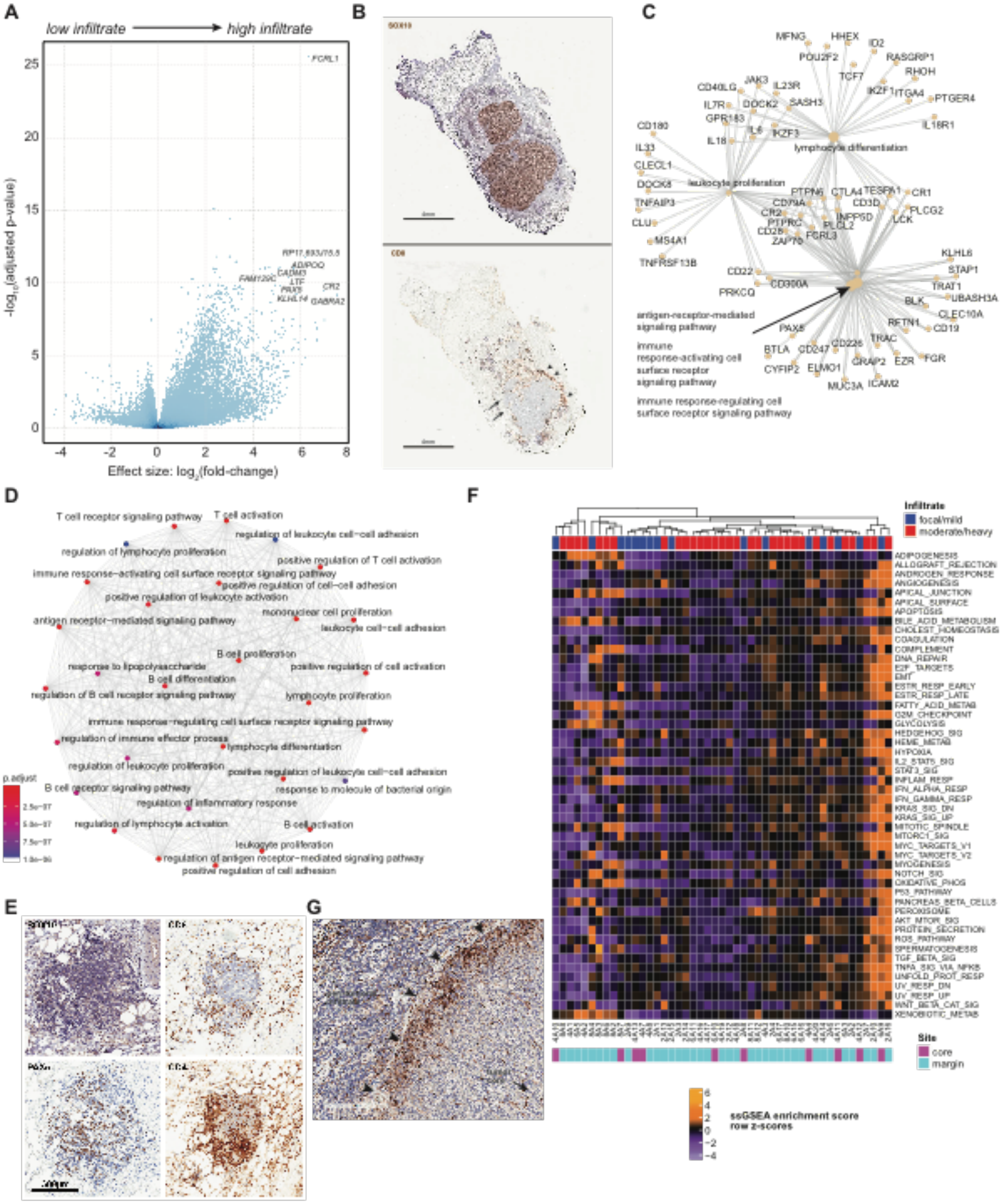

A
low infiltrate
high infiltrate
-log10(adjusted p-value)
Effect size: log2(fold-change)
25
20
15
10
5
0
-4
-2
0
2
4
6
8
FCRL1
RP11_693J15.5
ADIPOQ
CADM3
FAM129C
LTF
PAX5
KLHL14
CR2
GABRA2
B
SOX10
CD8
C
MFNG
HHEX
ID2
POU2F2
RASGRP1
RHOH
TCF7
CD40LG
JAK3
IL23R
IKZF1
ITGA4
PTGER4
IL7R
DOCK2
SASH3
GPR183
IL18R1
CD180
IL6
IKZF3
IL18
lymphocyte differentiation
IL33
CLECL1
DOCK8
leukocyte proliferation
PTPN6
CTLA4
TESPA1
CR1
TNFAIP3
CD79A
CD3D
PLCG2
CR2
INPP5D
CLU
PTPRC
PLCL2
LCK
CD28
FCRL3
ZAP70
MS4A1
TNFRSF13B
KLHL6
STAP1
CD22
CD300A
TRAT1
PRKCQ
UBASH3A
BLK
CLEC10A
antigen-receptor-mediated signaling pathway
RFTN1
CD19
PAX5
immune response-activating cell surface receptor signaling pathway
TRAC
BTLA
CD247
CD226
CYFIP2
ELMO1
GRAP2
EZR
FGR
MUC3A
ICAM2
immune response-regulating cell surface receptor signaling pathway
D
T cell receptor signaling pathway
T cell activation
regulation of leukocyte cell−cell adhesion
regulation of lymphocyte proliferation
positive regulation of T cell activation
immune response−activating cell surface receptor signaling pathway
positive regulation of cell−cell adhesion
positive regulation of leukocyte activation
mononuclear cell proliferation
antigen receptor−mediated signaling pathway
leukocyte cell−cell adhesion
B cell proliferation
response to lipopolysaccharide
positive regulation of cell activation
B cell differentiation
regulation of B cell receptor signaling pathway
lymphocyte proliferation
immune response−regulating cell surface receptor signaling pathway
regulation of immune effector process
lymphocyte differentiation
regulation of leukocyte proliferation
positive regulation of leukocyte cell−cell adhesion
p.adjust
B cell receptor signaling pathway
response to molecule of bacterial origin
2.5e−07
regulation of inflammatory response
5.0e−07
regulation of lymphocyte activation
B cell activation
7.5e−07
leukocyte proliferation
1.0e−06
regulation of antigen receptor−mediated signaling pathway
positive regulation of cell adhesion
E
SOX10
CD8
PAX5
CD4
200μm
G
F
Infiltrate
focal/mild
moderate/heavy
ADIPOGENESIS
ALLOGRAFT_REJECTION
ANDROGEN_RESPONSE
ANGIOGENESIS
APICAL_JUNCTION
APICAL_SURFACE
APOPTOSIS
BILE_ACID_METABOLISM
CHOLEST_HOMEOSTASIS
COAGULATION
COMPLEMENT
DNA_REPAIR
E2F_TARGETS
EMT
ESTR_RESP_EARLY
ESTR_RESP_LATE
FATTY_ACID_METAB
G2M_CHECKPOINT
GLYCOLYSIS
HEDGEHOG_SIG
HEME_METAB
HYPOXIA
IL2_STAT5_SIG
STAT3_SIG
INFLAM_RESP
IFN_ALPHA_RESP
IFN_GAMMA_RESP
KRAS_SIG_DN
KRAS_SIG_UP
MITOTIC_SPINDLE
MTORC1_SIG
MYC_TARGETS_V1
MYC_TARGETS_V2
MYOGENESIS
NOTCH_SIG
OXIDATIVE_PHOS
P53_PATHWAY
PANCREAS_BETA_CELLS
PEROXISOME
AKT_MTOR_SIG
PROTEIN_SECRETION
ROS_PATHWAY
SPERMATOGENESIS
TGF_BETA_SIG
TNFA_SIG_VIA_NFKB
UNFOLD_PROT_RESP
UV_RESP_DN
UV_RESP_UP
WNT_BETA_CAT_SIG
XENOBIOTIC_METAB
Site
core
margin
6
4
2
0
−2
−4
ssGSEA enrichment score
row z-scores

Figure 3: Transcriptomic analysis demonstrates considerable ITH underlying convergent immune phenotypes. (**A**) Volcano plot of differentially expressed genes comparing high versus low immune infiltrate regions across the tumor core and margins. Vertical red lines indicate a minimum 2-fold change in expression value; horizontal red lines indicate the adjusted p value threshold of <0.05. (**B**) Representative IHC sections demonstrating matched tumor content (SOX10 stain, above), and immune infiltration (CD8 stain, below) illustrating substantial variation of local CD8 T cell content ranging from low (arrows) to high (arrowheads). (**C**) Gene connection network of genes upregulated in immune-infiltrated samples that contribute to highly enriched GO terms/pathways, showing substantial connectivity. (**D**) Functional enrichment network showing diverse representation of immune cell pathways and functions in the immune infiltrate-derived differentially expressed genes, dominated by highly inter-connected T and B lymphocyte-related terms. (**E**) Representative IHC images of a para-tumoral lymphoid structure present in section 1B, demonstrating absence of tumor cells (SOX10) but mixed populations of CD4+, CD8+ and PAX5+ lymphocytes. Magnification 10X. (**F**) Single-sample gene set enrichment analysis demonstrating spatially discontiguous enrichment of functionally relevant gene sets throughout the tumor. IHC-based estimate of immune infiltrate (top) and sample location (bottom) are indicated. (**G**) Immunostained tumor tissue revealing restriction of tumor cell phospho-ERK1/2 expression (brown staining) to cells located at or immediately subjacent to the tumor cell surface. Arrowheads: tumor-surrounding tissue interface. Magnification: 10×.

Figure 4.

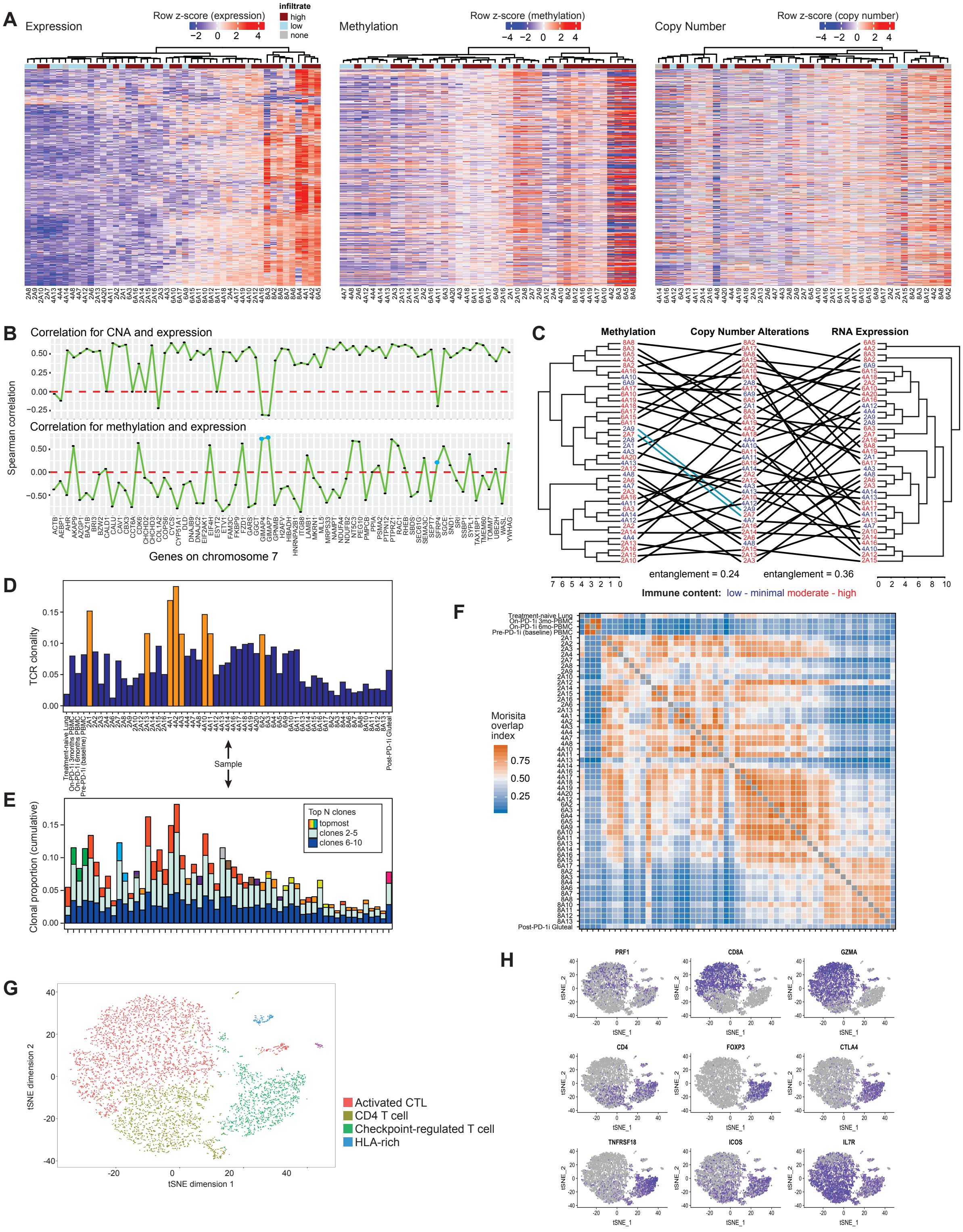

A
Expression
Row z-score (expression)
−2 0 2 4
infiltrate
high
low
none
Methylation
Row z-score (methylation)
−4 −2 0 2 4
Copy Number
Row z-score (copy number)
−4 −2 0 2 4
B
Correlation for CNA and expression
Correlation for methylation and expression
Spearman correlation
Genes on chromosome 7
C
Methylation
Copy Number Alterations
RNA Expression
entanglement = 0.24
entanglement = 0.36
Immune content: low - minimal moderate - high
D
TCR clonality
Sample
Treatment-naive Lung
On-PD-1i 3months PBMC
On-PD-1i 6months PBMC
Pre-PD-1i (baseline) PBMC
Post-PD-1i Gluteal
E
Clonal proportion (cumulative)
Top N clones
topmost
clones 2-5
clones 6-10
F
Morisita overlap index
0.75
0.50
0.25
G
tSNE dimension 1
tSNE dimension 2
Activated CTL
CD4 T cell
Checkpoint-regulated T cell
HLA-rich
H
PRF1
CD8A
GZMA
CD4
FOXP3
CTLA4
TNFRSF18
ICOS
IL7R
tSNE_1
tSNE_2

Figure 4: T cell repertoire dynamics reveal high ITH, potential for long-term clonal persistence and irregular intra-tumoral movement. (**A**) Heatmaps of the 560 most variably expressed genes (based on median absolute deviation) across the transcriptome dataset for which matched methylation and copy number data were available indicating distinct sample-wise clustering patterns within each dataset and generally unidirectional methylation patterns within samples. Data are log2-transformed counts (gene expression), beta values (methylation) and log2(probe intensity=observed intensity/reference intensity) (copy number), z-scored within each data type. (**B**) Correlation of copy number (upper panel) and methylation (lower panel) with transcript expression for the most variable genes on chromosome 7 showing mostly positive correlations for CNA and mostly negative correlations for methylation as expected. Three genes (indicated in blue, lower panel) showed discordant correlations for both CNA and methylation. (**C**) Tanglegram showing relationships between the sample clustering obtained independently from each of the copy number, methylation and transcriptome datasets for samples represented in all datasets. Entaglement values (range 0-1) indicate moderate lack of cluster structure concordance, indicative in this context of significant additional (unmeasured) factors contributing to the regulation of mRNA expression. Immune status of each subregion sample is indicated in color. (**D**) T cell receptor-Vβ (TCR) clonality (range 0-1) varied considerably between clinically relevant time points, between tumors and the peripheral blood, and spatially within the on-PD-1 inhibitor tumor. Samples with clonality > 0.1 are highlighted in orange. (**E**) Top 10 most abundant TCR clonotype proportions (i.e.: fraction of total identified TCR clonotypes) in each sample are represented as stacked bar plots. The topmost abundant clone is colored at the top of the bar, with each color representing a unique clonotype that may be shared between samples. Clonotypes 2-5 and 6-10 are colored in light blue and deep blue, respectively. (**F**) Morisita overlap index (MOI, range 0-1) values of TCR repertoires comparing the pre-treatment sample, peripheral blood samples, on-PD-1 inhibitor sample (each sub-region) and a post-PD-1 inhibitor sample were used to compare the overlap in shared nucleotypes in the TCR repertoire identified in each sample. Higher MOI indicates a greater proportion of shared TCR sequences. Within the on-PD-1 inhibitor sample,

TCR clonotypes present in section 8 were largely restricted to this geographic location, which was notably highly necrotic. There was considerable sharing of clonotypes between sections 4 and 6, and to a lesser degree between 2, 4 and 6, suggesting a greater degree of physical movement of T cells between these sections. (**G**) tSNE plot of TIL populations. The majority of the cells fell into an activated cytotoxic T cell lymphocyte, CD4 T cell and checkpoint-inhibited T cell phenotype. (**H**) Marker gene expression levels across TIL clusters. Relative expression of key marker genes associated with a cytotoxic T-cell phenotype (CD8, GZMA and PRF1), CD4 phenotype (CD4 and IL7R) and a multiply checkpoint inhibited phenotype (FOXP3, CTLA4, GITR and ICOS) are overlaid on the tSNE clusters.

Figure 5.

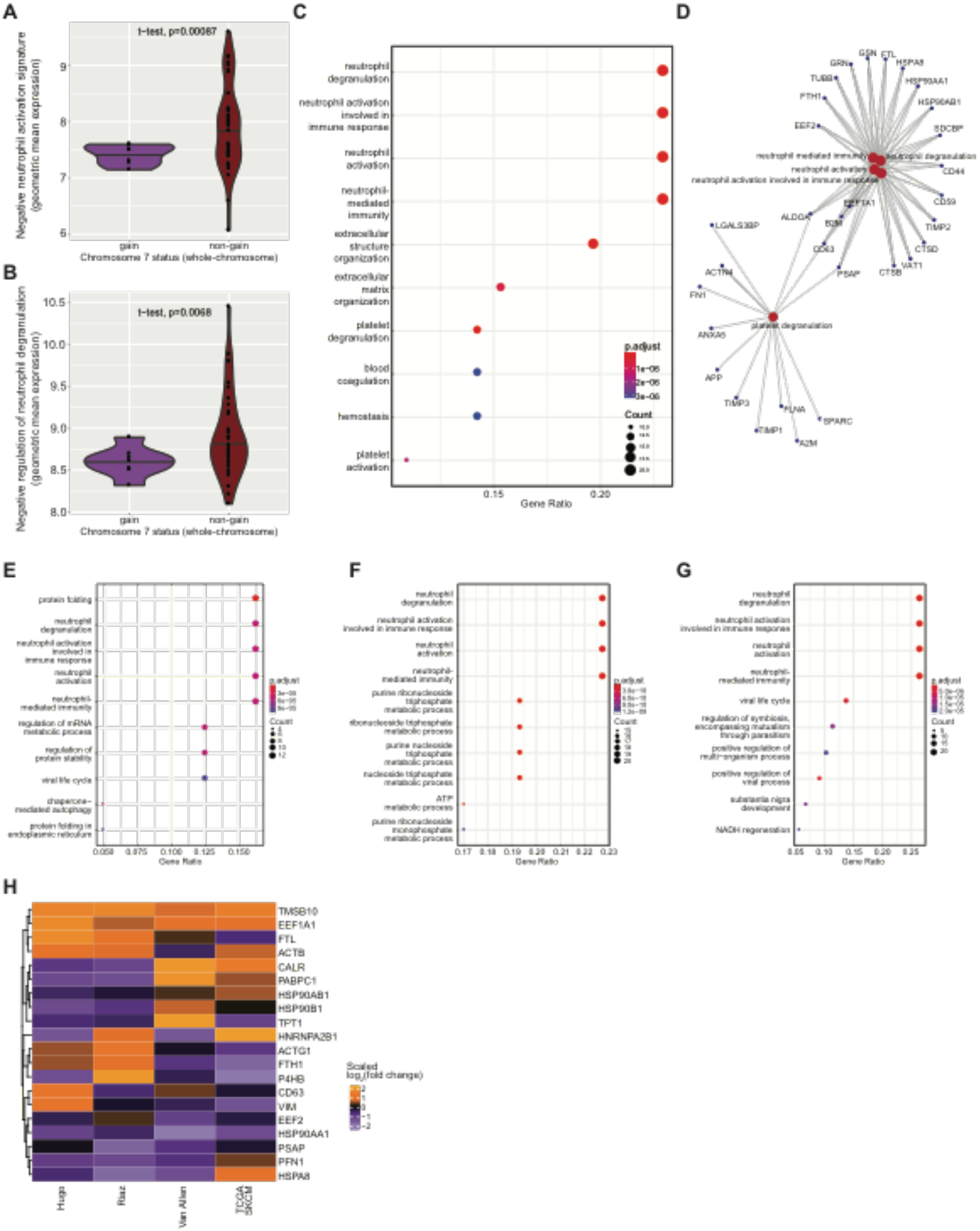

A
t-test, p=0.00087
Negative neutrophil activation signature
(geometric mean expression)
9
8
7
6
gain
non-gain
Chromosome 7 status (whole-chromosome)
B
t-test, p=0.0068
Negative regulation of neutrophil degranulation
(geometric mean expression)
10.5
10.0
9.5
9.0
8.5
8.0
gain
non-gain
Chromosome 7 status (whole-chromosome)
C
neutrophil degranulation
neutrophil activation involved in immune response
neutrophil activation
neutrophil-mediated immunity
extracellular structure organization
extracellular matrix organization
platelet degranulation
blood coagulation
hemostasis
platelet activation
0.15
0.20
Gene Ratio
p.adjust
1e-06
2e-06
3e-06
Count
D
GSN
FTL
GRN
HSPA8
TUBB
HSP90AA1
FTH1
HSP90AB1
EEF2
SDCBP
neutrophil mediated immunity
neutrophil degranulation
neutrophil activation
CD44
neutrophil activation involved in immune response
EEF1A1
CD59
ALDOA
B2M
TIMP2
LGALS3BP
CD63
CTSD
PSAP
CTSB
VAT1
ACTN4
FN1
platelet degranulation
ANXA5
APP
TIMP3
FLNA
SPARC
TIMP1
A2M
E
protein folding
neutrophil degranulation
neutrophil activation involved in immune response
neutrophil activation
neutrophil-mediated immunity
regulation of mRNA metabolic process
regulation of protein stability
viral life cycle
chaperone-mediated autophagy
protein folding in endoplasmic reticulum
0.050
0.075
0.100
0.125
0.150
Gene Ratio
p.adjust
Count
F
neutrophil degranulation
neutrophil activation involved in immune response
neutrophil activation
neutrophil-mediated immunity
purine ribonucleoside triphosphate metabolic process
ribonucleoside triphosphate metabolic process
purine nucleoside triphosphate metabolic process
nucleoside triphosphate metabolic process
ATP metabolic process
purine ribonucleoside monophosphate metabolic process
0.17
0.18
0.19
0.20
0.21
0.22
0.23
Gene Ratio
p.adjust
Count
G
neutrophil degranulation
neutrophil activation involved in immune response
neutrophil activation
neutrophil-mediated immunity
viral life cycle
regulation of symbiosis, encompassing mutualism through parasitism
positive regulation of multi-organism process
positive regulation of viral process
substantia nigra development
NADH regeneration
0.05
0.10
0.15
0.20
0.25
Gene Ratio
p.adjust
Count
H
TMSB10
EEF1A1
FTL
ACTB
CALR
PABPC1
HSP90AB1
HSP90B1
TPT1
HNRNPA2B1
ACTG1
FTH1
P4HB
CD63
VIM
EEF2
HSP90AA1
PSAP
PFN1
HSPA8
Scaled
log2(fold change)
2
1
0
−1
−2
Hugo
Riaz
Van Allen
TCGA SKCM

Figure 5: Chromosome 7 gain is associated with an unfavorable immune environment driven by neutrophil signatures that characterize non-responders to immune checkpoint blockade. (**A-B**) Chromosome 7 gain sub-regions of the melanoma mass progressing during anti-PD-1 therapy revealed lesser suppressive neutrophil signatures compared with sub-regions unaffected by chromosome 7 gains. Scores shown represent geometric mean log2-transformed counts of genes within the GO terms indicated (see also Supplementary Data 7). Plots include two-sample t test comparison with accompanying significance values. (**C**) Samples affected by whole-chromosome 7 gains within The Cancer Genome Atlas melanoma (SKCM) dataset revealed prominent differential upregulation of genes involved in neutrophil activation signatures revealed by gene ontology term enrichments. Gene ratios indicate the ratio of representation of input genes within the indicated GO term gene set relative to all queried GO term gene sets. Top enriched pathways are displayed after Benjamini-Hochberg correction for multiple testing. (**D**) Connected gene network of genes involved in major GO term enrichments within chromosome 7 gain TCGA SKCM samples. (**E-G**) Differentially expressed genes contrasting responders versus non-responders to immune checkpoint blockade agents in publicly available datasets revealed consistent GO term enrichments for neutrophil activation signatures; Top enriched pathways are displayed after Benjamini-Hochberg correction for multiple testing. (**E**) Van Allen, *et al.*, anti-CTLA-4 dataset, (**F**) Hugo, *et al.*, anti-PD-1 dataset, (**G**) Riaz, *et al.*, anti-PD-1 dataset. (**H**) Heatmap of genes found recurrently enriched across TCGA SKCM chromosome 7 gain and clinical non-responder samples showing relative enrichment of each gene across the datasets as a scaled value.

TABLES

Table 1. Differentially expressed genes confirm the activity of multiple immune subsets in regions of heavy immune infiltration. Top 50 most differentially expressed genes (n=25) upregulated, n=25 downregulated in heavy versus low immune infiltrate) between tumor sub-regions based on extent of immune infiltrate.

Gene	BaseMean	log$_2$FoldChange	lfcSE	stat	pvalue	padj
RP11.280H21.1	2.22	-4.34	1.65	-2.63	8.54E-03	4.95E-02
RP11.376M2.2	3.45	-3.93	1.12	-3.50	4.63E-04	5.46E-03
AL109763.2	1.75	-3.61	1.13	-3.18	1.47E-03	1.36E-02
AC093850.1	3.34	-3.51	1.18	-2.97	2.96E-03	2.29E-02
RP11.114H23.3	2.91	-3.43	0.97	-3.52	4.34E-04	5.16E-03
CTD.2651C21.3	1.51	-3.41	1.03	-3.31	9.34E-04	9.59E-03
RP11.307L14.2	1.88	-3.34	1.08	-3.10	1.91E-03	1.65E-02
ICAM5	1.66	-3.19	1.10	-2.89	3.86E-03	2.80E-02
RP11.29P20.1	12.71	-3.06	0.68	-4.51	6.46E-06	1.58E-04
RNF208	1.39	-2.89	1.07	-2.70	6.99E-03	4.30E-02
IL1RAPL2	12.62	-2.80	0.65	-4.32	1.54E-05	3.28E-04
EFNA3	1.98	-2.74	0.91	-3.00	2.71E-03	2.14E-02
GPR115	6.36	-2.59	0.69	-3.76	1.67E-04	2.42E-03
MAST1	9.13	-2.58	0.60	-4.33	1.50E-05	3.22E-04
RNU6.850P	2.17	-2.57	0.91	-2.82	4.79E-03	3.27E-02
RP11.191L17.1	5.79	-2.53	0.91	-2.79	5.31E-03	3.53E-02
AC007091.1	3.29	-2.40	0.88	-2.71	6.68E-03	4.17E-02
RP11.67M1.1	6.98	-2.39	0.58	-4.15	3.39E-05	6.41E-04
LINC00919	3.72	-2.33	0.84	-2.77	5.60E-03	3.67E-02
AC018742.1	16.53	-2.33	0.80	-2.91	3.59E-03	2.66E-02
SMYD1	4.20	-2.29	0.68	-3.35	8.14E-04	8.65E-03
RN7SL151P	4.88	-2.23	0.77	-2.88	3.99E-03	2.86E-02
RLBP1	27.57	-2.09	0.43	-4.82	1.46E-06	4.65E-05
CTD.3064H18.4	8.82	-2.01	0.70	-2.86	4.20E-03	2.97E-02
NDUFAF4P3	9.37	-2.00	0.46	-4.40	1.11E-05	2.49E-04
RP1.153P14.5	3.39	5.48	1.20	4.58	4.74E-06	1.22E-04
LTF	351.40	5.52	0.73	7.60	2.89E-14	4.03E-11
LINC00086	6.09	5.52	1.26	4.39	1.11E-05	2.50E-04
IGLV3.21	18.88	5.62	0.91	6.20	5.73E-10	8.57E-08
ADH1B	128.01	5.67	1.05	5.38	7.39E-08	4.12E-06
EPPK1	9.80	5.68	0.89	6.40	1.55E-10	3.18E-08
hsa.mir.5195	4.71	5.70	1.05	5.45	5.12E-08	3.03E-06
SLC16A9	6.97	5.73	1.16	4.92	8.62E-07	3.03E-05
ADIPOQ	300.86	5.73	0.73	7.86	3.72E-15	9.45E-12
RP11.89M16.1	4.91	5.75	1.00	5.76	8.42E-09	6.92E-07
KLHL14	19.27	5.83	0.82	7.10	1.23E-12	5.96E-10
PGBD4P1	5.38	5.90	1.05	5.62	1.96E-08	1.37E-06
MDS2	4.65	5.93	1.10	5.39	7.06E-08	3.99E-06
SAA2	11.85	5.99	0.97	6.17	6.69E-10	9.56E-08
DSC3	5.00	6.06	1.25	4.85	1.25E-06	4.11E-05
PCK1	26.46	6.19	1.08	5.75	9.17E-09	7.40E-07

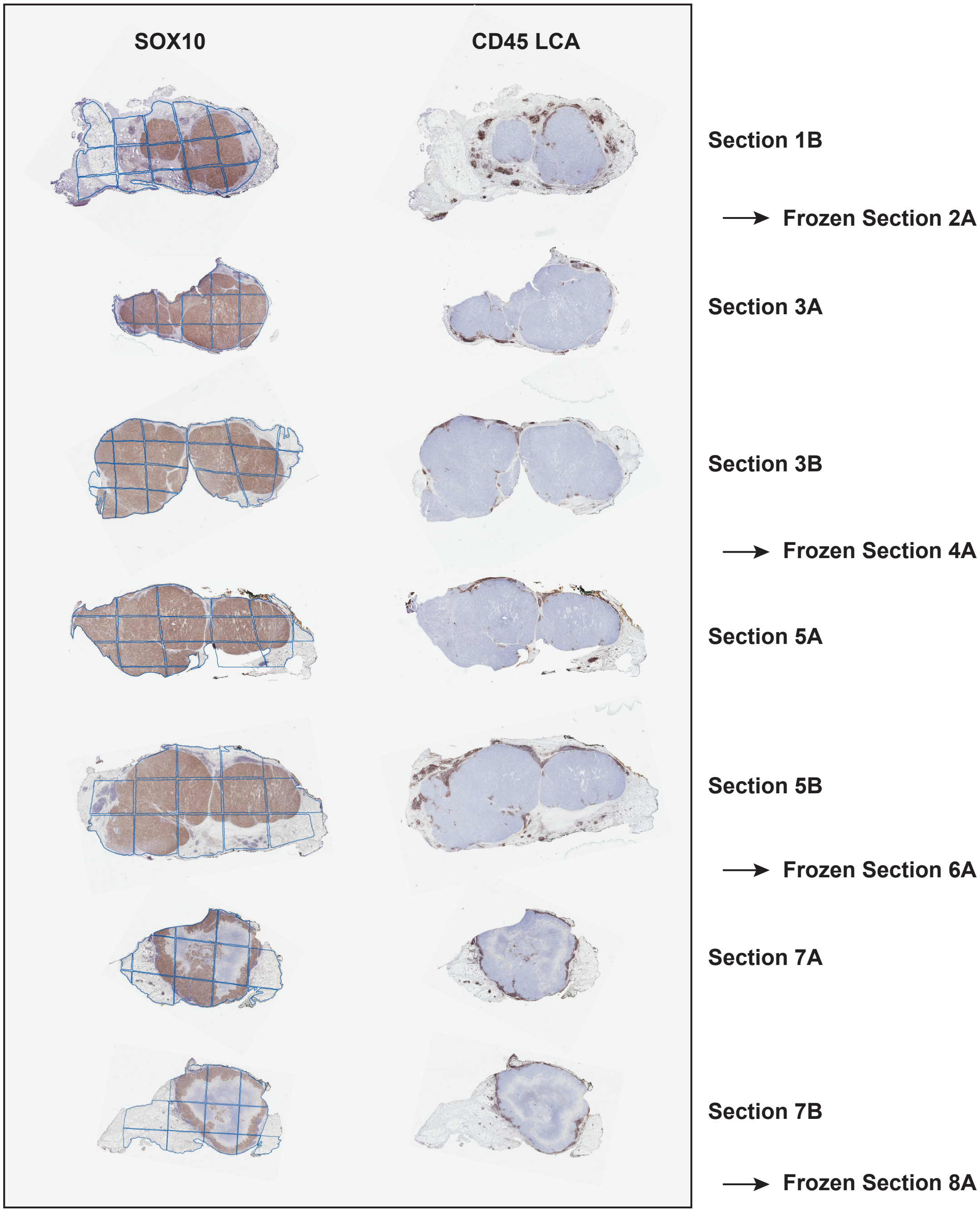
RP11.693J15.5	42.65	6.21	0.76	8.16	3.42E-16	1.52E-12
TNNT3	9.87	6.22	1.22	5.10	3.31E-07	1.40E-05
FCRL1	50.74	6.33	0.55	11.53	9.37E-31	2.50E-26
MAL2	7.46	6.64	1.05	6.30	2.91E-10	5.03E-08
CAPN6	7.91	6.72	1.10	6.11	9.98E-10	1.33E-07
RBP4	9.54	6.72	1.18	5.72	1.07E-08	8.49E-07
CR2	574.56	7.01	0.97	7.23	4.99E-13	3.24E-10
MFSD2A	17.06	7.02	1.10	6.38	1.82E-10	3.49E-08
GABRA2	39.51	7.54	1.07	7.07	1.54E-12	6.71E-10

Table 2: GO term enrichment reveals prominent T and B lymphocyte activation in immune infiltrated tumor samples. Top 20 most-enriched GO-BP terms within differentially expressed genes between samples displaying high versus low levels of leukocytic infiltrate.

GO Term	Description	Gene Ratio	Background Ratio	pvalue	p.adjust	qvalue
GO:0030098	lymphocyte differentiation	36/376	344/18493	7.01E-16	2.63E-12	2.13E-12
GO:0070661	leukocyte proliferation	31/376	283/18493	2.26E-14	4.25E-11	3.44E-11
GO:0002429	immune response-activating cell surface receptor signaling pathway	37/376	414/18493	4.14E-14	5.18E-11	4.20E-11
GO:0002768	immune response-regulating cell surface receptor signaling pathway	38/376	445/18493	7.79E-14	6.85E-11	5.55E-11
GO:0050851	antigen receptor-mediated signaling pathway	29/376	259/18493	9.12E-14	6.85E-11	5.55E-11
GO:0042113	B cell activation	31/376	303/18493	1.46E-13	9.14E-11	7.41E-11
GO:0050854	regulation of antigen receptor-mediated signaling pathway	15/376	57/18493	3.23E-13	1.73E-10	1.40E-10
GO:0046651	lymphocyte proliferation	28/376	264/18493	9.51E-13	4.47E-10	3.62E-10
GO:0032943	mononuclear cell proliferation	28/376	266/18493	1.14E-12	4.78E-10	3.87E-10
GO:0042100	B cell proliferation	17/376	91/18493	3.53E-12	1.33E-09	1.08E-09
GO:0042110	T cell activation	35/376	451/18493	1.22E-11	4.16E-09	3.37E-09
GO:0045785	positive regulation of cell adhesion	32/376	397/18493	3.63E-11	1.14E-08	9.20E-09
GO:1903039	positive regulation of leukocyte cell-cell adhesion	23/376	214/18493	8.22E-11	2.38E-08	1.92E-08
GO:0007159	leukocyte cell-cell adhesion	28/376	327/18493	1.59E-10	4.13E-08	3.35E-08
GO:0050870	positive regulation of T cell activation	22/376	202/18493	1.65E-10	4.13E-08	3.35E-08
GO:0050852	T cell receptor signaling pathway	19/376	150/18493	2.23E-10	5.24E-08	4.24E-08
GO:0050867	positive regulation of cell activation	30/376	384/18493	3.25E-10	7.18E-08	5.82E-08
GO:0022409	positive regulation of cell-cell adhesion	24/376	251/18493	3.62E-10	7.57E-08	6.13E-08
GO:0050855	regulation of B cell receptor signaling pathway	9/376	24/18493	5.39E-10	1.07E-07	8.64E-08
GO:0030183	B cell differentiation	17/376	125/18493	6.55E-10	1.22E-07	9.88E-08

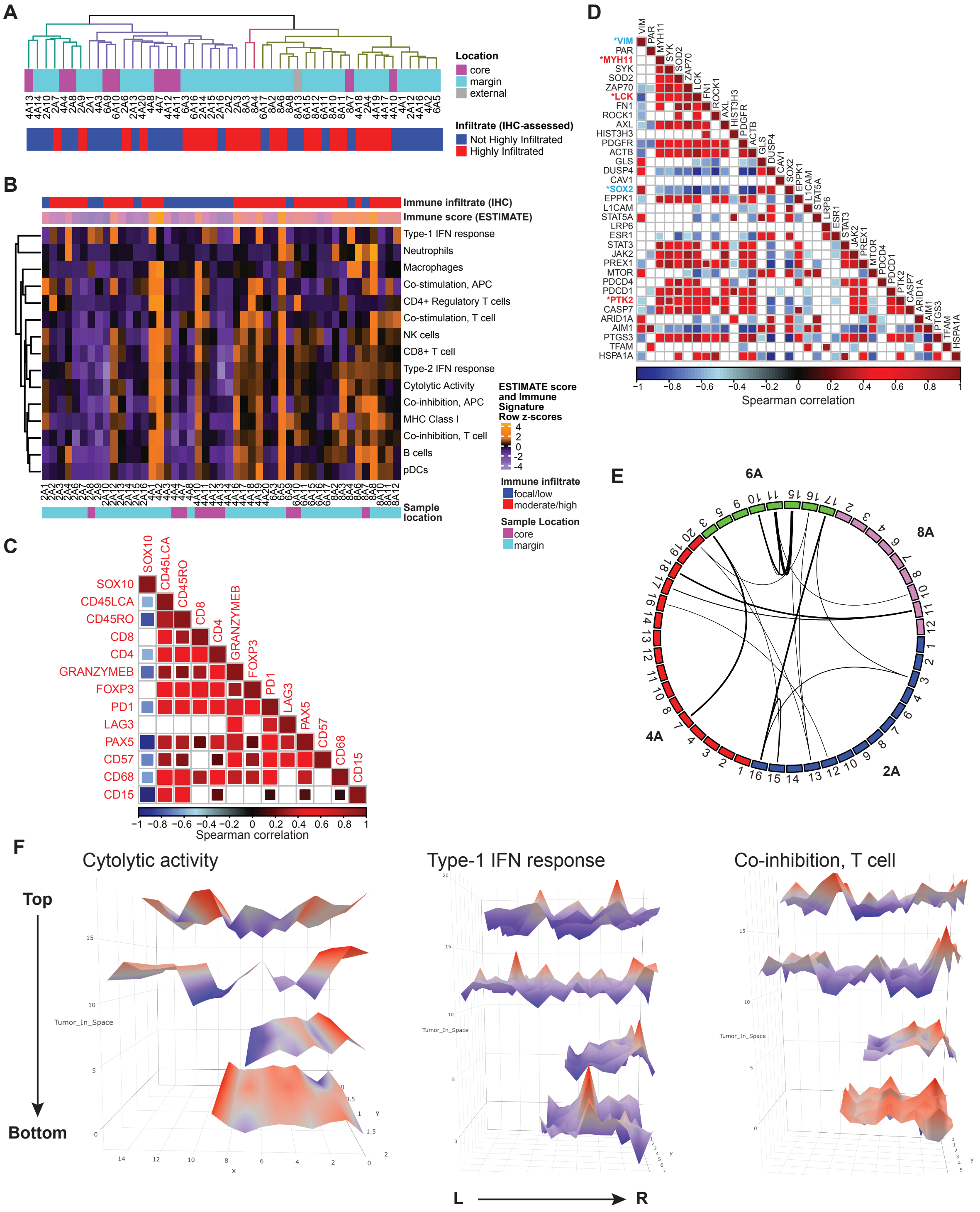
Supplemental Information

Supplementary Figure 1.


SOX10
CD45 LCA
Section 1B
Frozen Section 2A
Section 3A
Section 3B
Frozen Section 4A
Section 5A
Section 5B
Frozen Section 6A
Section 7A
Section 7B
Frozen Section 8A

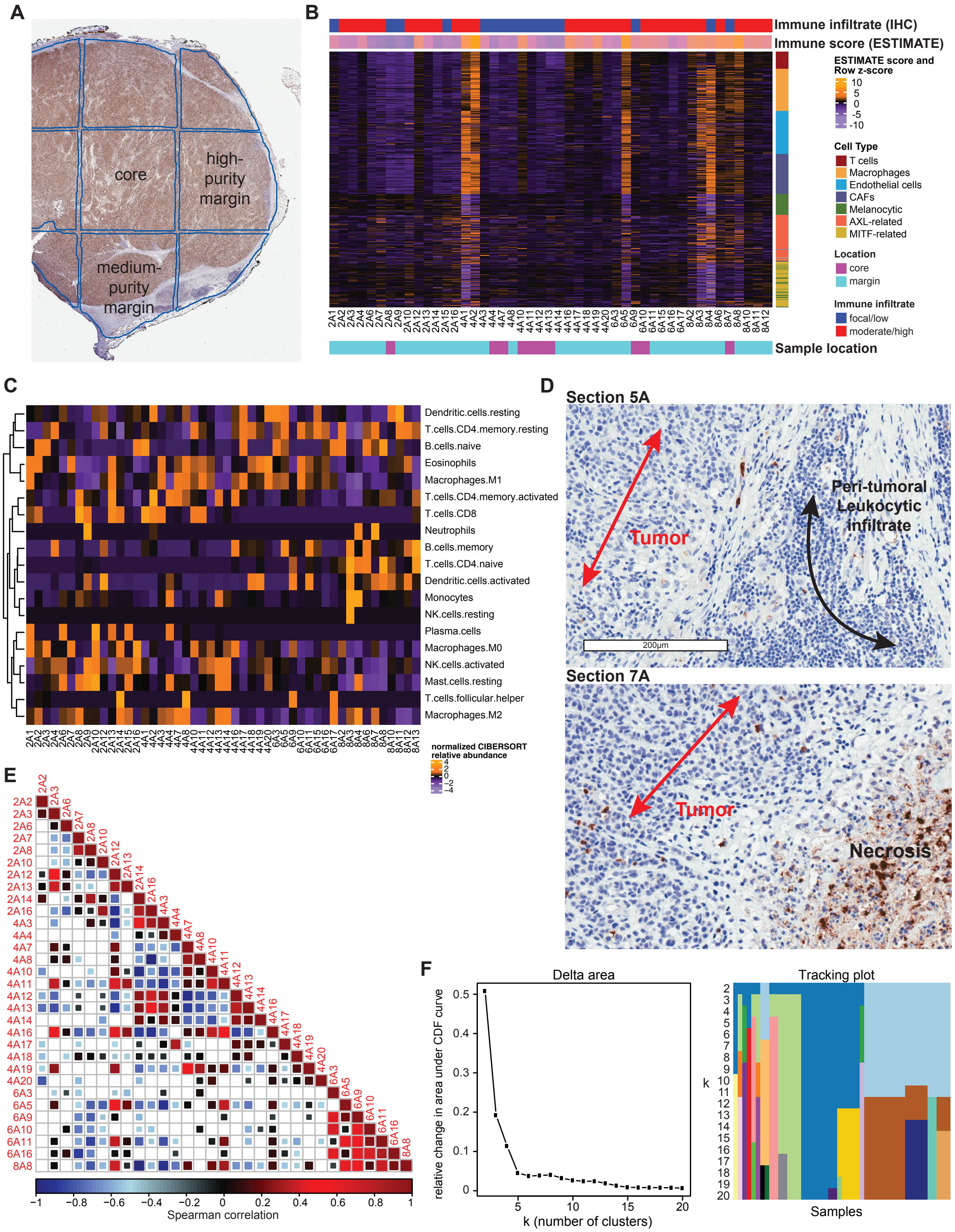
Supplementary Figure 1: Matched immunohistochemistry sections were used to infer tumor and immune composition of genomic sections. Shown are IHC stains for tumor (SOX10) and leukocytes (CD45LCA) from FFPE (odd-numbered) tumor sections, providing information from above and below each intercalating frozen (even-numbered) tumor section in order to infer tumor microenvironment composition based on multiple IHC stains. Sub-division of sections into sub-regions in order to study intratumoral heterogeneity was performed on a section- by-section basis on frozen sections, and analogous boundaries applied to adjacent FFPE sections as indicated by blue lines on the SOX10-stained sections.

Figure 2.


A
Location
core
margin
external
Infiltrate (IHC-assessed)
Not Highly Infiltrated
Highly Infiltrated
B
Immune infiltrate (IHC)
Immune score (ESTIMATE)
Type-1 IFN response
Neutrophils
Macrophages
Co-stimulation, APC
CD4+ Regulatory T cells
Co-stimulation, T cell
NK cells
CD8+ T cell
Type-2 IFN response
Cytolytic Activity
Co-inhibition, APC
MHC Class I
Co-inhibition, T cell
B cells
pDCs
ESTIMATE score and Immune Signature Row z-scores
4
2
0
-2
-4
Sample location
Immune infiltrate
focal/low
moderate/high
Sample Location
core
margin
C
SOX10
CD45LCA
CD45RO
CD8
CD4
GRANZYMEB
FOXP3
PD1
LAG3
PAX5
CD57
CD68
CD15
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
Spearman correlation
D
*VIM
PAR
*MYH11
SYK
SOD2
ZAP70
*LCK
FN1
ROCK1
AXL
HIST3H3
PDGFR
ACTB
GLS
DUSP4
CAV1
*SOX2
EPPK1
L1CAM
STAT5A
LRP6
ESR1
STAT3
JAK2
PREX1
MTOR
PDCD4
PDCD1
*PTK2
CASP7
ARID1A
AIM1
PTGS3
TFAM
HSPA1A
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1
Spearman correlation
E
6A
8A
4A
2A
F
Cytolytic activity
Type-1 IFN response
Co-inhibition, T cell
Top
Bottom
Tumor_In_Space
L
R

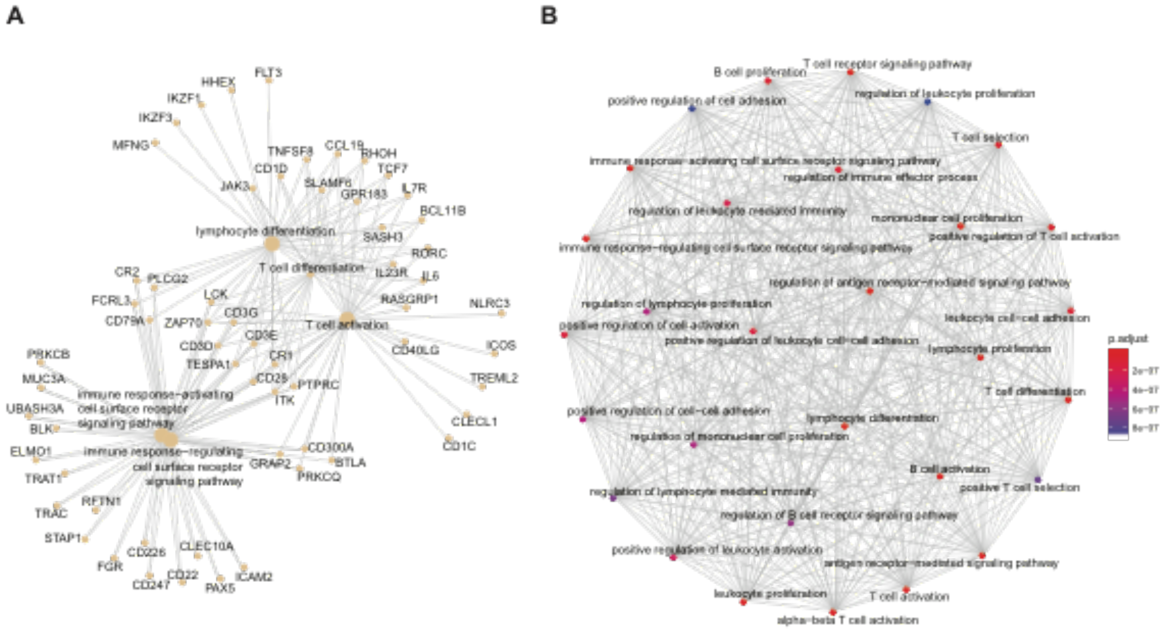
Supplementary Figure 2: Inter-tumoral analyses reveal longitudinal copy number evolution. (A) Copy number changes across the on-PD-1 inhibitor lesion demonstrate notable intra-tumor heterogeneity, and additional longitudinal/inter-tumoral heterogeneity in comparison to the treatment-naïve and post-PD-1 inhibitor lesions. (B) Chromosome 10 copy number alterations spanning metachronous tumors indicate step-wise losses (indicated by red bar, relative to a zero baseline) in tumors sampled at treatment-naïve, on-PD-1 inhibitor (regions 2A1, 2A16, 4A12, demonstrating ITH) and post-PD-1 inhibitor time points. Data are log2(R) where R is the probe intensity expressed as (observed intensity/reference intensity).

Supplementary Figure 3.



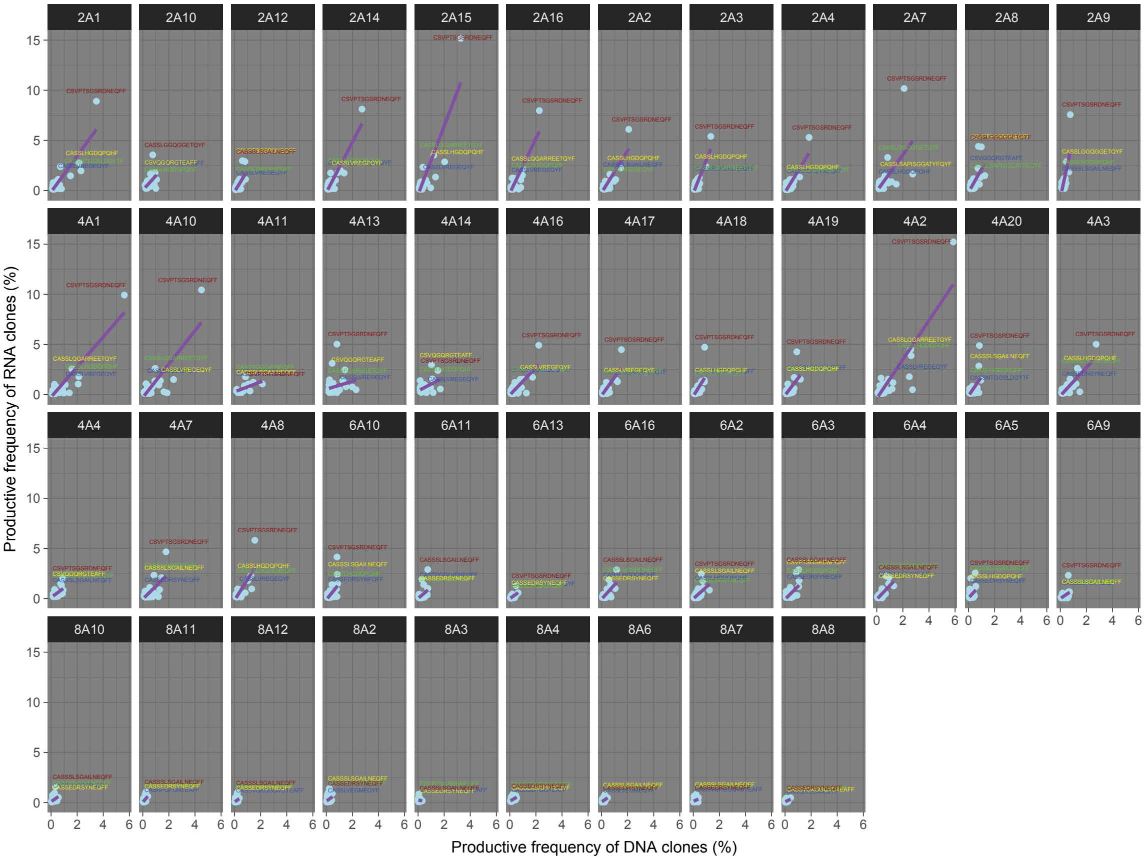
Supplementary Figure 3: Deconvolution of melanoma and immune cell transcriptomic and proteomic signatures. (A) Example histologic appearances (SOX10 stain) of subregions designated as core, high-purity margin and medium-purity margin. (B) Supervised clustering (in spatial order) of the transcriptomic profile of samples derived from the on-PD-1 inhibitor tumor lesion, highlighting tumor-derived and immune cell subset-derived signatures. IHC-based immune infiltrate and ESTIMATE immune scores (top) and IHC-based sample location (bottom) are indicated. (C) Heatmap of immune infiltrate composition deconvoluted from transcriptomic data using CIBERSORT. (D) Example of low-density (top) and high-density (bottom) CD15+ staining as a marker for the presence of neutrophils. The majority of CD15 staining was observed within necrotic regions of FFPE slice 7. (E) Sample inter-correlation matrix based on reverse-phase protein array quantification of protein expression, illustrating dispersed protein expression patterns throughout sub-regions of the on-PD-1 inhibitor tumor. (F) Delta area (left) and tracking plot (right) of consensus clustering of samples using transcriptome data to identifying the optimal number of sample clusters that captures the majority of intra-tumoral heterogeneity signatures. A total of 4 clusters provided the greatest balance between information and similar cluster size, whilst >4 cluster-solutions contributed lower-yield clusters containing few samples each.

Supplementary Figure 4.


A
B
FLT3
HHEX
IKZF1
IKZF3
MFNG
TNFSF8
CCL19
RHOH
CD1D
TCF7
JAK3
SLAMF6
GPR183
IL7R
BCL11B
lymphocyte differentiation
SASH3
RORC
CR2
PLCG2
T cell differentiation
IL23R
IL6
FCRL3
LCK
RASGRP1
NLRC3
CD79A
ZAP70
CD3G
T cell activation
CD3E
CD3D
CD40LG
ICOS
PRKCB
TESPA1
CR1
MUC3A
CD28
PTPRC
TREML2
immune response−activating cell surface receptor signaling pathway
ITK
UBASH3A
BLK
CLECL1
CD300A
CD1C
ELMO1
immune response−regulating cell surface receptor signaling pathway
GRAP2
BTLA
TRAT1
PRKCQ
RFTN1
TRAC
STAP1
CD226
CLEC10A
FGR
CD22
ICAM2
CD247
PAX5
B cell proliferation
T cell receptor signaling pathway
positive regulation of cell adhesion
regulation of leukocyte proliferation
T cell selection
immune response−activating cell surface receptor signaling pathway
regulation of immune effector process
regulation of leukocyte mediated immunity
mononuclear cell proliferation
positive regulation of T cell activation
immune response−regulating cell surface receptor signaling pathway
regulation of antigen receptor−mediated signaling pathway
regulation of lymphocyte proliferation
leukocyte cell−cell adhesion
positive regulation of cell activation
positive regulation of leukocyte cell−cell adhesion
lymphocyte proliferation
p.adjust
2e−07
4e−07
6e−07
8e−07
T cell differentiation
positive regulation of cell−cell adhesion
lymphocyte differentiation
regulation of mononuclear cell proliferation
B cell activation
positive T cell selection
regulation of lymphocyte mediated immunity
regulation of B cell receptor signaling pathway
positive regulation of leukocyte activation
antigen receptor−mediated signaling pathway
leukocyte proliferation
T cell activation
alpha−beta T cell activation

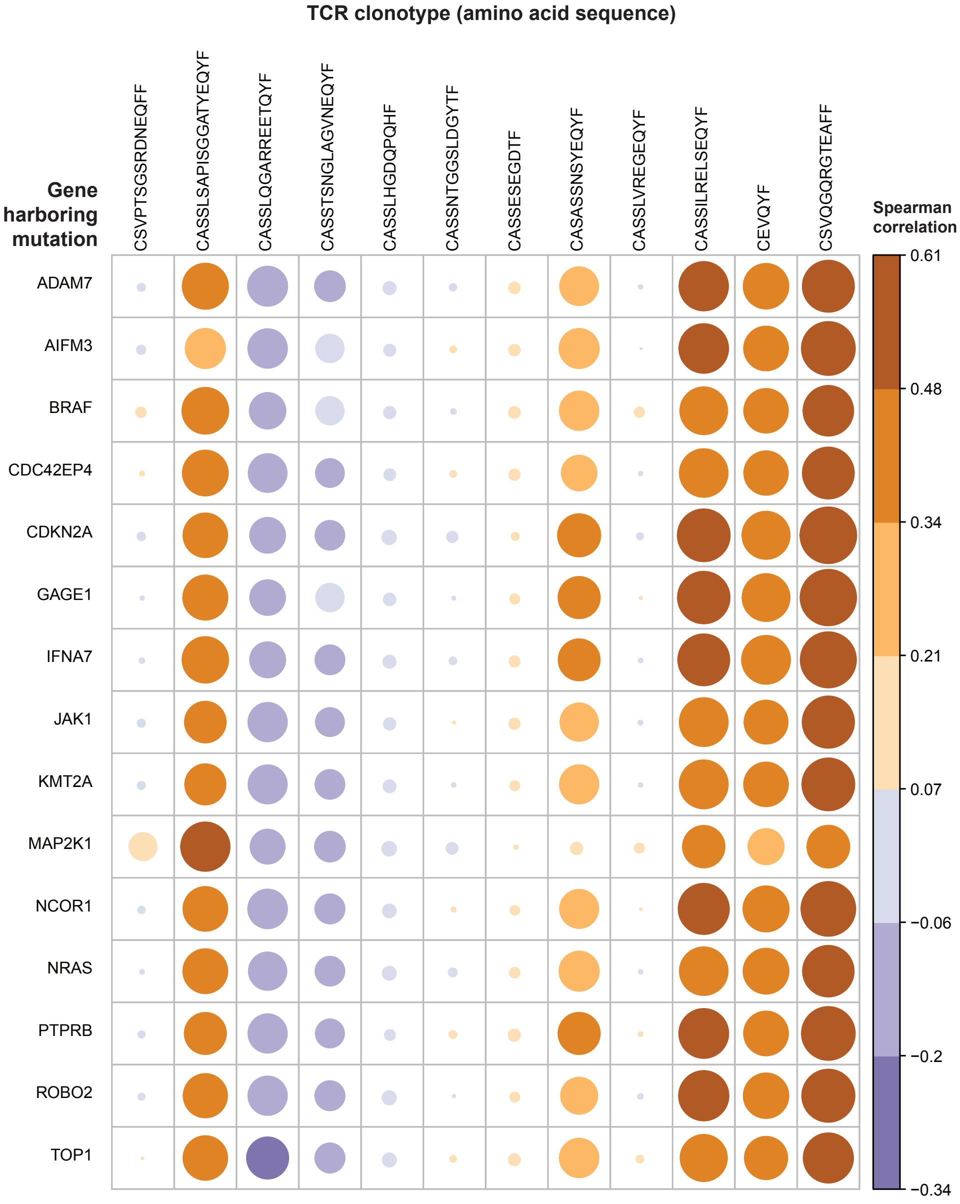
Supplementary Figure 4: Transcriptional heterogeneity at margin sites. (A) Gene connection network of differentially-expressed genes between samples located at the tumor margin comparing those with high versus low immune infiltrates. (B) Functional annotation network of differentially-expressed genes between high versus low immune infiltrate tumor margin samples, showing enrichment for T and B lymphocyte pathways and functions.

Supplementary Figure 5.



Supplementary Figure 5: Identification and distribution of dominant TCR clonotypes. TCR clonal activation inferred from matched TCR-Vβ DNA and RNA sequencing identifies dominant clonotypes and their spatial distribution. Data represent productive frequencies within the total TCR repertoire derived from either DNA-based or RNA-based sequencing, expressed as percentages.

Supplementary Figure 6.



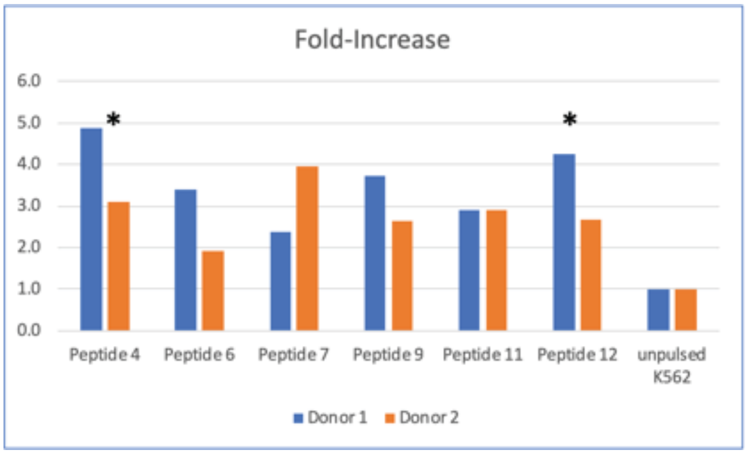
Supplementary Figure 6: Associations between high-abundance TCR clonotypes and truncal tumor mutations reveal probable distinct tumor-specific and bystander populations. Correlations between abundance of high-frequency TCR clonotypes (defined at the amino acid level) and a core set of truncal mutations (variant allele frequency) identified throughout all sub-regions of the on-PD-1 inhibitor tumor. The strength and direction of correlation is indicated by color, whilst the size of each circle indicates the statistical significance (larger indicates more highly significant), with all values meeting the criterion of adjusted p-value < 0.05

Supplementary Figure 7

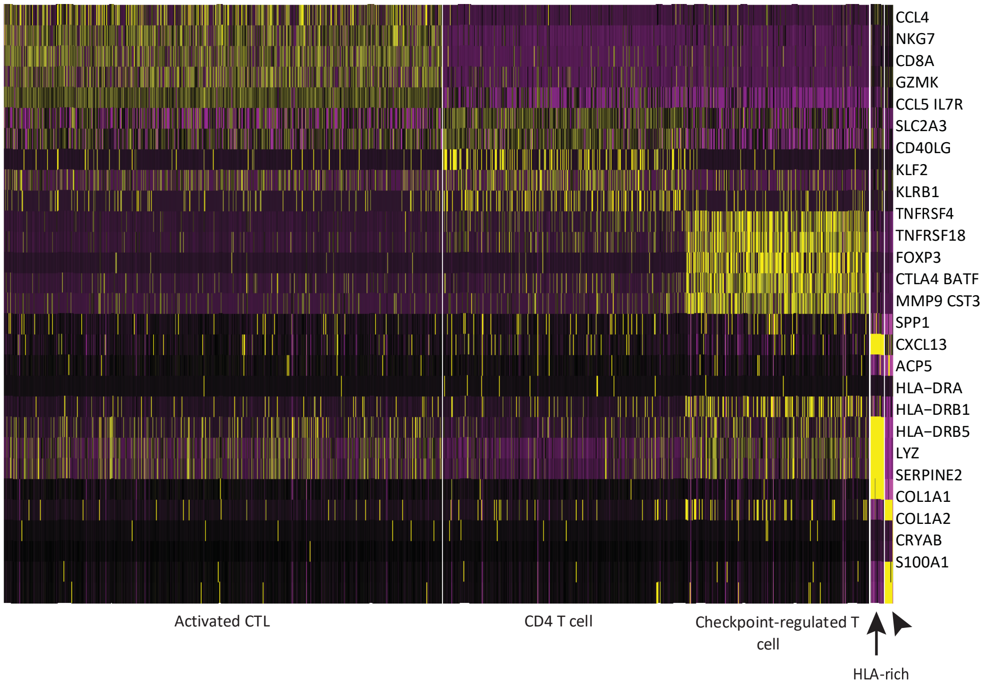
A

	Donor 1	Donor 2
Peptide 4	1.9	0.74
Peptide 6	1.33	0.46
Peptide 7	0.93	0.95
Peptide 9	1.45	0.63
Peptide 11	1.13	0.70
Peptide 12	1.66	0.64
unpulsed K562	0.39	0.24

B

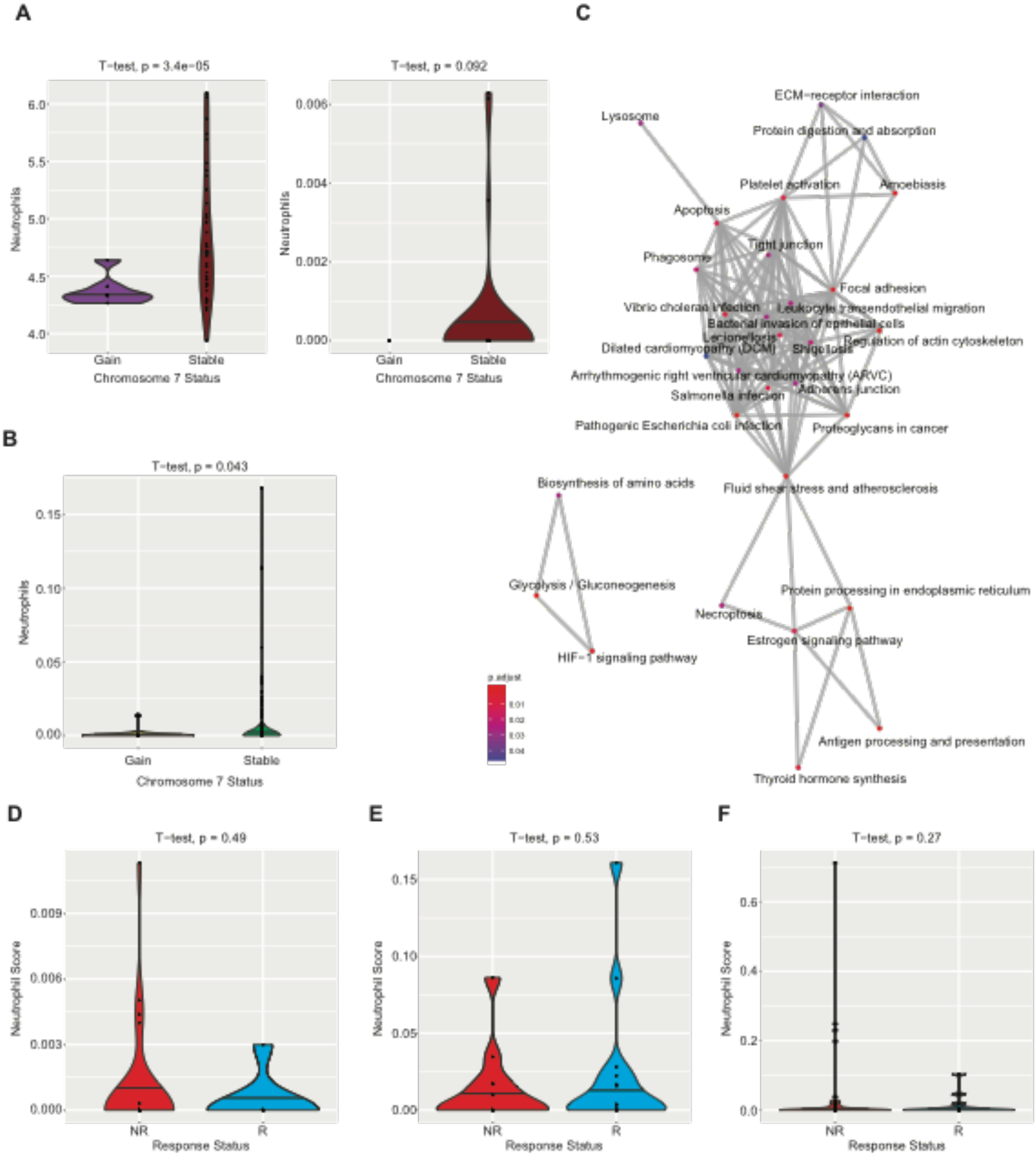


C



Supplementary Figure 7: Single cell sequencing identifies distinct T cell phenotypes. (A-B) Detection of epitope specific CD8+ T cells in the PBMC of HLA-A*0301 healthy donors. Potential epitope specific CD8+ T cells in PBMC were stimulated three times at 7-day intervals by co-culturing with HLA-A*0301 expressing K562 (A3-K562) pulsed with overlapping peptides (peptides 1-12) in the presence of IL-21. Epitope specificity of sensitized CD8+ T cells was determined by co-culturing cells with peptide pulsed K562 overnight and performing a standard intracellular IFN-γ production assay. Cells co-cultured with non-pulsed K562 cells served as background (negative) controls. Treatment with PMA and ionomycin served as a positive control for the assay. The table (A) lists the highest responding peptide candidates as % IFN-γ+ CD8+ T cells among total CD8+ T cells (baseline of 0.39% and 0.24% from Donors 1 and 2, respectively). Fold-increase in peptide reactivity above baseline (baseline = 1.0) is shown in the column chart (B) at right. (C) Heatmap of the most differentially-expressed genes between tumor infiltrating lymphocyte clusters defined by tSNE analysis.

Supplementary Figure 8.



Supplementary Figure 8: Neutrophil estimation and related Gene Ontology (GO)-term and KEGG pathway enrichments in clinical samples based on chromosome 7 gains. (A). Neutrophil quantification within the on-PD-1 inhibitor samples (lesion 2) comparing sites by chromosome 7 copy number status as determined from MCP counter (left) and CIBERSORT (right). (B) Neutrophil content derived from CIBERSORT within TCGA SKCM samples grouped by chromosome 7 copy number status. (C). KEGG pathways enriched in melanoma samples harboring chromosome 7 gains revealed numerous pathways involved in reaction to bacterial infections, phagosome and lysosome formation, and antigen processing, consistent with the highly neutrophil activation-dominant GO term enrichments observed. (D-F) Neutrophil content derived from CIBERSORT within clinical immune checkpoint blockade-treated samples grouped by reported response status (NR=non-responder, R=responder) from the Van Allen anti-CTLA-4 cohort (D) and the Hugo (E) and Riaz (F) anti-PD-1 cohorts.

Supplemental Data

Supplemental Data 1

Slice	**Sub-Region**	**Targeted Seq**	**WES**	**CNA**	**Methylation**	**RNAseq**	**RNAseq Total Library Size**	**TCRseq**	**TCRseq Total Templates**	**RPPA**	**IHC Slice Above**	**IHC Slice Below**
2	1	1	0	1	1	1	1.91E+08	1	34466	0	1	1
2	2	1	1	1	1	1	1.73E+08	1	9866	1	1	1
2	3	1	0	1	1	1	1.59E+08	1	6185	1	1	1
2	4	0	0	0	1	1	1.52E+08	1	41867	0	1	1
2	6	1	0	1	1	1	1.5E+08	1	2853	1	0	1
2	7	1	0	1	1	1	1.74E+08	1	3765	1	1	1
2	8	1	0	1	1	1	1.41E+08	1	3320	1	1	1

2	9	1	0	1	1	1	1.65E+08	1	4720	0	1	1
2	10	1	0	1	1	1	1.66E+08	1	2128	1	1	1
2	12	1	0	1	1	1	1.65E+08	1	17206	1	0	1
2	13	1	0	1	1	1	1.64E+08	1	7358	1	1	1
2	14	1	0	1	1	1	1.53E+08	1	4499	1	1	1
2	15	1	0	1	1	1	1.52E+08	1	16519	0	1	1
2	16	1	0	1	1	1	1.59E+08	1	6926	1	1	0
4	1	0	0	0	0	1	1.72E+08	1	74724	0	1	1
4	2	1	0	1	1	1	1.86E+08	1	54121	0	1	1
4	3	1	0	1	1	1	1.78E+08	1	23599	1	1	1
4	4	1	0	1	1	1	1.6E+08	1	10111	1	1	1
4	7	1	0	1	1	1	1.84E+08	1	11896	1	1	1
4	8	1	0	1	1	1	1.18E+08	1	12110	1	1	0
4	10	1	0	1	1	1	1.24E+08	1	35207	1	1	1
4	11	1	1	1	1	1	1.59E+08	1	27593	1	1	1
4	12	1	0	1	1	1	1.45E+08	1	7882	1	1	1
4	13	1	0	1	1	1	1.26E+08	1	3072	1	1	1
4	14	1	0	1	1	1	1.45E+08	1	3870	1	1	1
4	16	1	0	1	1	1	1.23E+08	1	39623	1	1	1
4	17	1	0	1	1	1	1.36E+08	1	58254	1	1	1
4	18	1	0	1	1	1	1.67E+08	1	86355	1	1	1
4	19	1	0	1	1	1	1.21E+08	1	103809	1	1	1
4	20	1	0	1	1	1	1.2E+08	1	29635	1	1	1
6	2	1	0	1	0	0	NA	1	184376	0	1	1
6	3	1	1	1	1	1	1.23E+08	1	14280	1	1	1
6	4	0	0	0	0	0	NA	1	54659	0	1	1
6	5	1	0	1	1	1	1.6E+08	1	28469	1	1	1

6	9	1	0	1	1	1	1.54E+08	1	47906	1	1	1
6	10	1	0	1	1	1	1.31E+08	1	19808	1	1	1
6	11	1	0	1	1	1	1.23E+08	1	82951	1	1	1
6	13	0	0	0	0	0	NA	1	21161	0	0	0
6	14	0	0	0	0	0	NA	1	9708	0	0	0
6	15	1	0	1	1	1	1.3E+08	1	55440	0	1	1
6	16	1	1	1	1	1	1.33E+08	1	6861	1	1	1
6	17	1	0	1	1	1	1.41E+08	1	65087	0	1	1
8	2	1	0	1	1	1	1.83E+08	1	21750	0	1	NA
8	3	1	0	1	1	1	1.57E+08	1	25144	0	1	NA
8	4	0	0	0	0	1	1.62E+08	1	26248	0	1	NA
8	6	0	0	0	0	1	1.52E+08	1	57751	0	1	NA
8	7	0	0	0	0	1	1.09E+08	1	34666	0	1	NA
8	8	1	0	1	1	1	1.52E+08	1	65498	1	1	NA
8	10	0	0	0	0	1	1.53E+08	1	26098	0	1	NA
8	11	0	0	0	0	1	1.35E+08	1	63907	0	1	NA
8	12	1	0	1	0	1	1.23E+08	1	46807	0	1	NA
8	13	0	0	0	0	0	NA	1	109336	0	0	NA

Supplemental Data 2

ABL1	ACVR1B	ACVR2A	AKT1
AKT2	AKT3	ALK	AMER1
APC	AR	ARAF	ARID1A
ARID1B	ARID2	ASXL1	ATM
ATR	ATRX	AURKA	AURKB
AXIN1	BAP1	BCL2	BCOR
BIRC2	BRAF	BRCA1	BRCA2
BTK	CARD11	CASP8	CBL
CCND1	CCND3	CCNE1	CD79A
CD79B	CDC27	CDC73	CDH1
CDK4	CDK6	CDKN1B	CDKN2A
CEBPA	CHEK1	CHEK2	CIC
CREBBP	CSF1R	CTCF	CTNNB1
CYLD	DAXX	DDR2	DICER1
DNMT3A	EGFR	ELF3	EP300
EPCAM	EPHA3	EPHA5	ERBB2
ERBB3	ERBB4	ERCC2	ERCC3
ERCC4	ERCC5	ESR1	ETV1
EZH2	FADD	FAM123B	FANCA
FANCD2	FBXW7	FGFR1	FGFR2
FGFR3	FGFR4	FH	FLT1
FLT3	FLT4	FOXA1	FOXL2
GABRA6	GATA1	GATA2	GATA3
GNA11	GNAQ	GNAS	H3F3A
HIST1H3B	HNF1A	HRAS	IDH1
IDH2	IGF1R	JAK1	JAK2
JAK3	KDM5C	KDM6A	KDR
KEAP1	KIT	KRAS	MAP2K1
MAP2K2	MAP2K4	MAP3K1	MAP3K4
MAPK1	MCL1	MDM2	MED12
MEN1	MET	MITF	MLH1
MLL2	MLL3	MPL	MSH2
MSH6	MTOR	MYD88	NBN
NCOR1	NF1	NF2	NFE2L2
NKX2-1	NOTCH1	NOTCH2	NOTCH3
NOTCH4	NPM1	NRAS	NSD1
NTRK3	PALB2	PAX5	PBRM1
PDGFRA	PDGFRB	PHF6	PIK3CA

PIK3CG	PIK3R1	POLE	PPP2R1A
PRDM1	PRG4	PTCH1	PTEN
PTPN11	RAC1	RAF1	RB1
RET	RICTOR	RNF43	ROS1
RUNX1	SDHB	SDHC	SDHD
SETBP1	SETD2	SF3B1	SMAD2
SMAD3	SMAD4	SMARCB1	SMO
SOCS1	SOX9	SPEN	SPOP
SRC	SRSF2	STAG2	STAT3
STK11	SUFU	SYK	TERT
TET2	TET2	TGFBR2	TNFAIP3
TOP1	TOP2A	TP53	TSC1
TSC2	TSHR	U2AF1	VEGFA
VHL	WHSC1L1	WT1	XPO1
AJUBA	ABL2	AKTip	AS160 = TBC1D4
AXIN2	AXL	B2M	BCL11A
CCND2	CD274/PDL1	CDK12	CDKN2C
COL2A1	CTLA4	CYP2C19	DDX3X
EPHA2	FAK = PTK2	FTO	GSK3B
H3F3B	HLA-A	HSP90AB1	IL7R
LRP1B	MAP3K13	MLL	MST1
MST1R	MUTYH	NTRK1	PDCD1
PLCG1	PMS2	PPM1D	PPP1R3A
PREX2	PTPRB	RAD51	RAD51C
RARA	RPS6KB1	RPTOR	RUNX1T1
SMARCA2	SMARCA4	SMARCD1	SMC1A
SMC3	SOS1	STK19	TBX3
TGFb1	TGFBR1	TNF	ZRSR2

Supplemental Data 3

gene	chr	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
RPS6KA6	X	37.4999506	-2.1415285	0.50671639	-4.2262861	2.38E-05	0.02845718
MPZ	1	125.857082	-1.6860535	0.3875135	-4.3509542	1.36E-05	0.02059679
OLFML2A	9	218.745644	-1.2022087	0.28511412	-4.2165879	2.48E-05	0.02845718
TUBB4B	9	540.121254	-0.524472	0.13023308	-4.0271799	5.64E-05	0.04130038
PARGP1	10	630.347187	0.29025227	0.07285094	3.98419412	6.77E-05	0.04612185
SHOC2	10	7167.3413	0.29700997	0.07547847	3.93502905	8.32E-05	0.04977879
ABI1	10	10852.9821	0.33369281	0.07932704	4.2065454	2.59E-05	0.02845718
PLEKHA1	10	2732.7107	0.352468	0.07942257	4.43788208	9.08E-06	0.01660221
OPTN	10	7387.2808	0.35745134	0.07818184	4.57205047	4.83E-06	0.01192585
LCOR	10	6014.13066	0.38570136	0.09829803	3.9237954	8.72E-05	0.04977879
FRA10AC1	10	2210.49936	0.39211203	0.06462872	6.06714841	1.30E-09	8.57E-06
UPF2	10	14157.0275	0.41163467	0.09174381	4.48678414	7.23E-06	0.01587042

KAT6B	10	5048.8543	0.42190343	0.09618556	4.38634898	1.15E-05	0.01897522
DCLRE1C	10	911.20574	0.42336868	0.10860386	3.89828374	9.69E-05	0.04977879
PFKFB3	10	5546.28019	0.45343985	0.10826334	4.18830454	2.81E-05	0.02921993
RSU1	10	4459.76695	0.46707871	0.07595307	6.14957002	7.77E-10	7.67E-06
EPC1	10	4705.87388	0.49195946	0.10107725	4.86716331	1.13E-06	0.0037273
FAM21A	10	229.288176	0.4928333	0.10575672	4.66006605	3.16E-06	0.00892057
ZNF22	10	1746.6713	0.49524957	0.12022173	4.11946814	3.80E-05	0.03572162
RUFY2	10	2014.46815	0.52835951	0.0967796	5.45941016	4.78E-08	0.00023592
C10ORF68	10	216.023248	0.80654431	0.19493223	4.13756256	3.51E-05	0.03466973
RP11.849H4.4	NA	43.9883109	1.40544588	0.36085134	3.89480577	9.83E-05	0.04977879
RP11.819M15.1	NA	54.0768145	1.65746976	0.40561605	4.08630223	4.38E-05	0.03764449
BEX5	X	31.5451148	2.80981838	0.71366661	3.93715828	8.25E-05	0.04977879
WHAMMP2	15	55.3270381	2.87030043	0.71035836	4.04063723	5.33E-05	0.04050038

STK32B	4	319.491492	2.92180985	0.65893799	4.43411959	9.24E-06	0.01660221
RGS14	5	10.6096607	2.92843768	0.72474123	4.04066658	5.33E-05	0.04050038
RP11.876N24.2	NA	8.7388639	3.50522458	0.89706869	3.9074205	9.33E-05	0.04977879
ABCA10	17	58.0983132	3.54853389	0.87604967	4.05060811	5.11E-05	0.04050038
GABRE	X	80.5302861	3.78448369	0.96496585	3.92188354	8.79E-05	0.04977879
RP11.214O1.3	NA	22.4697341	4.07825492	1.0374881	3.93089318	8.46E-05	0.04977879
BLK	8	62.7432839	4.20400218	1.02423221	4.10454012	4.05E-05	0.03637609
SELM	22	15.7193338	4.42435986	1.02089901	4.33378798	1.47E-05	0.02068037
AC093642.3	NA	27.770349	4.87752991	0.96285377	5.06570162	4.07E-07	0.00160758
GRIP2	3	11.5799925	5.07934722	1.26702613	4.0088733	6.10E-05	0.04304195
NLRP7	19	9.88063102	5.33458412	1.3629872	3.91389158	9.08E-05	0.04977879
MUC12	7	9.76824145	5.76379138	1.4763363	3.90411817	9.46E-05	0.04977879
TNNT3	11	20.9550967	7.38027558	1.74818742	4.22167298	2.42E-05	0.02845718

MT1B	16	37.923291	22.480718	2.49498124	9.01037556	2.05E-19	4.06E-15

Supplemental Data 4

Gene	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
FCRL1	50.74110637	6.333461098	0.5493271	11.5294896	9.37E-31	2.50E-26
DGKE	230.3780947	2.34026609	0.255627	9.15500366	5.44E-20	7.24E-16
GCNT4	72.35086778	3.041550888	0.34028067	8.93835925	3.95E-19	3.51E-15
RP11.553K8.5	103.3282115	2.988765476	0.36406353	8.20946125	2.22E-16	1.46E-12
RP11.347P5.1	202.3039542	2.425309003	0.29634874	8.18396925	2.75E-16	1.46E-12
RP11.693J15.5	42.64742901	6.213668502	0.76170599	8.15756812	3.42E-16	1.52E-12
PLCB4	110.1348642	3.276392776	0.40742825	8.04164365	8.86E-16	3.37E-12
CADM3	149.2069446	5.174369166	0.65266017	7.92812152	2.22E-15	7.41E-12
CLU	1280.537517	4.685860565	0.59390808	7.88987505	3.02E-15	8.95E-12
MS4A1	398.7408259	4.50496317	0.57329504	7.85801872	3.90E-15	9.45E-12
ADIPOQ	300.8602216	5.734927332	0.72925982	7.86403852	3.72E-15	9.45E-12

ICOS	238.4482965	2.325699191	0.29710299	7.82792258	4.96E-15	9.99E-12
FAM129C	28.35170317	5.067201566	0.64711925	7.83039843	4.86E-15	9.99E-12
S1PR1	1684.920618	2.409973953	0.30815183	7.82073563	5.25E-15	9.99E-12
TNFRSF10A	141.0750804	2.465683708	0.31621401	7.79751578	6.31E-15	1.05E-11
CNR1	47.24102267	4.289220422	0.54973768	7.8023039	6.08E-15	1.05E-11
SIDT1	161.2179634	2.360184583	0.3060229	7.71244441	1.23E-14	1.93E-11
CNR2	22.68480867	4.789065596	0.62438696	7.67002818	1.72E-14	2.55E-11
LTF	351.397333	5.519492242	0.72593991	7.60323572	2.89E-14	4.03E-11
PARP15	256.3799532	2.607730938	0.34324857	7.59720854	3.03E-14	4.03E-11
RFTN1	566.5286103	1.427788836	0.18898587	7.55500333	4.19E-14	5.31E-11
CD22	213.1755724	2.178713116	0.29367134	7.41888242	1.18E-13	1.37E-10
SELL	890.6279646	2.532260226	0.34107436	7.42436401	1.13E-13	1.37E-10
CD79A	20.13130195	4.932978895	0.66555367	7.41184243	1.25E-13	1.38E-10

STAP1	140.3040336	3.475002228	0.4708526	7.38023367	1.58E-13	1.45E-10
CYTIP	657.9512588	2.741192662	0.37046327	7.39936423	1.37E-13	1.45E-10
ELMO1	647.2066776	1.914306698	0.25919163	7.38568112	1.52E-13	1.45E-10
PAX5	94.25515507	5.171992882	0.6996783	7.39195843	1.45E-13	1.45E-10
FAM177B	62.65566997	2.533010392	0.34312772	7.38212115	1.56E-13	1.45E-10
MPP7	97.33979575	3.215726769	0.43606341	7.37444757	1.65E-13	1.47E-10
RP1.225E12.2	94.51533767	2.645008889	0.35939944	7.35952417	1.85E-13	1.59E-10
RUNX1T1	410.2516983	1.00100036	0.13666249	7.32461683	2.40E-13	1.92E-10
SLC38A1	2741.805105	2.86862742	0.39180463	7.32157611	2.45E-13	1.92E-10
HOOK1	84.23349955	3.412609281	0.46564342	7.32880387	2.32E-13	1.92E-10
CLEC2D	591.1158297	1.658893563	0.22690636	7.31091703	2.65E-13	2.02E-10
PZP	136.0011632	2.896779643	0.39719387	7.29311274	3.03E-13	2.24E-10
SNORA16.1	20.93666302	3.264808224	0.44835825	7.28169547	3.30E-13	2.37E-10

CD52	110.8822963	2.833434142	0.38963422	7.27203626	3.54E-13	2.44E-10
SPIB	48.26475437	3.555411978	0.48897632	7.27113329	3.56E-13	2.44E-10
KCND3	113.658807	3.169719761	0.43645069	7.26249226	3.80E-13	2.53E-10
CR2	574.5642824	7.012442134	0.97048954	7.22567511	4.99E-13	3.24E-10
ZNF483	92.75829031	2.520261019	0.34913933	7.21849655	5.26E-13	3.33E-10
MMRN1	186.0122621	4.824234917	0.66871	7.21424076	5.42E-13	3.36E-10
GVINP1	1521.47005	2.624681239	0.36512253	7.1884944	6.55E-13	3.97E-10
ANK3	232.7765494	2.857011549	0.39790727	7.18009383	6.97E-13	4.12E-10
RHOH	342.0863444	2.444313617	0.34104189	7.1671946	7.66E-13	4.43E-10
ABI3BP	2170.72449	2.62599135	0.36759919	7.14362661	9.09E-13	5.15E-10
PLAC8	88.40519034	3.399873219	0.47640194	7.13656464	9.57E-13	5.20E-10
TTC14	760.6838645	1.228459621	0.17206791	7.13938845	9.37E-13	5.20E-10
TTN	3419.891732	2.670656684	0.37439638	7.13323317	9.80E-13	5.22E-10

TNFRSF13B	127.1373614	3.488220704	0.48923612	7.12993288	1.00E-12	5.25E-10
FGR	32.71257859	2.803587637	0.39417327	7.1125768	1.14E-12	5.80E-10
ST18	21.88110126	4.032706441	0.56713085	7.11071607	1.15E-12	5.80E-10
ATF7IP2	175.29353	2.489300785	0.3501966	7.10829511	1.17E-12	5.80E-10
KLHL14	19.2738004	5.829742339	0.82087785	7.10183903	1.23E-12	5.96E-10
MCTP2	206.0448561	2.389619806	0.33675547	7.09600887	1.28E-12	6.00E-10
RP11.73O6.3	9.716468342	4.877665351	0.68733726	7.09646578	1.28E-12	6.00E-10
LINC00426	95.70967539	2.936893166	0.41427685	7.08920419	1.35E-12	6.20E-10
MYH11	566.9184405	1.767478937	0.24977473	7.07629202	1.48E-12	6.69E-10
INADL	298.2740165	2.85218986	0.4032341	7.07328531	1.51E-12	6.71E-10
GABRA2	39.50820215	7.542918448	1.06672098	7.07112599	1.54E-12	6.71E-10
IL7R	3942.315487	2.27318749	0.32193089	7.06110396	1.65E-12	7.10E-10
SOCS3	1664.839979	2.126691932	0.30220104	7.03734164	1.96E-12	8.29E-10

XIST	5676.406336	2.196880741	0.31316939	7.01499185	2.30E-12	9.57E-10
MEOX2	480.4176607	2.547336915	0.36343502	7.00905737	2.40E-12	9.83E-10
TMEM156	51.30874163	3.037131007	0.43401365	6.99777758	2.60E-12	1.02E-09
ARHGAP30	2164.953163	2.148877084	0.30704863	6.99849105	2.59E-12	1.02E-09
F5	197.4752133	2.013309795	0.28751532	7.00244356	2.52E-12	1.02E-09
RP1.111C20.4	60.81214619	2.612409571	0.37429353	6.97957456	2.96E-12	1.14E-09
RP11.53B2.2	74.435796	2.631179847	0.37736127	6.97257523	3.11E-12	1.18E-09
FAM169A	107.1422404	1.666912239	0.23919073	6.96896673	3.19E-12	1.20E-09
LBH	60.33778957	2.410893233	0.34636404	6.96057601	3.39E-12	1.25E-09
RASGRP2	83.19469228	3.489566626	0.50239822	6.94581803	3.76E-12	1.36E-09
CD48	357.1688026	2.676074017	0.38531648	6.94513258	3.78E-12	1.36E-09
CH25H	38.43907194	3.08651817	0.44504959	6.935223	4.06E-12	1.44E-09
TACR1	74.03975447	2.819917953	0.4104905	6.86963031	6.44E-12	2.23E-09

RP11.570L15.1	15.16813625	3.027611425	0.44071166	6.86982372	6.43E-12	2.23E-09
COL19A1	40.13482469	4.75106041	0.69253876	6.86035306	6.87E-12	2.33E-09
AFF3	135.2458246	2.749348862	0.40080892	6.85950027	6.91E-12	2.33E-09
MST4	285.4442332	2.356706017	0.34472098	6.83656111	8.11E-12	2.70E-09
KBTBD8	161.7661447	2.292637935	0.33667595	6.80962787	9.79E-12	3.22E-09
TRABD2A	68.1864891	2.259888392	0.33234899	6.79974511	1.05E-11	3.41E-09
NOVA2	35.37003833	3.086970492	0.45563575	6.77508402	1.24E-11	3.99E-09
ANK1	24.56248667	3.156782296	0.46620657	6.77120935	1.28E-11	4.05E-09
SPOCK2	1280.394485	2.326934351	0.34475022	6.74962403	1.48E-11	4.65E-09
RP11.436I24.1	21.41330974	3.080692074	0.45684604	6.74339236	1.55E-11	4.79E-09
TRAF3IP3	302.393281	1.920845002	0.28532896	6.73203662	1.67E-11	5.12E-09
SCARA5	43.91479913	3.975186924	0.59172023	6.71801758	1.84E-11	5.52E-09
NEURL1B	162.6996444	2.454707148	0.36539996	6.71786378	1.84E-11	5.52E-09

GPR183	738.2996802	1.875869214	0.27943308	6.71312517	1.90E-11	5.64E-09
BIRC3	6105.467362	1.891433896	0.28230489	6.6999685	2.08E-11	6.05E-09
SYT15	194.5348939	2.789165782	0.41641421	6.69805624	2.11E-11	6.05E-09
C14orf64	54.03787264	3.366448744	0.50249843	6.69942148	2.09E-11	6.05E-09
IRF8	778.5105336	2.113708057	0.31608661	6.6871168	2.28E-11	6.45E-09
ENPP2	1875.877502	1.690646168	0.25301757	6.68193196	2.36E-11	6.61E-09
EDN1	41.08132814	3.238727181	0.48529319	6.67375367	2.49E-11	6.92E-09
PCSK5	271.6040531	2.168165528	0.32524583	6.66623621	2.62E-11	7.21E-09
ARHGAP9	261.8492867	2.270009616	0.34106733	6.65560551	2.82E-11	7.67E-09
PRG4	157.4224403	4.121402613	0.61968133	6.65084204	2.91E-11	7.84E-09
TMC8	118.5583202	2.709245729	0.40751381	6.6482305	2.97E-11	7.90E-09

Chapter Four: Predictors of Response and Resistance to CICB

Contents of this chapter is based on:

Akash Mitra, Neeta Somaiah, Anthony P. Conley, Behrang Amini4, Heather Lin, Edwin R. Parra6, Luisa Solis, Ruth Salazar, Swati Gite, HongLei Chen, Luis Acosta-alderon, Samantha Tippen, Josh Baguley, Xingshi Song, Miles C. Andrews, Latasha D. Little, Taylor Tate, Beatriz Sanchez-Espiridion, Celia Garcia-Prieto6, Grace Mathew, Emily Keung, Hannah Beird, Jun Kim, Shaojun Zhang, Curtis Gumbs, Jianhua Zhang, Zachary Cooper, Cara Haymaker, Michael Oberst, Chantale Bernatchez, Vinod Ravi, Dejka Araujo, Maria Zarzour, John A. Livingston, Christina L. Roland, Najat Daw, Jaime Rodriguez-Canales, Weilin Wang, Linghua Wang, Hussein Tawbi, Alexander J Lazar, Ignacio I Wistuba, Jean-Charles Soria, Robert Benjamin, Shreyaskumar Patel, Patrick Hwu, P. Andrew Futreal. Immunogenomic correlates of response to combination immune checkpoint blockade in advanced sarcoma. Submitted, In review.

4.1 Introduction

An estimated 13,130 new cases of soft tissue sarcoma (STS) and 3,600 new cases of bone sarcomas were diagnosed in 2020 in the United States(Street, 2020). While chromosomal-translocation driven tumors such as synovial sarcoma and alveolar soft part sarcoma have a limited number of mutations and therefore fewer potential neoantigens, genetically complex tumors such as undifferentiated pleomorphic sarcomas and leiomyosarcomas typically have multiple mutations, which may in turn generate significantly more neoantigens (Groisberg et al., 2020, Cancer Genome Atlas Research Network. Electronic address and Cancer Genome Atlas Research, 2017). As a group, sarcomas are genomically heterogeneous with diverse phenotypes and tumor microenvironments influencing tumor growth and disease progression.

Given that the overall five-year survival for patients diagnosed with metastatic sarcoma is only 16%, there is a critical unmet need for more effective systemic therapy(2020). Immune checkpoint blockade (ICB) has led to promising results in patients with metastatic cancers(Hodi et al., 2010, Snyder et al., 2014b, Le et al., 2015, Hellmann et al., 2019, Motzer et al., 2018). However, compared to melanoma or non-small-cell lung cancer, sarcomas are characterized by less genomic instability with low tumor mutation burden (TMB) and a more immunosuppressive tumor microenvironment TME(Cancer Genome Atlas Research Network. Electronic address and Cancer Genome Atlas Research, 2017). One potential approach to overcoming these features that may limit ICB activity is combining ICB agents that target different aspects of the adaptive immune response. For instance, anti-CTLA-4 enhances T cell priming in secondary lymphoid organs while PD-1/PD-L1 blockade reverses inhibition of T cells within the tumor and improves antigen-specific responses; combination therapy results in the activation of previously phenotypically exhausted CD8 T-cells (Ribas, 2012, DR Leach, 1996, Wei et al., 2019). However, responses and mechanisms of response and resistance to combination ICB (C-ICB) appear distinct from those seen with ICB monotherapies and have not been detailed in metastatic sarcoma(Wei et al., 2019). To this end, we extensively profiled pre- and on-treatment tumor biopsies acquired from 57 patients with metastatic sarcoma enrolled to a phase II clinical trial (NCT02815955) evaluating combined durvalumab (anti-PD-L1) and tremelimumab (anti-CTLA-4) (Somaiah, submitted). Genomic, transcriptomic and immune-based molecular correlates were analyzed alongside relevant clinical data. Here we report the evaluation of genomic and transcriptomic features underlying clinical benefit as well as in-depth characterization of the TME underlying the mechanisms of C-ICB response in multiple sarcoma histologies.

4.2 Results

Patient Profile

Fifty-seven adult patients with metastatic sarcoma were treated across eight study arms based on their histologic subtype. All patients received combination durvalumab 1500 mg and tremelimumab 75 mg every 4 weeks for 4 cycles followed by durvalumab 1500 mg every 4 weeks for up to 8 additional cycles (Fig 1A and Supplemental Table 1). Patients underwent a pre-treatment biopsy and blood draw prior to initiation of combination therapy, followed by an on-treatment biopsy and blood draw at the 6-week timepoint. DNA and RNA were extracted from these longitudinal tumor biopsies, resulting in a total of 79 RNA samples which were analyzed by RNA-sequencing and 118 DNA samples which were analyzed by whole exome sequencing and T-cell receptor(TCR) sequencing.

Response was assessed using the immune-related response criteria (irRC) comparing baseline staging (CT or MRI-based) scans with repeat imaging at the 12-week time point. In total, 26 (46%) patients were classified as having received clinical benefit (18SD + 8 PR). One (2%) patient had no follow up scan on therapy, 7 (12%) patients had unconfirmed PD (i.e. did not have a sufficient period of follow-up at the time of analysis to confirm PD), and 23 (40%) patients had confirmed PD (Fig S1A, Supplemental Table 2). Patient responses in this analysis were coded based on best response criteria at the 12-week time point based on irRC criteria, with partial responders (PR) and patients with stable disease (SD) characterized as responders and patients with progressive disease (PD) as non-responders unless otherwise noted.

Copy number alterations acquired during ICB exposure may lead to therapy resistance

We detected a total of 9610 exonic mutations across our cohort after strict filtering criteria. The overall somatic TMB was low, with a median of 1.63 mutations per Mb (Fig 1B), similar to previous reports(Cancer Genome Atlas Research Network. Electronic address and Cancer

Genome Atlas Research, 2017). There was no significant correlation of TMB and response (Fig S1C). Non-synonymous mutations were compared against the COSMIC database and the most commonly mutated cancer gene was found to be *KMT2C* (21%) followed by *EP400* (19%), *TP53* and *POLE* (11%) (Fig 1B). We observed an enrichment in recurrent mutations involved in chromatin remodeling and transcriptional coregulation (p = 4.06 e-12) including *KMT2C*, *EP400*, *KMT2D* and *ARID1B*, as well as genes involved in chromosome regulation (p=1.89 e-9) including *POLE* and *POLD1*. While samples with mutations in *POLE*/*POLD1* had a higher mutational burden than samples without *POLE*/*POLD1* mutations, only 3 of 5 patients with these mutations responded to C-ICB. Next, we analyzed copy number alterations (CNA) across the cohort (Fig S1C) and found frequent CNA of genes in the MDM2-p53, p16-CDK4-RB1, PTPRB and CDKN2A pathways, as previously reported(Cancer Genome Atlas Research Network. Electronic address and Cancer Genome Atlas Research, 2017). *MDM2* amplifications were present in 85% of LPS cases, which were also predominantly non-responders. Additionally, we detected *CDK4* and *PTPRB* amplifications in all *MDM2*-amplified samples, suggesting a role for recurrent focal amplifications at 12q14-15 in these cases. Deep deletions in *TP53* were observed in four samples, mutually exclusive of any significant amplifications in other genes.

Globally across the exome, we found higher levels of copy number gain events in non-responders at both the on-treatment and inclusion of pre and on-treatment time points, compared to responders (p = 0.0093 and p = 0.013, respectively) (Fig S1E Fig 1C). Additionally, we noted acquired copy number losses in on-treatment samples of two patients affecting genes involved in antigen presentation machinery including *B2M* and *JAK2*/*JAK3*. Losses were also observed for other key immune genes including; *CD209*, required for antigen presentation on immature dendritic cells; *CD79A*, involved in the B lymphocyte antigen receptor complex(Gordon et al., 2003, Engering et al., 2002); *PDCD1*, encoding the PD-1 protein and *CD274* , encoding the PD-L1 protein. All of the patients with these treatment-emergent immune gene CNAs were non-responders, thus these CNA events may be means of acquired resistance to C-ICB(Zaretsky et al., 2016, Sharma et al., 2017). Additionally, non-responders frequently displayed shallow deletions of *RB1*, or gains of *TERT*, *S100A7* and *SDHA* compared with responders.

We then integrated focal CNAs into chromosomal segments in order to determine if response status was associated with specific segmental amplifications or deletions through the use of GISTIC(Mermel et al., 2011). Responders were characterized by a pre-treatment amplification of 14q11.2 (Fig S1D). Further examination revealed an amplification event spanning *DAD1*, a negative regulator of programmed cell death, previously associated with enhanced T-cell proliferation(N. A. Hong, 1999). Interestingly, the most significantly amplified peak found in non-responders at the pre-treatment time-point involved the 6p21.32 locus which contains several HLA and TAP genes essential for Major Histocompatibility Complex Class II (MHC-II) antigen presentation machinery (q = 0.0025)(Fig 1D). Additionally, 75% (9/12) of non-responders also contained amplifications spanning *PDGFA* on chromosome 7p22.3 which may activate MAPK signaling pathways.

Co-expression networks modulating response to C-ICB in sarcoma

To delineate the strongest underlying predictors of response to C-ICB, we compared the pre-treatment gene expression profiles of patients that went on to respond (partial response, n=3 and stable disease n = 13) against those with objective tumor growth (progressive disease, n=22). After accounting for confounding factors such as batch effects and gender in our differential expression model, we found response-associated upregulation of several genes modulating the immune response, including *CD22*, *FCRL1*, *FCRL2*, *MS4A1*, *HLA-V and HLA-G* (Fig 2A). Interestingly, these genes all play roles in the regulation of B lymphocyte immune responses. Non responders displayed significant upregulation of both *CXCL10* and *CXCL11*, chemokines involved in an activated T-cell response along with *MAGEC2*, a cancer testis antigen.

An active immune response involves complex interactions between the many cell types present within tumors, including contributions of many genes simultaneously. In order to estimate the network of interactions in pre-treatment samples and relate these to ICB response, we performed single sample network reconstruction implemented in LIONESS(Kuijjer et al., 2019). Through our interaction network-based analysis, we found nodes enriched in responders including genes involved with antigen presentation machinery (*B2M*), the innate immune system (*CD59* and *CD63*), genes involved in Th1 differentiation (*IL6ST*) as well as the pre-B cell growth stimulating factor *CXCL12* (Fig 2B). *EIF4G2*, which has been shown to be involved in interferon gamma signaling, was determined to be a hub node in responders(NAOMI LEVY-STRUMPF, 1997). *EIF4G2* was found to be connected to *MXRA8*, previously shown to be positively correlated with viral infectivity, with links to *CD59* and *CD63* involved in immune activation(Zhang et al., 2018). Non-responding patient samples had co-expression for multiple collagen family genes which may contribute towards an epithelial to mesenchymal transition-like phenotype(Shintani et al., 2008). Non-responders were also characterized by an upregulation of MKI67 and MAP4K4 which contribute towards elevated network activity in cellular proliferation and MAPK signaling. Through our co-expression-based analysis, we detected elevated network activity in innate immune activity in responders and MAPK signaling non-responders.

Distinct immune microenvironments present in sarcomas

To better characterize the features of the TME at the pre-treatment time point, we utilized a suite of immune deconvolution tools to quantify the immune and stromal features present in the TME, based on mRNA expression in each sample(Newman et al., 2015, Aran et al., 2017, Li et al., 2016b, Becht et al., 2016). We then applied consensus clustering methods and found three distinct clusters of pre-treatment samples based on the similarity of cell types present within the TME (Fig S4A, Fig 2C). An immune “hot” cluster was uniformly enriched for pan-immune features, an immune “intermediate” cluster was enriched for multiple immune cell sub-types but not CTLs and B cells, while an immune “cold” cluster lacked any particular immune enrichment. Interestingly, fibroblasts were the only cell type found at consistent levels across all pre-treatment samples (Fig S2).

Tumors classified as immune hot or intermediate prior to treatment almost exclusively (11/13 patients) remained hot or intermediate at the on-treatment timepoint (Fig 2D). Interestingly the proportion of responders was similar within all three immune phenotype clusters at both time-points, suggesting the existence of other factors that influence response to C-ICB.

To characterize the interactions present within the TME, we evaluated pairwise relationships in of multiple TME components evaluated from the transcriptome (Fig 3A). We examined absolute fractions of multiple immune compartments and subset these groups against major immune cell types. Macrophages and regulatory T-cells were correlated in non-responders. B-cells however, were positively correlated with multiple effector-immune phenotypes, including CD8 T-cells and eosinophils.

B-cells predict response to C-ICB

While the majority of immune cell sub-types were not found to be significantly associated with response, B-cells were significantly increased in responders to C-ICB at the pre-treatment time point ($p = 0.05$).To further stratify which immune cell sub-types may be underlying objective response to C-ICB, we compared patients with partial response (PR) and progressive disease (PD) (Fig S3). While CD8 T, CTLs and NK-cells were not significantly associated with irRC based response metrics, B-cells were the only immune subpopulation found to be significantly enriched in pre-treatment samples of patients who achieved a PR as compared to with PD ($p = 0.005$).

In order to validate our expression-based findings at the protein level, we performed IHC staining of CD20 and found elevated levels of CD20+ B cells in responders compared with non-responders to ICB ($p = 0.02$) (Fig 3C and Fig 3D). Specifically surveying the pre-treatment time point we found higher levels of B cells in responders ($p=0.07$), however response was not particularly correlated with the presence of identifiable tertiary lymphoid structures (Fig 3C Fig S4B Fig S4C). In order to gain more insight into the repertoire of B-cells that contribute to ICB response, we employed the TRUST algorithm to probe the RNA-seq data for BCR templates(Li et al., 2017). We found higher levels of Immunoglobulin Heavy (IGH) diversity ($p = 0.03$) and lower levels of IGL diversity ($p = 0.053$) in responders to ICB at the on-treatment time-point (Fig 3E). Additionally, we also found elevated levels of dominant IGL clones (clonotype contribution > 1% of repertoire) at the on-treatment time point in patients that responded to therapy ($p = 0.02$) (Fig 3F). Additionally, CD56+ NK cells were more abundant at the on-treatment time point in patients that responded to ICB ($p = 0.037$) (Fig S4D, Fig S4E and Fig S4F). Overall, these findings imply that C-ICB may specifically expand sub-populations of immune cell-types and antigen reactivities that assist in tumor reduction.

B cell enrichment significantly associated with response across multiple cohorts

As overall immune infiltrate has been shown to correlate with survival and response to anti-PD-1 and anti-CTLA-4 immunotherapy, we next investigated the reproducibility of our B-cell driven findings in three publicly available immunotherapy-treated melanoma cohorts (n = 168)(Van Allen et al., 2015a, Hugo et al., 2016, Riaz et al., 2017a). While transcriptomic deconvolution revealed elevated levels of multiple immune cell sub-types in responders, the most significantly associated cell type associated with response were driven through a clear B-cell signature (p = 0.0058) (Fig 4A and Fig 4B). Increased levels of B-cells were correlated with an infiltrate comprising of other immune cell subtypes however cytotoxic lymphocytes and CD8+ T-cells were not as tightly correlated with response across these cohorts (Fig 4B). Given that the Riaz anti-PD-1 dataset represents the largest cohort of pre and on-treated samples, we explored this dataset further and found a significantly higher immune rich infiltrate detected in the responders, compared to non-responders to anti-PD-1 therapy (Fig S5A). Deconvoluting the multiple subtypes present in the immune compartment confirmed that while several immune cell subtypes were enriched in responders, the most significant immune cell type were B cells (p = 0.003) (Fig S5B).

To validate our findings, we obtained gene expression data from 79 samples collected through a similar pre- and on-treatment biopsy sampling strategy from patients enrolled in SARC028 (NCT-2301039) and the expansion cohorts, evaluating the efficacy of single agent anti-PD-1, pembrolizumab, in patients with metastatic sarcoma(Tawbi et al., 2017, Keung et al., 2020). Longitudinal tumors biopsies were available from 79 patients, of whom 1 patient achieved a complete response, 12 PR, 30 SD and 36 PD; patients with tumor response by RECIST primarily had UPS or LPS histology. We performed similar immune deconvolution on samples passing our QC metrics and found that B-cells were the most significant immune cell subset amongst responders as compared to non-responders ($p = 0. 043$) (Fig 4C). An overlap of the gene sets between the original SARC028 and expansion cohorts revealed elevated levels of a gene essential for B-cell activation, *BANK1*, in responders to ICB ($p = 0.032$ Fig S5C)

Baseline and on-treatment TCR diversity correlates with response to C-ICB

In order to characterize T-cell contributions in this cohort interactions, we conducted multiplexed immunofluorescence(mIF) profiling and TCR sequencing of the pre-treatment and on-treatment samples under study . Pre-treatment biopsy slides were stained using a panel of lymphocytic effector and regulatory markers (see Methods). While CTL levels ($CD3^+$ $CD8^+$) strongly correlated with levels of PD-L1 expressing tumor cells, antigen experienced T cells ($CD3^+$ $PD1^+$) and antigen experienced CTLs ($CD3^+$ $CD8^+$ $PD1^+$) none of these populations correlated with response, at the pre-treatment time point (Fig 5A and Fig S6A). A second panel of markers including activated CTLs ($CD3^+$ $CD8^+$ $GZMB^+$), Effector/Memory CTLs ($CD3^+$ $CD8^+$ $CD45RO^+$) and regulatory T-cells ($CD3^+FOXP3^+CD8^-$) were strongly correlated with one another however they were not significantly positively correlated with response (Fig 5B and Fig S6B). While levels of CTLs and regulatory T-cells were relatively equal between responders and non-responders at both time points, we found a significantly stronger positive correlation in lymphocytic infiltrate with a regulatory T-cell phenotype in non-responders at the pre-treatment time point ($\rho = 0.83$, $p < 0.0001$). This implies that while certain sarcomas may be infiltrated with effector or activated CTLs that are necessary for tumor recognition, these tumors are also heavily infiltrated with regulatory T-cells that likely contributing to an immunosuppressive microenvironment. Effector immune cell infiltrates were similar between responders and non-responders at the on-treatment time-point.

Using genomic DNA isolated from pre- and on-treatment samples, we explored the relationship of the TCR repertoire and response to C-ICB. We compared the maximum productive frequency to quantify the clones that were most highly prevalent across both time-points and found that the frequency of the top productive TCR rearrangements to be higher in non-responders, compared to responders (Fig 5C, $p = 0.04$). We also examined dominant clones (clonotype contribution > 1% of repertoire) and found no significantly elevated levels of large clonotypes present at baseline in responders compared to non-responders (Fig S6C, $p = 0.22$). Interestingly, when comparing the TCR repertoire over time, we observed a higher proportion of dominant clonotypes ($p = 0.059$) prior to therapy, while on-treatment samples were dominated by rare clonotypes potentially suggesting diversification of the TCR repertoire over the course of C-ICB (Fig S6D and Fig S6E, $p = 0.083$).

Given that the highest frequency TCR clonotype was not positively correlated with response, we then characterized the diversity and richness of the entire repertoire present across our cohort. Evaluation of Simpson Diversity revealed a more diverse TCR repertoire in responders as compared to non-responders, across both baseline and on-treatment timepoints ($p = 0.047$) (Fig 5D). Additionally, levels of Simpson Evenness, depicting the richness of the repertoire, revealed a more evenly distributed repertoire in responders compared to non-responders ($p = 0.042$) (Fig S5F). Finally, we sought to compare TCR-sequencing metrics along with multiple immune cell phenotypes derived from mIF of effector and regulatory T cells. Using our consolidated analysis, we correlated the effect of multiple immune cell types on tumor shrinkage/growth as a continuous variable measured at the 12-week time point from baseline.

Tumor shrinkage was found to be significantly inversely correlated with both maximum productive frequency and Simpsons clonality ($p = 0.029$ and $p = 0.049$) (Fig 5E). Thus, a less clonal TCR repertoire at baseline correlated with tumor shrinkage within our cohort. Overall, these findings imply that a diverse TCR repertoire, may be beneficial for creating a favorable immunogenic response to C-ICB.

4.3 Discussion

The delineation of molecular and phenotypic correlates of response and resistance to immune checkpoint blockade can be used to stratify patients for therapy as well as provide insights into new therapeutic approaches. The work focused on sarcoma subtypes presented here reveals a picture of substantial heterogeneity at the level of the tumors themselves and the attendant microenvironment constituents. The histologies enrolled on trial ranged mutationally, from those with quiet genomes largely comprised of translocation driven cancers (ASPS, synovial sarcoma) to those with heavily rearranged and mutated genomes, exemplified by osteosarcoma. No significant correlation was seen from tumor genomic features, including tumor mutation burden on response to therapy. Indeed, the best responses noted were in ASPS, a canonical translocation-driven tumor. Additionally, PD-L1 staining at baseline proved to be uninformative as a predictor of response. The context of uniform treatment couple with a mixed response not obviously driven by tumor genome features per se, framed a deeper interrogation of the tumor microenvironmental impact on response. Total immune infiltrate as delineated by RNAseq deconvolution and tissue staining was not found to correlate with response. Whilst overall levels of immune infiltrate proved inadequate in providing robust response correlates, further investigation of the immunocyte populations revealed B-cell infiltrates as the most significant predictor of response in the cohort. These increased levels of

B-cell infiltrates and IGH diversity was established at the expression level and further validated through immunohistochemical staining. The involvement of B-cells in immune checkpoint response has been recently highlighted, including in a sarcoma trial, in the single agent anti-PD1/PD-L1 context(Petitprez et al., 2020). This work extends the association to additional sarcoma types as well as into the combined checkpoint inhibitor realm. Alongside this, whilst quantitative T-cell infiltrates were not predictive of response, a more diverse TCR repertoire was found to correlate with response. The role of B-cells and, importantly, B-cell/T-cell crosstalk in checkpoint blockade is not yet well understood. The detailed molecular and cellular cross-talk of these populations of immune cells prior to and upon exposure to therapy in the context of relatively somatic mutationally bland genomes provides a rich opportunity to further explore the activity of these powerful therapies in the highly varied sarcoma landscape, with a goal of further understanding critical immune biology and improving therapeutic responses in this area of largely unmet clinical need.

4.4 Figures

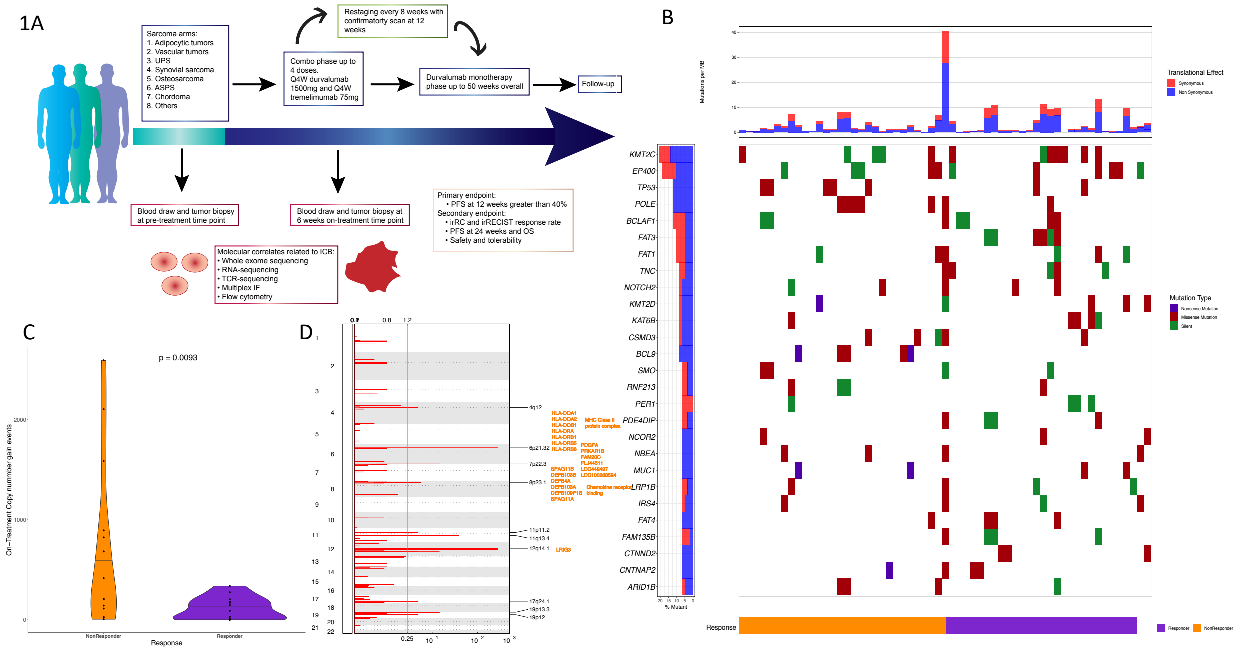


Fig1 A: 57 patients were enrolled across eight arms of a clinical trial encompassing multiple sarcoma histologies. Sample collection is indicated along the treatment timeline. B: Clinical and genetic profile of the cohort with TMB indicated at top bar plot. Co-mutation plot below indicates the most recurrently mutated genes with response and histology annotated in annotation tracks below. C: Density plot of copy number gain level events detected between responders and non-responders at the on-treatment time point (two-sided T-test) D: GISTIC arm level amplifications recurrently detected in non-responders to ICB. Genes annotated in orange represent genes amplified in non-responders along with significant pathways level enrichments (if any). Green vertical line indicates FDR corrected p-values of most significantly amplified regions in the genome in non-responders.

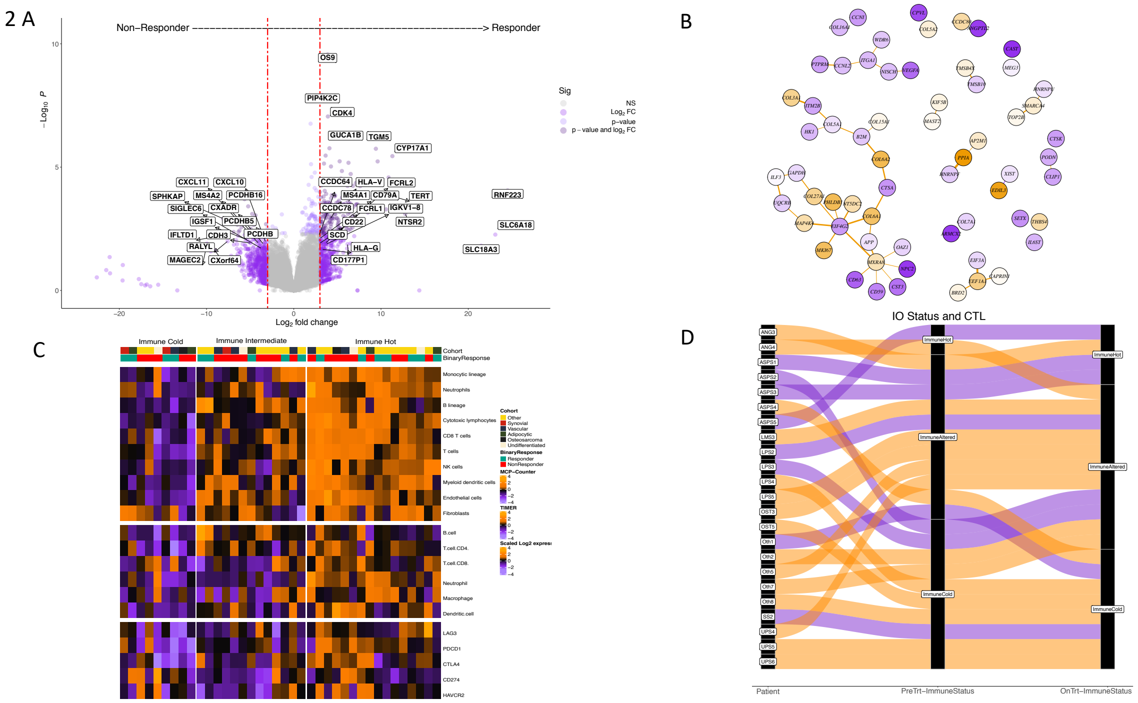


Fig2 A: Volcano plot of most differentially expressed genes at the pre-treatment time-point between patients with PR and PD. X-axis represents Log2 fold change indicated from differential expression analysis while Y-axis represents FDR adjusted p-value. Vertical red lines indicate absolute log fold change of 4 while horizontal red line indicates p-values less than 0.005 B: Single sample network-based correlation maps of significantly differentially expressed genes upregulated in responders as compared to non-responders at both time-points. Edges are colored based on whether they have higher weights in responding patients (purple) or non-responding patients (orange). Thicker edges represent higher log fold changes. Nodes (genes) are colored based on the t-statistic from the differential expression analysis. Nodes with absolute t-statistic < 1.5 are shown in white, nodes in red/blue have higher expression in patients with response/non-response, respectively C: Heatmap of consensus calls from immune deconvolution tools used to infer the varying TMEs present in the cohort. Top annotation track indicates histology followed by response. Heatmaps below show consensus clustering based

on an optimal cluster count of k=3 of varying immune/stromal sub-populations as well as gene-expression of various immuno-modulatory genes. Group 1 (left to right) indicates an immune cold, group 2 indicates an immune intermediate and group 3 indicates an immune hot phenotype D: Sankey plot depicting changes in immune phenotypes of patients with paired biopsies. Patient ID recorded in first axis, followed by pre and on-treatment immune status based on immune activity. Responders and non-responders annotated in purple and orange, respectively.

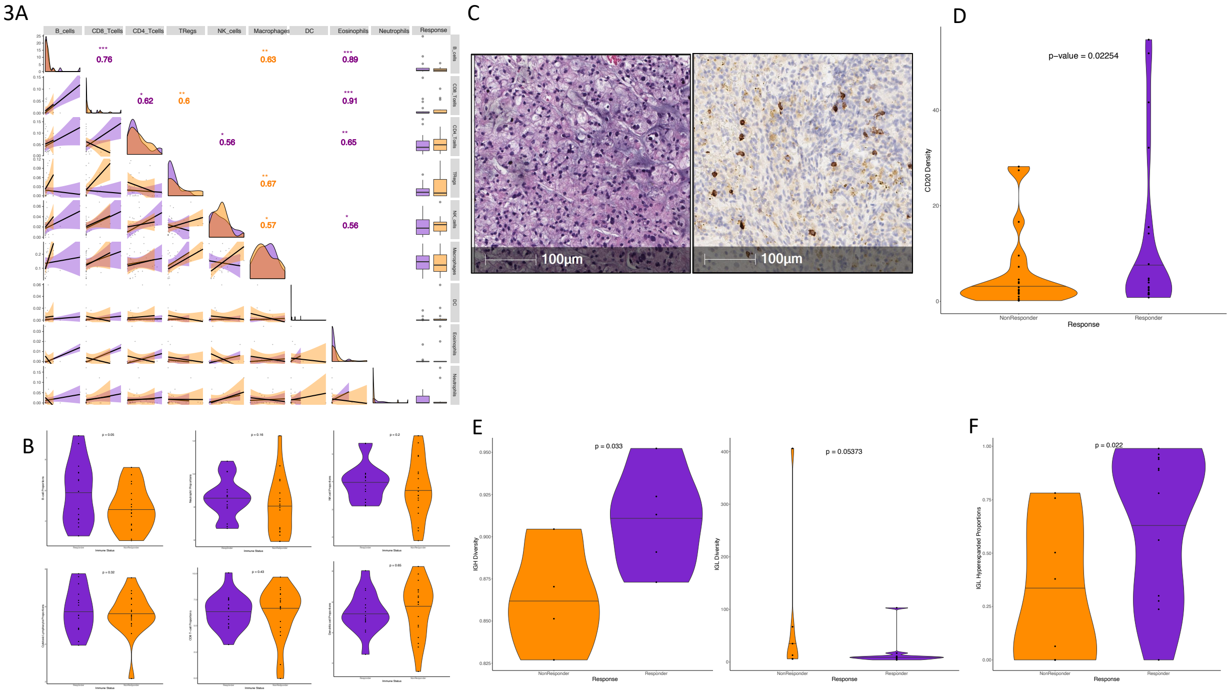


Fig3 A: Correlation density and scatter plots of various TME features ascertained from the absolute computation of CIBERSORT to ascertain which features associate with response. Values in boxes indicate correlation between responders (in purple) and non-responders (in orange) along with absolute frequencies of each group in the last column. Significance indicated with ” *** "if the p-value is < 0.001, ” ** "if the p-value is < 0.01, ” * ” if the p-value is < 0.05, else “ “. B: MCP-counter enrichment of various TME elements at the pre-treatment time-point. Each TME feature is represented in box plots with p-values (t-test) indicated between groups. C: CD20+ staining was used to stain B-lymphocyte enrichments. Bottom panel shows CD20+ staining of a patient that responded to therapy. D: CD20+ staining was higher in responders as compared to non-responders. E: IGH diversity was inferred using TRUST and found to be higher in responders as compared to responders to ICB at the on-treatment time-point. Similar IGL based inference indicated higher levels of IGL diversity non-responders compared to responders at the on-treatment time-point. F: Increased levels of

dominant IGL clones were detected in responders as compared to non-responders. (Two-sided T-test for all)

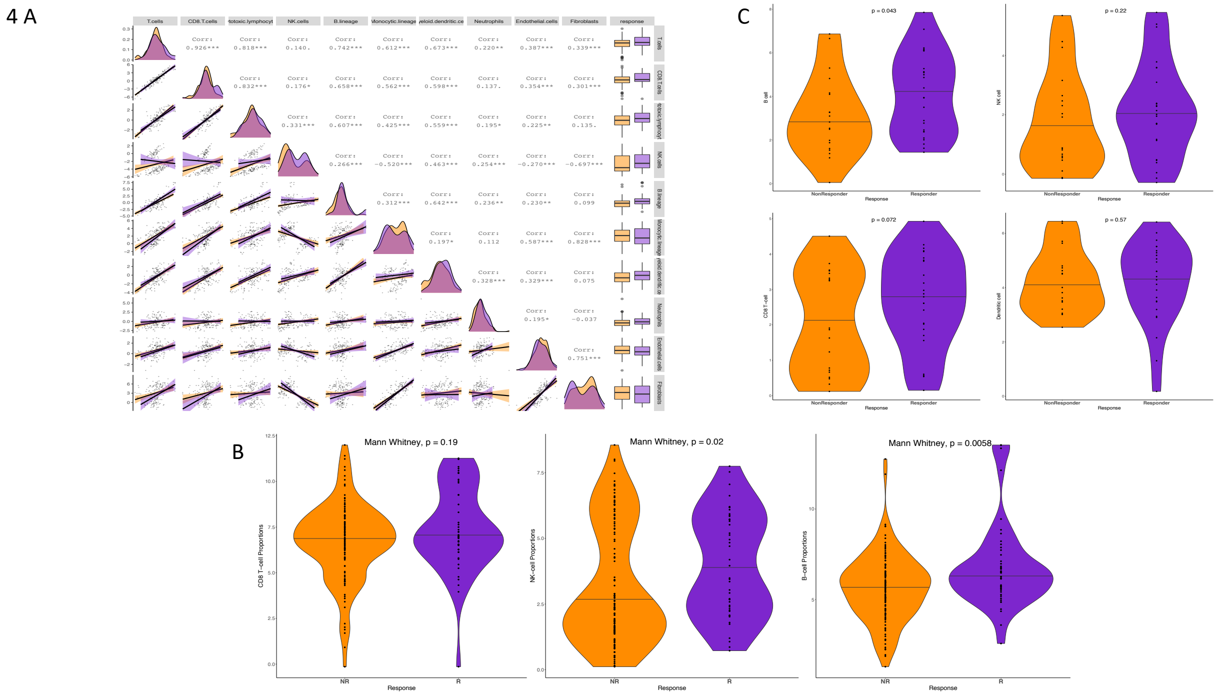


Fig4 A: Paired density and correlation plots of multiple TME features deconvoluted across the Riaz, Van-Allen and Hugo datasets using MCP counter. Value in each box depicts correlation value along with ” *** "if the p-value is < 0.001, ” ** "if the p-value is < 0.01, ” * ” if the p-value is < 0.05, else “ “. B: NK-cell, CD8 T-cell and B-cell proportions respectively, represented in responders vs non-responders using the combined cohorts (Multiple comparisons tested using Mann Whitney U test) C: NK-cell, CD8 –cell, Dendritic and B-cell proportions respectively, inferred using the SARC028 cohort of metastatic sarcoma patients receiving anti-PD-1 therapy. (two-sided T-test)

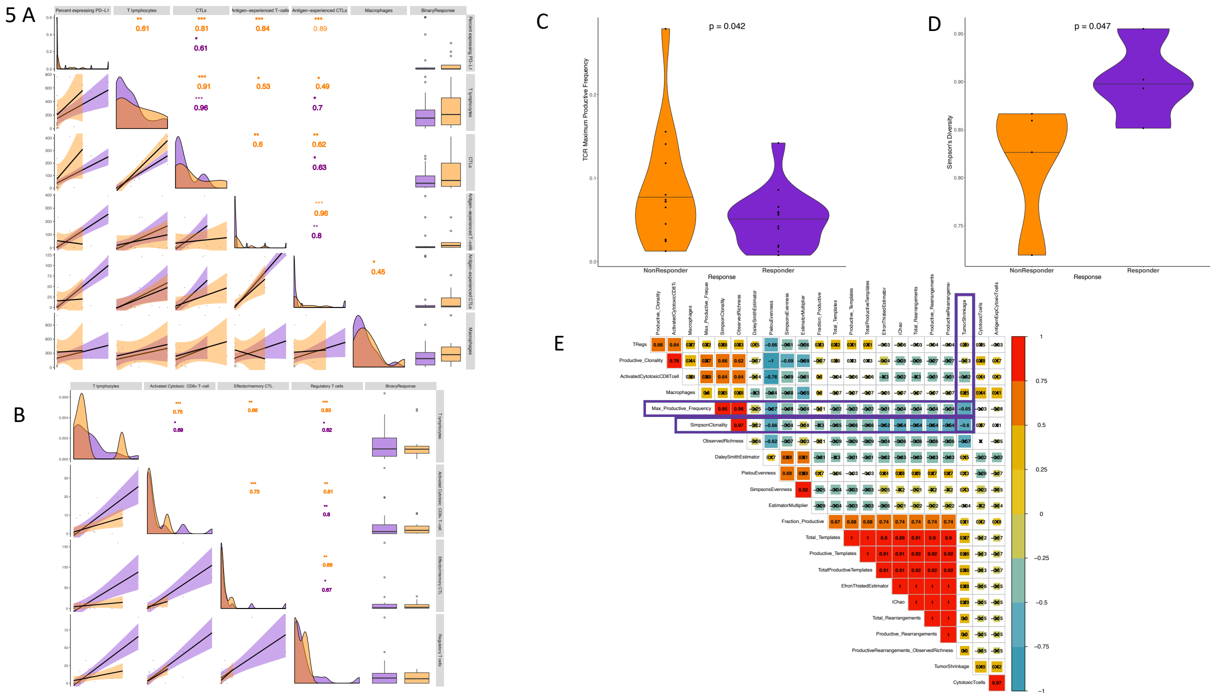


Fig5 A: Multiplexed IF displayed no significant enrichment of multiple effector immune subtypes in responders vs non-responders to ICB. B: Inclusion of regulatory markers as well as activated cytotoxic T-lymphocytes marked no significant enrichment in responders. Both panels A and B include paired density and scatter plots of the immune phenotypes of each marker. Values in boxes indicate correlation between responders (in purple) and non-responders (in orange) along with absolute frequencies of each group in the last column. Significance indicated with ” *** "if the p-value is < 0.001, ” ** "if the p-value is < 0.01, ” * ” if the p-value is < 0.05, else “ “. C: TCR dynamics of the maximum productive frequency, a surrogate for TCR clonality in responders as compared to non-responders. D: TCR level metrics of the diversity present in the repertoires of responding vs non-responding patients. E: Correlation plot of multiple TCR metrics along with multiplex immunofluorescence data containing various lymphoid markers. Tumor shrinkage was found to be most significantly inversely correlated with the maximum productive frequency of the repertoire along with TCR

clonality. Red indicates stronger positive correlation while blue indicates negative correlation. Correlation values are indicated within the box with “X” marking off correlations that were not found to be significantly associated with one-another.

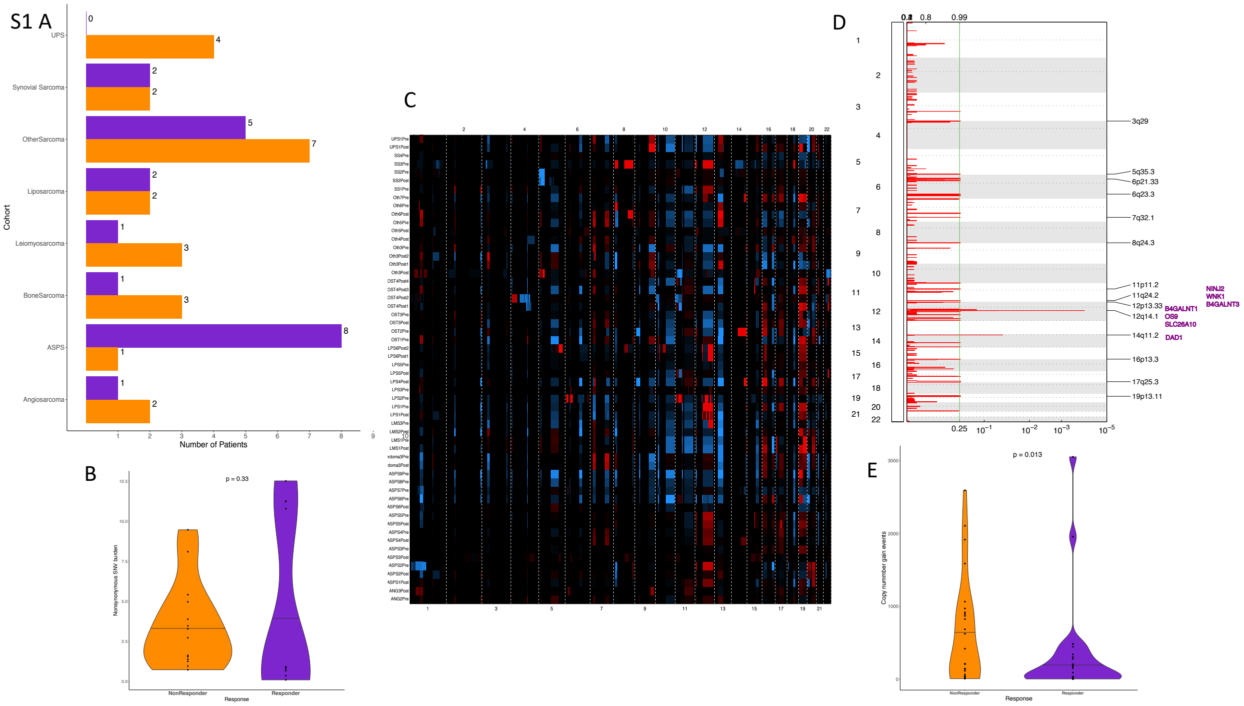
No. of patients	57 (100%)
Median age (range)	48 (22 -77)
Sex	
Male	31 (54%)
Female	26 (46%)
Performance status	
1	35 (61%)
2	21 (37%)
3	1 (2%)
Prior lines of therapy	
Median (range)	2 (0 -6)
None	5 (9)
1 -2	24 (42)
≥3	28 (49)
Cohorts	
Adipocytic tumors	6 (10)
Alveolar soft part sarcoma	10 (17)
Chordoma	5 (8)
Osteosarcoma	5 (8)
Undifferentiated pleomorphic sarcoma	5 (8)
Synovial sarcoma	5 (8)
Vascular tumors	10 (17)
Other tumors	11 (19)

Supplemental Table 1: Patient characteristics across cohorts

Cohort	N	PR (irRC) (%)	PR (irRecist) (%)	SD (irRC) (%)	SD (irRecist) (%)	PD (irRC) (%)	PD (irRecist) (%)	uPD (irRC) (%)	uPD (irRecist) (%)
All Cohorts	57	8(14)	7(12)	18(32)	20(35)	23(40)	22(39)	7(12)	7(12)
Adipocytic Tumors	6	0(0)	0(0)	3(50)	3(50)	3(50)	3(50)	0(0)	0(0)
ASPS	10	5(50)	4(40)	4(40)	5(50)	1(10)	1(10)	0(0)	0(0)
Chordoma	5	1(20)	1(20)	3(60)	3(60)	1(20)	1(20)	0(0)	0(0)
Osteosarcoma	5	0(0)	0(0)	1(20)	1(20)	3(60)	3(60)	*	*
Other	11	0(0)	0(0)	3(27)	3(27)	5(45)	5(45)	3(27)	3(27)
Synovial Sarcoma	5	0(0)	0(0)	2(40)	3(60)	1(20)	0(0)	2(40)	2(40)
UPS	5	1(20)	1(20)	1(20)	1(20)	2(40)	2(40)	1(20)	1(20)
Vascular Tumors	10	1(10)	1(10)	1(10)	1(10)	7(70)	7(70)	1(10)	1(10)

* Patient developed clinical progression and came off study before imaging was performed to confirm radiographic progression

Supplemental Table 2: Response metrics across cohorts



FigS1 A: Patient responses across all samples with a pre-treatment biopsy lesion present, spread across each histology. B: Copy number alteration plot depicts copy number gains and losses across the cohort. X-axis represents chromosomal locations; Y-axis indicates samples and copy number intensity of gains to losses depicted in red to blue scale respectively. C: Non-synonymous single nucleotide variant/TMB burden at the pre-treatment time-point was not found to be significantly different between responders and non-responders to ICB. D: GISTIC arm level amplifications recurrently detected in responders to ICB. Genes annotated in purple represent genes amplified in responders. Green vertical line indicates FDR corrected p-values of most significantly amplified regions in the genome in responders. E: Copy number gain events at the pre-treatment time point was slightly elevated in non-responders, however this did not reach significance thresholds (Two-sided T-test for all)

S2

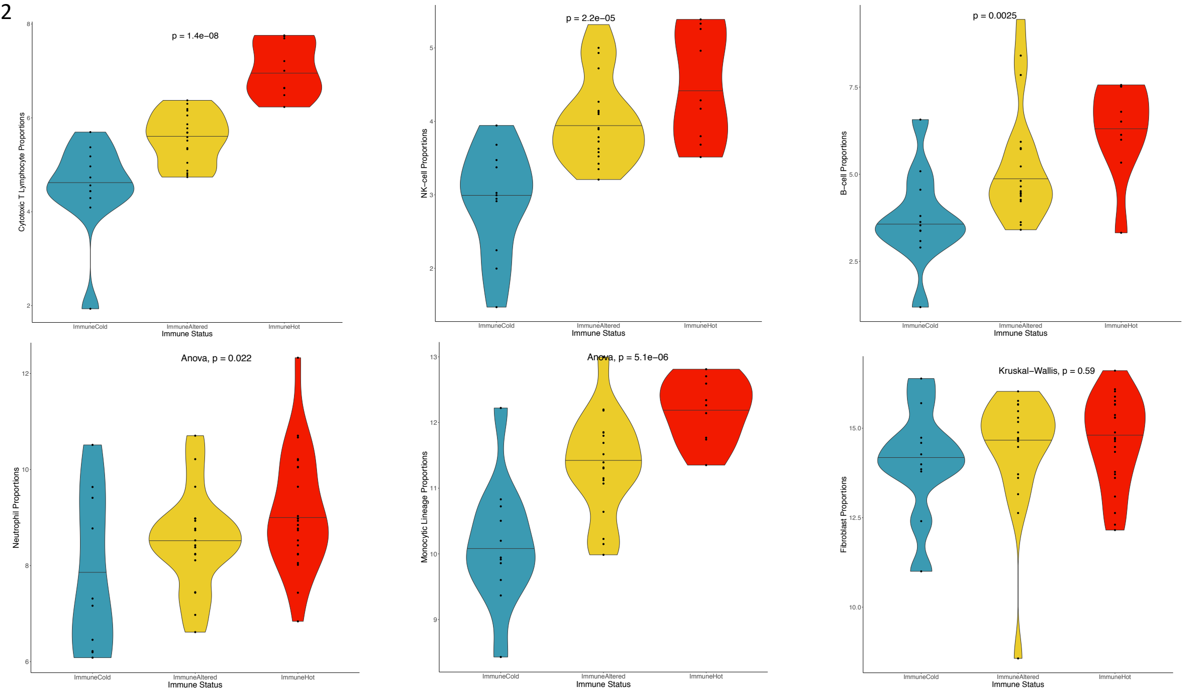


Fig S2: Enrichment of various elements present in the TME at the pre-treatment time-point. Each feature is indicated using density plots with ANOVA p-values indicated between the immune hot, immune altered and immune cold subgroups.

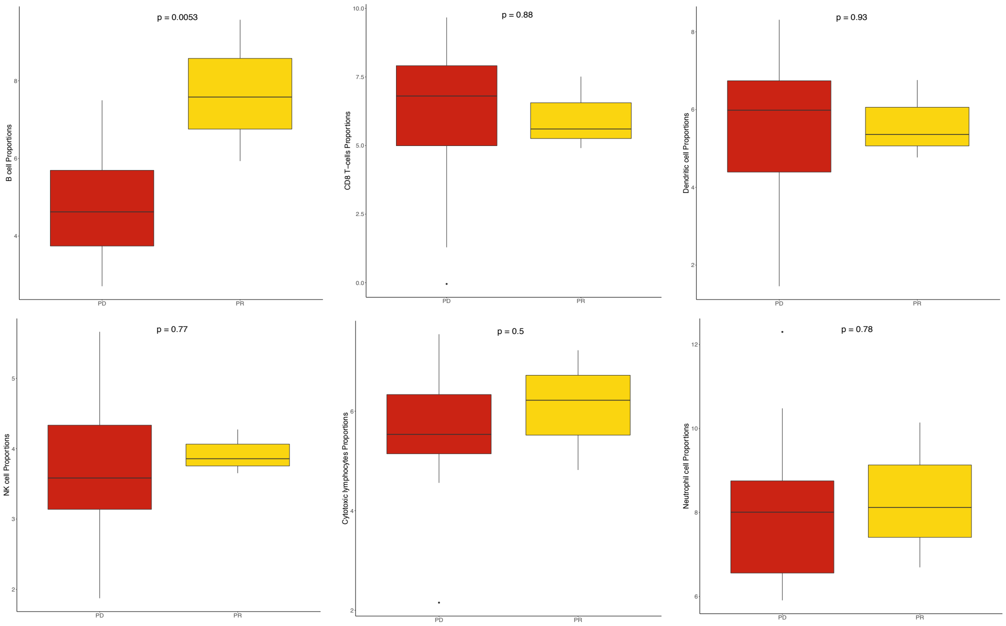
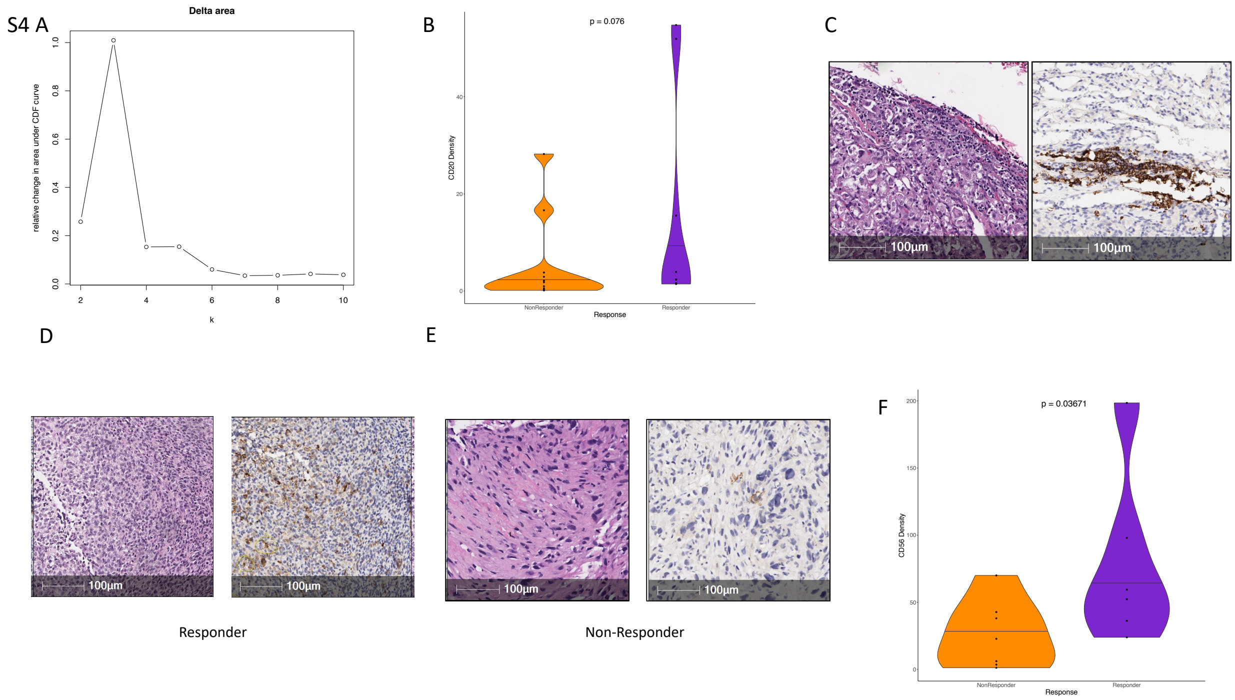


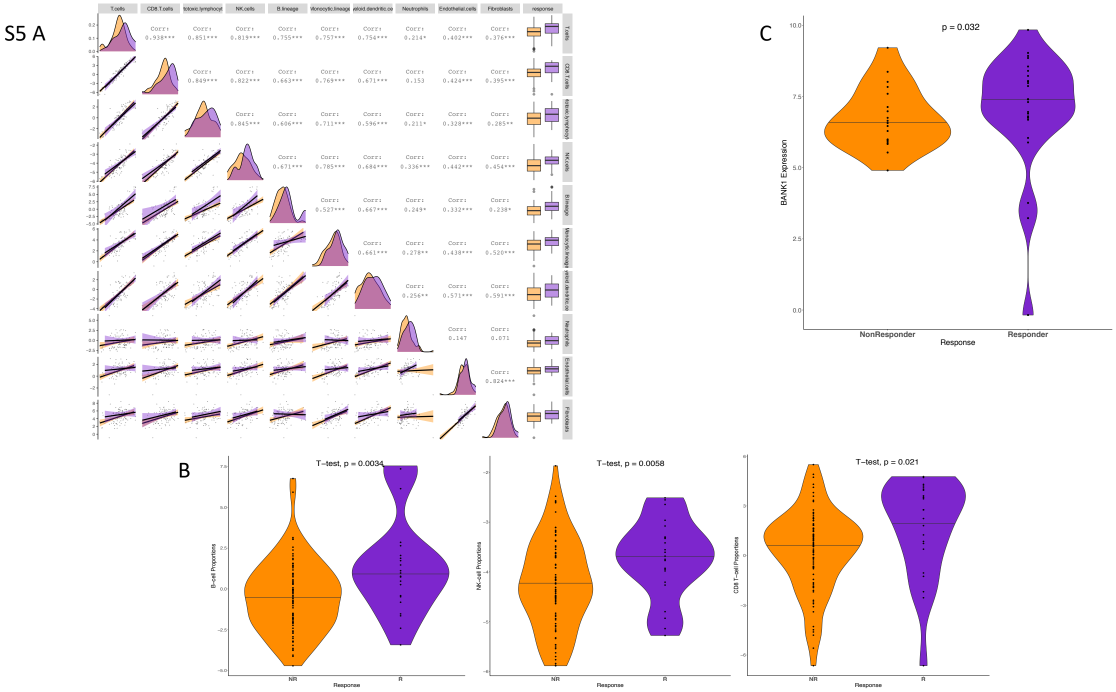
Fig S3: MCP-counter enrichment of various TME elements at the pre-treatment time-point. Each TME feature is represented in box plots with p-values (t-test) indicated between groups after stratifying only patients with tumor reduction (PR) and tumor growth (PD) within the cohort.



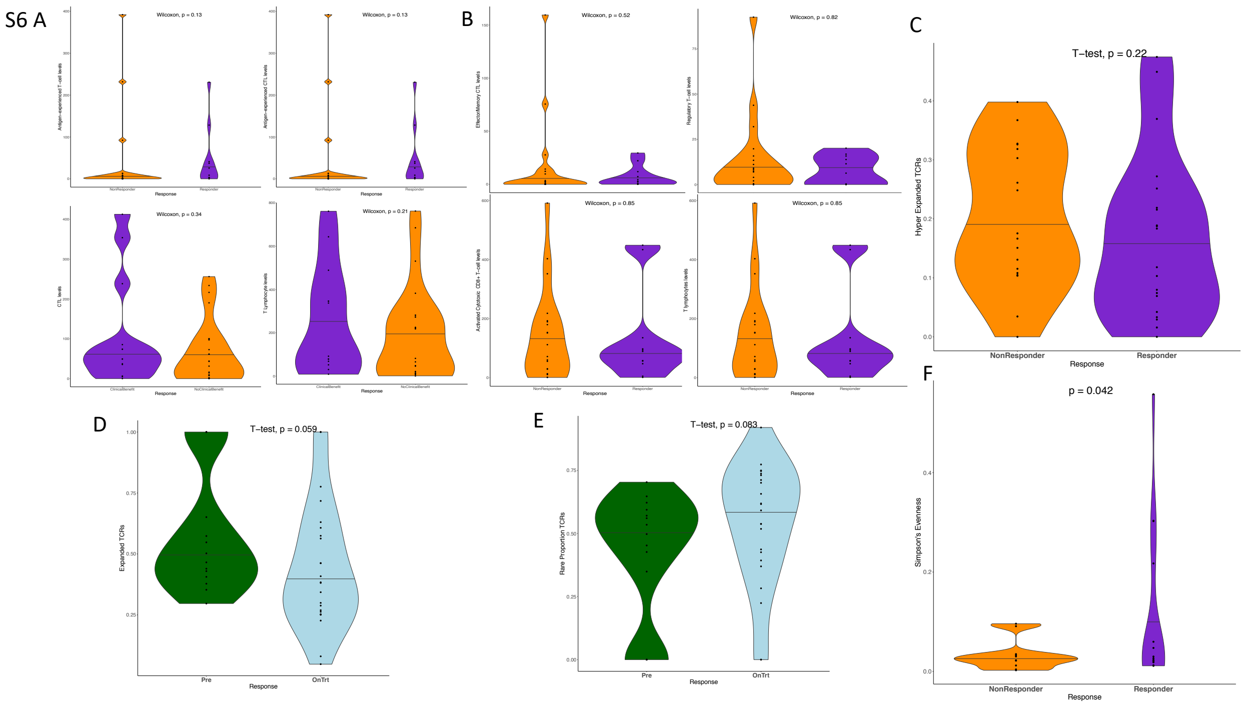
Responder

Non-Responder

FigS4 A: Consensus clustering of the most variable genes in the transcriptome revealed an optimal cluster solution of 3 as depicted by the relative change in area under the cumulative distribution function (CDF) curve which explains the maximum concentration of the consensus distribution. The largest K (optimal number of clusters) was determined at the level that selects the largest K to induce a large enough increase in area under the corresponding CDF. B: CD20 density levels were slightly elevated in responders compared to non-responders at the pre-treatment time-point, however this did not reach statistical significance. C: CD20+ staining for a patient deriving clinical benefit indicates peritumoral staining of B-cells at the tumor periphery D: CD56+ staining for a patient deriving clinical benefit indicates significant infiltration of NK-cells with low staining on malignant cells. E: CD56+ staining for a patient with progressive disease depicts lack of NK-cells within the tumor. F: CD56 density levels were significantly elevated at the on-treatment time point in responders compared to non-responders to ICB. (T-test and Wilcoxon rank-sum test for significance)



FigS5 A: Paired density and correlation plots of multiple TME features deconvoluted across the Riaz dataset with matched pre and on-treatment samples using MCP counter. Value in each box depicts correlation value along with " *** "if the p-value is < 0.001, " ** "if the p-value is < 0.01, " * " if the p-value is < 0.05, else " ". B: B-cell, NK-cell and CD8 T-cell proportions respectively, represented in responders vs non-responders using the Riaz cohort. C: : Using consensus gene panels across both the original SARC028 and expansion cohorts, single gene expression of BANK1 delineated response in the cohort.



FigS6 A: mIF density staining of effector immune-markers for antigen experienced T-cells, Cytotoxic T-Lymphocytes (CTL), Antigen-experienced CTLs and T-Lymphocyte levels were not significantly different between responders and non-responders in the cohort. B: mIF density staining of effector and regulatory immune-markers for effector/memory CTLs, activated cytotoxic CD8 T-cells, regulatory T-cells and T-Lymphocyte levels were not significantly different between responders and non-responders in the cohort. (T-test and Wilcoxon rank-sum test for significance). C: Dominant TCRs were not found to be predictive of response at the pre-treatment time-point. D: A trend of a higher proportion of dominant TCRs were observed at the pre-treatment time-point as compared to the on-treatment samples E: Rare TCRs were found to be non-significantly elevated at the on-treatment time point as compared to the pre-treatment samples. F: Simpsons evenness, a TCR-metric of diversity was found to be elevated in responders as compared to non-responders to ICB. (Two-sided T-test for all)

Chapter Five: Immune Spatial Microenvironment

Contents of this chapter is based on:

Akash Mitra*, Anuj Verma *, Edwin Parra Cuentas, Alicia Mejia, Zhimin Tong, Omkara Veeranki, Ana Garcia, Riham Katkhuda, Latasha Little, Curtis Gumbs, Wayne Hofstetter, Cara Haymaker, Ignacio Wistuba, Andrew Futreal, and Dipen Maru. "Deciphering the Immune Microenvironment of Barrett's Esophagus and Esophageal Adenocarcinoma." In Review

5.1 Introduction

The incidence of esophageal adenocarcinoma (EAC) in the United States increased 57-fold between 1973 and 2010 (Rosemurgy et al., 2019). Most EAC patients present with advanced disease and have very poor survival outcomes; patients with locoregional disease have a 5-year survival rate of 22%, and those with distant metastasis have a median survival duration of less than 20 months. Decreasing the incidence of EAC and improving EAC patients' survival outcomes are hindered primarily by a lack of validated biomarkers for predicting either the progression of Barrett esophagus (BE) to EAC or the response of locoregional EAC to adjuvant or neoadjuvant chemotherapy or chemoradiotherapy.

BE, defined as columnar mucosa with intestinal metaplasia of the tubular esophagus, is present in up to 15% of individuals with frequent symptoms of gastroesophageal reflux disease, and about 2% of the general adult population with the risk of BE progressing to high-grade dysplasia and/or EAC being low (absolute risk, 24 cases per 10,000 person-years) (Cook et al., 2018). The BE epithelium is genetically heterogeneous and clonally diverse (Li et al., 2014, Lai et al., 2007). The low risk of progression and high genetic diversity of BE pose significant challenges in identifying targetable mutations and in focusing preventive resources on patients with a high risk of progression to EAC. Prior studies have demonstrated differences in T-helper cell subtypes between BE and EAC and have analyzed the roles of critical immune mediators (e.g., IL-2, IL-4, IL-6, IL-8, TNF-alpha, and interferon-gamma) in the development of BE (Smith et al., 2009, Kavanagh et al., 2014, Kavanagh et al., 2016, Quante et al., 2012, Buas et al., 2017, Souza et al., 2009). However, achieving success with immune checkpoint inhibitors in the treatment and prevention of EAC requires an improved understanding of the characteristics of the T-cell repertoire in the microenvironments of BE and EAC. Identifying the immune microenvironment attributes that signal a need to induce stronger anti-tumor immune response by immune checkpoint inhibitors or other agents will help determine the risk of BE progressing to EAC and help in designing a biomarker strategy for immunomodulation in BE and EAC.

In this study, we characterized the immune microenvironment of BE and EAC using multiplex immunophenotyping and T-cell receptor (TCR) sequencing. We found that EAC has an immunosuppressive T-cell repertoire and shares TCR clones and V-gene usage with BE; however, EAC lacks an expansion of TCR clones, which would be expected in a tumor whose neoantigen burden is higher than that of a precancerous lesion like BE.

5.2 Results

Patient characteristics

The study population included 10 patients (8 men and 2 women) with median age of 65 years (range, 50 -75 years). Seven patients had a history of gastroesophageal reflux disease and received proton pump inhibitor therapy. Six patients were either current or past tobacco users. Three patients had a normal body mass index, 3 patients were overweight, and 4 patients were obese. BE samples tested for mIF and TCR were procured adjacent to EAC in 7 patients and away from EAC in 3 patients. Six patients had poorly differentiated adenocarcinoma, and 4 patients had moderately differentiated adenocarcinoma. Tumor was confined to the submucosa (T1b) in 5 patients, invaded up to the muscularis propria (T2) in 3 patients, and invaded the adventitia (T3) in 2 patients. Four patients had stage I disease, 3 had stage II disease, and 3 had stage III disease.

Multiplex immunofluorescence

Comparing quantitative changes in population of total T-cells and of T-cell subtypes in different histologic regions revealed significant, immunologically relevant differences in the distributions of T-cells subtypes between EAC and BE.

Intraepithelial compartment**:** Compared with BE, EAC had significantly more median total CD3+ T-cells (EAC= 353/mm^2, BE= 127/mm^2, P= 0.004), cytotoxic T-cells (CD3+CD8+) (EAC= 44/mm^2, BE= 17/mm^2, P= 0.011) (Fig. 1A), activated cytotoxic T-cells (CD3+CD8+granzyme B+) (EAC=6.3/mm^2, BE= 1.1/mm^2, P= 0.023) (Fig. 1B), and T regulatory cells (Tregs; CD3+FOXP3+CD8-) (EAC= 72.8/mm^2, BE= 9.2/mm^2, P= <0.001) (Fig. 1C), but no significantly different numbers of memory T-cells (CD3+CD45RO+) (EAC=43.7/mm^2, BE=14.6/mm^2, P=0.24) or effector/memory T-cells (CD3+CD8+CD45RO+) (EAC=1.7/mm^2, BE=1.7/mm^2, P=0.79). EAC had a significantly higher median Treg/total T-cell ratio than BE (EAC= 0.25, BE= 0.07, P= 0.035), whereas ratios of activated cytotoxic T-cells/total T-cells, memory T-cells/total T-cells, and memory-effector T-cells/total T-cells did not differ significantly between the two lesions, which suggests that the epithelial compartment of EAC has immunosuppressive T-cell infiltrate.

Stromal compartment: Similar to the epithelial compartment, higher median number of cytotoxic T cells (EAC= 210/mm^2, BE= 56/mm^2, P= 0.023) (Fig. 1D), activated cytotoxic T cells (EAC= 24.9/mm^2, BE= 2.5/mm^2, P= 0.007) (Fig. 1E), and T regulatory cells (EAC= 184/mm^2, BE= 79.9/mm^2, P= 0.015) (Fig. 1F) in the stromal compartment of EAC were observed as compared to matched BE. However, there was no significant difference in total T cells in the stromal compartments between BE and EAC (EAC= 1041/mm^2, BE= 528/mm^2, P= 0.063). Moreover, in contrast to the epithelial compartment, the ratio of activated T cells /total T cells in the stromal compartment was higher in EAC as compared to BE (EAC= 0.02, BE= <0.01, P= 0.04), while ratio of Tregs/total T cells (EAC= 0.21, BE= 0.15, P= 0.28) was not significantly different in the stromal compartments of these lesions. Few rare phenotypes were noted in the stroma of BE and EAC. CD3+ CD8+ FOXP3+ cells were significantly seen more in the EAC than in BE (EAC= 7.3/mm^2, BE= 0.7/mm, P= 0.042). Phenotype CD3+ CD45RO+ Granzyme B+ was uniquely seen only in the stroma of BE in 8 cases and not in any EAC sample.

Total (intraepithelial and stromal compartment): Comparing T-cell subtypes in the intraepithelial and stromal compartments in conjunction, we observed more median cytotoxic T-cells (EAC=125/mm^2, BE=41/mm^2, P=0.035) (Fig. 1G), activated cytotoxic T-cells (EAC=17.8/mm^2, BE=1.8/mm^2, P=0.009) (Fig. 1H), and Tregs (EAC=121.4/mm^2, BE=43.7/mm^2, P=0.015) (Fig. 1I-K) in EAC than those in BE. The ratio of activated cytotoxic T-cells/ total T-cells (EAC=.02, BE=.01, P=0.05) in the entire region of analysis was higher in EAC than in BE, while ratio of Tregs/total T-cells (EAC=0.19, BE=0.15, P=0.165) was not significantly different in the entire region of these lesions. Multiplex immunofluorescent images in show higher T-cells, cytotoxic T-cells, memory T-cells and T regulatory cells in EAC than in BE (Fig. S1,S2). Overall, the differences in the quantity and distribution of the T-cell subtypes in BE and EAC indicate T-cell response with higher Tregs in the intraepithelial compartment and higher activated cytotoxic T-cells in stroma suggesting a more suppressive microenvironment at play in EAC. Comparison of clinicopathologic features with distribution of T- cell subtypes did not show any significant correlation in EAC samples, while median number of activated cytotoxic T-cells (intraepithelial: tobacco users= 4.08/mm^2, non-users= 0.32/mm^2, P=0.02), ratio of activated cytotoxic T-cells/T-cells(intraepithelial: tobacco users= 0.028, non-users= 0.001, P=0.01; stroma: tobacco users= 0.011, non-users= 0.002, P=0.02) and ratio of activated cytotoxic T-cells/cytotoxic T-cells(intraepithelial: tobacco users= 0.26, non-users= 0.01, P=0.05; stroma: tobacco users= 0.09, non-users= 0.02, P=0.02) were higher in stromal and intraepithelial compartments of BE in tobacco users as compared to those in tobacco non-users.

TCR sequencing analysis

Given the differences in T cell subtypes' infiltration observed between BE and EAC, we went on to study the attributes of the TCR repertoire between BE and EAC. To this end, we sequenced the CDR3 region of the variable chain of the T-cell Receptor (TCR) beta-chain essential in antigen binding in a paired manner across our cohort. EAC had significantly higher mean numbers of TCR-beta productive templates (11,089 [range, 1,727-27,425]) and productive rearrangements (6,094 [range, 1,359-16,352]) than normal esophagus (2,497 [1,285-6,464] templates; 1,677 [926-4,252] rearrangements) or BE (8,396 [1035-41,480] templates; 4,024 [901-17,923] rearrangements) (p = 0.03 and 0.012 for templates and rearrangements respectively). Unlike normal esophagus or BE, EAC had uniformly distributed productive templates and rearrangements across our cohort, indicating a higher quantitative T-cell response in EAC as compared to BE and normal esophagus (Fig. 2A, B). Using Simpson clonality as a measure of the focus of the TCR repertoire, we found no significant differences between the clonality of the various sites surveyed (Fig. 2C). However, multiple measures of sample richness and abundance revealed d higher TCR-beta diversity in EAC as compared to normal esophagus and BE (p = 0.023, Daley Smith; p = 0.022, Inverse Chao; p = 0.0096, Efron Thisted). The higher levels of diversity in EAC were inversely correlated with an expanded repertoire (Fig. 2D-F).

Using the clonal space occupied by a rearrangement, we computed the representation of various clonotypes within specific proportions. Compared with normal esophagus and BE, EAC had lower levels of expanded clones (p = 0.067) (Fig. 3A, B). For each patient, we found lower levels of largely expanded clones in EAC than in BE or normal esophagus. Top clonotypes had a varied range of clonality, with clonotypes from normal esophagus having greater clonal proportions than BE and EAC. However, the top 10 clonotypes extracted from EAC or BE overlapped remarkably, implying potential shared antigenicity (p = 0.0273) (Fig. 3C). These expanded clonotypes were shared in a public manner among BE, EAC, and normal esophagus samples derived from each patient, but not among samples derived from different patients. These results may also indicate a lack of tumor-driven clonal expansion.

Assessing the overlap in the T-cell repertoire revealed significant sharing of TCR-beta clones among patients (Morisita Overlap Index ≥0.5). A higher degree of sharing was observed between normal esophagus and BE, with fewer samples sharing BE and EAC, and the lowest sharing observed between EAC and normal esophagus (Fig. 3D). An analysis of matched samples revealed remarkably similar V-gene usage among normal esophagus, BE, and EAC. V-gene usage did not differ significantly between any of the samples, with BE and EAC samples demonstrating V-gene usage for an allele that was absent in matched normal esophagus samples in only 2 patients (Pat 5 and 6). Moreover, except for those from patient 6, EAC and BE samples had no quantitative differences in V-gene usage for any alleles (Figure. 4A, B). Shannon–Johnson entropy analysis for V-gene usage also demonstrated no significant difference in entropy between BE and EAC (Figure. 4C). Similarly, mapping recurrent amino acid motifs to build k-mers using the TCR rearrangements revealed that normal esophagus, BE, and EAC had broadly similar amino acid motifs in the CDR3 sequence (Figure. 4D-F). In summary, our findings show that EAC has a more diverse TCR-beta repertoire than BE but shares clones and V-gene usage with BE. However, EAC lacks the clonal expansion and hyper-expanded clones that have been observed in other solid malignancies with high neoantigen burden.

Through a composite analysis of the TCR-sequencing and mIF data, we found diversity of the TCR repertoire correlated with infiltrating T-lymphocytes ($\rho = 0.68$ and $p = 0.002$ for Daley Smith Estimator and $\rho = 0.63$ and $p = 0.0007$ for inverse Chao) (Figure 4G). Additionally, all the metrics for repertoire richness indicated the highest level of correlation with a regulatory T-cell phenotype ($\rho = 0.71$ and $p = 0.0004$ for Daley Smith Estimator, $\rho = 0.63$ and $p = 0.002$ for inverseChao and $\rho = 0.57$ and $p = 0.009$ for Efron thisted estimator). While the clonality of the repertoire indicated a complementary enrichment for an activated immune phenotype ($\rho = 0.32$ and $p = 0.16$) as compared to a regulatory phenotype ($\rho = -0.39$ and $p = 0.08$). When analyzing the BE and EAC cohorts separately, we see the clonality of the repertoire being linked to an activated immune phenotype and the diversity of the TCR repertoire being correlated with a regulatory immune phenotype in both groups, however these did not hold to our significance testing thresholds (supplementary figure 3). These findings may indicate that while both BE and EAC lesions are infiltrated with T- lymphocytes, the majority of the infiltrate may be acting in an inhibitory role by regulating the immune system with a lack of T-cell expansion correlating with low levels of cytotoxic T-cell killing.

5.3 Discussion

In the present study, mIF and image analysis based immunophenotyping revealed a greater T-cell response in EAC than in BE. However, BE and EAC had significantly different distributions of T-cell sub-populations. EAC had higher ratios of cytotoxic T-cells/total cells and activated cytotoxic T-cells/cytotoxic T-cells in the stromal compartment than in the intraepithelial compartment, indicating a cytotoxic T-cell immune response that is less targeted towards the tumor cells. The presence of more Tregs in EAC as compared to BE, especially in the epithelial component, suggest that the higher quantitative T-cell response in EAC is immunosuppressive and elicits limited tumor cell cytotoxicity.

Previous studies in other solid tumors and preneoplastic lesions have demonstrated that FOXP3+ Tregs have a role in determining clinicopathologic features and/or patient outcomes (Ishibashi et al., 2006, Vacchelli et al., 2015, van der Linden et al., 2018). Kahraman et al. demonstrated that malignant ovarian tumors have higher Treg infiltration than benign lesions and borderline tumors (Kahraman et al., 2018). In estrogen receptor–negative breast cancer, an increased abundance of Tregs is associated with a higher histologic grade (Mahmoud et al., 2011) while in patients with metastatic colorectal cancer, high numbers of Tregs and CD8+ T-cells expressing chemokine receptor 7 are associated with better survival outcomes (Correale et al., 2012). High FOXP3+ Treg infiltration is associated with high T stage and poor survival outcomes in gastric cancer patients and with a higher recurrence rate in hepatocellular carcinoma patients (Li et al., 2019, Sasaki et al., 2008). Noble et al., in a study of EAC patients treated with neoadjuvant therapy and surgery, demonstrated that patients who had a high proportion of FOXP3+ Tregs had better survival; however, having more FOXP3+ Tregs was not associated with improved pathologic response to neoadjuvant therapy (Noble et al., 2016).

Our study was the first to assess the T-cell repertoires of matched BE and EAC samples. We found that EAC had more diverse, TCR-beta productive templates and rearrangements than BE, indicating that the TCR-beta repertoire in EAC has the potential to respond to a significantly larger variety of antigens. However, the TCR-beta clonal compositions of normal esophagus, BE, and EAC not only shared passenger TCR-beta clones but also shared the most expanded clones across each lesion. Limited diversity in the V-gene usage and tumor-specific clonal expansion, including a lack of expansion in hyperexpanded clones in EAC, further support limited T-cell response that targets antigens specific to EAC. Chen et al. demonstrated that the TCR-beta repertoire in squamous carcinoma differed significantly from those in normal tissue and peripheral blood, whereas the repertoires in different regions of the same tumor were similar (Chen et al., 2016b). Other groups reported similar findings; in particular, ovarian cancer had a distinct but homogenous repertoire compared with normal tissue, indicating a T-cell response more specific to the tumor antigens (Emerson et al., 2013, Sherwood et al., 2013).

Given our study's small sample size and limited evidence of the impact of heterogeneity on the immune microenvironment of BE and EAC, we consider our findings to be preliminary. Futures studies should investigate whether field effects arising from the procurement of some BE samples adjacent to EAC result in the higher degree of TCR sharing between BE and EAC. Although, immunophenotyping data showing significant differences between the immune subtypes in adjacent BE and EAC samples potentially argues against field effects having a significant impact in this respect.

In summary, our findings indicate that compared with BE, EAC has a more infiltrated and diverse but immunosuppressive T-cell infiltrate, the clonal expansion of which is limited and non-targeted. Thus, in EAC, the quantitative increase in immune infiltrate is not associated with a qualitative antitumor immune response against tumor cell antigens, supporting a need to modulate the immune microenvironment to induce tumor cell directed T-cell response.

5.4 Figures and Tables

Figure 1

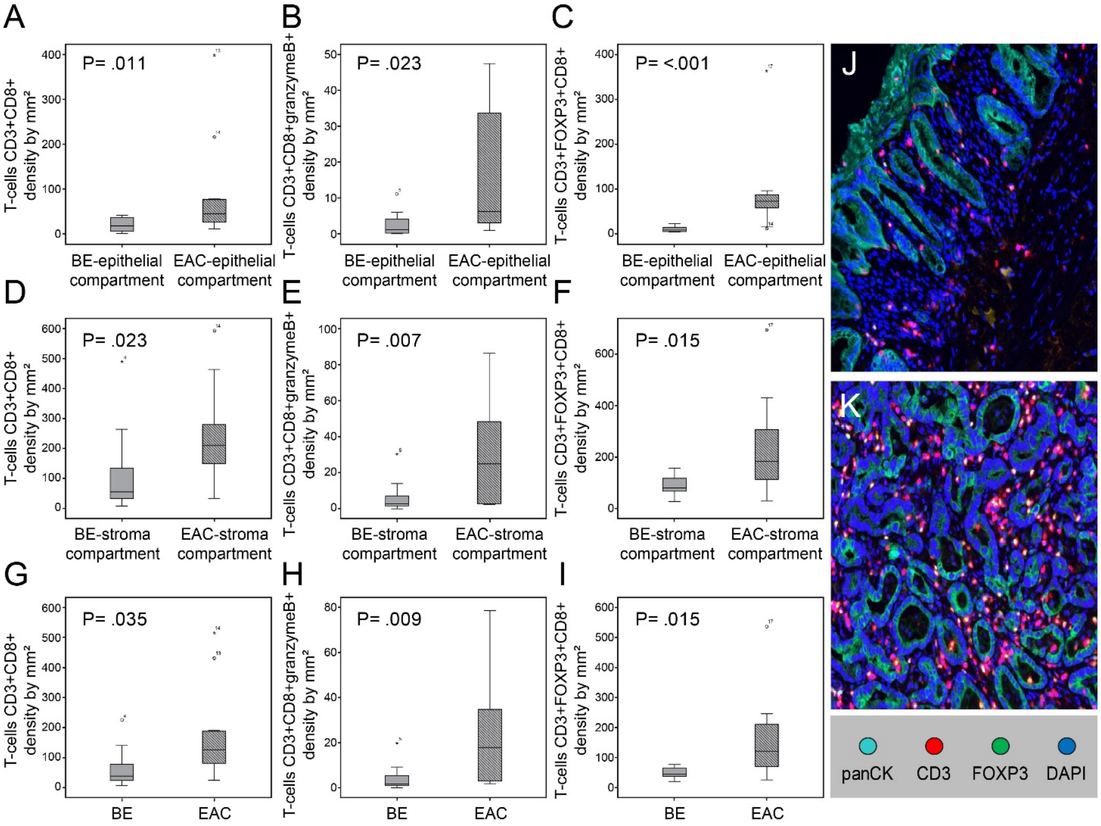


Figure-1: Box plot showing significantly higher median values of CD3+CD8+, CD3+CD8+Granzyme B+, and CD3+FOXP3+CD8- cells in EAC when compared with BE in the epithelial compartment (A to C), in the stromal compartment (D to F) and the entire BE and EAC (G-I). Multiplex immunofluorescence images (x200) showing a lower density of T regulatory cells (CD3+ FOXP3+) seen as yellow nuclear staining in the stroma and epithelial component of BE (J) compared to EAC (K) EAC- Esophageal Adenocarcinoma BE-Barrett's Esophagus

Figure 2

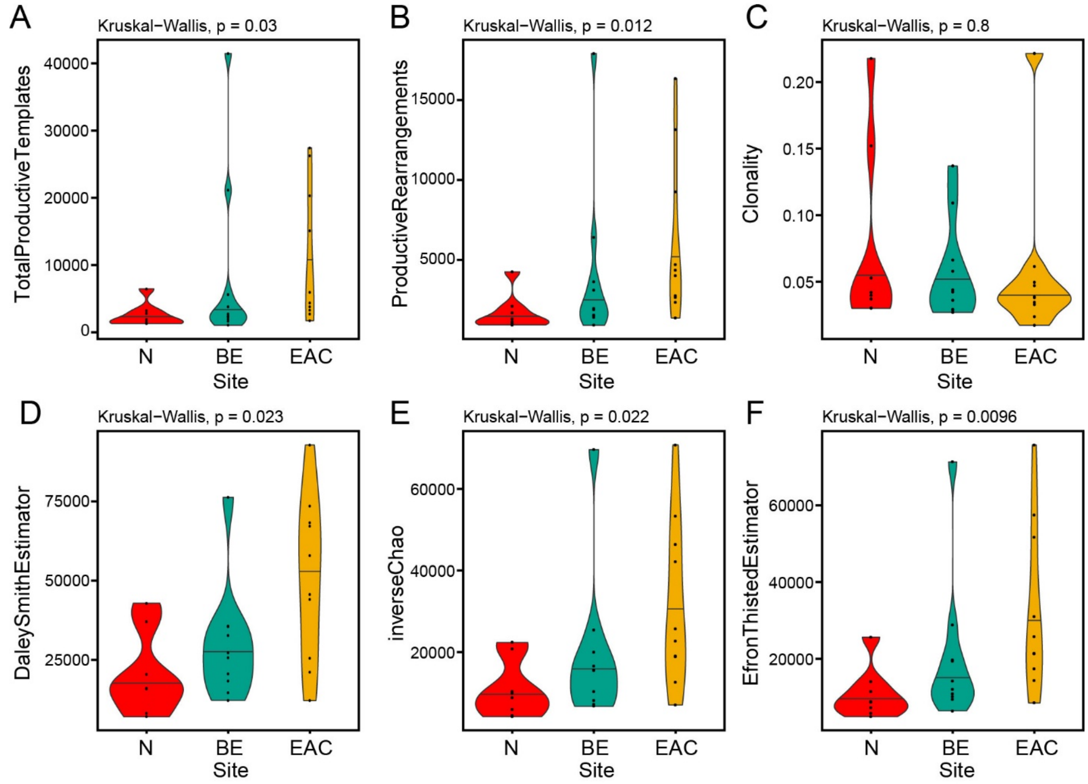


Figure-2: The CDR3 sequence of the TCR-beta chain was subject to high throughput sequencing of paired normal esophagus (N), Barrett's esophagus (B) and esophageal adenocarcinoma (T) across our cohort. A and B) Higher levels of productive templates as well as productive rearrangements were detected in T as compared to B and N. C) Simpsons clonality indicated non-significant enrichment of a clonal TCR repertoire present in N as compared to B and T. D-F) TCR-beta diversity (as a measure of richness and abundance) was computed and significantly higher levels of diversity were found in T as compared to B and N (D: Daley Smith, E: inverse Chao and F: Efron Thisted estimator).

Figure 3

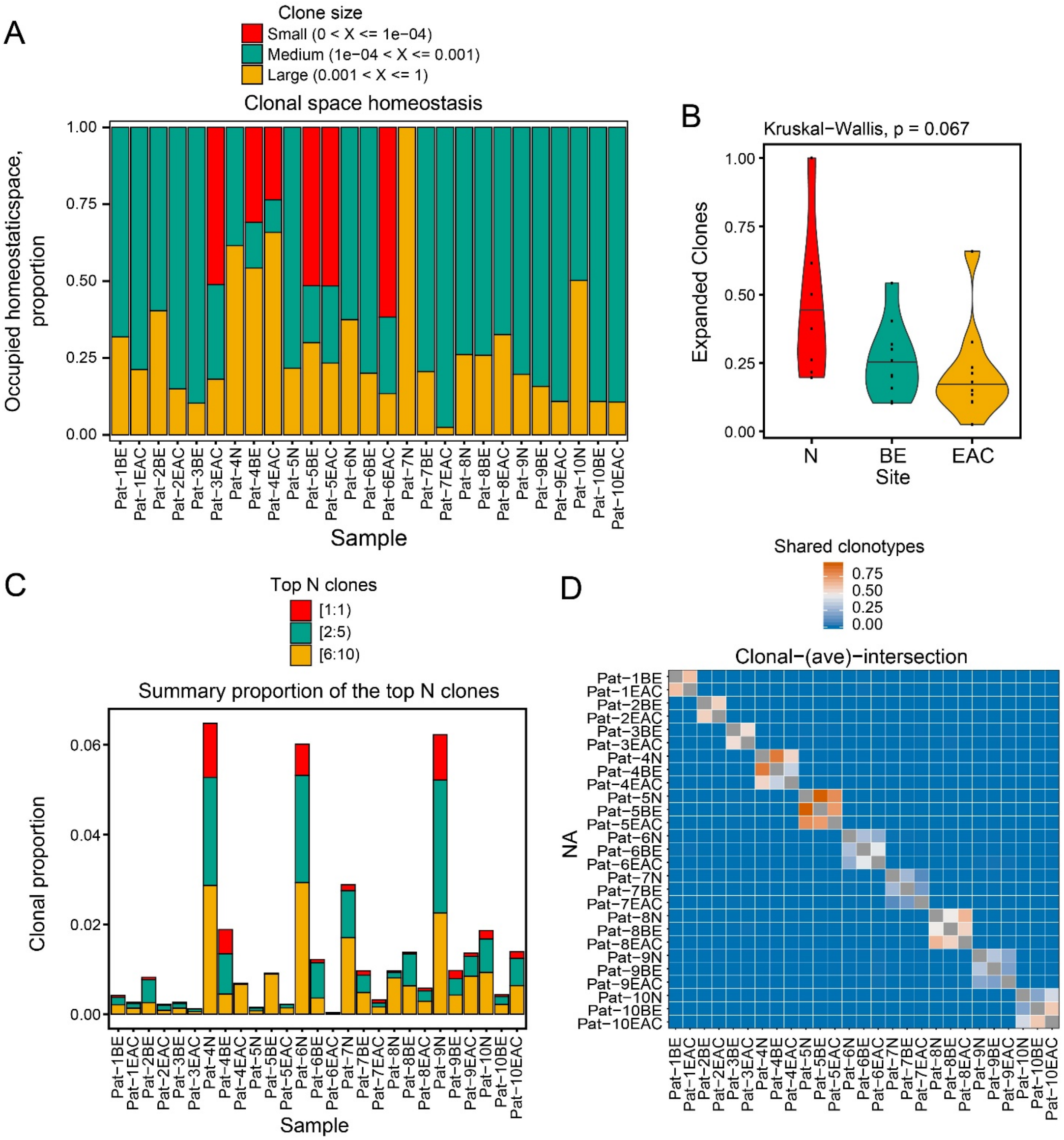


Figure-3: Sample IDs include patient indicators, following by site of sample. A) Examining the TCR repertoire, we characterized individual clonotypes based on their contributions towards the clonal space (indicated in figure) occupied by a given sample. B) Samples derived from N displayed higher levels of largely expanded clones as compared to B and T proportions. C) The top 10 clonotypes from each sample were computed with the clonality of the top clone, clones 2-5 and clones 6-10 from each sample showcasing a higher level of clonality present in N as compared to matched B and T. D) The Morisita Overlap index is a population overlap metric relating the dispersion of clones in our sample pairings. Clear levels of intra-patient sharing are depicted along the diagonal, however the patterns of within patient site-based sharing varied between each patient case.

Figure 4

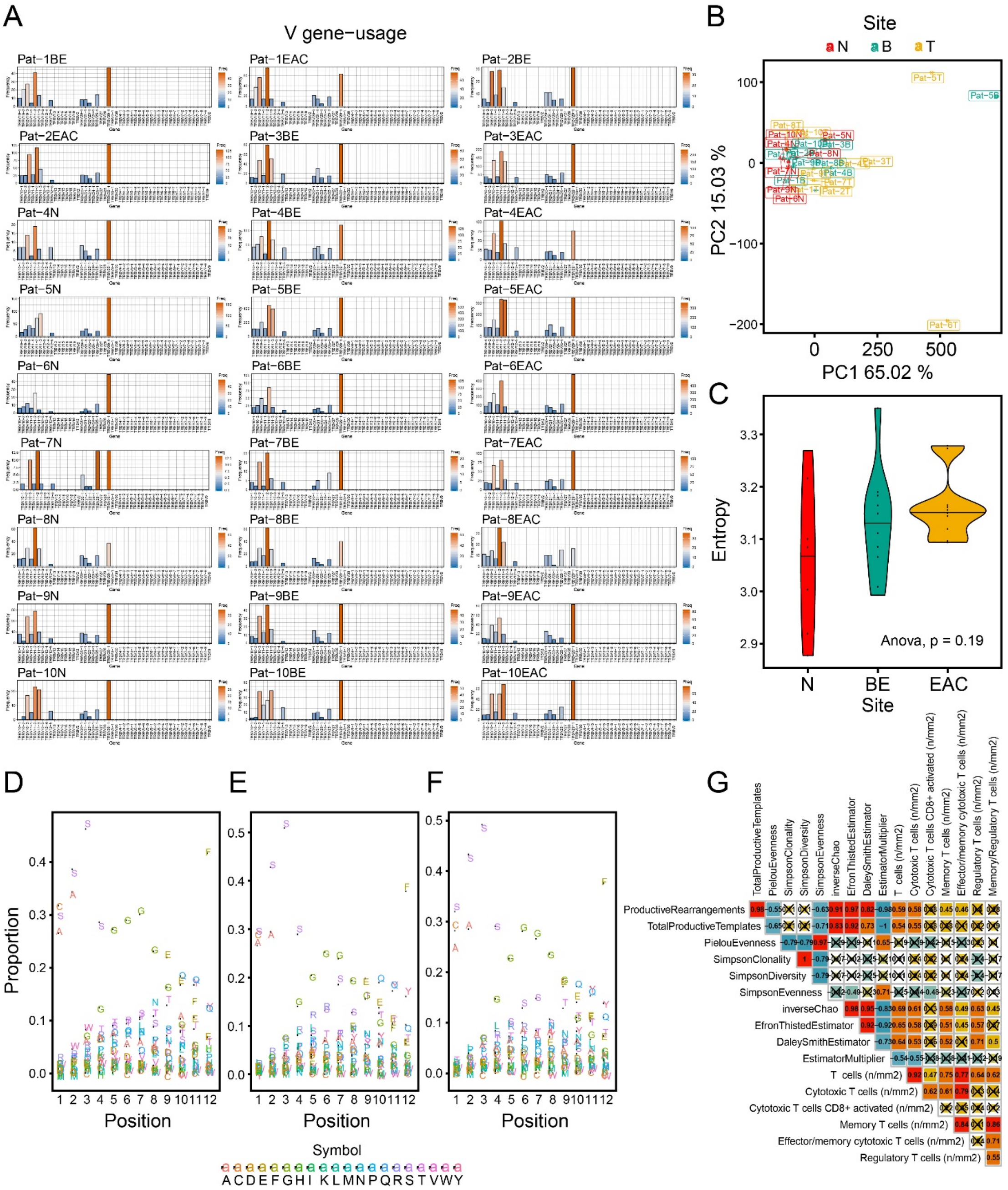


Figure-4: Variable (V), Diversity (D) and Joining (J) gene alleles were computed using high throughput sequencing. A) Bar graphs depict V gene usage in an allele specific manner with V-gene alleles plotted on the x-axis and frequency of each allele on the y-axis. Most paired N, B and T samples depict similar V-gene usage apart from Pat5T and Pat5B compared to Pat5N, and Pat6B and Pat6T. B) PCA-plot with PC1 on x-axis and PC-2 on y-axis depict primarily clustering of all samples based on V-gene usage, apart from samples previously identified. C) Shannon Johnson entropy analysis for V-gene indicates minor differences between B and T, however differences between them and the paired N. D-E) Consensus 12-mers were generated along each site displaying similar amino acid motifs in the CDR3 sequence of D) Normal esophagus, E) Barrett's Esophagus and F) Esophageal adenocarcinoma G) Correlation matrix showing the correlation values between TCR diversity/clonality metrics and immune infiltrate (color scale red to blue implying higher positive to negative correlation respectively; boxes with "x" did not hold to significance testing thresholds of $p < 0.05$.

Supplemental Information

Antibodies			
Target antigen	Clone	Dilution	Company (location)
AE1/AE3		1:300	Dako (Carpentaria, CA)
CD3		1:100	Dako
CD8	C8/144B	1:20	Thermo Fisher Scientific (Waltham, MA)
CD45RO	UCHL	Ready to use	Leica Biosystems (Wetzlar, Germany)
Granzyme B	F1	Ready to use	Leica Biosystems
FOXP3	206D	1:50	BioLegend (San Diego, CA)

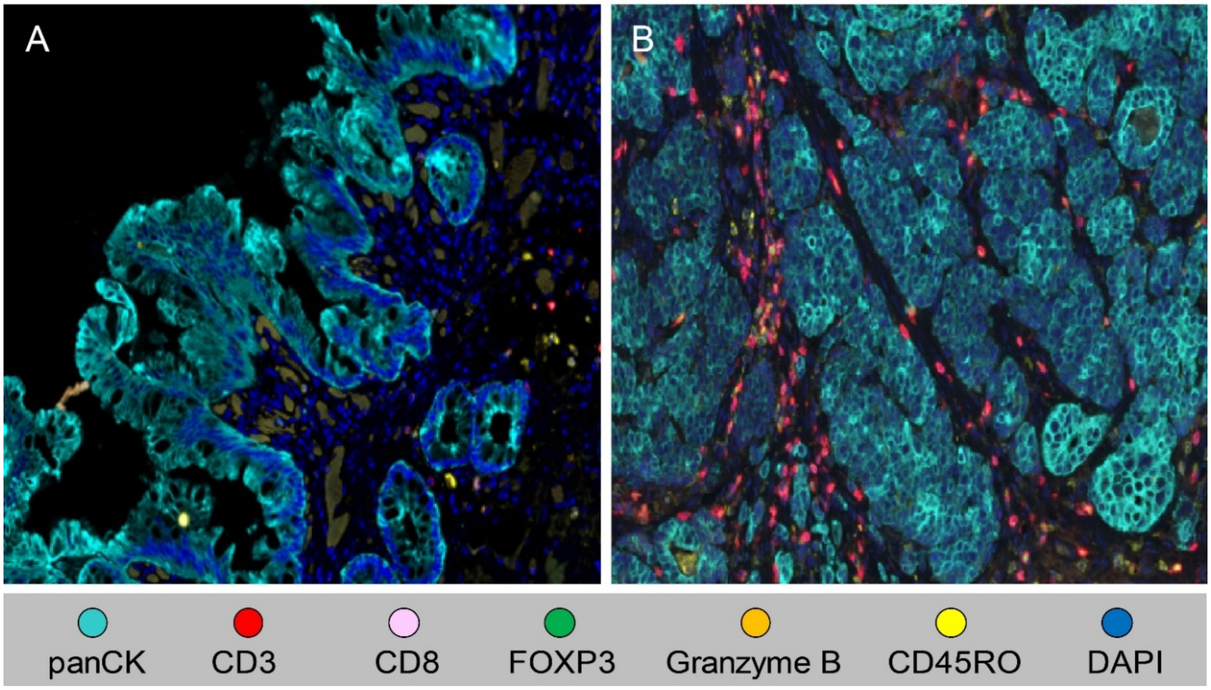
Cells

Cell type	Phenotype
EAC or BE epithelial cells	AE1/AE3+
T lymphocytes	CD3+
Cytotoxic T lymphocytes	CD3+ CD8+
Activated cytotoxic T lymphocytes	CD3+ CD8+ Granzyme B+

Memory T lymphocytes	CD3+ CD45RO+
Effector/memory cytotoxic T lymphocytes	CD3+ CD8+ CD45RO+
Regulatory T lymphocytes	CD3+ FOXP3+ CD8-
Memory/regulatory T lymphocytes	CD3+ FOXP3+ CD45RO+ CD8-

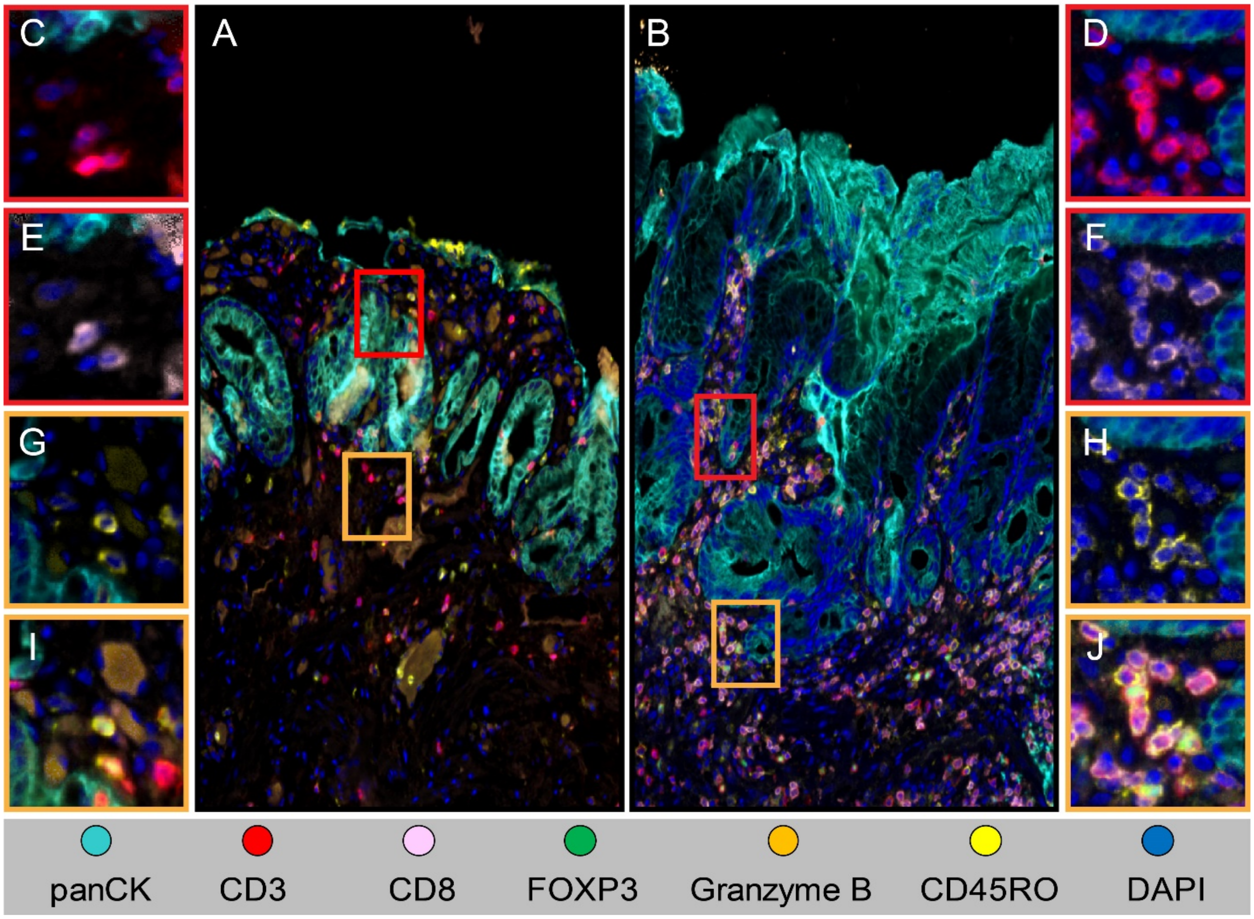
Supplemental Table 1: Details of the antibodies used for cell staining and the phenotypes identified using the antibodies

Supplemental Figure 1


A
B
panCK
CD3
CD8
FOXP3
Granzyme B
CD45RO
DAPI

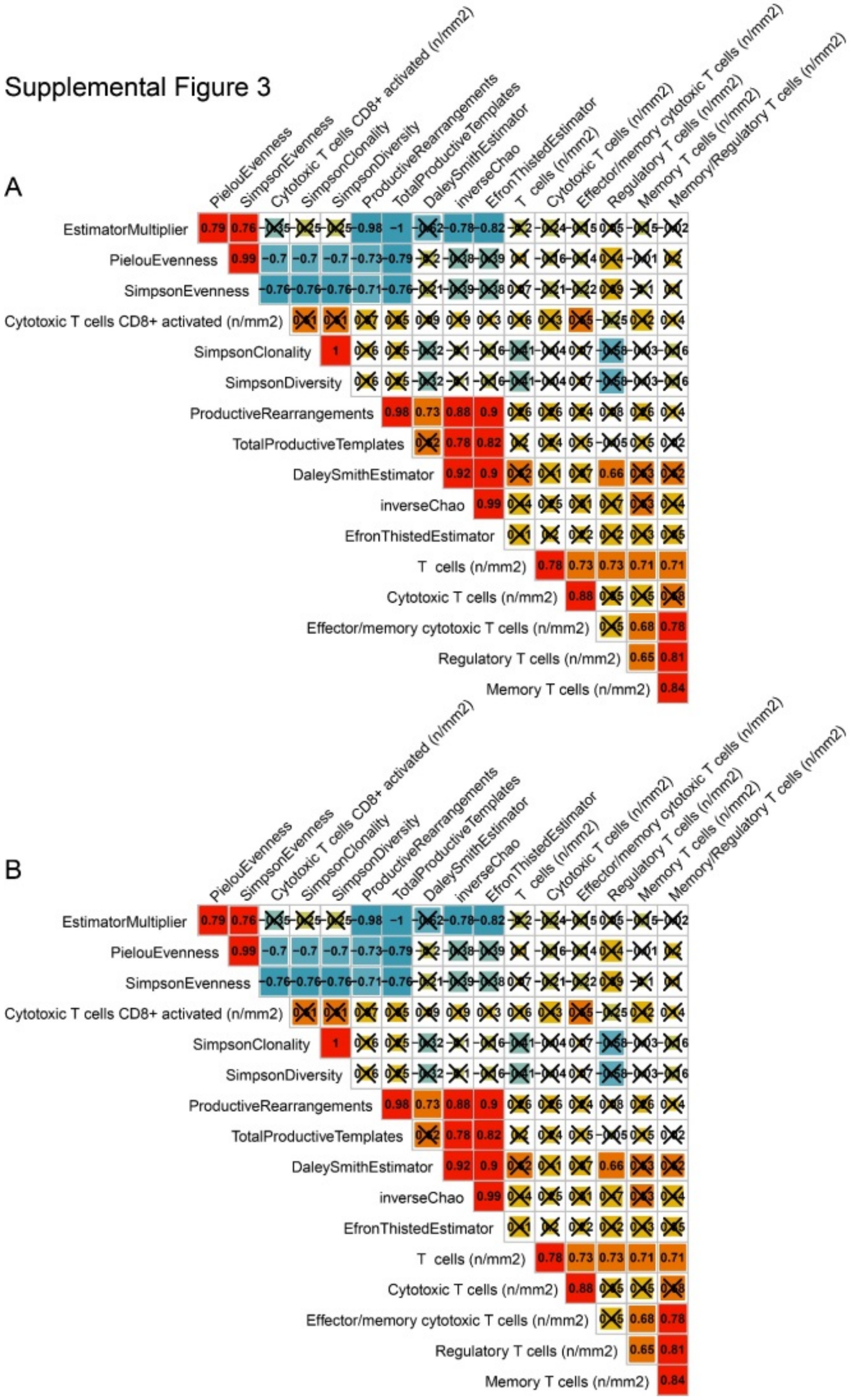
Supplemental Figure 1: Multiplex immunofluorescence (x200) images showing lower density of immune cells infiltrating in Barrett's Esophagus (A) than in Esophageal Adenocarcinoma (B).

Supplemental Figure 2


C
A
B
D
E
F
G
H
I
J
panCK
CD3
CD8
FOXP3
Granzyme B
CD45RO
DAPI

Supplemental Figure 2: Composite mIF low power(x200) images showing lower density of immune cells infiltrating the BE than in EAC (A and B). EAC shows a higher infiltrate of CD3+ T-cells and CD3+CD8+ Cytotoxic T-cells than BE (C-F). Lower CD45RO+ Memory T-cells in BE (G) than in EAC (H). Higher density of CD3+ FOXP3+ CD8- T regulatory cells in EAC (J) than in BE (I). mIF- Multiplex Immunofluorescence, BE- Barrett's Esophagus, EAC- Esophageal Adenocarcinoma

Supplemental Figure 3



Supplemental Figure 3: Correlation matrix between TCR repertoire characteristics and immune infiltrate in Barrett's esophagus(A) and Esophageal Adenocarcinoma(B)

Chapter Six: Discussion, Conclusions and Future Directions

6.1 Discussion and Conclusions:

With recent advances in genomic and spatial molecular profiling, we have an unprecedented lens through which we can characterize tumors and their surrounding neighborhoods. Our understanding of tumor heterogeneity from the times of Nowell have made significant advances however, certain central questions remain. Tumors evolve through space and time, in combination with positive and negative signals from the surrounding microenvironmental niche. While several studies have highlighted heterogeneity in the context of time and through therapeutic intervention, only a subset of studies have attempted to uncover the three-dimensional nature of tumors and their microenvironment(Snyder et al., 2014a, McGranahan et al., 2016, Rizvi et al., 2016, Jiménez-Sánchez et al., 2017, Hellmann et al., 2018, Chen et al., 2016a, Roh et al., 2017, Gerlinger et al., 2012a, Gerlinger et al., 2014, Jamal-Hanjani et al., 2017b, Joshi et al., 2019, Thrane et al., 2018, Tirosh et al., 2016). Given that tumor heterogeneity does not occur in silos, there is a growing demand to characterize tumor heterogeneity as it interplays with the immune (and stromal) microenvironment in the context of therapeutic response. ICB has demonstrated prolonged survival and durable responses in a subset of patients. However, given the variable response rates and occurrence of immune-related toxicity/adverse events, we require better molecular predictors for efficient patient stratification. In addition to biomarker discovery, characterization of the underlying mechanisms of resistance to ICB may help us in furthering the development of immunotherapies and the rationale for other combinatory agents. Through my work, I attempted to characterize a spatially aware understanding of the tumor microenvironment while retaining three-dimensional information, to elucidate genomic and non-genomic features and how their relationships may govern response and resistance to ICB. Further, through a longitudinal biopsy collection strategy I uncovered novel pre-treatment biomarkers and resistance mechanisms to combination ICB (CTLA-4 and PD-L1 blockade) through multidimensional profiling of tumor biopsies of patients with advanced sarcoma. Finally, to better understand disease progression and the interplay with the immune system, I characterized the immune contexture surrounding Barrett's Esophagus and high-grade dysplasia or esophageal adenocarcinoma.

In chapter 3, I used multiplatform profiling of genetic, epigenetic, transcriptomic and immune features to characterize the tumor and its surrounding microenvironment. As expected in the case of metastatic melanoma, we found sparse mutational intratumor heterogeneity. However, through the copy number profiling, we found distinct gains of chromosome 7 and chromosome 13 and loss of chromosome 10 in a spatially non-contiguous format. This finding of non-adjacent tumor regions showcasing similar genomic features is consistent with other studies showcasing evidence of parallel evolution within a tumor mass(Birkeland et al., 2018). Through transcriptomic profiling, we uncovered the upregulation of several key immune modulators including an effector T-cell response and B-cell activation however, this effector immune population was primarily located at the tumor margins and not within the core of the tumor. This is consistent with previous studies highlighting non-responding patients to have evidence immune exclusion at the periphery of the tumor bed(Tumeh et al., 2014). Sequencing of the beta region of the T-cell receptor revealed a dominant TCR clonotype that was found to be expanded at high frequencies within distinct regions of the tumor as well as other tumor lesions across a time span of eight years having received multiple targeted and immunotherapeutic regimens. This expanded clonotype was immunophenotyped using paired single cell RNA and TCRA/TCRB sequencing to reveal both activated and exhausted phenotypes suggesting a potential rise from multiple independent T-cell priming events. Moreover, we integrated our genetic and immune related findings to find distinct regions with low immune infiltrate (not affected through tumor core versus margin analysis) accompanied by gains in chromosome 7. While we validated this finding across multiple cohorts, other studies have recently shown a lower copy number burden being correlated with better overall survival across other tumor types(Lu et al., 2020). These studies have also shown that this copy number loss is associated with activated immune engagement, in particular lymphocyte activation and interferon signaling(Chen et al., 2016a, Roh et al., 2017, Lu et al., 2020).

However, there are caveats in interpreting the immunogenomic relationships in this study. Given that we primarily studied one given time-point, we were unable to ascertain whether the neutrophil activation/recruitment was due to anti-tumor-immune responses or whether their presence was a result of cellular destruction and necrosis through other means. The extensive immunogenomic heterogeneity is inherently limited through the multidimensional profiling of one tumor lesion. Thus, it is difficult to ascertain how typical this heterogeneity would be reflected in a broader population setting, particularly those with differing clinical regimens. Additionally, with the advent of spatial single-cell sequencing, studies have demonstrated a higher resolution view of the tumor-immune cross talk while retaining 2-dimensional information in several tumor types(Moncada et al., 2020, Berglund et al., 2018, Casasent et al., 2018).

In chapter four, we utilized a longitudinal biopsy sampling strategy to profile the dynamic changes tumors undergo through combination ICB and to delineate predictors of response to ICB at the pre-treatment setting in sarcoma. This study demonstrated adaptive mechanisms of resistance to ICB that have been shown in other tumor types but not within the context of sarcoma. Interestingly, the primary immune cell subset that was enriched in responders compared to non-responders at the pre-treatment setting were B-cells. While Fridman and colleagues demonstrated robust profiling of the tumor microenvironment and multiple immune cells that were found to be enriched within certain sarcomas, our work went on to show a sole B-cell infiltrate that correlated with response(Petitprez et al., 2020). Additionally, we did not observe instances of tertiary lymphoid structures or T-cell derived immune engagement within responders. The presence of B-cells at the pre-treatment setting may assist in the activation or recruitment of other immune effector cells. Through the secretion of cytokines including TNF, IL-2 and IFNG, B-cells can alter immune constituents and result in active immune engagement. In parallel to these findings, we also observed an increase in the TCR repertoire diversity at the pre-treatment setting in responders. The increased levels of diversity may provide an advantage to combination ICB through the availability of multiple TCR clonotypes, having a higher probability of a particular clonotype being tumor-directed. The combination of higher B-cell proliferation and T-cell diversity may work in conjunction to create a more robust effector immune response, however further studies would be required in characterizing their interplay in the setting of combination ICB.

However, this study provided significant challenges in interpretation, in particular dealing with multiple histologies of sarcoma subtypes that were enrolled on trial. While genomic and transcriptional analysis included the confounders of batch effects and the histological subtypes, some of these findings may still be driven through factors aside from response. Additionally, given that sarcomas represent rare tumors, the power for predicting response in any one particular subtype was low. In order to curb the low sample size, we included the only other sarcoma ICB trial to date with matched molecular data(Tawbi et al., 2017). While we sequenced the TCR repertoire, our findings on the BCR repertoire included an inference of the BCR from the transcriptomic data. While several algorithms attempt to infer the BCR from RNA-seq data, biological challenges in terms of immune rearrangements as well as technical challenges as a result of read depth may influence these results.

Chapter five highlighted the relationship of multiple immune cell subsets in the context of Barrett's esophagus (BE) and high-grade dysplasia. While studies have highlighted the role of the inflammatory microenvironment on T-cell phenotypes, which may lead to the development of various tumors, we still lack a complete understanding of the microenvironmental niche and the role of the immune repertoire present in patients with BE and EAC(Quante et al., 2012, Kavanagh et al., 2014, Kavanagh et al., 2016, Noble et al., 2016, Buas et al., 2017). Through our findings, we characterized an increase in effector and regulatory immune populations in the context of esophageal adenocarcinoma (EAC). While we did detect elevated levels of cytotoxic T-lymphocyte in EAC, these were mostly restricted to the stroma. Parallel identification of an increase in regulatory T-cells may suggest an inflamed yet immunosuppressed tumor microenvironment. Previous studies have reported an increase in macrophage and regulatory T-cell components as well as an increased levels of exhausted T-cells in the progression of EAC(Zheng et al., 2020). Additionally, we found significantly higher measures of TCR diversity present in the EAC of these patients, potentially due to a larger proportion of mutations and neoantigens through disease progression. Surprisingly, we observed high levels of overlap in the expanded clonotypes at both disease sites. Through a composite analysis, we found that an increase in the TCR repertoire richness was most highly correlated with an immunosuppressive phenotype which may indicate that while both BE and EAC lesions are infiltrated with T-lymphocytes, the majority of the infiltrate may be acting in a suppressive role by regulating the immune system with a lack of T-cell expansion correlating with low levels of cytotoxic T-cell killing.

However, the majority of these findings are limited to analysis that are solely focused on the immune repertoire and the varying tumor microenvironments between BE and EAC. While we were able to phenotype the immune populations within these two distinct regions, the use of transcriptomic modalities such as bulk or RNA-sequencing would provide us a global picture of the various immune and stromal cell types that characterize the different microenvironmental niches in Barrett's and esophageal cancer. While inflammation has been shown to be a significant predictive factor in the development of EAC, a more robust characterization of effector, exhausted and regulatory immune cells is needed. Additionally, a high degree of TCR overlap between the distinct regions may suggest similar immunogenic peptides being secreted at both sites. This may be due to mutations that are clonal and found in early stages of BE and remain present through the development of EAC. Genomic characterization of these regions may help us better understand the clonal nature of mutations and more importantly neoantigens in the context of tumor development in EAC.

6.2 Future Directions

Through Chapter 3, I attempted to characterize the varying levels of spatial ITH that we observe in the context of metastatic melanoma along with the interplay of the TME. Significant work has gone into furthering our understanding of spatial ITH, both at the bulk tumor level through umbrella studies such as TRACERx as well as through advances in single cell spatial sequencing(Satija et al., 2015, Casasent et al., 2018, Moncada et al., 2020, Jamal-Hanjani et al., 2017b, Biswas et al., 2019, Joshi et al., 2019).

Firstly, this study profiled one patient (with three biopsies) at unprecedented resolution. While our integrative analysis was validated across public cohorts (at the bulk level), we would obtain a finer understanding through prospective collection and the establishment of a larger cohort with similar clinical and pathological features. Through the collection of these large multi-region sequencing datasets, we can gain an understanding of the molecular underpinnings of the disease at play. Seminal work by Swanton and colleagues, has established the role of clonal neoantigens in the development and targeting of therapeutics towards lung cancer(McGranahan et al., 2016). A similar approach in the context of melanoma, a tumor type in which neoantigen derived long peptide vaccines play a significant role, would be worthwhile(Hu et al., 2021).

Given the advances in spatial transcriptomic single cell sequencing, we could better deconvolute the immune architecture and the various components of the TME at play while retaining spatial information. Through paired transcriptomic and locational information, we could then construct clusters of cells to most accurately phenotype multiple cell subtypes and use this information downstream to construct cellular neighborhoods by characterizing the neighbors of any given cell. Through combined cell type and cellular neighborhood reconstruction, we could then observe how these cellular clusters function within each particular neighborhood and how neighborhood compositions can in turn affect cellular functions. Additionally, with advances in single cell DNA sequencing, the inference of mutations and copy number profiles would allow us to better integrate the genomic and transcriptomic data(Wang et al., 2014, Minussi et al., 2021). While our work primarily showcased integration of multi-region sequencing, the single cell resolution of genomic features, in particular the amplification of chromosome 7 and how it affects local immune contextures could be further interrogated.

In Chapter 4 we performed longitudinal collection of biopsies of patients with advanced sarcoma on exposure to CICB. While our cohort represents the largest collection of sarcoma patients treated with CICB, our cohort was restricted to 57 patients with varying histologies. While a subset of patients did show response, this was most significantly noted in the group with alveolar soft part sarcoma (ASPS). Other unpublished studies have highlighted the high objective response rate within this histology but further interrogation of the pathognomonic ASPSCR1-TFE3 fusion is required. This unbalanced recurrent translocation can lead to the generation of neoantigens and through the induction of CICB, may result in tumor shrinkage. Additional means to expand analysis could include the incorporation of liquid biopsy to survey adaptive immune signatures that may be present in the blood. Through transcriptomic analysis, we may be able to identify peripheral immune subsets associated with favorable response.

Moreover, we uncovered a sole B-cell infiltrate to be strongly predictive of response in our cohort. However, further studies would be needed to better understand the interplay of T- and B-cells in response to CICB in sarcoma. These immune cell-types, especially in conjunction with tertiary lymphoid structures, may create sites for the generation of an active anti-tumor immune response. We also observed increased levels of TCR and IGH diversity present at the pre-treatment time point in responders. Validation of IGH diversity and a deeper characterization of the BCR repertoire through BCR-sequencing would provide us with a more robust understanding of the immune repertoire on exposure to CICB.

In Chapter 5, we attempted to uncover the differing immune microenvironments between BE and EAC. While the majority of this analysis was focused on the immune contexture, our findings were restricted to regulatory and immune cell phenotypes. Through the incorporation of multiple marker panels, we could extend our findings to include other immune subpopulations including the dendritic cells, NK-cells, macrophages as well as exhausted immune cell-types. A deeper understanding of the spatial neighborhoods within these multiple marker panels may assist in better understanding the crosstalk between multiple cell sub-types. Additionally, comparing transcriptional states between the immune cell located at the periphery versus infiltrated cells may also further our understanding of the immune microenvironment within BE and EAC.

Finally, the addition of genomic data would greatly increase our understanding of the shared overlap in the TCR repertoire. Through the use of whole genome sequencing and neoantigen prediction algorithms, we could uncover potential neoantigens that are conserved and shared between the BE and EAC. Given that neoantigen intratumor heterogeneity may have a significant impact on anti-tumor immunity, a thorough integrated analysis of ITH, along with neoantigen burden and how this pertains to the immune repertoire may improve therapeutic opportunities for individuals with esophageal adenocarcinoma.

Through my work, I've established the multiple layers of intratumor heterogeneity and how the tumor-immune crosstalk in the tumor microenvironment can shape disease progression and response to therapy. While these studies were undertaken across multiple tumor types, it's clear that studying the tumor or immune heterogeneity in silos do not paint a complete picture of the molecular changes that cells undergo within the tumor and its surrounding microenvironment. Through the spatial incorporation of the tumor along with its neighboring cells, we can understand how different cell-types and cellular-neighborhoods influence tumor progression and response to therapy. Through in-depth longitudinal collection, we can better understand temporal changes in tumor, in particular the dynamic changes tumors undergo in response to a therapeutic agent. While the pre-existing genetic heterogeneity has been shown to be important through multiple studies in the context of ICB, recent work has shown the importance of immune and other host-immune factors that are essential in response. Surveying the immune compartment, studies have shown TCR clonality to be predictive of response in melanoma, however the same findings did not extend to the cohort of sarcomas that we studied(Tumeh et al., 2014, Roh et al., 2017). The immune heterogeneity that can be sampled through repertoire analysis highlights the diversity that we observe in immune repertoires and a better understanding of tumor-type specific immune heterogeneity is essential. A deeper understanding of the mechanistic links between biomarkers may provide new insights into cancer immunology and ultimately therapeutic intervention. Immune microenvironments and their relationship in space and time are critical to further our understanding of disease progression. While we observed significantly higher levels of lymphocytic infiltrate in matched cases with EAC and BE, this infiltrate was primarily composed of regulatory T-cells, with a sparse effector immune infiltrate. Understanding how tumors employ immune exclusion is essential for multiple tumor types. Through molecular characterization of immune cell sub-types paired with spatial information pertaining to cellular neighborhoods, we may better understand the cross talk of cancer and immune cells in the tumor microenvironment and ultimately leverage this for therapeutic benefit.

References

2020. *SOFT TISSUE AND BONE TUMOURS,* GENEVA, WORLD HEALTH ORGANIZATION.

AKBANI, R., AKDEMIR, K. C., AKSOY, B. A., ALBERT, M., ALLY, A., AMIN, S. B., ARACHCHI, H., ARORA, A., AUMAN, J. T., AYALA, B., BABOUD, J., BALASUNDARAM, M., BALU, S., BARNABAS, N., BARTLETT, J., BARTLETT, P., BASTIAN, B. C., BAYLIN, S. B., BEHERA, M., BELYAEV, D., BENZ, C., BERNARD, B., BEROUKHIM, R., BIR, N., BLACK, A. D., BODENHEIMER, T., BOICE, L., BOLAND, G. M., BONO, R., BOOTWALLA, M. S., BOSENBERG, M., BOWEN, J., BOWLBY, R., BRISTOW, C. A., BROCKWAY-LUNARDI, L., BROOKS, D., BRZEZINSKI, J., BSHARA, W., BUDA, E., BURNS, W. R., BUTTERFIELD, Y. S. N., BUTTON, M., CALDERONE, T., CAPPELLINI, G. A., CARTER, C., CARTER, S. L., CHERNEY, L., CHERNIACK, A. D., CHEVALIER, A., CHIN, L., CHO, J., CHO, R. J., CHOI, Y. L., CHU, A., CHUDAMANI, S., CIBULSKIS, K., CIRIELLO, G., CLARKE, A., COONS, S., COPE, L., CRAIN, D., CURLEY, E., DANILOVA, L., D'ATRI, S., DAVIDSEN, T., DAVIES, M. A., DELMAN, K. A., DEMCHOK, J. A., DENG, Q. A., DERIBE, Y. L., DHALLA, N., DHIR, R., DICARA, D., DINIKIN, M., DUBINA, M., EBROM, J. S., EGEA, S., ELEY, G., ENGEL, J., ESCHBACHER, J. M., FEDOSENKO, K. V., FELAU, I., FENNELL, T., FERGUSON, M. L., FISHER, S., FLAHERTY, K. T., FRAZER, S., FRICK, J., FULIDOU, V., GABRIEL, S. B., GAO, J., GARDNER, J., GARRAWAY, L. A., GASTIER-FOSTER, J. M., GAUDIOSO, C., GEHLENBORG, N., GENOVESE, G., GERKEN, M., GERSHENWALD, J. E., GETZ, G., et al. 2015. Genomic Classification of Cutaneous Melanoma. *Cell*.

ALMAND, B., CLARK, J. I., NIKITINA, E., VAN BEYNEN, J., ENGLISH, N. R., KNIGHT, S. C., CARBONE, D. P. & GABRILOVICH, D. I. 2001. Increased production of immature myeloid cells in cancer patients: a mechanism of immunosuppression in cancer. *J Immunol,* 166**,** 678-89.

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*.

ARAN, D., HU, Z. & BUTTE, A. J. 2017. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology*.

AXELSON, H., FREDLUND, E., OVENBERGER, M., LANDBERG, G. & PAHLMAN, S. 2005. Hypoxia-induced dedifferentiation of tumor cells--a mechanism behind heterogeneity and aggressiveness of solid tumors. *Semin Cell Dev Biol,* 16**,** 554-63.

BALKWILL, F., ALBERTO MANTOVANI 2001. "Inflammation and cancer: back to Virchow?." *Lancet,* 357**,** 539-545.

BARBIE, D. A., TAMAYO, P., BOEHM, J. S., KIM, S. Y., MOODY, S. E., DUNN, I. F., SCHINZEL, A. C., SANDY, P., MEYLAN, E., SCHOLL, C., FRÖHLING, S., CHAN, E. M., SOS, M. L., MICHEL, K., MERMEL, C., SILVER, S. J., WEIR, B. A., REILING, J. H., SHENG, Q., GUPTA, P. B., WADLOW, R. C., LE, H., HOERSCH, S., WITTNER, B. S., RAMASWAMY, S., LIVINGSTON, D. M., SABATINI, D. M., MEYERSON, M., THOMAS, R. K., LANDER, E. S., MESIROV, J. P., ROOT, D. E., GILLILAND, D. G., JACKS, T. & HAHN, W. C. 2009. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*.

BARRET SCHLOERKE, D. C., JOSEPH LARMARANGE, FRANCOIS BRIATTE, MORITZ MARBACH, EDWIN THOEN, AMOS ELBERG AND JASON CROWLEY 2021. GGally: Extension to 'ggplot2'. *R package version 2.1.1*.

BAYSAL, B. E., SHARMA, S., HASHEMIKHABIR, S. & JANGA, S. C. 2017. RNA Editing in Pathogenesis of Cancer. *Cancer Res,* 77**,** 3733-3739.

BEAUGERIE, L., SVRCEK, M., SEKSIK, P., BOUVIER, A. M., SIMON, T., ALLEZ, M., BRIXI, H., GORNET, J. M., ALTWEGG, R., BEAU, P., DUCLOS, B., BOURREILLE, A., FAIVRE, J., PEYRIN-BIROULET, L., FLEJOU, J. F., CARRAT, F. & GROUP, C. S. 2013. Risk of colorectal high-grade dysplasia and cancer in a prospective observational cohort of patients with inflammatory bowel disease. *Gastroenterology,* 145**,** 166-175 e8.

BECHT, E., GIRALDO, N. A., LACROIX, L., BUTTARD, B. N. D., ELAROUCI, N., PETITPREZ, F., SELVES, J., LAURENT-PUIG, P., SAUTÏ¿½S-FRIDMAN, C., FRIDMAN, W. H. & DE REYNIÏ¿½S, A. L. 2016. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology*.

BERGLUND, E., MAASKOLA, J., SCHULTZ, N., FRIEDRICH, S., MARKLUND, M., BERGENSTRAHLE, J., TARISH, F., TANOGLIDI, A., VICKOVIC, S., LARSSON, L., SALMEN, F., OGRIS, C., WALLENBORG, K., LAGERGREN, J., STAHL, P., SONNHAMMER, E., HELLEDAY, T. & LUNDEBERG, J. 2018. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nat Commun,* 9**,** 2419.

BIRKELAND, E., ZHANG, S., PODUVAL, D., GEISLER, J., NAKKEN, S., VODAK, D., MEZA-ZEPEDA, L. A., HOVIG, E., MYKLEBOST, O., KNAPPSKOG, S. & LONNING, P. E. 2018. Patterns of genomic evolution in advanced melanoma. *Nat Commun,* 9**,** 2665.

BISWAS, D., BIRKBAK, N. J., ROSENTHAL, R., HILEY, C. T., LIM, E. L., PAPP, K., BOEING, S., KRZYSTANEK, M., DJUREINOVIC, D., LA FLEUR, L., GRECO, M., DOME, B., FILLINGER, J., BRUNNSTROM, H., WU, Y., MOORE, D. A., SKRZYPSKI, M., ABBOSH, C., LITCHFIELD, K., AL BAKIR, M., WATKINS, T. B. K., VEERIAH, S., WILSON, G. A., JAMAL-HANJANI, M., MOLDVAY, J., BOTLING, J., CHINNAIYAN, A. M., MICKE, P., HACKSHAW, A., BARTEK, J., CSABAI, I., SZALLASI, Z., HERRERO, J., MCGRANAHAN, N., SWANTON, C. & CONSORTIUM, T. R. 2019. A clonal expression biomarker associates with lung cancer mortality. *Nat Med,* 25**,** 1540-1548.

BOLOTIN, D. A., POSLAVSKY, S., MITROPHANOV, I., SHUGAY, M., MAMEDOV, I. Z., PUTINTSEVA, E. V. & CHUDAKOV, D. M. 2015. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods,* 12**,** 380-1.

BONI, A., COGDILL, A. P., DANG, P., UDAYAKUMAR, D., NJAUW, C.-N. J., SLOSS, C. M., FERRONE, C. R., FLAHERTY, K. T., LAWRENCE, D. P., FISHER, D. E., TSAO, H. & WARGO, J. A. 2010. Selective BRAF <sup>V600E</sup> Inhibition Enhances T-Cell Recognition of Melanoma without Affecting Lymphocyte Function. *Cancer Research,* 70**,** 5213-5219.

BROWN, S. D., WARREN, R. L., GIBB, E. A., MARTIN, S. D., SPINELLI, J. J., NELSON, B. H. & HOLT, R. A. 2014. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome Research*.

BRUNSON, J. 2020. ggalluvial: Layered Grammar for Alluvial Plots. *Journal of Open Source Software,* 5.

BUAS, M. F., HE, Q., JOHNSON, L. G., ONSTAD, L., LEVINE, D. M., THRIFT, A. P., GHARAHKHANI, P., PALLES, C., LAGERGREN, J., FITZGERALD, R. C., YE, W., CALDAS, C., BIRD, N. C., SHAHEEN, N. J., BERNSTEIN, L., GAMMON, M. D., WU, A. H., HARDIE, L. J., PHAROAH, P. D., LIU, G., IYER, P., CORLEY, D. A., RISCH, H. A., CHOW, W. H., PRENEN, H., CHEGWIDDEN, L., LOVE, S., ATTWOOD, S., MOAYYEDI, P., MACDONALD, D., HARRISON, R., WATSON, P., BARR, H., DECAESTECKER, J., TOMLINSON, I., JANKOWSKI, J., WHITEMAN, D. C., MACGREGOR, S., VAUGHAN, T. L. & MADELEINE, M. M. 2017. Germline variation in inflammation-related pathways and risk of Barrett's oesophagus and oesophageal adenocarcinoma. *Gut,* 66**,** 1739-1747.

CAMPOLI, M. R., CHANG, C.-C., KAGESHITA, T., WANG, X., MCCARTHY, J. B. & FERRONE, S. 2004. Human high molecular weight-melanoma-associated antigen (HMW-MAA): a melanoma cell surface chondroitin sulfate proteoglycan (MSCP) with biological and clinical significance. *Crit Rev Immunol*.

CANCER GENOME ATLAS RESEARCH NETWORK. ELECTRONIC ADDRESS, E. D. S. C. & CANCER GENOME ATLAS RESEARCH, N. 2017. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. *Cell,* 171**,** 950-965 e28.

CARLSON, C. S., EMERSON, R. O., SHERWOOD, A. M., DESMARAIS, C., CHUNG, M. W., PARSONS, J. M., STEEN, M. S., LAMADRID-HERRMANNSFELDT, M. A., WILLIAMSON, D. W., LIVINGSTON, R. J., WU, D., WOOD, B. L., RIEDER, M. J. & ROBINS, H. 2013. Using synthetic templates to design an unbiased multiplex PCR assay. *Nature Communications*.

CARTER, S. L., EKLUND, A. C., KOHANE, I. S., HARRIS, L. N. & SZALLASI, Z. 2006. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet,* 38**,** 1043-8.

CASASENT, A. K., SCHALCK, A., GAO, R., SEI, E., LONG, A., PANGBURN, W., CASASENT, T., MERIC-BERNSTAM, F., EDGERTON, M. E. & NAVIN, N. E. 2018. Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell,* 172**,** 205-217 e12.

CE, S. 1963. The mathematical theory of communication. *MD Comput.,* 1997**,** 306-317.

CHASSERIAU, J., RIVET, J., BILAN, F., CHOMEL, J. C., GUILHOT, F., BOURMEYSTER, N. & KITZIS, A. 2004. Characterization of the different BCR-ABL transcripts with a single multiplex RT-PCR. *Journal of Molecular Diagnostics*.

CHEN, B., KHODADOUST, M. S., LIU, C. L., NEWMAN, A. M. & ALIZADEH, A. A. 2018a. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol,* 1711**,** 243-259.

CHEN, D. S., IRVING, B. A. & HODI, F. S. 2012. Molecular pathways: next-generation immunotherapy--inhibiting programmed death-ligand 1 and programmed death-1. *Clin Cancer Res,* 18**,** 6580-7.

CHEN, D. S. & MELLMAN, I. 2013. *Oncology meets immunology: The cancer-immunity cycle*.

CHEN, H., LI, C., PENG, X., ZHOU, Z., WEINSTEIN, J. N., CANCER GENOME ATLAS RESEARCH, N. & LIANG, H. 2018b. A Pan-Cancer Analysis of Enhancer Expression in Nearly 9000 Patient Samples. *Cell,* 173**,** 386-399 e12.

CHEN, L., LI, Y., LIN, C. H., CHAN, T. H., CHOW, R. K., SONG, Y., LIU, M., YUAN, Y. F., FU, L., KONG, K. L., QI, L., LI, Y., ZHANG, N., TONG, A. H., KWONG, D. L., MAN, K., LO, C. M., LOK, S., TENEN, D. G. & GUAN, X. Y. 2013. Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. *Nat Med,* 19**,** 209-16.

CHEN, P. L., ROH, W., REUBEN, A., COOPER, Z. A., SPENCER, C. N., PRIETO, P. A., MILLER, J. P., BASSETT, R. L., GOPALAKRISHNAN, V., WANI, K., DE MACEDO, M. P., AUSTIN-BRENEMAN, J. L., JIANG, H., CHANG, Q., REDDY, S. M., CHEN, W. S., TETZLAFF, M. T., BROADDUS, R. J., DAVIES, M. A., GERSHENWALD, J. E., HAYDU, L., LAZAR, A. J., PATEL, S. P., HWU, P., HWU, W. J., DIAB, A., GLITZA, I. C., WOODMAN, S. E., VENCE, L. M., WISTUBA, I. I., AMARIA, R. N., KWONG, L. N., PRIETO, V., ERIC DAVIS, R., MA, W., OVERWIJK, W. W., SHARPE, A. H., HU, J., ANDREW FUTREAL, P., BLANDO, J., SHARMA, P., ALLISON, J. P., CHIN, L. & WARGO, J. A. 2016a. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discovery,* 6**,** 827-837.

CHEN, Z., ZHANG, C., PAN, Y., XU, R., XU, C., CHEN, Z., LU, Z. & KE, Y. 2016b. T cell receptor beta-chain repertoire analysis reveals intratumour heterogeneity of tumour-infiltrating lymphocytes in oesophageal squamous cell carcinoma. *J Pathol,* 239**,** 450-8.

CHIU, C. H., WANG, Y. T., WALTHER, B. A. & CHAO, A. 2014. An improved nonparametric lower bound of species richness via a modified good-turing frequency formula. *Biometrics,* 70**,** 671-82.

CIBULSKIS, K., LAWRENCE, M. S., CARTER, S. L., SIVACHENKO, A., JAFFE, D., SOUGNEZ, C., GABRIEL, S., MEYERSON, M., LANDER, E. S. & GETZ, G. 2013. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnology,* 31**,** 213-219.

CLOUGHESY, T. F., MOCHIZUKI, A. Y., ORPILLA, J. R., HUGO, W., LEE, A. H., DAVIDSON, T. B., WANG, A. C., ELLINGSON, B. M., RYTLEWSKI, J. A., SANDERS, C. M., KAWAGUCHI, E. S., DU, L., LI, G., YONG, W. H., GAFFEY, S. C., COHEN, A. L., MELLINGHOFF, I. K., LEE, E. Q., REARDON, D. A., O'BRIEN, B. J., BUTOWSKI, N. A., NGHIEMPHU, P. L., CLARKE, J. L., ARRILLAGA-ROMANY, I. C., COLMAN, H., KALEY, T. J., DE GROOT, J. F., LIAU, L. M., WEN, P. Y. & PRINS, R. M. 2019. Neoadjuvant anti-PD-1 immunotherapy promotes a survival benefit with intratumoral and systemic immune responses in recurrent glioblastoma. *Nature Medicine*.

COATES, P. J., RUNDLE, J. K., LORIMORE, S. A. & WRIGHT, E. G. 2008. Indirect macrophage responses to ionizing radiation: Implications for genotype-dependent bystander signaling. *Cancer Research*.

COIT, D. G., THOMPSON, J. A., ALGAZI, A., ANDTBACKA, R., BICHAKJIAN, C. K., CARSON, W. E., DANIELS, G. A., DIMAIO, D., ERNSTOFF, M., FIELDS, R. C., FLEMING, M. D., GONZALEZ, R., GUILD, V., HALPERN, A. C., HODI, F. S., JOSEPH, R. W., LANGE, J. R., MARTINI, M. C., MATERIN, M. A., OLSZANSKI, A. J., ROSS, M. I., SALAMA, A. K., SKITZKI, J., SOSMAN, J., SWETTER, S. M., TANABE, K. K., TORRES-ROCA, J. F., TRISAL, V., URIST, M. M., MCMILLIAN, N. & ENGH, A. 2016. Melanoma, version 2.2016 clinical practice guidelines in oncology. *JNCCN Journal of the National Comprehensive Cancer Network*.

COOK, M. B., COBURN, S. B., LAM, J. R., TAYLOR, P. R., SCHNEIDER, J. L. & CORLEY, D. A. 2018. Cancer incidence and mortality risks in a large US Barrett's oesophagus cohort. *Gut,* 67**,** 418-529.

CORREALE, P., ROTUNDO, M. S., BOTTA, C., DEL VECCHIO, M. T., GINANNESCHI, C., LICCHETTA, A., CONCA, R., APOLLINARI, S., DE LUCA, F., TASSONE, P. & TAGLIAFERRI, P. 2012. Tumor infiltration by T lymphocytes expressing chemokine receptor 7 (CCR7) is predictive of favorable outcome in patients with advanced colorectal carcinoma. *Clin Cancer Res,* 18**,** 850-7.

DALEY, T. & SMITH, A. D. 2013. Predicting the molecular complexity of sequencing libraries. *Nature Methods,* 10**,** 325-327.

DANG, H. X., WHITE, B. S., FOLTZ, S. M., MILLER, C. A., LUO, J., FIELDS, R. C. & MAHER, C. A. 2017. ClonEvol: Clonal ordering and visualization in cancer sequencing. *Annals of Oncology*.

DEMIRCIOGLU, D., CUKUROGLU, E., KINDERMANS, M., NANDI, T., CALABRESE, C., FONSECA, N. A., KAHLES, A., LEHMANN, K. V., STEGLE, O., BRAZMA, A., BROOKS, A. N., RATSCH, G., TAN, P. & GOKE, J. 2019. A Pan-cancer Transcriptome Analysis Reveals Pervasive Regulation through Alternative Promoters. *Cell,* 178**,** 1465-1477 e17.

DOUBROVINA, E., CARPENTER, T., PANKOV, D., SELVAKUMAR, A., HASAN, A. & O'REILLY, R. J. 2012. Mapping of novel peptides of WT-1 and presenting HLA alleles that induce epitope-specific HLA-restricted T cells with cytotoxic activity against WT-1+ leukemias. *Blood,* 120**,** 1633-1646.

DOUGLAS HANAHAN, R. A. W. 2000. The Hallmarks of Cancer. *Cell,* 100**,** 57-70.

DR LEACH, M. K., JP ALLISON 1996. Enhancement of antitumor immunity by CTLA-4 blockade. *Science,* 271 1734-1736.

EFRON, B. & THISTED, R. 1976. Estimating the number of unseen species: How many words did Shakespeare know? *Biometrika,* 63**,** 435-447.

EMERSON, R. O., SHERWOOD, A. M., RIEDER, M. J., GUENTHOER, J., WILLIAMSON, D. W., CARLSON, C. S., DRESCHER, C. W., TEWARI, M., BIELAS, J. H. & ROBINS, H. S. 2013. High-throughput sequencing of T-cell receptors reveals a homogeneous repertoire of tumour-infiltrating lymphocytes in ovarian cancer. *J Pathol,* 231**,** 433-440.

ENGERING, A., GEIJTENBEEK, T. B., VAN VLIET, S. J., WIJERS, M., VAN LIEMPT, E., DEMAUREX, N., LANZAVECCHIA, A., FRANSEN, J., FIGDOR, C. G., PIGUET, V. & VAN KOOYK, Y. 2002. The dendritic cell-specific adhesion receptor DC-SIGN internalizes antigen for presentation to T cells. *J Immunol,* 168**,** 2118-26.

FAN, J., LIANG, H., SHEN, T., WANG, S., JI, X., YEE, C., LU, F. & SHAO, Y. 2018. Early Env-specific CTLs effectively suppress viral replication in SHIV controller macaques. *Cellular Immunology,* 331**,** 30-37.

FAVERO, F., JOSHI, T., MARQUARD, A. M., BIRKBAK, N. J., KRZYSTANEK, M., LI, Q., SZALLASI, Z. & EKLUND, A. C. 2015. Sequenza: Allele-specific copy number and mutation profiles from tumor sequencing data. *Annals of Oncology*.

FIDLER, I. J. 1978. Tumor Heterogeneity and the Biology of Cancer Invasion and Metastasis. *Cancer Research*.

FINOTELLO, F., MAYER, C., PLATTNER, C., LASCHOBER, G., RIEDER, D., HACKL, H., KROGSDAM, A., LONCOVA, Z., POSCH, W., WILFLINGSEDER, D., SOPPER, S., IJSSELSTEIJN, M., BROUWER, T. P., JOHNSON, D., XU, Y., WANG, Y., SANDERS, M. E., ESTRADA, M. V., ERICSSON-GONZALEZ, P., CHAROENTONG, P., BALKO, J., DE MIRANDA, N. & TRAJANOSKI, Z. 2019. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med,* 11**,** 34.

FLAVAHAN, W. A., GASKELL, E. & BERNSTEIN, B. E. 2017. Epigenetic plasticity and the hallmarks of cancer. *Science,* 357.

FRANCISZKIEWICZ, K., BOISSONNAS, A., BOUTET, M., COMBADIERE, C. & MAMI-CHOUAIB, F. 2012. Role of chemokines and chemokine receptors in shaping the effector phase of the antitumor immune response. *Cancer Res,* 72**,** 6325-32.

FREDERICK, D. T., PIRIS, A., COGDILL, A. P., COOPER, Z. A., LEZCANO, C., FERRONE, C. R., MITRA, D., BONI, A., NEWTON, L. P., LIU, C., PENG, W., SULLIVAN, R. J., LAWRENCE, D. P., HODI, F. S., OVERWIJK, W. W., LIZEE, G., MURPHY, G. F., HWU, P., FLAHERTY, K. T., FISHER, D. E. & WARGO, J. A. 2013. BRAF Inhibition Is Associated with Enhanced Melanoma Antigen Expression and a More Favorable Tumor Microenvironment in Patients with Metastatic Melanoma. *Clinical Cancer Research,* 19**,** 1225-1231.

GABRILOVICH, D. I., OSTRAND-ROSENBERG, S. & BRONTE, V. 2012. *Coordinated regulation of myeloid cells by tumours*.

GABRILOVICH, D. I., VELDERS, M. P., SOTOMAYOR, E. M. & KAST, W. M. 2001. Mechanism of immune dysfunction in cancer mediated by immature Gr-1+ myeloid cells. *J Immunol,* 166**,** 5398-406.

GALILI, T. 2015. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics,* 31**,** 3718-3720.

GARRAWAY, L. A., WIDLUND, H. R., RUBIN, M. A., GETZ, G., BERGER, A. J., RAMASWAMY, S., BEROUKHIM, R., MILNER, D. A., GRANTER, S. R., DU, J., LEE, C., WAGNER, S. N., LI, C., GOLUB, T. R., RIMM, D. L., MEYERSON, M. L., FISHER, D. E. & SELLERS, W. R. 2005. Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature,* 436**,** 117-122.

GERLINGER, M., HORSWELL, S., LARKIN, J., ROWAN, A. J., SALM, M. P., VARELA, I., FISHER, R., MCGRANAHAN, N., MATTHEWS, N., SANTOS, C. R., MARTINEZ, P., PHILLIMORE, B., BEGUM, S., RABINOWITZ, A., SPENCER-DENE, B., GULATI, S., BATES, P. A., STAMP, G., PICKERING, L., GORE, M., NICOL, D. L., HAZELL, S., FUTREAL, P. A., STEWART, A. & SWANTON, C. 2014. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature Genetics,* 46**,** 225-233.

GERLINGER, M., QUEZADA, S. A., PEGGS, K. S., FURNESS, A. J. S., FISHER, R., MARAFIOTI, T., SHENDE, V. H., MCGRANAHAN, N., ROWAN, A. J., HAZELL, S., HAMM, D., ROBINS, H. S., PICKERING, L., GORE, M., NICOL, D. L., LARKIN, J. & SWANTON, C. 2013. Ultra-deep T cell receptor sequencing reveals the complexity and intratumour heterogeneity of T cell clones in renal cell carcinomas: Ultra-deep sequencing of T cell repertoires in renal cancer. *The Journal of Pathology,* 231**,** 424-432.

GERLINGER, M., ROWAN, A. J., HORSWELL, S., LARKIN, J., ENDESFELDER, D., GRONROOS, E., MARTINEZ, P., MATTHEWS, N., STEWART, A., TARPEY, P., VARELA, I., PHILLIMORE, B., BEGUM, S., MCDONALD, N. Q., BUTLER, A., JONES, D., RAINE, K., LATIMER, C., SANTOS, C. R., NOHADANI, M., EKLUND, A. C., SPENCER-DENE, B., CLARK, G., PICKERING, L., STAMP, G., GORE, M., SZALLASI, Z., DOWNWARD, J., FUTREAL, P. A. & SWANTON, C. 2012a. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *New England Journal of Medicine,* 366**,** 883-892.

GERLINGER, M., ROWAN, A. J., HORSWELL, S., LARKIN, J., ENDESFELDER, D., GRONROOS, E., MARTINEZ, P., MATTHEWS, N., STEWART, A., TARPEY, P., VARELA, I., PHILLIMORE, B., BEGUM, S., MCDONALD, N. Q., BUTLER, A., JONES, D., RAINE, K., LATIMER, C., SANTOS, C. R., NOHADANI, M., EKLUND, A. C., SPENCER-DENE, B., CLARK, G., PICKERING, L., STAMP, G., GORE, M., SZALLASI, Z., DOWNWARD, J., FUTREAL, P. A. & SWANTON, C. 2012b. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *New England Journal of Medicine*.

GORDON, M. S., KANEGAI, C. M., DOERR, J. R. & WALL, R. 2003. Somatic hypermutation of the B cell receptor genes B29 (Igbeta, CD79b) and mb1 (Igalpha, CD79a). *Proc Natl Acad Sci U S A,* 100**,** 4126-31.

GREAVES, M. & MALEY, C. C. 2012. Clonal evolution in cancer. *Nature,* 481**,** 306-13.

GREAVES, P. & GRIBBEN, J. G. 2013. The role of B7 family molecules in hematologic malignancy. *Blood,* 121**,** 734-44.

GROISBERG, R., ROSZIK, J., CONLEY, A. P., LAZAR, A. J., PORTAL, D. E., HONG, D. S., NAING, A., HERZOG, C. E., SOMAIAH, N., ZARZOUR, M. A., PATEL, S., BROWN, R. E. & SUBBIAH, V. 2020. Genomics, Morphoproteomics, and Treatment Patterns of Patients with Alveolar Soft Part Sarcoma and Response to Multiple Experimental Therapies. *Mol Cancer Ther,* 19**,** 1165-1172.

GROUP, P. T. C., CALABRESE, C., DAVIDSON, N. R., DEMIRCIOGLU, D., FONSECA, N. A., HE, Y., KAHLES, A., LEHMANN, K. V., LIU, F., SHIRAISHI, Y., SOULETTE, C. M., URBAN, L., GREGER, L., LI, S., LIU, D., PERRY, M. D., XIANG, Q., ZHANG, F., ZHANG, J., BAILEY, P., ERKEK, S., HOADLEY, K. A., HOU, Y., HUSKA, M. R., KILPINEN, H., KORBEL, J. O., MARIN, M. G., MARKOWSKI, J., NANDI, T., PAN-HAMMARSTROM, Q., PEDAMALLU, C. S., SIEBERT, R., STARK, S. G., SU, H., TAN, P., WASZAK, S. M., YUNG, C., ZHU, S., AWADALLA, P., CREIGHTON, C. J., MEYERSON, M., OUELLETTE, B. F. F., WU, K., YANG, H., GROUP, P. T. W., BRAZMA, A., BROOKS, A. N., GOKE, J., RATSCH, G., SCHWARZ, R. F., STEGLE, O., ZHANG, Z. & CONSORTIUM, P. 2020. Genomic basis for RNA alterations in cancer. *Nature,* 578**,** 129-136.

GU, Z., EILS, R. & SCHLESNER, M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*.

GU, Z., GU, L., EILS, R., SCHLESNER, M. & BRORS, B. 2014. Circlize implements and enhances circular visualization in R. *Bioinformatics*.

HÄNZELMANN, S., CASTELO, R. & GUINNEY, J. 2013. GSVA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*.

HELLMANN, M. D., NATHANSON, T., RIZVI, H., CREELAN, B. C., SANCHEZ-VEGA, F., AHUJA, A., NI, A., NOVIK, J. B., MANGARIN, L. M. B., ABU-AKEEL, M., LIU, C., SAUTER, J. L., REKHTMAN, N., CHANG, E., CALLAHAN, M. K., CHAFT, J. E., VOSS, M. H., TENET, M., LI, X. M., COVELLO, K., RENNINGER, A., VITAZKA, P., GEESE, W. J., BORGHAEI, H., RUDIN, C. M., ANTONIA, S. J., SWANTON, C., HAMMERBACHER, J., MERGHOUB, T., MCGRANAHAN, N., SNYDER, A. & WOLCHOK, J. D. 2018. Genomic Features of

Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. *Cancer Cell*.

HELLMANN, M. D., PAZ-ARES, L., BERNABE CARO, R., ZURAWSKI, B., KIM, S. W., CARCERENY COSTA, E., PARK, K., ALEXANDRU, A., LUPINACCI, L., DE LA MORA JIMENEZ, E., SAKAI, H., ALBERT, I., VERGNENEGRE, A., PETERS, S., SYRIGOS, K., BARLESI, F., RECK, M., BORGHAEI, H., BRAHMER, J. R., O'BYRNE, K. J., GEESE, W. J., BHAGAVATHEESWARAN, P., RABINDRAN, S. K., KASINATHAN, R. S., NATHAN, F. E. & RAMALINGAM, S. S. 2019. Nivolumab plus Ipilimumab in Advanced Non-Small-Cell Lung Cancer. *N Engl J Med,* 381**,** 2020-2031.

HEPPNER, G. H. 1984. Tumor Heterogeneity. *Perspectives in Cancer Research*.

HODI, F. S., O'DAY, S. J., MCDERMOTT, D. F., WEBER, R. W., SOSMAN, J. A., HAANEN, J. B., GONZALEZ, R., ROBERT, C., SCHADENDORF, D., HASSEL, J. C., AKERLEY, W., VAN DEN EERTWEGH, A. J. M., LUTZKY, J., LORIGAN, P., VAUBEL, J. M., LINETTE, G. P., HOGG, D., OTTENSMEIER, C. H., LEBBÉ, C., PESCHEL, C., QUIRT, I., CLARK, J. I., WOLCHOK, J. D., WEBER, J. S., TIAN, J., YELLIN, M. J., NICHOL, G. M., HOOS, A. & URBA, W. J. 2010. Improved survival with ipilimumab in patients with metastatic melanoma. *The New England journal of medicine*.

HU, X., ZHANG, J., WANG, J., FU, J., LI, T., ZHENG, X., WANG, B., GU, S., JIANG, P., FAN, J., YING, X., ZHANG, J., CARROLL, M. C., WUCHERPFENNIG, K. W., HACOHEN, N., ZHANG, F., ZHANG, P., LIU, J. S., LI, B. & LIU, X. S. 2019. Landscape of B cell immunity and related immune evasion in human cancers. *Nat Genet,* 51**,** 560-567.

HU, Z., LEET, D. E., ALLESOE, R. L., OLIVEIRA, G., LI, S., LUOMA, A. M., LIU, J., FORMAN, J., HUANG, T., IORGULESCU, J. B., HOLDEN, R., SARKIZOVA, S., GOHIL, S. H., REDD, R. A., SUN, J., ELAGINA, L., GIOBBIE-HURDER, A., ZHANG, W., PETER, L., CIANTRA, Z., RODIG, S., OLIVE, O., SHETTY, K., PYRDOL, J., UDUMAN, M., LEE, P. C., BACHIREDDY, P., BUCHBINDER, E. I., YOON, C. H., NEUBERG, D., PENTELUTE, B. L., HACOHEN, N., LIVAK, K. J., SHUKLA, S. A., OLSEN, L. R., BAROUCH, D. H., WUCHERPFENNIG, K. W., FRITSCH, E. F., KESKIN, D. B., WU, C. J. & OTT, P. A. 2021. Personal neoantigen vaccines induce persistent memory T cell responses and epitope spreading in patients with melanoma. *Nat Med,* 27**,** 515-525.

HUANG, D., SHERMAN, B. T., TAN, Q., COLLINS, J. R., ALVORD, W. G., ROAYAEI, J., STEPHENS, R., BASELER, M. W., LANE, H. C. & LEMPICKI, R. A. 2007. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology,* 8**,** R183.

HUGO, W., ZARETSKY, J. M., SUN, L., SONG, C., MORENO, B. H., HU-LIESKOVAN, S., BERENT-MAOZ, B., PANG, J., CHMIELOWSKI, B., CHERRY, G., SEJA, E., LOMELI, S., KONG, X., KELLEY, M. C., SOSMAN, J. A., JOHNSON, D. B., RIBAS, A. & LO, R. S. 2016. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell,* 165**,** 35-44.

INC, P. T. *Collaborative data science,* Montréal, QC, Plotly Technologies Inc.

ISHIBASHI, Y., TANAKA, S., TAJIMA, K., YOSHIDA, T. & KUWANO, H. 2006. Expression of Foxp3 in non-small cell lung cancer patients is significantly higher in tumor tissues than in normal tissues, especially in tumors smaller than 30 mm. *Oncol Rep,* 15**,** 1315-9.

JAFFE, A. E., MURAKAMI, P., LEE, H., LEEK, J. T., FALLIN, M. D., FEINBERG, A. P. & IRIZARRY, R. A. 2012. Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International Journal of Epidemiology,* 41**,** 200-209.

JAMAL-HANJANI, M., WILSON, G. A., MCGRANAHAN, N., BIRKBAK, N. J., WATKINS, T. B. K., VEERIAH, S., SHAFI, S., JOHNSON, D. H., MITTER, R., ROSENTHAL, R., SALM, M., HORSWELL, S., ESCUDERO, M., MATTHEWS, N., ROWAN, A., CHAMBERS, T., MOORE, D. A., TURAJLIC, S., XU, H., LEE, S.-M., FORSTER, M. D., AHMAD, T., HILEY, C. T., ABBOSH, C. & FALZON, M. 2017a. Tracking the Evolution of Non–Small-Cell Lung Cancer. *New England Journal of Medicine*.

JAMAL-HANJANI, M., WILSON, G. A., MCGRANAHAN, N., BIRKBAK, N. J., WATKINS, T. B. K., VEERIAH, S., SHAFI, S., JOHNSON, D. H., MITTER, R., ROSENTHAL, R., SALM, M., HORSWELL, S., ESCUDERO, M., MATTHEWS, N., ROWAN, A., CHAMBERS, T., MOORE, D. A., TURAJLIC, S., XU, H., LEE, S.-M., FORSTER, M. D., AHMAD, T., HILEY, C. T., ABBOSH, C., FALZON, M., BORG, E., MARAFIOTI, T., LAWRENCE, D., HAYWARD, M., KOLVEKAR, S., PANAGIOTOPOULOS, N., JANES, S. M., THAKRAR, R., AHMED, A., BLACKHALL, F., SUMMERS, Y., SHAH, R., JOSEPH, L., QUINN, A. M., CROSBIE, P. A., NAIDU, B., MIDDLETON, G., LANGMAN, G., TROTTER, S., NICOLSON, M., REMMEN, H., KERR, K., CHETTY, M., GOMERSALL, L., FENNELL, D. A., NAKAS, A., RATHINAM, S., ANAND, G., KHAN, S., RUSSELL, P., EZHIL, V., ISMAIL, B., IRVIN-SELLERS, M., PRAKASH, V., LESTER, J. F., KORNASZEWSKA, M., ATTANOOS, R., ADAMS, H., DAVIES, H., DENTRO, S., TANIERE, P., O'SULLIVAN, B., LOWE, H. L., HARTLEY, J. A., ILES, N., BELL, H., NGAI, Y., SHAW, J. A., HERRERO, J., SZALLASI, Z., SCHWARZ, R. F., STEWART, A., QUEZADA, S. A., LE QUESNE, J., VAN LOO, P., DIVE, C., HACKSHAW, A. & SWANTON, C. 2017b. Tracking the Evolution of Non–Small-Cell Lung Cancer. *New England Journal of Medicine,* 376**,** 2109-2121.

JANEWAY CA JR, T. P., WALPORT M, ET AL 2001. Immunobiology: The Immune System in Health and Disease. *New York: Garland Science,* 5.

JIANHUA, Z. 2018. *CNTools: Convert segment data into a region by sample matrix to allow for other high level computational analyses*.

JIMÉNEZ-SÁNCHEZ, A., MEMON, D., POURPE, S., VEERARAGHAVAN, H., LI, Y., VARGAS, H. A., GILL, M. B., PARK, K. J., ZIVANOVIC, O., KONNER, J., RICCA, J., ZAMARIN, D., WALTHER, T., AGHAJANIAN, C., WOLCHOK, J. D., SALA, E., MERGHOUB, T., SNYDER, A. & MILLER, M. L. 2017. Heterogeneous Tumor-Immune Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient. *Cell,* 170**,** 927-938.e20.

JOHNSON, W. E., LI, C. & RABINOVIC, A. 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics,* 8**,** 118-127.

JOSHI, K., DE MASSY, M. R., ISMAIL, M., READING, J. L., UDDIN, I., WOOLSTON, A., HATIPOGLU, E., OAKES, T., ROSENTHAL, R., PEACOCK, T., RONEL, T., NOURSADEGHI, M., TURATI, V., FURNESS, A. J. S., GEORGIOU, A., WONG, Y. N. S., BEN AISSA, A., SUNDERLAND, M. W., JAMAL-HANJANI, M., VEERIAH, S., BIRKBAK, N. J., WILSON, G. A., HILEY, C. T., GHORANI, E., GUERRA-ASSUNCAO, J. A., HERRERO, J., ENVER, T., HADRUP, S. R., HACKSHAW, A., PEGGS, K. S., MCGRANAHAN, N., SWANTON, C., CONSORTIUM, T. R., QUEZADA, S. A. & CHAIN, B. 2019. Spatial heterogeneity of the T cell receptor repertoire reflects the mutational landscape in lung cancer. *Nat Med,* 25**,** 1549-1559.

JOYCE, J. A. & POLLARD, J. W. 2009. *Microenvironmental regulation of metastasis*.

KABBARAH, O., NOGUEIRA, C., FENG, B., NAZARIAN, R. M., BOSENBERG, M., WU, M., SCOTT, K. L., KWONG, L. N., XIAO, Y., CORDON-CARDO, C., GRANTER, S. R., RAMASWAMY, S., GOLUB, T., DUNCAN, L. M., WAGNER, S. N., BRENNAN, C. & CHIN, L. 2010. Integrative genome comparison of primary and metastatic melanomas. *PLoS ONE*.

KAHLES, A., LEHMANN, K. V., TOUSSAINT, N. C., HUSER, M., STARK, S. G., SACHSENBERG, T., STEGLE, O., KOHLBACHER, O., SANDER, C., CANCER GENOME ATLAS RESEARCH, N. & RATSCH, G. 2018. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell,* 34**,** 211-224 e6.

KAHRAMAN, D. S., DINIZ, G., SAYHAN, S., SAYAR, C., AYAZ, D., GOKCU, M. & KARADENIZ, T. 2018. The prognostic significance of pdl1 and foxp3 expressions in tumor cells and the tumor microenvironment of ovarian epithelial tumors. *Int J Clin Exp Pathol,* 11**,** 3884-3890.

KALLURI, R. & ZEISBERG, M. 2006. Fibroblasts in cancer. *Nat Rev Cancer,* 6**,** 392-401.

KANDOTH, C., MCLELLAN, M. D., VANDIN, F., YE, K., NIU, B., LU, C., XIE, M., ZHANG, Q., MCMICHAEL, J. F., WYCZALKOWSKI, M. A., LEISERSON, M. D. M., MILLER, C. A., WELCH, J. S., WALTER, M. J., WENDL, M. C., LEY, T. J., WILSON, R. K., RAPHAEL, B. J. & DING, L. 2013. Mutational landscape and significance across 12 major cancer types. *Nature,* 502**,** 333-339.

KAROSIENE, E., RASMUSSEN, M., BLICHER, T., LUND, O., BUUS, S. & NIELSEN, M. 2013. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics,* 65**,** 711-724.

KAVANAGH, M. E., CONROY, M. J., CLARKE, N. E., GILMARTIN, N. T., O'SULLIVAN, K. E., FEIGHERY, R., MACCARTHY, F., O'TOOLE, D., RAVI, N., REYNOLDS, J. V., O'SULLIVAN, J. & LYSAGHT, J. 2016. Impact of the inflammatory microenvironment on T-cell phenotype in the progression from reflux oesophagitis to Barrett oesophagus and oesophageal adenocarcinoma. *Cancer Lett,* 370**,** 117-24.

KAVANAGH, M. E., O'SULLIVAN, K. E., O'HANLON, C., O'SULLIVAN, J. N., LYSAGHT, J. & REYNOLDS, J. V. 2014. The esophagitis to adenocarcinoma sequence; the role of inflammation. *Cancer Lett,* 345**,** 182-9.

KEUNG, E. Z., BURGESS, M., SALAZAR, R., PARRA, E. R., RODRIGUES-CANALES, J., BOLEJACK, V., VAN TINE, B. A., SCHUETZE, S. M., ATTIA, S., RIEDEL, R. F., HU, J., OKUNO, S. H., PRIEBAT, D. A., MOVVA, S., DAVIS, L. E., REED, D. R., REUBEN, A., ROLAND, C. L., REINKE, D., LAZAR, A. J., WANG, W. L., WARGO, J. A. & TAWBI, H. A. 2020. Correlative Analyses of the SARC028 Trial Reveal an Association Between Sarcoma-Associated Immune Infiltrate and Response to Pembrolizumab. *Clin Cancer Res,* 26**,** 1258-1266.

KHALILI, J. S., LIU, S., RODRIGUEZ-CRUZ, T. G., WHITTINGTON, M., WARDELL, S., LIU, C., ZHANG, M., COOPER, Z. A., FREDERICK, D. T., LI, Y., ZHANG, M., JOSEPH, R. W., BERNATCHEZ, C., EKMEKCIOGLU, S., GRIMM, E., RADVANYI, L. G., DAVIS, R. E., DAVIES, M. A., WARGO, J. A., HWU, P. & LIZEE, G. 2012. Oncogenic BRAF(V600E) Promotes Stromal Cell-Mediated Immunosuppression Via Induction of Interleukin-1 in Melanoma. *Clinical Cancer Research,* 18**,** 5329-5340.

KIM, D., PERTEA, G., TRAPNELL, C., PIMENTEL, H., KELLEY, R. & SALZBERG, S. L. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology,* 14**,** R36.

KONIECZKOWSKI, D. J., JOHANNESSEN, C. M., ABUDAYYEH, O., KIM, J. W., COOPER, Z. A., PIRIS, A., FREDERICK, D. T., BARZILY-ROKNI, M., STRAUSSMAN, R., HAQ, R., FISHER, D. E., MESIROV, J. P., HAHN, W. C., FLAHERTY, K. T., WARGO, J. A., TAMAYO, P. & GARRAWAY, L. A. 2014. A Melanoma Cell State Distinction Influences Sensitivity to MAPK Pathway Inhibitors. *Cancer Discovery,* 4**,** 816-827.

KUIJJER, M. L., HSIEH, P. H., QUACKENBUSH, J. & GLASS, K. 2019. lionessR: single sample network inference in R. *BMC Cancer,* 19**,** 1003.

LAFAVE, L. M., KARTHA, V. K., MA, S., MELI, K., DEL PRIORE, I., LAREAU, C., NARANJO, S., WESTCOTT, P. M. K., DUARTE, F. M., SANKAR, V., CHIANG, Z., BRACK, A., LAW, T., HAUCK, H., OKIMOTO, A., REGEV, A., BUENROSTRO, J. D. & JACKS, T. 2020. Epigenomic State Transitions Characterize Tumor Progression in Mouse Lung Adenocarcinoma. *Cancer Cell,* 38**,** 212-228 e13.

LAI, L. A., PAULSON, T. G., LI, X., SANCHEZ, C. A., MALEY, C., ODZE, R. D., REID, B. J. & RABINOVITCH, P. S. 2007. Increasing genomic instability during premalignant neoplastic progression revealed through high resolution array-CGH. *Genes Chromosomes Cancer,* 46**,** 532-42.

LARKIN, J., CHIARION-SILENI, V., GONZALEZ, R., GROB, J. J., RUTKOWSKI, P., LAO, C. D., COWEY, C. L., SCHADENDORF, D., WAGSTAFF, J., DUMMER, R., FERRUCCI, P. F., SMYLIE, M., HOGG, D., HILL, A., MARQUEZ-RODAS, I., HAANEN, J., GUIDOBONI, M., MAIO, M., SCHOFFSKI, P., CARLINO, M. S., LEBBE, C., MCARTHUR, G., ASCIERTO, P. A., DANIELS, G. A., LONG, G. V., BASTHOLT, L., RIZZO, J. I., BALOGH, A., MOSHYK, A., HODI, F. S. & WOLCHOK, J. D. 2019. Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *N Engl J Med,* 381**,** 1535-1546.

LAWRENCE, M. S., STOJANOV, P., MERMEL, C. H., ROBINSON, J. T., GARRAWAY, L. A., GOLUB, T. R., MEYERSON, M., GABRIEL, S. B., LANDER, E. S. & GETZ, G. 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature,* 505**,** 495-501.

LE, D. T., URAM, J. N., WANG, H., BARTLETT, B. R., KEMBERLING, H., EYRING, A. D., SKORA, A. D., LUBER, B. S., AZAD, N. S., LAHERU, D., BIEDRZYCKI, B., DONEHOWER, R. C., ZAHEER, A., FISHER, G. A., CROCENZI, T. S., LEE, J. J., DUFFY, S. M., GOLDBERG, R. M., DE LA CHAPELLE, A., KOSHIJI, M., BHAIJEE, F., HUEBNER, T., HRUBAN, R. H., WOOD, L. D., CUKA, N., PARDOLL, D. M., PAPADOPOULOS, N., KINZLER, K. W., ZHOU, S., CORNISH, T. C., TAUBE, J. M., ANDERS, R. A., ESHLEMAN, J. R., VOGELSTEIN, B. & DIAZ, L. A. 2015. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *New England Journal of Medicine*.

LI, B., LI, T., PIGNON, J. C., WANG, B., WANG, J., SHUKLA, S. A., DOU, R., CHEN, Q., HODI, F. S., CHOUEIRI, T. K., WU, C., HACOHEN, N., SIGNORETTI, S., LIU, J. S. & LIU, X. S. 2016a. Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat Genet,* 48**,** 725-32.

LI, B., LI, T., WANG, B., DOU, R., ZHANG, J., LIU, J. S. & LIU, X. S. 2017. Ultrasensitive detection of TCR hypervariable-region sequences in solid-tissue RNA-seq data. *Nat Genet,* 49**,** 482-483.

LI, B., SEVERSON, E., PIGNON, J.-C., ZHAO, H., LI, T., NOVAK, J., PENG, J., SHEN, H., ASTER JON, C., RODIG, S., SIGNORETTI, S., LIU JUN, S. & LIU, X. S. 2016b. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biology***,** 1-16.

LI, F., SUN, Y., HUANG, J., XU, W., LIU, J. & YUAN, Z. 2019. CD4/CD8 + T cells, DC subsets, Foxp3, and IDO expression are predictive indictors of gastric cancer prognosis. *Cancer Med,* 8**,** 7330-7344.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*.

LI, X., GALIPEAU, P. C., PAULSON, T. G., SANCHEZ, C. A., ARNAUDO, J., LIU, K., SATHER, C. L., KOSTADINOV, R. L., ODZE, R. D., KUHNER, M. K., MALEY, C. C., SELF, S. G., VAUGHAN, T. L., BLOUNT, P. L. & REID, B. J. 2014. Temporal and spatial evolution of somatic chromosomal alterations: a case-cohort study of Barrett's esophagus. *Cancer Prev Res (Phila),* 7**,** 114-27.

LI, Y., BLEAKLEY, M. & YEE, C. 2005. IL-21 Influences the Frequency, Phenotype, and Affinity of the Antigen-Specific CD8 T Cell Response. *The Journal of Immunology,* 175**,** 2261-2269.

LIN, S. M., DU, P., HUBER, W. & KIBBE, W. A. 2008. Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Research*.

LINCH, M., GOH, G., HILEY, C., SHANMUGABAVAN, Y., MCGRANAHAN, N., ROWAN, A., WONG, Y. N. S., KING, H., FURNESS, A., FREEMAN, A., LINARES, J., AKARCA, A., HERRERO, J., ROSENTHAL, R., HARDER, N., SCHMIDT, G., WILSON, G. A., BIRKBAK, N. J., MITTER, R., DENTRO, S., CATHCART, P., ARYA, M., JOHNSTON, E., SCOTT, R., HUNG, M., EMBERTON, M., ATTARD, G., SZALLASI, Z., PUNWANI, S., QUEZADA, S. A., MARAFIOTI, T., GERLINGER, M., AHMED, H. U. & SWANTON, C. 2017. Intratumoural evolutionary landscape of high-risk prostate cancer: the PROGENY study of genomic and immune parameters. *Annals of Oncology,* 28**,** 2472-2480.

LIPPITZ, B. E. 2013. Cytokine patterns in patients with cancer: a systematic review. *The Lancet Oncology,* 14**,** e218-e228.

LIU, C., YANG, X., DUFFY, B., MOHANAKUMAR, T., MITRA, R. D., ZODY, M. C. & PFEIFER, J. D. 2013. ATHLATES: Accurate typing of human leukocyte antigen through exome sequencing. *Nucleic Acids Research*.

LIU, C., YU, S., KAPPES, J., WANG, J., GRIZZLE, W. E., ZINN, K. R. & ZHANG, H. G. 2007. Expansion of spleen myeloid suppressor cells represses NK cell cytotoxicity in tumor-bearing host. *Blood,* 109**,** 4336-42.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*.

LU, Z., CHEN, H., LI, S., GONG, J., LI, J., ZOU, J., WU, L., YU, J., HAN, W., SUN, H., JIAO, X., ZHANG, X., PENG, Z., LU, M., WANG, Z., ZHANG, H. & SHEN, L. 2020. Tumor copy-number alterations predict response to immune-checkpoint-blockade in gastrointestinal cancer. *J Immunother Cancer,* 8.

LUNDEGAARD, C., LAMBERTH, K., HARNDAHL, M., BUUS, S., LUND, O. & NIELSEN, M. 2008. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic acids research,* 36.

MAHMOUD, S. M., PAISH, E. C., POWE, D. G., MACMILLAN, R. D., LEE, A. H., ELLIS, I. O. & GREEN, A. R. 2011. An evaluation of the clinical significance of FOXP3+ infiltrating cells in human breast cancer. *Breast Cancer Res Treat,* 127**,** 99-108.

MANDAL, R., SAMSTEIN, R. M., LEE, K.-W., HAVEL, J. J., WANG, H., KRISHNA, C., SABIO, E. Y., MAKAROV, V., KUO, F., BLECUA, P., RAMASWAMY, A. T., DURHAM, J. N., BARTLETT, B., MA, X., SRIVASTAVA, R., MIDDHA, S., ZEHIR, A., HECHTMAN, J. F., MORRIS, L. G. T., WEINHOLD, N., RIAZ, N., LE, D. T., DIAZ, L. A. & CHAN, T. A. 2019. Genetic diversity of tumors with mismatch repair deficiency influences anti–PD-1 immunotherapy response. *Science,* 364**,** 485-491.

MARKET, E. & PAPAVASILIOU, F. N. 2003. V(D)J recombination and the evolution of the adaptive immune system. *PLoS Biol,* 1**,** E16.

MCGRANAHAN, N., FURNESS, A. J. S., ROSENTHAL, R., RAMSKOV, S., LYNGAA, R., SAINI, S. K., JAMAL-HANJANI, M., WILSON, G. A., BIRKBAK, N. J., HILEY, C. T., WATKINS, T. B. K., SHAFI, S., MURUGAESU, N., MITTER, R., AKARCA, A. U., LINARES, J., MARAFIOTI, T., HENRY, J. Y., VAN ALLEN, E. M., MIAO, D., SCHILLING, B., SCHADENDORF, D., GARRAWAY, L. A., MAKAROV, V., RIZVI, N. A., SNYDER, A., HELLMANN, M. D., MERGHOUB, T., WOLCHOK, J. D., SHUKLA, S. A., WU, C. J., PEGGS, K. S., CHAN, T. A., HADRUP, S. R., QUEZADA, S. A. & SWANTON, C. 2016. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*.

MCGRANAHAN, N., ROSENTHAL, R., HILEY, C. T., ROWAN, A. J., WATKINS, T. B. K., WILSON, G. A., BIRKBAK, N. J., VEERIAH, S., VAN LOO, P., HERRERO, J., SWANTON, C. & CONSORTIUM, T. R. 2017. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell,* 171**,** 1259-1271 e11.

MELLMAN, I., COUKOS, G. & DRANOFF, G. 2011. Cancer immunotherapy comes of age. *Nature,* 480**,** 480-9.

MERMEL, C. H., SCHUMACHER, S. E., HILL, B., MEYERSON, M. L., BEROUKHIM, R. & GETZ, G. 2011. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biology,* 12**,** R41.

MINUSSI, D. C., NICHOLSON, M. D., YE, H., DAVIS, A., WANG, K., BAKER, T., TARABICHI, M., SEI, E., DU, H., RABBANI, M., PENG, C., HU, M., BAI, S., LIN, Y. W., SCHALCK, A., MULTANI, A., MA, J., MCDONALD, T. O., CASASENT, A., BARRERA, A., CHEN, H., LIM, B., ARUN, B., MERIC-BERNSTAM, F., VAN LOO, P., MICHOR, F. & NAVIN, N. E. 2021. Breast tumours maintain a reservoir of subclonal diversity during expansion. *Nature,* 592**,** 302-308.

MITRA, A., ANDREWS, M. C., ROH, W., DE MACEDO, M. P., HUDGENS, C. W., CARAPETO, F., SINGH, S., REUBEN, A., WANG, F., MAO, X., SONG, X., WANI, K., TIPPEN, S., NG, K. S., SCHALCK, A., SAKELLARIOU-THOMPSON, D. A., CHEN, E., REDDY, S. M., SPENCER, C. N., WIESNOSKI, D., LITTLE, L. D., GUMBS, C., COOPER, Z. A., BURTON, E. M., HWU, P., DAVIES, M. A., ZHANG, J., BERNATCHEZ, C., NAVIN, N., SHARMA, P., ALLISON, J. P., WARGO, J. A., YEE, C., TETZLAFF, M. T., HWU, W. J., LAZAR, A. J. & FUTREAL, P. A. 2020. Spatially resolved analyses link genomic and immune diversity and reveal unfavorable neutrophil activation in melanoma. *Nat Commun,* 11**,** 1839.

MONCADA, R., BARKLEY, D., WAGNER, F., CHIODIN, M., DEVLIN, J. C., BARON, M., HAJDU, C. H., SIMEONE, D. M. & YANAI, I. 2020. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol,* 38**,** 333-342.

MORRIS, L. G. T., RIAZ, N., DESRICHARD, A., ŞENBABAOĞLU, Y., HAKIMI, A. A., MAKAROV, V., REIS-FILHO, J. S. & CHAN, T. A. 2016. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget,* 7.

MORRIS, T. J., BUTCHER, L. M., FEBER, A., TESCHENDORFF, A. E., CHAKRAVARTHY, A. R., WOJDACZ, T. K. & BECK, S. 2014. ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics,* 30**,** 428-430.

MOTZER, R. J., TANNIR, N. M., MCDERMOTT, D. F., ARÉN FRONTERA, O., MELICHAR, B., CHOUEIRI, T. K., PLIMACK, E. R., BARTHÉLÉMY, P., PORTA, C., GEORGE, S., POWLES, T., DONSKOV, F., NEIMAN, V., KOLLMANNSBERGER, C. K., SALMAN, P., GURNEY, H., HAWKINS, R., RAVAUD, A., GRIMM, M.-O., BRACARDA, S., BARRIOS, C. H., TOMITA, Y., CASTELLANO, D., RINI, B. I., CHEN, A. C., MEKAN, S., MCHENRY, M. B., WIND-ROTOLO, M., DOAN, J., SHARMA, P., HAMMERS, H. J. & ESCUDIER, B. 2018. Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. *New England Journal of Medicine,* 378**,** 1277-1290.

N. A. HONG, N. H. K., S. N. HSIEH, D. CADO AND A. WINOTO 1999. In Vivo Overexpression of Dad1, the Defender Against Apoptotic Death-1, Enhances T Cell Proliferation But Does Not Protect Against Apoptosis. *J Immunol* 163.

NAOMI LEVY-STRUMPF, L. P. D., HANNA BERISSI, ADI KIMCHI 1997. DAP-5, a Novel Homolog of Eukaryotic Translation Initiation

Factor 4G Isolated as a Putative Modulator of Gamma

Interferon-Induced Programmed Cell Death. *MOLECULAR AND CELLULAR BIOLOGY***,** p 1615-1625.

NAZAROV, V. I., POGORELYY, M. V., KOMECH, E. A., ZVYAGIN, I. V., BOLOTIN, D. A., SHUGAY, M., CHUDAKOV, D. M., LEBEDEV, Y. B. & MAMEDOV, I. Z. 2015. tcR: An R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics*.

NEWMAN, A. M., LIU, C. L., GREEN, M. R., GENTLES, A. J., FENG, W., XU, Y., HOANG, C. D., DIEHN, M. & ALIZADEH, A. A. 2015. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*.

NEWMAN, A. M., STEEN, C. B., LIU, C. L., GENTLES, A. J., CHAUDHURI, A. A., SCHERER, F., KHODADOUST, M. S., ESFAHANI, M. S., LUCA, B. A., STEINER, D., DIEHN, M. & ALIZADEH, A. A. 2019. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol,* 37**,** 773-782.

NG, P. K. S., LI, J., JEONG, K. J., SHAO, S., CHEN, H., TSANG, Y. H., SENGUPTA, S., WANG, Z., BHAVANA, V. H., TRAN, R., SOEWITO, S., MINUSSI, D. C., MORENO, D., KONG, K., DOGRULUK, T., LU, H., GAO, J., TOKHEIM, C., ZHOU, D. C., JOHNSON, A. M., ZENG, J., IP, C. K. M., JU, Z., WESTER, M., YU, S., LI, Y., VELLANO, C. P., SCHULTZ, N., KARCHIN, R., DING, L., LU, Y., CHEUNG, L. W. T., CHEN, K., SHAW, K. R., MERIC-BERNSTAM, F., SCOTT, K. L., YI, S., SAHNI, N., LIANG, H. & MILLS, G. B. 2018. Systematic Functional Annotation of Somatic Mutations in Cancer. *Cancer Cell*.

NOBLE, F., MELLOWS, T., MCCORMICK MATTHEWS, L. H., BATEMAN, A. C., HARRIS, S., UNDERWOOD, T. J., BYRNE, J. P., BAILEY, I. S., SHARLAND, D. M., KELLY, J. J., PRIMROSE, J. N., SAHOTA, S. S., BATEMAN, A. R., THOMAS, G. J. & OTTENSMEIER, C.

H. 2016. Tumour infiltrating lymphocytes correlate with improved survival in patients with oesophageal adenocarcinoma. *Cancer Immunol Immunother,* 65**,** 651-62.

NOWELL, P. C. 1976. The clonal evolution of tumor cell populations. *Science***,** 23-28.

OCK, C. Y., HWANG, J. E., KEAM, B., KIM, S. B., SHIM, J. J., JANG, H. J., PARK, S., SOHN, B. H., CHA, M., AJANI, J. A., KOPETZ, S., LEE, K. W., KIM, T. M., HEO, D. S. & LEE, J. S. 2017. Genomic landscape associated with potential response to anti-CTLA-4 treatment in cancers. *Nature Communications*.

PARRA, E. R., FRANCISCO-CRUZ, A. & WISTUBA, II 2019. State-of-the-Art of Profiling Immune Contexture in the Era of Multiplexed Staining and Digital Analysis to Study Paraffin Tumor Tissues. *Cancers (Basel),* 11.

PARRA, E. R., URAOKA, N., JIANG, M., COOK, P., GIBBONS, D., FORGET, M. A., BERNATCHEZ, C., HAYMAKER, C., WISTUBA, II & RODRIGUEZ-CANALES, J. 2017. Validation of multiplex immunofluorescence panels using multispectral microscopy for immune-profiling of formalin-fixed and paraffin-embedded human tumor tissues. *Sci Rep,* 7**,** 13380.

PENG, W., CHEN, J. Q., LIU, C., MALU, S., CREASY, C., TETZLAFF, M. T., XU, C., MCKENZIE, J. A., ZHANG, C., LIANG, X., WILLIAMS, L. J., DENG, W., CHEN, G., MBOFUNG, R., LAZAR, A. J., TORRES-CABALA, C. A., COOPER, Z. A., CHEN, P. L., TIEU, T. N., SPRANGER, S., YU, X., BERNATCHEZ, C., FORGET, M. A., HAYMAKER, C., AMARIA, R., MCQUADE, J. L., GLITZA, I. C., CASCONE, T., LI, H. S., KWONG, L. N., HEFFERNAN, T. P., HU, J., BASSETT, R. L., BOSENBERG, M. W., WOODMAN, S. E., OVERWIJK, W. W., LIZEE, G., ROSZIK, J., GAJEWSKI, T. F., WARGO, J. A., GERSHENWALD, J. E., RADVANYI, L., DAVIES, M. A. & HWU, P. 2016. Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy. *Cancer Discovery,* 6**,** 202-216.

PETITPREZ, F., DE REYNIES, A., KEUNG, E. Z., CHEN, T. W., SUN, C. M., CALDERARO, J., JENG, Y. M., HSIAO, L. P., LACROIX, L., BOUGOUIN, A., MOREIRA, M., LACROIX, G., NATARIO, I., ADAM, J., LUCCHESI, C., LAIZET, Y. H., TOULMONDE, M., BURGESS, M. A., BOLEJACK, V., REINKE, D., WANI, K. M., WANG, W. L., LAZAR, A. J., ROLAND, C. L., WARGO, J. A., ITALIANO, A., SAUTES-FRIDMAN, C., TAWBI, H. A. & FRIDMAN, W. H. 2020. B cells are associated with survival and immunotherapy response in sarcoma. *Nature,* 577**,** 556-560.

PFEIFHOFER-OBERMAIR, C., TYMOSZUK, P., PETZER, V., WEISS, G. & NAIRZ, M. 2018. Iron in the Tumor Microenvironment—Connecting the Dots. *Frontiers in Oncology,* 8**,** 549.

POLLACK, S. M., JONES, R. L., FARRAR, E. A., LAI, I. P., LEE, S. M., CAO, J., PILLARISETTY, V. G., HOCH, B. L., GULLETT, A., BLEAKLEY, M., CONRAD, E. U., EARY, J. F., SHIBUYA, K. C., WARREN, E. H., CARSTENS, J. N., HEIMFELD, S., RIDDELL, S. R. & YEE, C. 2014. Tetramer guided, cell sorter assisted production of clinical grade autologous NY-ESO-1 specific CD8+ T cells. *Journal for ImmunoTherapy of Cancer,* 2**,** 36.

PYONTECK, S. M., AKKARI, L., SCHUHMACHER, A. J., BOWMAN, R. L., SEVENICH, L., QUAIL, D. F., OLSON, O. C., QUICK, M. L., HUSE, J. T., TEIJEIRO, V., SETTY, M., LESLIE, C. S., OEI, Y., PEDRAZA, A., ZHANG, J., BRENNAN, C. W., SUTTON, J. C., HOLLAND, E. C., DANIEL, D. & JOYCE, J. A. 2013. CSF-1R inhibition alters macrophage polarization and blocks glioma progression. *Nature Medicine*.

QIAN, B. Z. & POLLARD, J. W. 2010. Macrophage Diversity Enhances Tumor Progression and Metastasis. *Cell*.

QUANTE, M., BHAGAT, G., ABRAMS, J. A., MARACHE, F., GOOD, P., LEE, M. D., LEE, Y., FRIEDMAN, R., ASFAHA, S., DUBEYKOVSKAYA, Z., MAHMOOD, U., FIGUEIREDO, J. L., KITAJEWSKI, J., SHAWBER, C., LIGHTDALE, C. J., RUSTGI, A. K. & WANG, T. C. 2012. Bile acid and inflammation activate gastric cardia stem cells in a mouse model of Barrett-like metaplasia. *Cancer Cell,* 21**,** 36-51.

R CORE TEAM. 2019. *R: A language and environment for statistical computing. R Foundation for Statistical Computing* [Online]. Vienna, Austria. Available: https://www.R-project.org/ [Accessed].

RACLE, J., DE JONGE, K., BAUMGAERTNER, P., SPEISER, D. E. & GFELLER, D. 2017. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife,* 6.

REUBEN, A., SPENCER, C. N., PRIETO, P. A., GOPALAKRISHNAN, V., REDDY, S. M., MILLER, J. P., MAO, X., DE MACEDO, M. P., CHEN, J., SONG, X., JIANG, H., CHEN, P. L., BEIRD, H. C., GARBER, H. R., ROH, W., WANI, K., CHEN, E., HAYMAKER, C., FORGET, M. A., LITTLE, L. D., GUMBS, C., THORNTON, R. L., HUDGENS, C. W., CHEN, W. S., AUSTIN-BRENEMAN, J., SLOANE, R. S., NEZI, L., COGDILL, A. P., BERNATCHEZ, C., ROSZIK, J., HWU, P., WOODMAN, S. E., CHIN, L., TAWBI, H., DAVIES, M. A., GERSHENWALD, J. E., AMARIA, R. N., GLITZA, I. C., DIAB, A., PATEL, S. P., HU, J., LEE, J. E., GRIMM, E. A., TETZLAFF, M. T., LAZAR, A. J., WISTUBA, I. I., CLISE-DWYER, K., CARTER, B. W., ZHANG, J., FUTREAL, P. A., SHARMA, P., ALLISON, J. P., COOPER, Z. A. & WARGO, J. A. 2017. Genomic and immune heterogeneity are associated with differential responses to therapy in melanoma. *npj Genomic Medicine*.

RIAZ, N., HAVEL, J. J., MAKAROV, V., DESRICHARD, A., URBA, W. J., SIMS, J. S., HODI, F. S., MARTÍN-ALGARRA, S., MANDAL, R., SHARFMAN, W. H., BHATIA, S., HWU, W.-J., GAJEWSKI, T. F., SLINGLUFF, C. L., CHOWELL, D., KENDALL, S. M., CHANG, H., SHAH, R., KUO, F., MORRIS, L. G. T., SIDHOM, J.-W., SCHNECK, J. P., HORAK, C. E., WEINHOLD, N. & CHAN, T. A. 2017a. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell,* 0**,** 1-16.

RIAZ, N., HAVEL, J. J., MAKAROV, V., DESRICHARD, A., URBA, W. J., SIMS, J. S., HODI, F. S., MARTÍN-ALGARRA, S., MANDAL, R., SHARFMAN, W. H., BHATIA, S., HWU, W. J., GAJEWSKI, T. F., SLINGLUFF, C. L., CHOWELL, D., KENDALL, S. M., CHANG, H., SHAH, R., KUO, F., MORRIS, L. G. T., SIDHOM, J. W., SCHNECK, J. P., HORAK, C. E., WEINHOLD, N. & CHAN, T. A. 2017b. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell,* 171**,** 934-949.e15.

RIBAS, A. 2012. Tumor Immunotherapy Directed at PD-1. *New England Journal of Medicine,* 366**,** 2517-2519.

RIELLA, L. V., PATERSON, A. M., SHARPE, A. H. & CHANDRAKER, A. 2012. Role of the PD-1 pathway in the immune response. *Am J Transplant,* 12**,** 2575-87.

RITCHIE, M. E., PHIPSON, B., WU, D., HU, Y., LAW, C. W., SHI, W. & SMYTH, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research,* 43**,** e47-e47.

RIZVI, N. A., HELLMANN, M. D., SNYDER, A., KVISTBORG, P., MAKAROV, V., HAVEL, J. J., LEE, W., YUAN, J., WONG, P., HO, T. S., MILLER, M. L., REKHTMAN, N., MOREIRA, A. L., IBRAHIM, F., BRUGGEMAN, C., GASMI, B., ZAPPASODI, R., MAEDA, Y., SANDER, C., GARON, E. B. & MERGHOUB, T. 2016. Mutational landscape determines sensitivity to PD-1 blockade in non – small cell lung cancer. *Science,* 348**,** 124.

ROBINS, H. S., CAMPREGHER, P. V., SRIVASTAVA, S. K., WACHER, A., TURTLE, C. J., KAHSAI, O., RIDDELL, S. R., WARREN, E. H. & CARLSON, C. S. 2009a. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood*.

ROBINS, H. S., CAMPREGHER, P. V., SRIVASTAVA, S. K., WACHER, A., TURTLE, C. J., KAHSAI, O., RIDDELL, S. R., WARREN, E. H. & CARLSON, C. S. 2009b. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood,* 114**,** 4099-107.

ROH, W., CHEN, P.-L., REUBEN, A., SPENCER, C. N., PRIETO, P. A., MILLER, J. P., GOPALAKRISHNAN, V., WANG, F., COOPER, Z. A., REDDY, S. M., GUMBS, C., LITTLE, L., CHANG, Q., CHEN, W.-S., WANI, K., DE MACEDO, M. P., CHEN, E., AUSTIN-BRENEMAN, J. L., JIANG, H., ROSZIK, J., TETZLAFF, M. T., DAVIES, M. A., GERSHENWALD, J. E., TAWBI, H., LAZAR, A. J., HWU, P., HWU, W.-J., DIAB, A., GLITZA, I. C., PATEL, S. P., WOODMAN, S. E., AMARIA, R. N., PRIETO, V. G., HU, J., SHARMA, P., ALLISON, J. P., CHIN, L., ZHANG, J., WARGO, J. A. & FUTREAL, P. A. 2017. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Science translational medicine,* 9**,** 239-42.

ROONEY, M. S., SHUKLA, S. A., WU, C. J., GETZ, G. & HACOHEN, N. 2015. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell,* 160**,** 48-61.

ROPER, N., GAO, S., MAITY, T. K., BANDAY, A. R., ZHANG, X., VENUGOPALAN, A., CULTRARO, C. M., PATIDAR, R., SINDIRI, S., BROWN, A.-L., GONCEARENCO, A., PANCHENKO, A. R., BISWAS, R., THOMAS, A., RAJAN, A., CARTER, C. A., KLEINER, D. E., HEWITT, S. M., KHAN, J., PROKUNINA-OLSSON, L. & GUHA, U. 2019. APOBEC Mutagenesis and Copy-Number Alterations Are Drivers of Proteogenomic Tumor Evolution and Heterogeneity in Metastatic Thoracic Tumors. *Cell Reports,* 26**,** 2651-2666.e6.

ROSEMURGY, A., WILFONG, C., CRAIGG, D., CO, F., SUCANDY, I. & ROSS, S. 2019. The Evolving Landscape of Esophageal Cancer: A Four-Decade Analysis. *Am Surg,* 85**,** 944-948.

ROSENTHAL, R., CADIEUX, E. L., SALGADO, R., BAKIR, M. A., MOORE, D. A., HILEY, C. T., LUND, T., TANIC, M., READING, J. L., JOSHI, K., HENRY, J. Y., GHORANI, E., WILSON, G. A., BIRKBAK, N. J., JAMAL-HANJANI, M., VEERIAH, S., SZALLASI, Z., LOI, S., HELLMANN, M. D., FEBER, A., CHAIN, B., HERRERO, J., QUEZADA, S. A., DEMEULEMEESTER, J., VAN LOO, P., BECK, S., MCGRANAHAN, N., SWANTON, C. & CONSORTIUM, T. R. 2019. Neoantigen-directed immune escape in lung cancer evolution. *Nature,* 567**,** 479-485.

ROTH, A., KHATTRA, J., YAP, D., WAN, A., LAKS, E., BIELE, J., HA, G., APARICIO, S., BOUCHARD-CÔTÉ, A. & SHAH, S. P. 2014. PyClone: Statistical inference of clonal population structure in cancer. *Nature Methods*.

SAMSTEIN, R. M., LEE, C.-H., SHOUSHTARI, A. N., HELLMANN, M. D., SHEN, R., JANJIGIAN, Y. Y., BARRON, D. A., ZEHIR, A., JORDAN, E. J., OMURO, A., KALEY, T. J., KENDALL, S. M., MOTZER, R. J., HAKIMI, A. A., VOSS, M. H., RUSSO, P., ROSENBERG, J., IYER, G., BOCHNER, B. H., BAJORIN, D. F., AL-AHMADIE, H. A., CHAFT, J. E., RUDIN, C. M., RIELY, G. J., BAXI, S., HO, A. L., WONG, R. J., PFISTER, D. G., WOLCHOK, J. D., BARKER, C. A., GUTIN, P. H., BRENNAN, C. W., TABAR, V., MELLINGHOFF, I. K., DEANGELIS, L. M., ARIYAN, C. E., LEE, N., TAP, W. D., GOUNDER, M. M., D'ANGELO, S. P., SALTZ, L., STADLER, Z. K., SCHER, H. I., BASELGA, J., RAZAVI, P., KLEBANOFF, C. A., YAEGER, R., SEGAL, N. H., KU, G. Y., DEMATTEO, R. P., LADANYI, M., RIZVI, N. A., BERGER, M. F., RIAZ, N., SOLIT, D. B., CHAN, T. A. & MORRIS, L. G. T. 2019. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nature Genetics,* 51**,** 202-206.

SANGIOVANNI, A., DEL NINNO, E., FASANI, P., DE FAZIO, C., RONCHI, G., ROMEO, R., MORABITO, A., DE FRANCHIS, R. & COLOMBO, M. 2004. Increased survival of cirrhotic patients with a hepatocellular carcinoma detected during surveillance. *Gastroenterology,* 126**,** 1005-14.

SASAKI, A., TANAKA, F., MIMORI, K., INOUE, H., KAI, S., SHIBATA, K., OHTA, M., KITANO, S. & MORI, M. 2008. Prognostic value of tumor-infiltrating FOXP3+ regulatory T cells in patients with hepatocellular carcinoma. *Eur J Surg Oncol,* 34**,** 173-9.

SATHIRAPONGSASUTI, J. F., LEE, H., HORST, B. A. J., BRUNNER, G., COCHRAN, A. J., BINDER, S., QUACKENBUSH, J. & NELSON, S. F. 2011. Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV. *Bioinformatics (Oxford, England)*.

SATIJA, R., FARRELL, J. A., GENNERT, D., SCHIER, A. F. & REGEV, A. 2015. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*.

SCHATZ, D. G., MARJORIE A. OETTINGER, AND MARK S. SCHLISSEL 1992. V (D) J recombination: molecular biology and regulation. *Annual review of immunology* 10**,** 359-383.

SCHLIEP, K. P. 2011. phangorn: Phylogenetic analysis in R. *Bioinformatics*.

SCHNEIDER-HOHENDORF, T., MOHAN, H., BIEN, C. G., BREUER, J., BECKER, A., GORLICH, D., KUHLMANN, T., WIDMAN, G., HERICH, S., ELPERS, C., MELZER, N., DORNMAIR, K., KURLEMANN, G., WIENDL, H. & SCHWAB, N. 2016. CD8(+) T-cell pathogenicity in Rasmussen encephalitis elucidated by large-scale T-cell receptor sequencing. *Nat Commun,* 7**,** 11153.

SCHUMACHER, T. N. & SCHREIBER, R. D. 2015. Neoantigens in cancer immunotherapy. *Science,* 348**,** 69-74.

SHARMA, P., HU-LIESKOVAN, S., WARGO, J. A. & RIBAS, A. 2017. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell,* 168**,** 707-723.

SHERWOOD, A. M., EMERSON, R. O., SCHERER, D., HABERMANN, N., BUCK, K., STAFFA, J., DESMARAIS, C., HALAMA, N., JAEGER, D., SCHIRMACHER, P., HERPEL, E., KLOOR, M., ULRICH, A., SCHNEIDER, M., ULRICH, C. M. & ROBINS, H. 2013. Tumor-infiltrating lymphocytes in colorectal tumors display a diversity of T cell receptor sequences that differ from the T cells in adjacent mucosal tissue. *Cancer Immunol Immunother,* 62**,** 1453-61.

SHINTANI, Y., MAEDA, M., CHAIKA, N., JOHNSON, K. R. & WHEELOCK, M. J. 2008. Collagen I promotes epithelial-to-mesenchymal transition in lung cancer cells via transforming growth factor-beta signaling. *Am J Respir Cell Mol Biol,* 38**,** 95-104.

SHIRAISHI, Y., KATAOKA, K., CHIBA, K., OKADA, A., KOGURE, Y., TANAKA, H., OGAWA, S. & MIYANO, S. 2018. A comprehensive characterization of cis-acting splicing-associated variants in human cancer. *Genome Res,* 28**,** 1111-1125.

SHUKLA, S. A., ROONEY, M. S., RAJASAGI, M., TIAO, G., DIXON, P. M., LAWRENCE, M. S., STEVENS, J., LANE, W. J., DELLAGATTA, J. L., STEELMAN, S., SOUGNEZ, C., CIBULSKIS, K., KIEZUN, A., HACOHEN, N., BRUSIC, V., WU, C. J. & GETZ, G. 2015. Comprehensive analysis of cancer-associated somatic mutations in class i HLA genes. *Nature Biotechnology*.

SIMPSON, E. H. 1949. Measurement of Diversity. *Nature,* 163**,** 688-688.

SINHA, P., CLEMENTS, V. K. & OSTRAND-ROSENBERG, S. 2005. Reduction of myeloid-derived suppressor cells and induction of M1 macrophages facilitate the rejection of established metastatic disease. *J Immunol,* 174**,** 636-45.

SLOWIKOWSKI, K., SCHEP, A., HUGHES, S., LUKAUSKAS, S., IRISSON, J., KAMVAR, Z. N., RYAN, T., CHRISTOPHE, D., HIROAKI, Y. & GRAMME, P. 2018. *Automatically Position Non-Overlapping Text Labels with 'ggplot2'* [Online]. Available: https://cran.r-project.org/package=ggrepel [Accessed].

SMITH, K. J., O'BRIEN, S. M., GREEN, A. C., WEBB, P. M., WHITEMAN, D. C. & STUDY OF DIGESTIVE, H. 2009. Current and past smoking significantly increase risk for Barrett's esophagus. *Clin Gastroenterol Hepatol,* 7**,** 840-8.

SMYTH, G. K. 2004. Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology,* 3**,** 1-25.

SNYDER, A., MAKAROV, V., MERGHOUB, T., YUAN, J., ZARETSKY, J. M., DESRICHARD, A., WALSH, L. A., POSTOW, M. A., WONG, P., HO, T. S., HOLLMANN, T. J., BRUGGEMAN, C., KANNAN, K., LI, Y., ELIPENAHLI, C., LIU, C., HARBISON, C. T., WANG, L., RIBAS, A., WOLCHOK, J. D. & CHAN, T. A. 2014a. Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *The New England journal of medicine***,** 2189-2199.

SNYDER, A., MAKAROV, V., MERGHOUB, T., YUAN, J., ZARETSKY, J. M., DESRICHARD, A., WALSH, L. A., POSTOW, M. A., WONG, P., HO, T. S., HOLLMANN, T. J., BRUGGEMAN, C., KANNAN, K., LI, Y., ELIPENAHLI, C., LIU, C., HARBISON, C. T., WANG, L., RIBAS, A., WOLCHOK, J. D. & CHAN, T. A. 2014b. Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *New England Journal of Medicine,* 371**,** 2189-2199.

SOETART, K. 2013. plot3D: Plotting multi-dimensional data.

SOUZA, R. F., HUO, X., MITTAL, V., SCHULER, C. M., CARMACK, S. W., ZHANG, H. Y., ZHANG, X., YU, C., HORMI-CARVER, K., GENTA, R. M. & SPECHLER, S. J. 2009. Gastroesophageal reflux might cause esophagitis through a cytokine-mediated mechanism rather than caustic acid injury. *Gastroenterology,* 137**,** 1776-84.

SPRANGER, S., BAO, R. & GAJEWSKI, T. F. 2015. Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. *Nature*.

STEWART, T., TSAI, S.C., GRAYSON, H., HENDERSON, R. & OPELZ, G. 1995. Incidence of de-novo breast cancer in women chronically immunosuppressed after organ transplantation. *Lancet,* 346**,** 796–798.

STREET, W. 2020. Cancer Facts & Figures 2020.

STURM, G., FINOTELLO, F. & LIST, M. 2020. Immunedeconv: An R Package for Unified Access to Computational Methods for Estimating Immune Cell Fractions from Bulk RNA-Sequencing Data. *Methods Mol Biol,* 2120**,** 223-232.

TALHOUK, A., KOMMOSS, S., MACKENZIE, R., CHEUNG, M., LEUNG, S., CHIU, D. S., KALLOGER, S. E., HUNTSMAN, D. G., CHEN, S., INTERMAGGIO, M., GRONWALD, J., CHAN, F. C., RAMUS, S. J., STEIDL, C., SCOTT, D. W. & ANGLESIO, M. S. 2016. Single-Patient Molecular Testing with NanoString nCounter Data Using a Reference-Based Strategy for Batch Effect Correction. *PLoS One,* 11**,** e0153844.

TAWBI, H. A., BURGESS, M., BOLEJACK, V., VAN TINE, B. A., SCHUETZE, S. M., HU, J., D'ANGELO, S., ATTIA, S., RIEDEL, R. F., PRIEBAT, D. A., MOVVA, S., DAVIS, L. E., OKUNO, S. H., REED, D. R., CROWLEY, J., BUTTERFIELD, L. H., SALAZAR, R., RODRIGUEZ-CANALES, J., LAZAR, A. J., WISTUBA, I. I., BAKER, L. H., MAKI, R. G., REINKE, D. & PATEL, S. 2017. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *The Lancet Oncology,* 18**,** 1493-1501.

TEAM, I. 2019. immunarch: An R Package for Painless Bioinformatics Analysis of T-Cell and B-Cell Immune Repertoires. . *Zenodo*.

TESCHENDORFF, A. E., MARABITA, F., LECHNER, M., BARTLETT, T., TEGNER, J., GOMEZ-CABRERO, D. & BECK, S. 2013. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics,* 29**,** 189-196.

THRANE, K., ERIKSSON, H., MAASKOLA, J., HANSSON, J. & LUNDEBERG, J. 2018. Spatially resolved transcriptomics enables dissection of genetic heterogeneity in stage III cutaneous malignant melanoma.

TIROSH, I., IZAR, B., PRAKADAN, S. M., II, M. H. W., TREACY, D., TROMBETTA, J. J., ROTEM, A., RODMAN, C., LIAN, C., MURPHY, G., FALLAHI-SICHANI, M., DUTTON-REGESTER, K., LIN, J.-R., KAZER, S. W., GAILLARD, A. & KOLB, K. E. 2016. Dissecting the multicellular exosystem of metastatic melanoma by single-cell RNA-seq. *Science,* 352**,** 189-196.

TOPALIAN, S. L., DRAKE, C. G. & PARDOLL, D. M. 2015. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell,* 27**,** 450-61.

TSANG, Y. H., DOGRULUK, T., TEDESCHI, P. M., WARDWELL-OZGO, J., LU, H., ESPITIA, M., NAIR, N., MINELLI, R., CHONG, Z., CHEN, F., CHANG, Q. E., DENNISON, J. B., DOGRULUK, A., LI, M., YING, H., BERTINO, J. R., GINGRAS, M. C., ITTMANN, M., KERRIGAN, J., CHEN, K., CREIGHTON, C. J., ETEROVIC, K., MILLS, G. B. & SCOTT, K. L. 2016. Functional annotation of rare gene aberration drivers of pancreatic cancer. *Nature Communications*.

TUMEH, P. C., HARVIEW, C. L., YEARLEY, J. H., SHINTAKU, I. P., TAYLOR, E. J. M., ROBERT, L., CHMIELOWSKI, B., SPASIC, M., HENRY, G., CIOBANU, V., WEST, A. N., CARMONA, M., KIVORK, C., SEJA, E., CHERRY, G., GUTIERREZ, A. J., GROGAN, T. R., MATEUS, C., TOMASIC, G., GLASPY, J. A., EMERSON, R. O., ROBINS, H., PIERCE, R. H., ELASHOFF, D. A., ROBERT, C. & RIBAS, A. 2014. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*.

VACCHELLI, E., SEMERARO, M., ENOT, D. P., CHABA, K., POIRIER COLAME, V., DARTIGUES, P., PERIER, A., VILLA, I., RUSAKIEWICZ, S., GRONNIER, C., GOERE, D., MARIETTE, C., ZITVOGEL, L. & KROEMER, G. 2015. Negative prognostic impact of regulatory T cell infiltration in surgically resected esophageal cancer post-radiochemotherapy. *Oncotarget,* 6**,** 20840-50.

VAN ALLEN, E. M., MIAO, D., SCHILLING, B., SHUKLA, S. A., BLANK, C., ZIMMER, L., SUCKER, A., HILLEN, U., FOPPEN, M. H., GOLDINGER, S. M., UTIKAL, J., HASSEL, J. C., WEIDE, B., KAEHLER, K. C., LOQUAI, C., MOHR, P., GUTZMER, R., DUMMER, R., GABRIEL, S., WU, C. J., SCHADENDORF, D. & GARRAWAY, L. A. 2015a. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma.[Erratum appears in Science. 2015 Nov 13;350(6262):aad8366; PMID: 26564858]. *Science,* 350**,** 207-211.

VAN ALLEN, E. M., MIAO, D., SCHILLING, B., SHUKLA, S. A., BLANK, C., ZIMMER, L., SUCKER, A., HILLEN, U., FOPPEN, M. H. G., GOLDINGER, S. M., UTIKAL, J., HASSEL, J. C., WEIDE, B., KAEHLER, K. C., LOQUAI, C., MOHR, P., GUTZMER, R., DUMMER, R., GABRIEL, S., WU, C. J., SCHADENDORF, D. & GARRAWAY, L. A. 2015b. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science*.

VAN ALLEN, E. M., WAGLE, N., STOJANOV, P., PERRIN, D. L., CIBULSKIS, K., MARLOW, S., JANE-VALBUENA, J., FRIEDRICH, D. C., KRYUKOV, G., CARTER, S. L., MCKENNA, A., SIVACHENKO, A., ROSENBERG, M., KIEZUN, A., VOET, D., LAWRENCE, M., LICHTENSTEIN, L. T., GENTRY, J. G., HUANG, F. W., FOSTEL, J., FARLOW, D., BARBIE, D., GANDHI, L., LANDER, E. S., GRAY, S. W., JOFFE, S., JANNE, P., GARBER, J., MACCONAILL, L., LINDEMAN, N., ROLLINS, B., KANTOFF, P., FISHER, S. A., GABRIEL, S., GETZ, G. & GARRAWAY, L. A. 2014. Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. *Nature Medicine,* 20**,** 682-688.

VAN DER AUWERA, G. A., CARNEIRO, M. O., HARTL, C., POPLIN, R., DEL ANGEL, G., LEVY-MOONSHINE, A., JORDAN, T., SHAKIR, K., ROAZEN, D., THIBAULT, J., BANKS, E., GARIMELLA, K. V., ALTSHULER, D., GABRIEL, S. & DEPRISTO, M. A. 2013. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline: The Genome Analysis Toolkit Best Practices Pipeline. *In:* BATEMAN, A., PEARSON, W. R., STEIN, L. D., STORMO, G. D. & YATES, J. R. (eds.) *Current Protocols in Bioinformatics.* Hoboken, NJ, USA: John Wiley & Sons, Inc.

VAN DER LINDEN, M., VAN ESCH, E., BULTEN, J., DREEF, E., MASSUGER, L., VAN DER STEEN, S., BOSSE, T., DE HULLU, J. & VAN POELGEEST, M. 2018. The immune cell infiltrate in the microenvironment of vulvar Paget disease. *Gynecol Oncol,* 151**,** 453-459.

VAN DER MAATEN, L. & HINTON, G. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research*.

WAGGOTT, D., CHU, K., YIN, S., WOUTERS, B. G., LIU, F. F. & BOUTROS, P. C. 2012. NanoStringNorm: An extensible R package for the pre-processing of nanostring mRNA and miRNA data. *Bioinformatics,* 28**,** 1546-1548.

WANG, K., LI, M. & HAKONARSON, H. 2010. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*.

WANG, W., GREEN, M., CHOI, J. E., GIJÓN, M., KENNEDY, P. D., JOHNSON, J. K., LIAO, P., LANG, X., KRYCZEK, I., SELL, A., XIA, H., ZHOU, J., LI, G., LI, J., LI, W., WEI, S., VATAN, L., ZHANG, H., SZELIGA, W., GU, W., LIU, R., LAWRENCE, T. S., LAMB, C., TANNO, Y., CIESLIK, M., STONE, E., GEORGIOU, G., CHAN, T. A., CHINNAIYAN, A. & ZOU, W. 2019. CD8+ T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature*.

WANG, Y., WATERS, J., LEUNG, M. L., UNRUH, A., ROH, W., SHI, X., CHEN, K., SCHEET, P., VATTATHIL, S., LIANG, H., MULTANI, A., ZHANG, H., ZHAO, R., MICHOR, F., MERIC-BERNSTAM, F. & NAVIN, N. E. 2014. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature,* 512**,** 155-160.

WEI, S. C., ANANG, N. A. S., SHARMA, R., ANDREWS, M. C., REUBEN, A., LEVINE, J. H., COGDILL, A. P., MANCUSO, J. J., WARGO, J. A., PE'ER, D. & ALLISON, J. P. 2019. Combination anti-CTLA-4 plus anti-PD-1 checkpoint blockade utilizes cellular mechanisms partially distinct from monotherapies. *Proc Natl Acad Sci U S A,* 116**,** 22699-22709.

WICKHAM, H. 2007. Reshaping Data with the reshape Package. *Journal of Statistical Software,* 21**,** 20.

WICKHAM, H. 2011a. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*.

WICKHAM, H. 2011b. ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics,* 3**,** 180-185.

WICKHAM, H. 2017. *R: Package 'reshape2'*.

WICKHAM, H., FRANCOIS, R., HENRY, L., MÜLLER, K. & RSTUDIO. 2017a. *dplyr: A Grammar of Data Manipulation.* [Online]. Available: https://cran.r-project.org/package=dplyr [Accessed].

WICKHAM, H., HENRY, L. & RSTUDIO. 2017b. *tidyr: Easily Tidy Data with spread() and gather() Functions.* [Online]. Available: https://CRAN.R-project.org/package=tidyr [Accessed].

WILKERSON, M. D. & HAYES, D. N. 2010. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics*.

WOLF, Y., BARTOK, O., PATKAR, S., ELI, G. B., COHEN, S., LITCHFIELD, K., LEVY, R., JIMENEZ-SANCHEZ, A., TRABISH, S., LEE, J. S., KARATHIA, H., BARNEA, E., DAY, C. P., CINNAMON, E., STEIN, I., SOLOMON, A., BITTON, L., PEREZ-GUIJARRO, E., DUBOVIK, T., SHEN-ORR, S. S., MILLER, M. L., MERLINO, G., LEVIN, Y., PIKARSKY, E., EISENBACH, L., ADMON, A., SWANTON, C., RUPPIN, E. & SAMUELS, Y. 2019. UVB-Induced Tumor Heterogeneity Diminishes Immune Response in Melanoma. *Cell,* 179**,** 219-235 e21.

YARCHOAN, M., JOHNSON, B. A., 3RD, LUTZ, E. R., LAHERU, D. A. & JAFFEE, E. M. 2017. Targeting neoantigens to augment antitumour immunity. *Nat Rev Cancer,* 17**,** 209-222.

YE, K., SCHULZ, M. H., LONG, Q., APWEILER, R. & NING, Z. 2009. Pindel: A pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*.

YOSHIHARA, K., SHAHMORADGOLI, M., MARTÍNEZ, E., VEGESNA, R., KIM, H., TORRES-GARCIA, W., TREVIÑO, V., SHEN, H., LAIRD, P. W., LEVINE, D. A., CARTER, S. L., GETZ, G., STEMKE-HALE, K., MILLS, G. B. & VERHAAK, R. G. W. 2013. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications,* 4.

YU, G. 2018. *enrichplot: Visualization of Functional Enrichment Result*.

YU, G., WANG, L.-G., HAN, Y. & HE, Q.-Y. 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*.

YU, G., WANG, L. G., YAN, G. R. & HE, Q. Y. 2015. DOSE: An R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*.

ZARETSKY, J. M., GARCIA-DIAZ, A., SHIN, D. S., ESCUIN-ORDINAS, H., HUGO, W., HU-LIESKOVAN, S., TORREJON, D. Y., ABRIL-RODRIGUEZ, G., SANDOVAL, S., BARTHLY, L., SACO, J., HOMET MORENO, B., MEZZADRA, R., CHMIELOWSKI, B., RUCHALSKI, K., SHINTAKU, I. P., SANCHEZ, P. J., PUIG-SAUS, C., CHERRY, G., SEJA, E., KONG, X., PANG, J., BERENT-MAOZ, B., COMIN-ANDUIX, B., GRAEBER, T. G., TUMEH, P. C., SCHUMACHER, T. N. M., LO, R. S. & RIBAS, A. 2016. Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *New England Journal of Medicine,* 375**,** 819-829.

ZHANG, J. & FENG, B. “biocViews microarray. Package ‘cghMCR'” (2013); www.bioconductor.org/packages/release/bioc/html/cghMCR.html.

ZHANG, R., KIM, A. S., FOX, J. M., NAIR, S., BASORE, K., KLIMSTRA, W. B., RIMKUNAS, R., FONG, R. H., LIN, H., PODDAR, S., CROWE, J. E., JR., DORANZ, B. J., FREMONT, D. H. & DIAMOND, M. S. 2018. Mxra8 is a receptor for multiple arthritogenic alphaviruses. *Nature,* 557**,** 570-574.

ZHANG, X. C., XU, C., MITCHELL, R. M., ZHANG, B., ZHAO, D., LI, Y., HUANG, X., FAN, W., WANG, H., LERMA, L. A., UPTON, M. P., HAY, A., MÉNDEZ, E. & ZHAO, L. P. 2013. Tumor Evolution and Intratumor Heterogeneity of an Oropharyngeal Squamous Cell Carcinoma Revealed by Whole-Genome Sequencing. *Neoplasia,* 15**,** 1371-IN7.

ZHENG, G. X. Y., TERRY, J. M., BELGRADER, P., RYVKIN, P., BENT, Z. W., WILSON, R., ZIRALDO, S. B., WHEELER, T. D., MCDERMOTT, G. P., ZHU, J., GREGORY, M. T., SHUGA, J., MONTESCLAROS, L., UNDERWOOD, J. G., MASQUELIER, D. A., NISHIMURA, S. Y., SCHNALL-LEVIN, M., WYATT, P. W., HINDSON, C. M., BHARADWAJ, R., WONG, A., NESS, K. D., BEPPU, L. W., DEEG, H. J., MCFARLAND, C., LOEB, K. R., VALENTE, W. J., ERICSON, N. G., STEVENS, E. A., RADICH, J. P., MIKKELSEN, T. S., HINDSON, B. J. & BIELAS, J. H. 2017. Massively parallel digital transcriptional profiling of single cells. *Nature Communications*.

ZHENG, Y., CHEN, Z., HAN, Y., HAN, L., ZOU, X., ZHOU, B., HU, R., HAO, J., BAI, S., XIAO, H., LI, W. V., BUEKER, A., MA, Y., XIE, G., YANG, J., CHEN, S., LI, H., CAO, J. & SHEN, L. 2020. Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nat Commun,* 11**,** 6268.

Vita

Akash Mitra was born on November 10, 1992. After completing high school, he entered Rutgers University in New Brunswick, New Jersey. He received his Bachelor of Science degree with a major in biotechnology and bioinformatics from Rutgers in May 2016. In August of 2016 he entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences.

Permanent address:

1885 El Paseo St, Apt 35410 Houston, TX, 77054