Immune signatures and targets in human tumors

Borch, Annie

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

DTU Health Tech
Department of Health Technology

Technical University of Denmark

PhD Thesis

Immune signatures and targets in human tumors

Author:
Annie Borch

Main supervisor:
Sine Reker Hadrup
Institute of Health technology, Technical University of Denmark

Co-supervisor:
Lars Rønn Olsen
Institute of Health technology, Technical University of Denmark

Co-supervisor:
Ulrik Lassen
Department of Oncology, University Hospital of Copenhagen

December 23, 2022

PREFACE

This PhD thesis was conducted at the Department of Health Technology at the Technical University of Denmark (DTU) and two weeks of exchange at University College London (UCL) in Nicholas McGranhans group. The thesis was initiated in April 2019 and handed in December 2022. Supervision during the thesis was mainly conducted by Professor Sine Reker Hadrup with co-supervision from associate Professor Lars Rønn Olsen, both from the Department of Health Technology, and Professor Ulrik Lassen from the University Hospital of Copenhagen.

The thesis consists of a general introduction followed by four manuscripts and finalized with an epilogue discussing the concerns and interests points and perspectives of the thesis.

DTU, December 23, 2022

Annie Borch

ABSTRACT

Cancer is a global problem and causes many deaths every year. Treatment with immunotherapy has, in the past decades, helped decrease cancer mortality. High-throughput sequencing technologies have opened new research possibilities, and the advantages of using omics data in cancer research have improved immunotherapy strategies and, among others, contribute to novel findings regarding biomarker discovery and personalized cancer vaccines. Biomarkers can ideally be used to predict the possibility of a patient obtaining clinical benefit from immunotherapy; however, those currently known account poorly across all patients. Hence, the optimal biomarker has yet to be found. $CD8^+$ T cells are key players in the immune system to combat cancer. They have the ability to specifically recognize personalized mutation-derived epitopes (neoepitopes) on the surface of cancer cells and directly kill cancerous cells. Multiple approaches exist to efficiently predict possible neoepitope candidates that are presented at the surface of cancer cells, but only a small fraction of these is truly immunogenic, meaning that they are, in fact, recognized by $CD8^+$ T cells.

This thesis studies potential biomarkers to predict the outcome of treatment with immunotherapy, either through the use of omics data and/or through the characterization and identification of immunogenic neoepitopes. Four Manuscripts are included to address this: **Manuscript I** investigates predictive biomarkers across a diverse cohort and states that a combination of biomarker is more suited to predict patient survival than observing single biomarkers. The suggested combinations include T-cell signatures and cancer cell signatures. The Manuscript propose a combination of either neoepitope load and programmed cell death ligand 2 (PD-L2) or neoepitope, PD-L2, and cytolytic activity (CYT) as a combined potential biomarker. **Manuscript II** and **Manuscript III** study a large pool of neoeptiope candiadtes and validate their immunogenicity potential. Both Manuscripts suggest that the abundance and frequency of neoantigen reactive $CD88^+$ T cells (NARTs) can be used as a predictive biomarker in different setups. The tumor microenvironment (TME) also results in diversity when comparing patients with many detected NARTs to those with few. Furthermore, we observe differences in the characteristics of neoepitope candidates and describe how they can be used to distinguish immunogenic neoepitopes from non-immunogenic ones. The final Manuscript (**Manuscript IV**) investigates even more broadly the characteristics of immunogenic neoepitopes. A feature-based machine learning approach with random forest modeling using a dataset

of above 19000 validated neoepitope candidates are used to predict the immunogenicity of neoepitopes. From this analysis, the physicochemical properties of the neopeptides are found to be the most important features in predicting immunogenicity, and a high abundance of hydrophobic and aromatic residues are the most essential properties. Additionally, patient-specific TME features are implemented in the neoepitope prediction, resulting in a slight improvement.

Even though the statements in this thesis need further validation with more data, the four Manuscripts of this thesis add novel insight into the field of cancer immunotherapy through the discovery of potential biomarkers that can be used to predict survival probability from immunotherapy treatment and by improving the neoepitope prediction.

RESUMÉ

Kræft er et globalt problem og forårsager mange dødsfald hvert år. Behandling med immunterapi har i de seneste årtier været med til at mindske kræftdødeligheden. Sekventeringsteknologier har åbnet nye forskningsmuligheder, og fordelene ved at bruge omics-data i kræftforskning har forbedret immunterapistrategier og bidrager blandt andet til nye viden omkring biomarkører og personaliserede cancervacciner. Biomarkører kan ideelt set bruges til at forudsige muligheden for, at en patient opnår klinisk fordel ved immunterapi; dog er de kendte i øjeblikket ikke gældende for alle patienter. Derfor er den optimale biomarkør endnu ikke fundet. $CD8^+$ T celler er nøglespillere i immunsystemet til at bekæmpe kræft. De har evnen til specifikt at genkende personaliserede mutations fremkommet epitoper (neoepitoper) på overfladen af kræftceller og kan hermed direkte dræbe dem. Der findes flere muligheder til effektivt at forudsige mulige neoepitop kandidater, der kan blive præsenteret på overfladen af cancerceller, men kun en lille del af disse er virkelig immunogene, hvilket betyder, at de kan blive genkendt af $CD8^+$ T-celler.

Denne afhandling undersøger mulige biomarkører som kan bruges til at forudsige udfaldet af kræftbehandling for en kræftpatient, enten ved brug af omics-data eller igennem identifikationen af immunogene neoepitoper. Fire manuskripter er inkluderet for at for undersøge: **Manuskript I** undersøger prædiktive biomarkører på tværs af en forskellige kræft typer og foreslog, at en kombination af biomarkører er bedre til at forudsige patienters sandsynlighed for overlevelse end at observere enkelte biomarkører. Den foreslåede kombination indeholder T celle signaturer og cancer celle signaturer. Artiklen foreslog en kombination af enten mængden af neoepitoper og programmed cell death ligand 2 (PD-L2) eller mængden af neoepitoper, PD-L2 og cytolytisk aktivitet (CYT) som en kombineret potentiel biomarkør. **Manuskript II** og **Manuskript III** studerer en stor pulje af forudsete neoepitope kandidater og validerer deres immunogenicitets potentiale. Begge artikler tyder på, at mængden og hyppigheden af neoantigen-reaktive T celler (NART'er) kan bruges som en biomarkør til at forudse overlevelse på forskellige måder. Tumormikromiljøet (TME) viser også diversitet, når man sammenlignede patienter med mange påviste NART'er med dem med få. Desuden observerede vi forskellige neoepitop karakteristika og beskriver, hvordan de kan bruges til at skelne mellem immunogene neoepitoper fra ikke - immunogene . Det sidste manuskript (**Manuskript IV**) undersøger endnu bredere karakteristika af immunogene neoepitoper. Der blev lavet en maskinlærings model

indeholdende neoepitope karakteristika med en Random Forrest (RF) model ved brug af et datasæt med over 19000 valideret neoepitop kandidater, som blev brugt til at forudsige immunogeniciteten af neoepitoper. Ud fra denne analyse blev de fysisk-kemiske egenskaber af peptiderne, fundet til at være de vigtigste egenskaber til at forudsige immunogenicitet, og en høj andel af hydrofobe og aromatiske aminosyre i peptiderne, var de vigtigste egenskaber. Derudover blev patientspecifikke TME-karakteristika implementeret i neoepitop forudsigelsen, hvilket resulterer i en forbedret model.

Selvom udsagnene i denne afhandling har brug for yderligere validering med mere data, tilføjer de fire artikler i denne afhandling ny viden indenfor cancerimmunterapi på baggrund af opdagelsen af potentielle biomarkører, der kan bruges til at forudsige overlevelsessandsynlighed fra immunterapibehandling og ved at forbedre neoepitop forudsigelse.

ACKNOWLEDGMENTS

I am grateful that I made it that far, and I could not have done it without a bunch of amazing people. The journey started as a master's student, where I got to write the most amazing project I could dream of. I got hooked on the topic and got excellent supervision from my bioinformatic supervisor Anne-Mette; thank you. My current main supervisor, Sine, used all her fantastic application skills and got funding for my PhD. I am grateful that you, Sine, believed in me and allowed me this fantastic PhD opportunity, and thanks for all your guidance and encouragement during my PhD journey. As you know, it has not been easy for me to be a lonely bioinformatician in a big wet lab group with too many side projects. But you helped me prioritize, so I managed. Moreover, thanks to the Novo Nordisk foundation for breaking the curse and providing Sine money for my PhD.

Thanks to my co-supervisor Lars Rønn Olsen for great advice and support and Ulrik Lassen for giving me guidance and insides from a clinical viewpoint. Moreover, I would like to thank the doctors I have collaborated with, Vinicius and Simone. It has been a pleasure to work with you both. I would also like to express my deep gratitude to Morten Nielsen. Even though you officially aren't my supervisor, you have helped me a lot, especially during the final project for my PhD. Thanks for your guidance and patience every time I mixed up machine-learning terms. I would also like to thank the people in your group, especially Ibel and Birkir, for their excellent collaboration.

I am very grateful for every former and current member of SRHgroup for providing me advice, motivation and tons of "hyggelige" activities, including; lunch club, running club, and so forth. You have all been a big part of my journey, especially you, Sara. We have shared many frustrations and happy moments, and I am very thankful for sharing this journey with you. Particularly thanks to my former and current bioinformatic college, and an extended thanks to Kamilla for being a great friend and office neighbor with daily bioinformatics discussion, assistance and support. Also thanks to all my other office mates for their daily advice and joyful conversations, especially Amalie, for indispensable support and Mo for cheering me up and stabilizing my blood sugar. A special thanks to all the people who I have been collaborating with, thanks for the battle, and all the fantastic lab-work you have done. An enormous thanks to everyone who

have contributed to proofreading my thesis; I appreciate you all for taking the time and helping me.

Thanks to everyone at XTI for the motivating and critical questions during the internal seminar and encouraging smiles in the hallway. Mainly thanks to Isabel and Line for starting this PhD Journey together. Thanks for the everyday morning coffee with you, Line. It has been amazing to have you on my side during this journey which started at the master's student office with long nights and red wine, and I am glad we can continue the morning coffees for a while more.

Great thanks to all the patients who believed in science and agreed to participate in these trials. Without your cooperation, we had no chance of during this kind of exiting researching in cancer.

Thanks to my opponents, for taking the time to evaluate my precious work.

Finally, I would like to thank all my friends and family for love and support. Jesper, my lovely husband, for always believing in me and our fantastic daughter Elena "Eviggald". You both mean the world to me; endless love for you both.

PUBLICATIONS

Manuscripts included in the thesis

I **Annie Borch***, Anne-Mette Bjerregaard*, Vinicius Araújo Barbosa de Lima*, Olga Oestrup, Christina Westmose Yde, Aron Charles Eklund, Morten Mau-Sørensen, Carolina Barra Quaglia, Inge-Marie Svane, Finn Cilius Nielsen, Samuel A. Funts, Ulrik Lassen* and Sine Reker Hadrup*. "Neoepitope load, T cell signatures and PD-L2 as combined biomarker strategy for response to checkpoint inhibition."
Frontiers Genetic, In review, Submitted 30 September 2022

II Jeppe Sejerø Holm*, Samuel A. Funt*, **Annie Borch**, Kamilla Kjærgaard Munk, Anne-Mette Bjerregaard, James L. Reading, Colleen Maher, Ashley Regazzi, Phillip Wong, Hikmat Al-Ahmadie, Gopa Iyer, Tripti Tamhane, Amalie Kai Bentzen, Nana Overgaard Herschend, Susan De Wolf, Alexandra Snyder, Taha Merghoub, Jedd D. Wolchok, Morten Nielsen, Jonathan E. Rosenberg, Dean F. Bajorin* Sine Reker Hadrup*. "Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma"
Nature communications - 11 April 2022

III Nikolaj Pagh Kristensen*, Christina Heeke*, Siri A. Tvingsholm*, **Annie Borch**, Arianna Draghi, Michael D. Crowther, Ibel Carri, Kamilla K. Munk, Jeppe Sejerø Holm, Anne-Mette Bjerregaard, Amalie Kai Bentzen, Andrea M. Marquard, Zoltan Szallasi, Nicholas McGranahan, Rikke Andersen, Morten Nielsen, Göran B. Jönsson, Marco Donia, Inge Marie Svane, Sine Reker Hadrup. "Neoantigen-reactive CD8+ T cells affect clinical outcome of adoptive cell therapy with tumor-infiltrating lymphocytes in melanoma."
The Journal of clinical investigation - (JAN 2022), Vol. 132, no. 2

IV **Annie Borch**, Ibel Carri*, Birkir Reynisson*, Heli Magali García Álvarez, Kamilla K. Munk, Alessandro Montemurro, Nikolaj Pagh Kristensen, Siri A. Tvingsholm, Jeppe Sejerø Holm, Christina Heeke, Keith Henry Moss, Ulla Kring Hansen,Anna-Lisa Schaap-Johansen, Frederik Otzen Bagger, Carolina Barra, Morten Nielsen,* and Sine Reker Hadrup*. "IMPROVE, A feature model to predict neoepitope immunogenicity through broad-scale validation of T cell recognition."
Manuscript in preparation

* these author are contributed equally.

Manuscripts not included in the thesis

I Vinicius Araujo B. de Lima, **Annie Borch**, Morten Hansen, Arianna Draghi, Iben Spanggaard, Kristoffer Rohrberg, Sine Reker Hadrup, Ulrik Lassen and Inge Marie Svane. "Common phenotypic dynamics of tumor-infiltrating lymphocytes across different histologies upon checkpoint inhibition: impact on clinical outcome"
Cytotherapy - April 2020, Volume 22, Issue 4

II Anna-Lisa Schaap-Johansen, Milena Vujovic´ , **Annie Borch**, Sine Reker Hadrup and Paolo Marcatili. "T cell epitope prediction and its application to immunotherapy."
Frontiers in Immunology — 2021, Volume 12, pp. 712488

III Ulla Kring Hansen1†, Sofie Ramskov 1†, Anne-Mette Bjerregaard, **Annie Borch**, Rikke Andersen, Arianna Draghi , Marco Donia, Amalie Kai Bentzen, Andrea Marion Marquard, Zoltan Szallasi, Aron Charles Eklund, Inge Marie Svane and Sine Reker Hadrup. "Tumor-infiltrating T cells from clear cell renal cell carcinoma patients recognize neoepitopes derived from point and frameshift mutations"
Fontiers in Immunuology - 2 March 2020,

IV Sunil Kumar Saini, Andreas Due Ørskov, Anne-Mette Bjerregaard, Ashwin Unnikrishnan, Staffan Holmberg-Thydén, **Annie Borch**, Kathrine Valentini Jensen, Govardhan Anande, Amalie Kai Bentzen, Andrea Marion Marquard, Tripti Tamhane, Marianne Bach Treppendahl, Anne Ortved Gang, Inge Høgh Dufva, Zoltan Szallasi, Nicola Ternette, Anders Gorm Pedersen, Aron Charles Eklund, John Pimanda, Kirsten Grønbæk Sine Reker Hadrup. "Human endogenous retroviruses form a reservoir of T cell targets in hematological cancers"
Volume 132, Issue 2 on January 18, 2022

V Simone Maarup, Benedikte Hasselbalch, **Annie Borch**, Filip Mundt, Sille Hendriksen, Arianna Draghi, Mario Presti, Marco Donia, Signe Koggersbøl Skadborg, Ib Jarle Christensen, Vilde Pedersen, Jane Skjøth-Rasmussen, Christina Westmose Yde6, Matthias Mann, Bjarne Winther Kristensen, Hans Skovgaard Poulsen, Inge Marie Svane, Sine Reker Hadrup, Ulrik Lassen. "Paired primary and recurrent tumor analyses identify a subgroup of long-term survivors after neoadjuvant nivolumab and bevacizumab for recurrent glioblastoma"
Manuscript in preparation

VI Keith Henry Moss, Ulla Kring Hansen, Vinicius Araújo Barbosa de Lima,Esteban Sanchez Marquez, **Annie Borch**, Anne-Mette Bjerregaard, Olga Oestrup, Amalie Kai Bentzen, Andrea Marion Marquard, Inge Marie Svane, Ulrik Lassen, Sine Reker Hadrup. "Characteristics of neoantigen-reactive $CD8^+$ T cells as an immune predictor for clinical outcome to checkpoint blockade therapy in a pan-cancer setting".
Manuscript in preparation

VII Signe Skadborg, SimoneMaarup, Arianna Draghi, **AnnieBorch**, Inge Marie Svane, Benedikte Hasselbalch, Hans Poulsen, Ulrik Lassen and Sine Reker Hadrup. "Evidence of T cell activation and intratumoral nivolumab-presence in glioblastoma patients treated with nivolumab and bevacizumab".
Manuscript in preparation

ABBREVIATIONS

ACT	Adoptive cell transfer
TIL-ACT	Adoptive cell transfer with Tumor infiltrated lymphocyte (TIL)
ANN	Artificial neural network
APC	Antigen presented cell
AUC	Area under receiver operating characteristic curve
AUC01	AUC 10%
BWA	Burrows-Wheeler aligner
CD	Cluster differentiation
CPI	Checkpoint inhibitors
CTLA-4	Cytotoxic T-lymphocyte–associated antigen 4
CYT	Cytolytic activity
DAI	Differential agretopicity index
DEA	Differential expression analysis
ER	Endoplasmatic Reticulum
FN	False negative
FP	False positive
FPR	False positive rate
GATK	Genome analysis toolkit
GSEA	Gene set enrichment analysis
GZMA	Granzyme A
HLA	Human leukocyte antigen
KIR	Killer cell immunoglobulin-like receptor
LOHHLA	Loss of hetrozygozity HLA
MHC	Major histocompatibility complex
ML	Machine learning
mUC	Metastatic urothelial carcinoma
MuPeXI	MUtant Peptide Extractor and Informer
NART	Neoantigen-reactive CD8+ T cell
NK	Natural killer cells
PBMC	Peripheral blood mononuclear cells
PD-1	Programmed cell Death 1
PD-L1	Programmed cell Death ligand 1
PD-L2	Programmed cell Death ligand 2
pMHC	peptide:Major histocompatibility complex
PoN	Panel of Normal

PRF1	Perforin 1
PVACseq	personalized Variant Antigens by Cancer Sequencing
RankEL	Eluted ligand % rank
RankBA	Binding affinity % rank
RF	Random Forrest
RNAseq	RNA sequencing
ROC	Receiver operating characteristic
SC	Single cell
TAP	Transporter associated with antigen processing
TCR	T cell receptor
TCV	Therapeutic cancer vaccines
TIL	Tumor infiltrated lymphocyte
TIminer	Tumor Immunology Miner
TMB	Tumor mutational burden
TME	Tumor microenviroment
TN	True negative
TNR	True negative rate
TP	True positive
TPR	True positive rate
VEP	Variant effect predictor
WGS	Whole genome sequencing
WXS	Whole exome sequencing

TABLE OF CONTENTS

1 INTRODUCTION

1.1 CANCER OMICS

Cancer development is a rising global health problem, with an estimated 28.4 million new cases projected to occur in 2040 alone (Sung et al., 2021). The development of cancer is caused by genetic alterations, which can lead to abnormal and uncontrollable cell growth. Any cell in the body has the potential to turn into a cancer cell, resulting in many different types of cancer that can behave and grow differently (Cooper and Hausman, 2000). The immune system has the capacity to eliminate cancer cells, (Gonzalez et al., 2018), resulting in a great focus in using the immunesytem role in cancer elimination and developing novel immunotherapies (Huang and Chang, 2019). However, genetically, all individuals are unique, and thus their immune systems differ, making new uniform treatment strategies challenging to develop. Cancer heterogeneity, host genetics, as well as environmental and other factors, make cancer extremely challenging to combat (McGranahan and Swanton, 2017). Researchers worldwide took advantage of high-throughput sequencing to understand cancer heterogeneity which has also influenced the development of optimized treatment strategies (Kuksin et al., 2021). Different sequencing technologies can be applied to produce different types of omics data, which can, be used to investigate cancer cells and the tumor microenvironment (TME) from different angles. Whole genome sequencing (WGS) and whole exome sequencing (WXS) produce genomics data studying genes at the DNA level, whereas transcriptomics analyzes based on RNA sequencing (RNAseq) data allows for analysis of etc., the gene expression in cancer dynamics. Both strategies individually have been used to reveal a lot of information regarding cancer diagnostic and prognostics and have been applied to drive decisions on treatment options. Better information coverage can be obtained by multi-omics approaches integrating combinations of different omics strategies, which this thesis uses to explore cancer omics (Chakraborty et al., 2018).

1.2 TUMOR HETEROGENEITY

The evolutionary nature of malignant tumors and their heterogeneity have been known for centuries. From the moment that sequencing approaches became accessible, a broader knowledge of tumor heterogeneity has been established, via the identification of specific mutations in high-throughput studies (McGranahan and Swanton, 2017).

1.2.1 Mutations types

Mutations in cancer cells can occur in various manners leading to different consequences. For example, mutations that cause an amino acid change within the protein sequence are non-synonymous mutations, whereas synonymous mutations do not lead to an amino acid change (Chu and Wei, 2019). This thesis will focus on the non-synonymous mutations, and the sum of all non-synonymous mutations is defines as the tumor mutational burden (TMB). Non-synonymous mutation can be caused by different errors resulting in different mutation consequences. Missense mutations are single nucleotide replacements. Frameshift mutations are caused either by short insertions or deletions, resulting in a changed reading frame of the nucleotide sequence. If an insertion or deletion appears within the reading frame, it is called in-frame insertion or in-frame deletion, respectively (Bjerregaard et al., 2017; McLaren et al., 2016).

1.2.2 Clonal evolution

During cell division of cancer cells, they can pass along present mutations and concurrently develop new mutations. The early-arrived and highly prevalent mutations are classified as clonal mutations and the later-arrived and less prevalent mutations are classified as sub-clonal mutations (Figure 1). PyClone is a computational program that can identify clonal and sub-clonal mutations by grouping mutations into clusters and estimating the mutation prevalence (Roth et al., 2014). Accurate estimations are highly dependent on knowing the purity of the tumor, which describes the fraction of tumor and non-tumor cells in a sample. The purity can be obtained from other bioinformatic tools such as allele-specific copy number analysis of tumours (ASCAT) and Sequenza (Favero et al., 2015; Raine et al., 2016). Break-through research has been conducted studying the clonal and sub-clonal mutations and will likely improve the precision of drug targets in personalized cancer vaccines (McGranahan et al., 2016; McGranahan and Swanton, 2017).

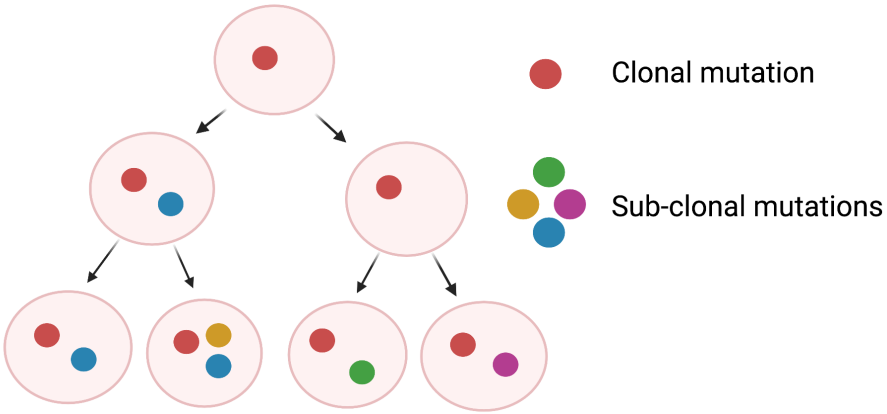


Figure 1: Clonal and sub-clonal mutations. During division of cancer cells, the cell's original and early arrived mutations are defined as clonal mutations. Clonal mutations will be prevalent in many of the cancer cells. Contrary sub-clonal mutations are the later-arrived mutations and are less prevalent. The red dot represents an example of a clonal mutation, whereas the other colors represent an example of sub-clonal mutations. Modified from McGranahan et al. (2016)

1.3 TUMOR MICROENVIRONMENT

Transcriptomics data are usually used to analyze tumor cells and the surrounding TME. The infiltration of immune cells in the TME is essential for tumor regression. Besides immune and tumor cells, the TME consists of stromal cells, blood vessels, and extracellular matrix. The TME can be divided into "hot" and "cold" TMEs, defined by the environment's concentration of essential immune cells, chemokines, and cytokines. Many versions of how to characterize a "hot" and "cold" TME have been suggested. In general, a "hot" has high infiltration of effector cells, namely $CD8^+$ T cells and natural killer (NK) cells, which are able to directly kill cancer cells and secrete pro-inflammatory cytokines and chemokines further contributing to the generation of a cytotoxic TME. In addition, these "hot" tumors are generally characterise by a high TMB. Cancer patients with a "hot" TME are found to be favorable in obtaining clinical benefit from treatment with immunotherapy. Contrary, a "cold" TME consists of an immunosuppressive environment resulting in low tumor-infiltrating lymphocytes (TILs) with low chemokine, and cytokine production. Patients with a "cold" TME are known to be unfavorable in responding to immunotherapy (Gajewski et al., 2017; Nagarsheth et al., 2017). Recently Ren at al. introduced an intermediate TME between "cold" and "hot", used to explain TME variance.

For example, a tumor with a high TMB does not necessarily have many TILs (Figure 2) (Ren et al., 2022).

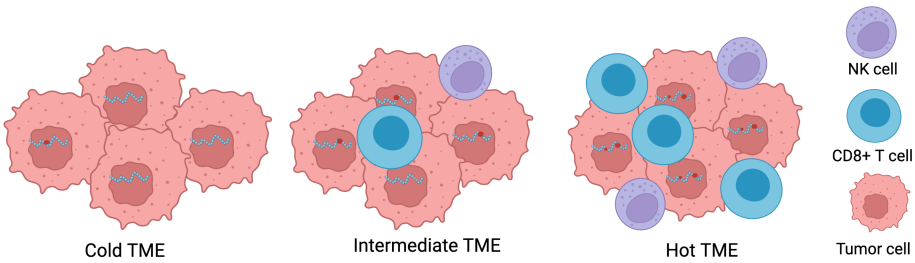


Figure 2: Tumor microenvironment (TME) categories. A "cold" TME is defined as a TME with no or few mutations and few or no infiltrating immune cells, such as NK cells and $CD8^+$ T cells. On the contrary, a "hot" TME has a high tumor mutational burden and many infiltrating immune cells. The "intermediate" TME is defined as an environment between the "cold" and "hot" TME. Modified from Ren et al. (2022)

Different sequencing strategies can be used to explore TME from transcriptomics data. RNAseq can be conducted on a single-cell (sc) level where it is possible to distinguish different cell types and study the effect of the TME on a cell-specific level. Contrary to sc-RNAseq, bulk RNA sequencing data give a picture of the mixed tumor sample, and the TME can be interpreted as one combined environment. Bulk sequencing technology was invented before the sc-RNAseq technology and is cheaper and faster (Li and Wang, 2021). Therefore, many studies are based on data from bulk RNAseq. However, bioinformatic tools have have been developed, enabling cell type estimations also known as pseudo sc and estimates cell populations from expression data obtained from bulk RNAseq. As an example, Microenvironment Cell Populations-counter (MCP-counter) estimates the abundance of 10 populations, including; endothelial cells, fibroblasts, myeloid dendritic cells, monocytic lineage cells, neutrophils, NK cells, B cell lineage, cytotoxic lymphocytes, T cells, and $CD8^+$ T cells (Becht et al., 2016). Another approach to studying differences in TME between two groups is differential expression analysis (DEA), typically followed by gene set enrichment analysis (GSEA). DEA can calculate gene regulation between two groups and find differentially expressed genes between the two groups (Love et al., 2014). The GSEA can identify if a cluster of upregulated genes is within the same gene set or gene pathway and calculate whether the given pathway is differentially expressed in the patient group (Yu et al., 2012; Zito et al., 2021)

1.4 T CELLS IN CANCER

One of the most important immune cells in cancer elimination are Cluster differentiation (CD)8^+ T cells ($CD8^+$ T cells), as they can kill cancer cells directly by recognizing cancer-antigens presented on major histocompatibility complex (MHC) class I molecules. CD4+ T cells, also known as the helper T cells, recognize peptides presented on MHC class II molecules and can stimulate $CD8^+$ T cells towards tumor cell killig (Ferris et al., 2020). $CD8^+$ T cells and the MHC-I restricted cancer-derived epitopes will be the main focus of this thesis.

1.4.1 Antigen presentation and human leucocyte antigen

T cells recognize cancer epitopes presented on the surface of antigen-presenting cells (APCs) or cancer cells thanks to the MHC complex, referred to as human leucocite antigen (HLA) molecules in humans. For the purpose of this thesis, I will only focus on antigen processing and MHC-I presentation to $CD8^+$ T cells, which is a multiple-step pathway. First, proteasomes cleave the protein into peptides in the cytosol. Next, the peptides will be transported to the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP) to meet the MHC-I molecules. MHC-I is held stable in the ER by ER chaperone proteins until the peptide binds into the binding cleft of the MHC-I complex. Finally, the peptide-bound MHC-I complex (pMHC) is transported through the golgi apparatus to the cell surface and presented to $CD8^+$ T cells (Figure 3A) (Neefjes et al., 2011). The HLA-encoding genes are placed on chromosome six, one of the most polymorphic regions of the human genome. The HLA class I exists as three classes of classical HLAs (HLA-A, HLA-B, and HLA-C) where each individual express six HLA molecules (three from mom and three from dad). There exists 5000 different alleles for each HLA class and this allelic difference results in most people being heterozygous in each HLA class (Robinson et al., 2020). The different HLA properties also give diversity in peptide binding. The anchor residues are responsible for peptide binding and can vary for each HLA molecule, although they are mainly located at positions 2 and 9 of the peptide sequence. The remaining amino acids of the peptide are responsible for T cell recognition (Figure 3B) (Fritsch et al., 2014). The knowledge of the HLA composition of an individual plays an essential role in epitope discovery in cancer. Bioinformatic tools have allowed the identification of a person's HLA alleles in a high-throughput manner from sequencing data. Precise HLA typing is complicated as many HLAs share many similarities, but a bioinformatic tool called Optitype, is able to can type HLA alleles from sequencing data with high accuracy, and is the tool used in this thesis for HLA typingn(Kiyotani et al., 2017; Szolek et al., 2014).

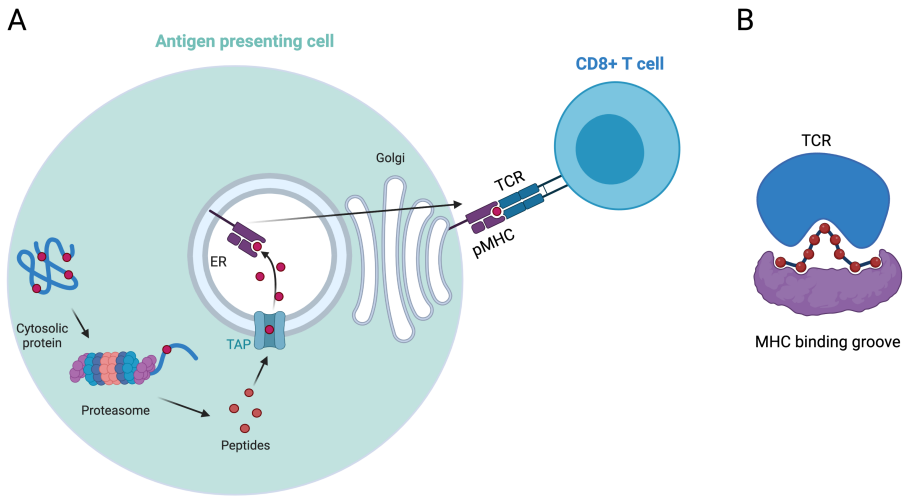


Figure 3: Peptide processing and binding groove of MHC-I. A) Proteins in the cytosol are cleaved into small peptides by the proteasome, and the peptides are then transported to the endoplasmic reticulum (ER) by the transporter associated with antigen presentation (TAP). The peptides can bind to the MHC molecule in the ER and be transported through the Golgi apparatus to the cell surface. B) MHC-I binding groove binds to the peptides with two anchor positions, mainly 2 and 9, where the rest of the peptide is presented to the T cell receptor (TCR). A is Modified from Neefjes et al. (2011) and B from Fritsch et al. (2014)

1.4.2 CD8^+ T cell activation

Upon CD8^+ T cell antigen recognition on the surface of APCs, CD8^+ T cells differentiate from an initial naive stage to an activated phenotype. Activation of naïve T cells happens in the lymph node where the T cell receptor (TCR) recognizes pMHC complexes on APCs. The activation rely on the costimulatory interaction of the CD28 receptor on T cells to its ligand B7 on the APC, and the secretion of cytokines by the APC determines which effector cell it becomes (Figure 4A). The effector CD8^+ T cells can proliferate and travel to the tumor site and kill cancer cells by secreting cytolytic molecules such as such as Granzyme A (GZMA) and perforin 1 (PRF1) after recognizing the tumor antigen (Figure 4B) (Raskov et al., 2021; Rooney et al., 2015).

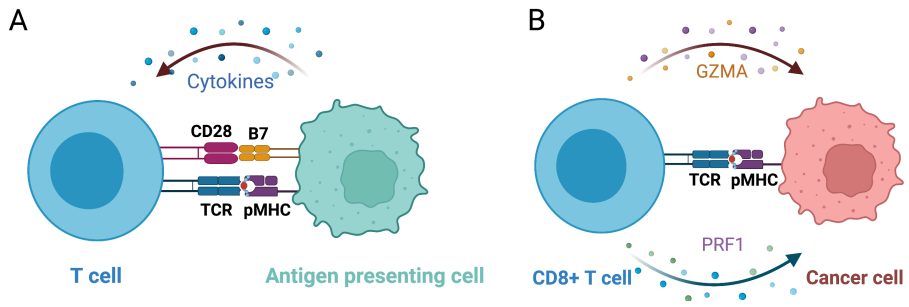


Figure 4: T cells in cancer. A) T cell activation depends on three signals. Firstly the T cell receptor (TCR) recognizes peptide-bound MHC (pMHC) on the cell surface of antigen-presenting cells (APC). Then the binding of CD28 co-receptor on the T cells to its ligand B7 on APC. Lastly, cytokine secretion which determines which effector cell it becomes. B) An activated $CD8^+$ T cell can kill the cancer cells only by recognizing the pMHC molecule on the cancer cell and then secreting enzymes such as granzyme A (GZMA) and perforin 1 (PRF1). Modified from Murphy and Weaver (2017).

1.5 NEOEPITOPES

Neopeptides are tumor specific peptides carrying an amino acid alteration derived from mutational changes in the cancer genome. When presented in the context of MHC-I molecules to $CD8^+$ T cells they are referred to as neoepitopes, and neoepitopes which can elicit CD8+ T cell reactivity are identified as immunogenic neoepitopes. Owing to the neoepitopes mutation origin, they are highly variable in between tumors and patients, conferring them a great potential for the development of personalized immunotherapies (e.g., cancer vaccines) (Saxena et al., 2021).

1.5.1 Prediction of neoepitopes

Several tools have been developed in recent years to predict patient-specific neoepitope candidates from sequencing data, such as Tumor Immunology miner (TIminer), antigen.garnish, personalized Variant Antigens by Cancer Sequencing (PVACseq) and MUtant Peptide Extractor and Informer (MuPeXI) (Bjerregaard et al., 2017; Hundal et al., 2016; Richman et al., 2019; Tappeiner et al., 2017). The overall strategy is to use WGS or WXS sequencing data of tumors and compared to sequencing data from blood to detect tumor specific mutations. Additionally, patient-specific HLA alleles can be used for selecting peptides that bind to their particular HLA, and tumor RNAseq data can be used for filtering

the subset of expressed mutations (Hackl et al., 2016; Schaap-Johansen et al., 2021).

In this thesis, MuPeXI was used to predict patient-specific neoepitope candidates. The best practice guidelines from the genome analysis tool kit (GATK) (Van der Auwera et al., 2013) were used pre-process data and to detect tumor specific mutations. An overview of the and the prediction pipeline used for this thesis is illustrated in Figure 5.

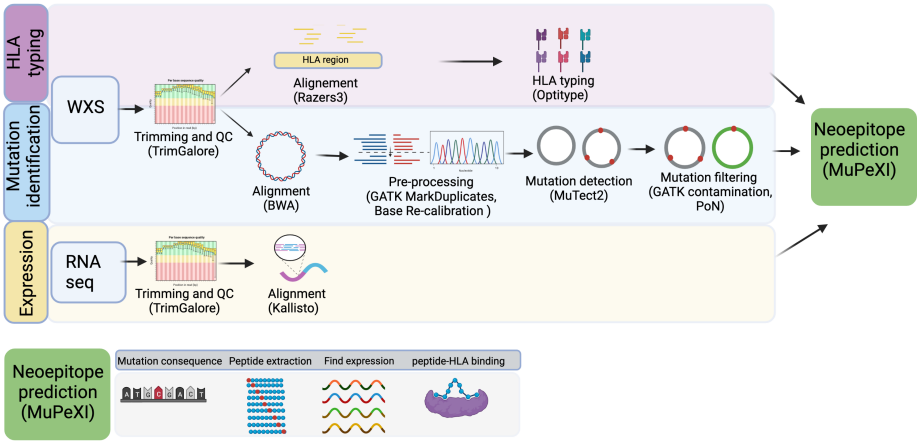


Figure 5: Neoepitope prediction pipeline. WXS is trimmed, and quality checked with TrimGalore. For HLA typing, the trimmed reads are aligned with Razers3, and the HLA alleles are then typed with OptiType. For mutation identification, the trimmed WXS reads are aligned with (BWA) and further pre-processed by GATKs best practice with MarkDuplicate and base re-calibration. Mutect2 is applied to call somatic mutations, which are further filtered with a contamination filter and a panel of normal (PoN). RNAseq is also trimmed with TrimGalore, whereas the alignment and estimation of transcript expression are conducted with Kallisto. The typed HLAs, filtered mutations, and transcript expression are inputs to MuPeXI. MuPeXI has incorporated variant effect predictor (VEP) to find mutation consequences and extract possible peptides from the relevant mutations. The transcript where the mutation arrives from is found in the expression data, and the peptide-MHC binding is found using netMHCpan. Modified from Bjerregaard et al. (2017) and Hackl et al. (2016)

Firstly, the reads were trimmed, quality-checked (Felix Krueger, 2021), and aligned to a reference genome (Li and Durbin, 2009). GATK pre-processing steps were used to further filter the reads that were then used as input to MuTect2, which identifies the somatic mutations (Benjamin et al., 2019). GATK filters the identified mutations with a contamination filter and a panel of normal (PoN) to remove technical artifacts. MuPeXI has incorporated variant ef-

fect predictor (VEP) to identify mutation consequences. Regarding neopeptide selection, the NetMHCpan predicted binding affinity to HLA was used (Jurtz et al., 2017), for which the typed patient specific HLA alleles were obtained with OptiType (Szolek et al., 2014). Additionally, the expression of transcripts was found using Kallisto (Bray et al., 2016) and used to filter out non-expressed mutations.

1.5.2 Neoepitope screening

Currently, neoepitope prediction tools can successfully identify relevant neoepitope candidates out of a tumor's sequencing data. However, they still fail to identify which of the neoepitope candidates can be recognized by $CD8^+$ T cells. The $CD8^+$ T cells that recognize these immunogenic neoepitopes are referred to as neoantigen-reactive $CD8^+$ T cells (NARTs) (Schaap-Johansen et al., 2021; Schumacher and Schreiber, 2015). To distinguish true immunogenic neoepitopes from the non-immunogenic ones, mapping neoepitopes experimentally in samples from cancer patients is crucial. Bentzen and Hadrup *et al.* have developed a technology to detect immunogenic neoepitopes from a large collection of pMHC candidates (Bentzen et al., 2016), which has been used in this thesis to validate NARTs. The overall procedure of this technique is illustrated in Figure 6. First, predicted neoepitope candidates are folded with their corresponding HLA molecule, and patient specific fluorescently-tagged pMHC multimers are generated and labeled with a unique DNA barcode. These barcode-labeled pMHC multimers are combined with patients' samples, for example peripheral blood mononuclear cell (PBMC) or TILs, and any $CD8^+$ T cells from these samples that bind to a pMHC complex are sorted for sequecing of the pMHC associated barcodes. A computational tool, "Barracoda" (`https://services.healthtech.dtu.dk/service.php?Barracoda-1.8`), can process the sequencing data and with computational enrichment analysis identify which pMHC complexes give rise to NARTs.

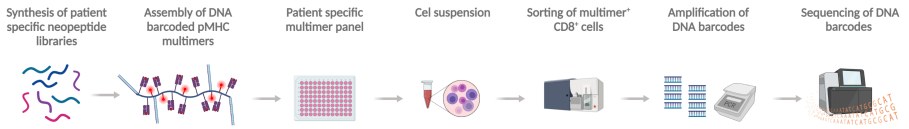


Figure 6: Neoepitope screening technique. Peptides are combined with MHC, and the pMHC multimers are constructed and labeled with a barcode. Then, the multimers are combined with the patient's samples, and $CD8^+$ T cells bound pMHCs are sorted and sent for sequencing. Modified from Kristensen et al. (2022).

Previous studies have been searching for NARTs from predicted neoepitopes, but typically only 1% to 3 % of these will give rise to NARTs (Schumacher and Schreiber, 2015). Therefore, there is a high need to identify the characteristics of immunogenic neoepitopes to further improve neoepitope prediction tools.

1.5.3 Neoepitope immunogenicity

The characterization of neoepitope immunogenicity is poorly understood, which can be caused by cancer heterogeneity and additionally by the limited amount of validated neoepitopes. The limited amount of data available challenges researchers' possibilities to define characteristics of immunogenic neoepitopes and hence the development of accurate neoepitope prediction tools. In that regard, researchers have used both cancer-derived epitopes and pathogen-derived epitopes to discover the characteristics of immunogenicity. For example in 2017, Luksza *et al.* created a fitness model representing the foreignness score of epitopes, by investigating both pathogenic epitopes and neoepitopes. They discovered that the foreignness score could be used to distinguish immunogenic neoeptiopes from the non-immunogenic ones (Łuksza et al., 2017). This score has later been validated by Richman *et al.* and they additionally found that hydrophobicity provides insight into neoepitope immunogenicity (Richman et al., 2019). Another study investigated the physical properties of the peptide core, which is the part of the peptide responsible for TCR recognition, and used machine learning (ML) to predict antigen immunogenicity. They revealed that a high fraction of hydrophobic and aromatic residues in the core gave rise to immunogenic epitopes (Schmidt et al., 2021). However, studies focusing on only cancer epitopes are based on limited data. In 2016 a study revealed that mutation clonality was important in eliciting a T cell response compared to neoepitopes with sub-clonal mutations, but only based on a few neoepitopes (McGranahan et al., 2016). Wells *et al.* studied the immunogenicity restricted to neoepitopes derived from cancers by collecting data from different studies. They studied the characteristics of 286 validated neoepitope candidates, where 29 were found to be immunogenic. They discovered that binding affinity, binding stability, tumor abundance, and hydrophobicity could be used to distinguish immunogenic from non-immunogenic neoepitopes independently. They also studied the foreignness score and the differential agretopicity index (DAI). DAI is the pMHC binding affinity ratio between the mutated and corresponding normal peptides, which has also been proposed by other studies to determine immunogenicity (Case et al., 2019; Coelho et al., 2020; Ghorani et al., 2018). Finally, they proposed a combination strategy to select immunogenic neoepitopes with high specificity, representing the complexity to define immunogenicity owing to the many possible parameters involved (Wells et al., 2020).

Altogether there is evidence that immunogenic neoepitopes share some characteristics. However, the evidence is made across small datasets and requires validation and improvement of the neoepitope prediction tools to predict neoepitopes with higher accuracy.

1.6 TUMOR ESCAPE MECHANISMS

All the characteristics mentioned in Chapter 1.5.3 can help to identify possible immunogenic neoepitopes. However, even the best predicted candidate might not raise a T cell response, because tumors can evolve to avoid immune recognition through different processes known as tumor escape mechanisms. One type of escape mechanism is limited expression of HLA molecules (Figure 7A). As described in Chapter 1.4.1, the pathway for presenting an antigen on the cell surface is complex, and several mechanisms can cause this type of escape. The down regulation of HLA can result in limited presentation of neoepitopes and hereby hide for TCR mediated recognizing. However, in that scenario NK cells will be activated because the inhibitory NK receptors (for example, killer cell immunoglobulin-like receptor (KIR)), can not bind to their HLA ligand (Poggi and Zocchi, 2006). McGranahan *et al.* discovered that the loss of heterozygosity HLA (LOHHLA) happens in 40 % of non-small-cell lung cancer patients (McGranahan et al., 2017). Additionally, a limited amount of NKs and lack of immunological memory may also prevent NK cells from eliminating the tumor cells (Lisiecka and Kostro, 2016). Another mechanism of tumor escape is loss of immunogenicity (Figure 7B). As described in Chapter 1.5.3 only a small fraction of predicted neoepitopes is immunogenic and can trigger NART responses. Immunoselection mechanisms favour a selection of non-immunogenic neoepitopes that prevent T cell recognition (Lisiecka and Kostro, 2016; Vesely and Schreiber, 2013). As stated in Chapter 1.3 an inflamed TME is essential for tumor regression, whereas a non-inflamed immunosuppressive environment lacks tumor recognition. An immunosuppressive environment can also result in immune escape. This type of escape mechanism can be caused by many environmental factors, such as expression of immunosuppressive molecules by cancer cells, e.g. programmed cell death ligand 1 (PD-L1) or programmed cell death ligand 2 (PD-L2), which prevent T cell activation, or loss of costimulatory molecules by APCs, essential for $CD8^+$ T cell activation, such as B7-1 or B7-2 (Figure 7C) (Beatty and Gladney, 2015; Lisiecka and Kostro, 2016).

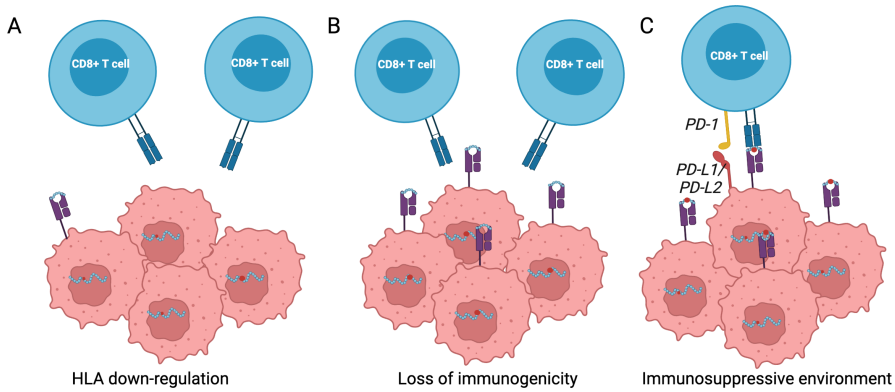


Figure 7: Tumor escape mechanisms. The tumor has the ability to escape from T cell recognition. A) One way is to down-regulate the expression of HLA molecules. B) Another escape mechanism is the loss of immunogenicity, whereas the presented neoantigen is not immunogenic and unable to raise an immune response. C) An immunosuppressive TME can also lead to tumor escape and prevent tumor recognition. Modified from Beatty and Gladney (2015).

1.7 IMMUNOTHERAPY

Cancer treatment with immunotherapy has made a paradigm shift in treating cancer patients with solid malignancies. Contrary to the classical treatment options, such as radiotherapy and chemotherapy, immunotherapy aims to boost the immune system's ability to recognize and kill cancer cells. Different immunotherapy treatment strategies have shown to benefit different patient groups (Tong et al., 2018), from which checkpoint inhibitors, adoptive cell transfer with TILs, and therapeutic cancer vaccines are briefly revised in the following sections.

1.7.1 Checkpoint inhibitors

One type of immunotherapy treatment is checkpoint inhibitors (CPI). CPIs are monoclonal antibodies that block inhibitory signaling pathways between APC and/or cancer cells towards $CD8^+$ T cells. In 2018 James P. Allison and Tasuku Honjo got awarded the Nobel prize for their discovery of the immunosuppressive molecules, Cytotoxic T-lymphocyte–associated antigen 4 (CTLA-4) and programmed cell death 1 (PD-1), respectively (Huang and Chang, 2019). CTLA-4 is a co-receptor on $CD8^+$ T cells that competes with CD28 for binding to B7-1 or B7-2. As described in chapter 1.4.2, $CD8^+$ T cell activation is dependent on

engagement of CD28 through binding of B7 expressed by APCs. However, activation of T cells will be inhibited if CTLA-4 binds B7 instead (Borghaei et al., 2009; Buchbinder and Desai, 2016). The great success of CTLA-4 therapy has led to the discovery of additional CPI molecules targeting other immunosuppressive interactions. For example, the interaction between PD-1 on T cells and PD-L1/PD-L2 on cancer cells inhibits T cell proliferation and cytokines production. To prevent this suppressive interaction, monoclonal antibodies have been developed to block PD-1 or PD-L1 (Figure 8). CPI treatments blocking the PD-1 to PD-L1 axis have been shown to be more efficient than anti-CTLA-4 (Fritz and Lenardo, 2019; Oiseth and Aziz, 2017). Therefore, CPI treatments have received great attention, and have demonstrated clinical benefit across multiple cancer types (Chowell et al., 2022; Liang et al., 2020). Another ligand for PD-1, which also suppresses T cell activation, is PD-L2. However, studies on PD-L2 are not as abundant as PD-L1 studies (Solinas et al., 2020).

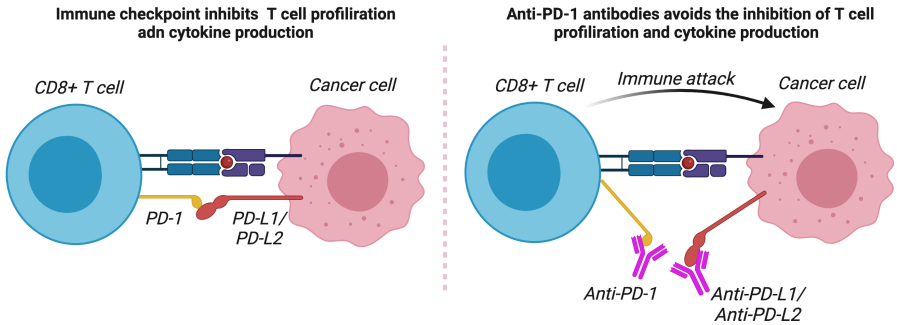


Figure 8: Immunotherapy with checkpoint inhibitors (CPI). The interaction between PD-1 and PD-L1 or PD-L2 suppresses the $CD8^+$ T cells' ability to kill cancer cells. CPI with anti-PD-1 or anti-PD-L1 blocks this interaction so that $CD8^+$ T cell can be activated and kill cancer cells. Modified from Oiseth and Aziz (2017).

1.7.2 Adoptive cell transfer with TILs

Adoptive cell transfer (ACT) with TILs (TIL-ACT) is a personalized type of immunotherapy, where TILs are expanded to great numbers from a patient tumor lesion *ex vivo* and infused back into the patient. In 1988, Rosenberg *et al.* were the first to show beneficial responses to TIL-ACT in patients with metastatic melanoma (Rosenberg et al., 1988). Since then, a range of clinical trials have proven the effectiveness of TIL-ACT, even in patients with poor prognosis and failed efficacy of CPI treatment (Besser et al., 2013; Rosenberg and Restifo, 2015; Rosenberg et al., 2011). Despite the clinical success of TIL-ACT in melanoma patients, 40-50 % of patients do not respond. Both TMB

and the amount of predicted neoepitopes (neoepitope load) have been suggested as potential biomarkers to identify the patients that will benefit from ACT treatment (Lauss et al., 2017).

1.7.3 Therapeutic cancer vaccines

Therapeutic cancer vaccines (TCVs) aim to expand patients´ endogenous T cells to kill cancer cells. By choosing vaccine targets specific for a patient´s particular tumor, TCVs can be a personalized immunotherapy strategy as ACT. However, demonstrating the efficacy of TCVs has proven challenging (Saxena et al., 2021). Multiple elements can be the reason behind the unsuccessful trials with TCVs, such as a non-effective adjuvant, an immunosuppressive TME, or lack of T cell infiltration (Shemesh et al., 2021). The availability of neoepitope prediction tools, as mentioned in Chapter 1.5.1, has made neopeptide-based vaccines a suitable immunotherapeutic option. These therapies have been proven safe and feasible as a therapeutic strategy in various phase 1 and phase 2 clinical trials (Saxena et al., 2021). However, the effectiveness of neopeptide-based vaccines is highly dependent on choosing the right target. For this, neoepitope prediction tools are needed to characterize immunogenic neoepitopes. Neopeptide-based vaccines are still new in cancer therapies, and further studies are warranted to improve the understanding of their efficacy (Blass and Ott, 2021).

1.8 BIOMARKERS

Different types of biomarkers exist, which can be classified based on their usage. Diagnostic biomarkers are used to detect diseases and provide a diagnosis. Prognostic biomarkers describe patient survival probability independent of treatment, whereas predictive biomarkers give an insight into patients' outcomes from a specific treatment setting (Khomiak et al., 2020; Winter et al., 2013). This thesis focuses on predictive biomarkers.

1.8.1 TMB and neoepitope load

Despite the successful survival improvement gained from CPI treatment, far from all patients benefit from this treatment; accordingly, there is a need to identify predictive biomarkers that can predict patients' probability of benefiting from this immunotherapy. TMB is one of the few approved predictive biomarkers and has been widely investigated across different cancer types (Chan et al., 2019; Goodman et al., 2017). Predicted neoepitopes can give more insight into which mutations are immunogenic, meaning that they are potentially shown to

the immune system and $CD8^+$ T cells. Therefore, the neoepitope load has also been considered as a potential biomarker and it has shown promising results across a diverse cohort of patients (Wells et al., 2020). However, other studies have also demonstrated the non-predictive power of using TMB and neoepitope load, as they do not work for all patients (Blank et al., 2018; Wood et al., 2020).

1.8.2 TME signatures

Besides the TMB and neoepitope load, different TME signatures have also been explored as potential biomarkers. Among others, expression of the CPI target PD-L1, have proven to work as predictive biomarkers for anti-PD-1 and anti-PD-L1 therapy (Doroshow et al., 2021; Reck et al., 2016) Yet again, they are not predictive for all patients (Burdett and Desai, 2020). As described in Chapter 1.7.1, another ligand to PD-1, although somewhat neglected, is PD-L2 and binds to PD-1 with a higher affinity than PD-L1 (Solinas et al., 2020). Latchman *et al.* observed PD-L2 as immunosuppressive, inhibits T cell profiliration, and has been correlated with pure survival probability (Latchman et al., 2001). On the contrary, Obeid *et al.* showed that high PD-L2 expression was correlated with increased $CD8^+$ T cell infiltration and was associated with better survival (Obeid et al., 2016).

Another important TME factor is TILs as they are essential to recognize neoepitopes and kill cancer cells. Studies have suggested different signatures describing TIL abundance estimated from bulk RNAseq as potential biomarkers. For example, cytolytic activity (CYT) is the geometric mean of GZMA and PRF1, also described in Chapter 1.4.2, and has been shown to predict patient outcomes across different cancer types (Rooney et al., 2015). Another described signature is T cell diversity, which can be utilized from the output obtained from a computational tool that map reads from RNAseq to TCR sequences, (Bolotin et al., 2015) and has previous been linked with positive response rate within anti-PD-1 CPI therapy (Zhao et al., 2019).

1.8.3 Combination strategies

Since many of the single biomarkers fail to predict patients' outcome alone, the better predictive capacity of combining several biomarkers have been demonstrated. One way to combine biomarkers is to calculate an immune score of some selected signatures (Cristescu et al., 2018). Another approach that allows the combination of multiple biomarkers is to apply ML. For this, large-scale data are needed to produce an ML model that can predict survival probability with high specificity and sensitivity without overfitting (Acharjee et al., 2020;

Chowell et al., 2022). Overall, the field of investigated biomarkers is enormous and has many inconsistencies, making it clear that more research within the field is required to clarify the question about which biomarkers should be considered when deciding if patients should be treated with CPI.

1.9 MACHINE LEARNING

In the past two decades, ML has increasingly been applied in cancer research, both for predicting the risk of developing cancer; and a patient's specific survival probability or response to a particular treatment. ML models can interpret multiple features and predict outcomes based on the data it is trained on. The input for these predictions can include patient specific data such as genomic, proteomic, and clinical data (Cruz and Wishart, 2006). Various ML algorithms have been used to make predictions within cancer research. Among these are the artificial neural network (ANN) which is constructed of multiple layers and nodes (Jurtz et al., 2017). However, ANNs are black box models where it is almost impossible to extract information about the feature space, and this makes it challenging to understand which features are important for the model performance. Contrary to ANN, Random Forrest (RF) gives a better understanding of the decisions made in the model and its feature space (Chen and Ishwaran, 2012; Dayhoff and DeLeo, 2001), and is the model used for this thesis.

1.9.1 Random forest model

RF models are based on a decision tree algorithm composed of sequential decisions and can be used in resolving regression and classification problems. Figure 9A illustrates a simple procedure of how a decision for a classification problem with two classes can be utilized. For each sequential decision within the algorithm, the data is split by the variables that most efficiently separate the two classes, here represented in red and blue (Somvanshi et al., 2017). The RF model consists of multiple randomly constructed decision trees, and the final output is an ensemble of all these trees using bagging (Figure 9B). The ensemble score from a RF model can be represented as a binary value predicting the class, or as a continuous value given by the probability (Ali et al., 2012).

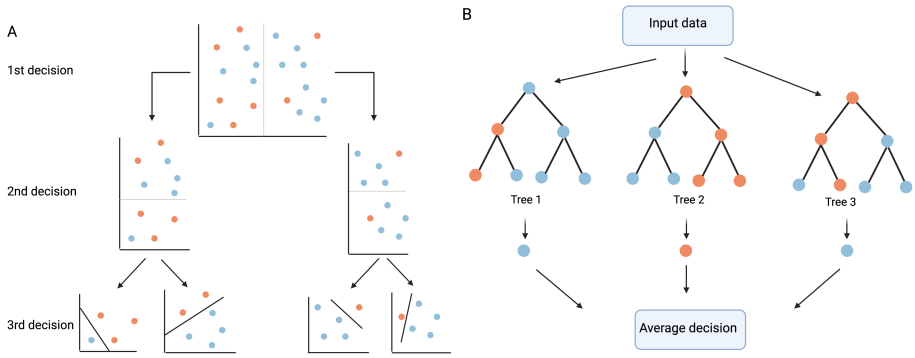


Figure 9: Decision tree and Random Forrest. A) An example of how the decision tree splits the data during multiple decision rounds. B) A simple random forest model with multiple trees predicting a binary outcome, where the final prediction from the model is an assemble of the prediction from all trees. A) Modified from Somvanshi et al. (2017) and B) Modified from Ali et al. (2012).

Features in RF models

All features applied in the RF model might not be of relevance for the predictive outcome, and the relevant ones can have different grades of importance. One of the strengths of the RF algorithm is that it can quantify the importance of each implemented feature (Gregorutti et al., 2017). There are many ways to calculate the feature importance, but the one used in this thesis is called Gini importance. The Gini importance calculates the importance of each feature as the number of times a feature is used to split a node (across all trees), weighted by the number of samples it splits (Boulesteix et al., 2012). However, calculating feature importance for highly correlated parameters can draw inconsistency according to the RF model's feature selection as it is unable to choose between correlated features and might choose differently when repeating the same model (Gregorutti et al., 2017).

Imbalanced data

Imbalanced data is a common problem in ML although it reflects most of the "real world" problems. In most cases, RF classification models can predict both balanced and imbalanced data. However, the RF has a decreased ability to predict accurately if more than 85 % of the data are belonging to one class (either positive or negative). There are different ways to overcome this imbalanced problem. One approach is to weigh the different classes differently. Another technique is to downsample the overrepresented class or oversample the underrepresented class (Perry and Bader-El-Den, 2015; Zhu et al., 2018).

The downsampling procedure, which involves sub-sampling data and using an ensemble score, has shown promising results in predicting disease risk across eight different disease categories (Khalilia et al., 2011).

1.9.2 Overfitting and redundancy

Overfitting is a common problem in ML. Overfitting is when a model performs well on the training data but fails to perform when the trained model is used to make predictions on a new dataset. This can occur when the model is tested and trained on the same or similar data. A way to overcome an overfitted model is to use K-fold cross-validation. Here K is the number of data subsets, also known as partitions. One of the most common K-fold cross-validation setups is the 5-fold cross-validation (Figure 10) (Refaeilzadeh et al., 2009). In a 5-fold cross-validation, the model is trained using data from four partitions and the model performance is then tested using data from the last partition, which was left out during training. This procedure is repeated five times, and the final prediction is obtained from the five test-sets. However, this technique can still result in an overfitted model due to data redundancy (Kohavi and Edu, 1993).

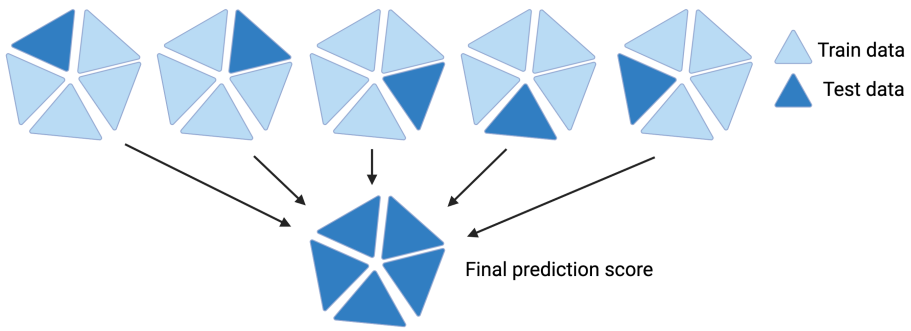


Figure 10: 5-fold cross-validation technique. Data is separated into 5 partitions, where 4/5 are used to train and the last 1/5 is used to test. This procedure is repeated 5 times, and the final output is obtained from the 5 test-sets, which cover all data points. Modified from Refaeilzadeh et al. (2009).

Redundancy occurs when data across different partitions share many similarities, and constructing the partitioning with the proper technique can help remove data redundancy. A common approach for making these partitions is using clustering algorithms to cluster similar data within the same partition and, with this, avoid overfitting (Hobohm et al., 1992; Jurtz et al., 2017). This could be,

for example, clustering peptides with similar sequence identity within the same partition to avoid overfitting.

Performance measurements

The performance of ML models can be evaluated with different measurements, mainly based on four categories: true positive (TP) indicating the positive data points which also are predicted to be positive. True negative (TN), given as the negative data points and truly predicted as negative. False positive (FP) is negative data points but predicted wrongly as positive. False negative (FN), these are the positive data points but wrongly predicted as negative. These four categories represent the content of a confusion matrix and can be used to calculate various performance measurements where sensitivity and specificity are two commonly used terms in ML. Both high sensitivity and specificity are essential for model performance. However, it is difficult to achieve a good performance in both. The calculation for sensitivity, also known as true positive rate (TPR) (Equation 1), describes how well the model finds the positives, regardless of how many FP are obtained.

$$Sensitivity\ (TPR) = \frac{TP}{TP + FN} \tag{1}$$

Contrary to the sensitivity, the specificity, also known as the true negative rate (TNR), calculates the number of TN out of all actual negatives in the dataset regardless of the FN pool (Equation 2) (Dwivedi, 2018).

$$Specificity\ (TNR) = \frac{TN}{TN + FP} \tag{2}$$

1-specificity is called the false positive rate (FPR), which, together with the sensitivity, are the calculations needed to illustrate the receiver operating characteristic (ROC) curves. The ROC curves represent, in general, how well the model performed, and from the ROC curve, it's possible to calculate the area under receiver operating characteristic curve (AUC) (Figure 11). AUC can be between 0 and 1, where 1 is a good-performing model, and 0.5 is a random prediction. AUC below 0.5 is not meaningful as it then describes the opposite of what's expected for a given model, and for that case, the target value can be inverted (Nahm, 2022). The performance at the beginning of the ROC curve focus on the high specificity. For that case, the partial AUC, AUC 10% (AUC01), can be calculated using only the area under the curve until the FPR reaches 0.1 (10%) (Figure 11) (Walter, 2005).

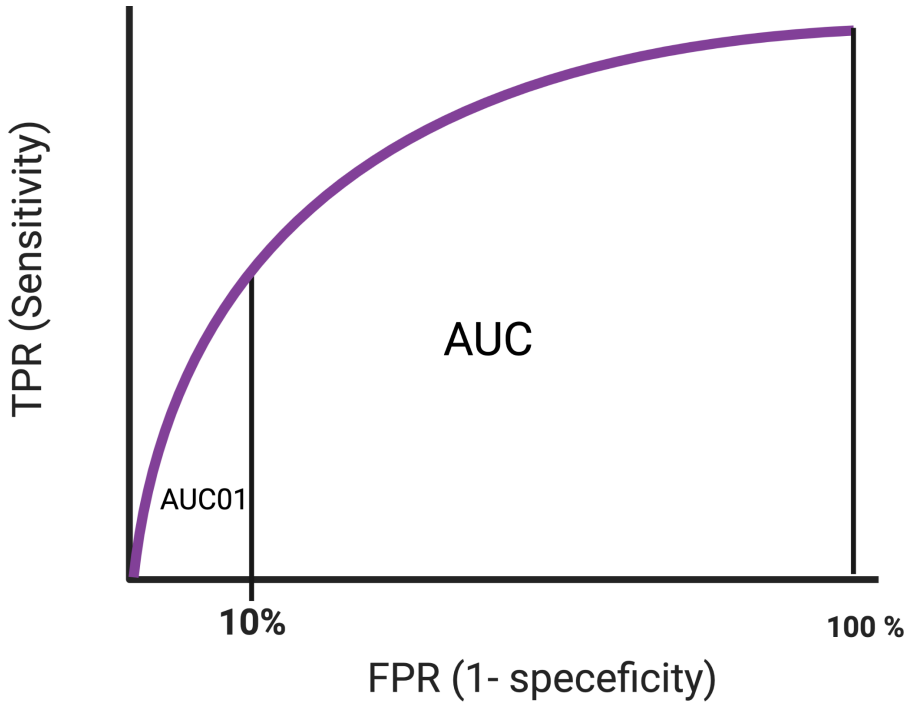


Figure 11: Area under the ROC curve. AUC is a performance measurement from 0 to 1, where 0.5 is random, and 1 describes a perfect performance. The AUC01 is calculated when FPR reaches 10% and represents the beginning of the curve, which focuses on high specificity. Modified from Walter (2005).

2 OBJECTIVES

Ground-breaking research in immunotherapy treatment has decreased the mortality rate in cancer patients. However, mortality rates are still high. Treatment options and guidance within the existing cancer treatment are still to be found. The overall aim of this PhD thesis is to identify immune signatures that can help optimize treatment strategies for cancer patients. The thesis uses bioinformatic tools to analyze omics data to explore immunue signatures and targets in human tumors driven by driven by the cytotoxic properties of $CD8^+$ T cells against cancer cells. The four included manuscripts cover novel understanding in the field of cancer immunotherapy.

Manuscript I
Neoepitope load, T cell signatures and PD-L2 as combined biomarker strategy for response to checkpoint inhibition
Immunotherapy with CPI only benefits a sub-group of patients, and it is, therefore, critical to discover biomarkers that can be used to predict patient's ability to obtain clinical benefit. This Manuscript explores potential biomarkers across different cancer types and CPI treatments. Additionally, a combination strategy of identified biomarkers is applied to discover the impact of using a combination of biomarkers rather than single biomarkers.

Manuscript II
Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma
Anti-PD-L1 CPI therapy has shown great success in Metastatic urothelial carcinoma (mUC) patients, but the outcome of these patients varies. This Manuscript investigates neoepitopes' potential to raise a $CD8^+$ T cell response prior to and during treatment with anti PD-L1. Furthermore, the Manuscript addresses the availability to use the NARTs as a biomarker to predict patient outcomes. Additionally, the TME was studied for differences in patients with high and low detected NARTs. Lastly, the neoepitope characteristics were observed to define distinctive features for immunogenic and non-immunogenic neoepitopes.

Manuscript III
Neoantigen-reactive $CD8^+$ T cells affect clinical outcome of adoptive cell therapy with tumor-infiltrating lymphocytes in melanoma
TIL-ACT in melanoma patients has shown great potential also in patients that

failed CPI treatment. This Manuscript aims to study NARTs in melanoma patients prior to and during treatment. This includes the possibility of adopting NART frequency as a predictor for patient outcomes and how the TME is related to high NART frequency. Furthermore, the Manuscript observes features that characterize of neoepitope immunogenicity.

Manuscript IV

IMPROVE, A feature model to predict neoepitope immunogenicity through broad-scale validation of T cell recognition

Neoepitope prediction tools have enabled the ability to identify patient-specific neoepitope candidates from high-throughput sequencing data. However, only a few percent of these candidates give rise to NARTs, and the limited data availability of validated neoepitopes makes it challenging to characterize immunogenic neoepitopes and improve the prediction tools. This Manuscript studies a large pool of validated neoepitopes and investigates which features can distinguish the immunogenic neoepitopes from the non-immunogenic ones. The Manuscript uses a feature-based machine learning model to improve the prediction of neoepitopes. Additionally, the Manuscript observed how patients' specific TME parameter affects the possibility of having NARTs and how implementing such features can improve the prediction of neoepitopes.

3 MANUSCRIPT I

Neoepitope load, T cell signatures and PD-L2 as combined biomarker strategy for response to checkpoint inhibition

Annie Borch*, Anne-Mette Bjerregaard*, Vinicius Araújo Barbosa de Lima*, Olga Oestrup, Christina Westmose Yde, Aron Charles Eklund, Morten Mau-Sørensen, Carolina Barra Quaglia, Inge-Marie Svane, Finn Cilius Nielsen, Samuel A. Funts, Ulrik Lassen* and Sine Reker Hadrup*.
* these author are contributed equally.
Frontiers Genetic, submitted 30 September 2022

Contribution: In this Manuscript, I have contributed to all the bioinformatic analyses and written the manuscript. The manuscript was developed in collaboration with Anne-Mette Bjeeregaard, who helped write the manuscript and provided guidance concerning some of the bioinformatic analysis. Additionally, Vinicius de lima was responsible for taking the patient samples.

NEOEPITOPE LOAD, T CELL SIGNATURES AND PD-L2 AS COMBINED BIOMARKER STRATEGY FOR RESPONSE TO CHECKPOINT INHIBITION IMMUNOTHERAPY

Annie Borch[1†], Anne-Mette Bjerregaard[1,2†], Vinicius Araújo Barbosa de Lima[3†], Olga Oestrup[4], Christina Westmose Yde[4], Aron Charles Eklund[5], Morten Mau-Sørensen[3], Carolina Barra[6,] Inge-Marie Svane[7], Finn Cilius Nielsen[4,8], Samuel A. Funts[9], Ulrik Lassen[3]*, Sine Reker Hadrup[1]*

[1] Department of Health Technology, Technical University of Denmark, Lyngby, Denmark
[2] Department of Bioinformatics and Datamining, Novo Nordisk, Denmark
[3] Department of Oncology, Phase 1 Unit, Rigshospitalet, Copenhagen, Denmark
[4] Center for Genomic Medicine, Rigshospitalet, Copenhagen, Denmark
[5] Clinical Microbiomics A/S, Copenhagen, Denmark
[6] Department of Health Technology, Section for bioinformatics, Technical University of Denmark, Lyngby, Denmark
[7]National Center for Cancer Immune Therapy, Copenhagen University Hospital, Herlev, Denmark
[8] Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
[9] Weill Cornell Medical College, New York, NY, USA

[†]these authors contributed equally
*Equal contribution and corresponding authors

Corresponding author: sirha@dtu.dk

Keywords: Tumor Mutational burden, Neoepitopes, T cell signatures, Immunotherapy, Immune Checkpoint Inhibition, Biomarker, Programmed cell death 1 ligand 2

ABSTRACT

Immune checkpoint inhibition for the treatment of cancer has provided a breakthrough in oncology, and several new checkpoint inhibition pathways are currently being investigated regarding their potential to provide additional clinical benefit. However, only a fraction of patients respond to such treatment modalities, and there is an urgent need to identify biomarkers to rationally select patients that will benefit from treatment. In this study, we explore different tumor associated characteristics for their association to favorable clinical outcome in a diverse cohort of cancer patients treated with checkpoint inhibitors. We studied 29 patients in a basket trial, comprising 12 different tumor types, treated with 10 different checkpoint inhibition regimens. Our analysis revealed that even across this diverse cohort, patients achieving clinical benefit had significantly higher neoepitope load, higher expression of T cell signatures and higher PD-L2 expression, which also correlated with improved progression-free and overall survival. Importantly, the combination of biomarkers serves as a better predictor than each of the biomarker alone. Basket trials are frequently used in modern immunotherapy trial design and here we identify a set of biomarkers of potential relevance across

multiple cancer types, allowing for selection patients that most likely will benefit from checkpoint inhibition therapy.

INTRODUCTION

Immune checkpoint inhibition (ICI) cancer immunotherapies are approved for treatment in several cancer types, and various novel combinations are being tested in a large number of clinical trials(Tang et al., 2018). Despite the success of this treatment modality, a substantial fraction of patients do not respond. Consequently, there is an urgent need to identify biomarkers that allow for the selection of patients that are most likely to benefit from ICI. Tumor mutational burden (TMB), defined as the number of non-synonymous mutations, has been demonstrated as a potential biomarker(Chan et al., 2018) also across a diverse set of cancers(Goodman et al., 2017). However, it is evident that TMB as a single parameter does not apply to all patient groups(Blank et al., 2018; Holm et al., 2022). The TMB is believed to drive the antigen recognition of tumors, and hence the predicted neoepitope load might serve as an even better parameter to determine the tumor immunogenicity. Neoepitope load consists of the number of predicted neopeptides originating from non-synonymous mutations, potentially presented by the human leucocyte antigen (HLA) class I molecules. Studies show that the clinical benefit of immunotherapy is associated with high neoepitope load across multiple cancer types (Wells et al., 2020) underlining the possibility to use this parameter as a biomarker, but not all studies agree with this assertion (Wood et al., 2020). Programmed cell death protein 1 (PD-1), Programmed death ligand 1 (PD-L1), and 2 (PD-L2) have been discovered as single biomarkers for ICI treatments that block the PD-L1/PD-L2 to PD-1 interaction, but the predictiveness of these biomarkers are complex and do not work for all patients (Latchman et al., 2001; Yearley et al., 2017; Yang et al., 2019; Burdett and Desai, 2020).

Evaluation of novel immunotherapeutic treatments for cancer is often initially pursued using a ‘basket trial’ design, as most of such therapies can potentially benefit patients with multiple different cancer diagnosis, and that features of responsiveness often appear to be similar across different cancer indications. This indeed holds true for both expression of PD-L1 (Herbst et al., 2014; Doroshow et al., 2021), T cell infiltration (Ros-Martínez et al., 2020), and TMB (Samstein et al., 2019; Sha et al., 2020). In the present study we evaluate a diverse patient group for potential genetic signatures that can be relevant for response to ICI. We utilize whole exome sequencing (WXS), RNA sequencing (RNAseq), and expression arrays from patients treated with ICI to investigate the impact of high TMB, neoepitope load, and transcriptional signatures in the tumor microenvironment (TME) on patients’ overall survival (OS) and progression-free survival (PFS). On this basis, we have identified combinations of tumor characteristics and immune signatures, that can strengthen the identification of patients with a clinical benefit following ICI.

MATERIAL AND METHODS

PATIENTS

Thirty-two patients with metastatic solid tumors referred to treatment with checkpoint inhibitors from December 2014 to February 2018 were included in the study. Patients in this cohort were, by the time of the first medical appointment, offered inclusion into the Copenhagen Prospective Personalized Oncology (COPPO) study at the Phase 1 Unit - Rigshospitalet, Copenhagen, Denmark (Tuxen et al., 2014, 2018), where prior to treatment initiation, patients are informed about the possibility of comprehensive genomic analysis, i.e. whole-exome sequencing and RNA sequencing of their tumors. This program is a feasibility study in phase I setting for patients with solid tumors and exhausted treatment options.

Inclusion criteria for immune therapy were defined by the protocols with available slots (EUDRACT number:2013-002844-10, 2014-002835-32, 2014-002605-38, 2014-000948-14, 2015-003771-30, and 2017-001147-13) and for two patients (pt no. 10 and 20) treatment off-label was given based on high mutational burden (> 1000 non-synonymous mutations). Response to treatment was assessed according to response evaluation criteria in solid tumors 1.1 (RECIST) criteria. The obtained best RECIST criteria (and lasting for at least 2 months) were used for this study's analysis.

Fresh tumor biopsies were obtained before initiation of therapy. In one case, a biopsy was obtained after treatment with ICI and for another patient, the WXS from blood was not available. Therefore, these patients were excluded from analyses. Demographic data for the remaining 29 patients can be seen in Supplementary Table S1. It should be noted that the RNAseq analysis for patient no. 19 did not succeed, but the patient is still included in the data analysis using the microarray data.

MOLECULAR ANALYSIS OF TISSUE BIOPSIES

Biopsies stored in RNAlater (Sigma-Aldrich) were used for comprehensive molecular profiling. Briefly, DNA and RNA were isolated using AllPrep DNA/RNA kit (Qiagen). Blood samples were collected in EDTA tubes, and genomic DNA was extracted using a Tecan automation workstation (Promega). Molecular profiling consists of whole-exome sequencing (Illumina platform) and mRNA expression arrays (Human U133 Plus2.0, Affymetrix).

DNA libraries were prepared from 200 ng of DNA. Fragmentation was done on Covaris S2 (Agilent) to approximately 300-bp fragments, and adaptor ligation was done using KAPA HTP Library Preparation Kit. Exomes were enriched with SureSelectXT Clinical Research Exome kit (Agilent). Sequencing was carried as paired-end sequencing, aiming at an average coverage of 50–100x using the HiSeq2500 and NextSeq500 platforms from Illumina. RNAseq libraries were prepared from 100 ng of total RNA using the Total RNA-Seq library Prep Kit (Illumina). Sequencing was done on the HiSeq2500 and NextSeq500 platforms.

Purified RNA was immediately analyzed on arrays. RNA was reverse transcribed and used for cRNA synthesis, labeling, and hybridization with GeneChip® *Human Genome U133 Plus 2.0* Array (Affymetrix) according to the manufacturer's protocol. The arrays were washed and stained with phycoerythrin conjugated streptavidin using the Affymetrix Fluidics Station 450, and the arrays were scanned in the Affymetrix GeneArray 3000 7G scanner to generate fluorescent images.

NEXT GENERATION SEQUENCING DATA ANALYSIS

WXS and RNAseq data were processed according to the Genome analysis tool kit best practice guidelines for somatic variant calling (Van der Auwera et al., 2013). Raw reads from both were quality trimmed using the wrapper tool Trim Galore 0.4.0 (Felix Krueger, 2021) combining Cutadapt (Martin, 2011) and FastQC(Andrews, 2010) trimming reads to an average phred score of 20 and a minimum length of 50 bp. Reads were aligned to the human genome (GRCh38) using the Burrows-Wheeler Aligner (Li and Durbin, 2009) version 0.7.16a with default mem options and with a read-group provided for each sample, thereby ensuring compatibility with the following steps. Reads were sorted using Samtools 1.6 (Li et al., 2009). Duplicated reads were marked using Picard-tools version 2.9.1 MarkDuplicates. To reduce false positive variant calls, base recalibration was performed with GATK version 4.0.1.1. SNV and indel calls were made using GATKs build-inn version of MuTect2(Cibulskis et al., 2013) designed to call somatic variants for both single nucleotide variants (SNVs) and indels from matched tumor and normal samples. HLA alleles of each patient were inferred from the WXS data using OptiType 1.2(Szolek et al., 2014) with default settings after filtering the reads aligning to the HLA region with RazerS version 3.4.0(Weese et al., 2012) Kallisto 0.42.1 (Bray et al., 2016) was used to determine the gene expression from RNAseq data.

DIFFERENTIAL EXPRESSION ANALYSIS, AND GENE SET ENRICHMENT ANALYSIS

Raw microarray data was imported into R and normalized by the Robust Multi-array Average (RMA) algorithm. The 'hgu133plus2.db' package version 4.1.0 was used to translate between probe set IDs and Human Gene Organization (HUGO) gene names. The 'limma' package (Smyth, 2005) (version 3.5.3) was used to test for differential expression between groups. P values were adjusted using the method of Benjamini and Hochberg (BH). The package 'ComplexHeatmap' version 2.13.1 was used to create heat maps (Yu et al., 2012; Gu et al., 2016) of the differential expressed genes with adjusted p-value < 0.05 and log foldchange > 0.5 and log foldchange < -0.5. Gene Set Enrichment Analysis (GSEA) is made from the differential expression analysis results in R with cluserProfiler (Yu et al., 2012) version 4.0.5 and enrichplot (Yu, 2021) version 1.13.2 with Gene Ontology pathway database.

ASSESSMENT OF TMB AND NEOEPITOPE LOAD

The total tumor mutational burden of all mutations acquired in each tumor was assessed by counting each entry passing the filtering criteria of GATK4's MuTect2 output VCF file. This VCF file was given as input to the neoepitope predictor, mutant peptide extractor and informer (MuPeXI) 1.2.0 (Bjerregaard et al., 2017) together with RNAseq expression values obtained from Kallisto in transcripts per million (TPM) and the HLA alleles detected by OptiType. The output neopeptides were selected for their potential to be neoepitopes by selecting peptides originating from genes with an expression above 0.1 TPM and a predicted binding eluted ligand percentile rank (EL %Rank) score < 2, evaluated by NetMCHpan 4.0 (Jurtz et al., 2017). The number of selected potential neoepitopes was used as the neoepitope load. Additionally, TMB of non-synonymous mutations were determined from the MuPeXI output logfile summarizing peptides originating from missense variant mutations, in-frame insertions, and deletions, together with frameshift mutations. Mutation types were

determined by Ensembl's variant effect predictor (VEP) version 87(McLaren et al., 2016) as a dependency of MuPeXI.

Determination of T cell diversity by CDR3 sequence identification from RNAseq

MiXCR(Bolotin et al., 2015) version 2.1.1 was used to determine complementarity-determining region 3 (CDR3) sequences from bulk RNAseq data with the optimized setting for this specific purpose (Brown et al., 2016). The quality trimmed reads from RNAseq were used as input following MiXCR's identification of specific clone identification from the IMGT database (Smyth, 2005) reference of known CDR3 sequences, together with the clone count of each clone detected referring the reads aligning to this specific clone of the CDR3 reference library. Shannon entropy(Shannon, 1948) was calculated as a T cell diversity measurement (Stewart et al., 1997).

Statistical analysis

Since the data analyzed in this study is not normally distributed, an unpaired Mann-Whitney /Wilcoxon rank-sum test was used, with normal approximation using continuity correction of the calculated p-value to calculate and determine statistically significant differences between groups using R statistical software version 4.1.1.

Validation Cohort

The validation cohort consists of 24 Metastatic Urothelial carcinoma (mUC) patients (Holm et al., 2022). WXS and RNAseq are preprocessed in the same manner as the sequencing data from this study and a detailed description can be found in (Holm et al., 2020). Shortly, variants are called with GATK 3.8, and TPM from each gene was found by Kallisto alignment.

RESULTS

PATIENT COHORT AND CLINICAL OUTCOME

We investigated different biomarkers for their potential to identify patients with a favorable clinical outcome in a diverse cohort of 29 cancer patients with 12 different tumor indications treated with 10 different ICI combinations. All treatments included blocking of the PD-1 to PD-L1/PD-L2 interaction (Table 1).

Table 1. Overview of diagnoses, given treatment and response pattern.

ID	Diagnosis	Treatment	RECIST (best obtained)	Biopsy site	Number of prior treatments
1	PAAD	Atezolizumab + CEA-IL2	PD	Liver	2
2	BRCA	Atezolizumab	PD	Liver	6
3	COAD	Atezolizumab + CEA-IL2	PD	Lung	3
4	READ	Atezolizumab + CEA-IL2	PD	Liver	3
5	CESC	Atezolizumab + CD40	PD	Lymph node	6
6	PAAD	Atezolizumab + CEA-IL2	PD	Lung	2
7	CCA-IG	Ipilimumab + Nivolumab	PD	Peritoneum	4
8	SKCN	Prembolizumab	PD	Liver	3
9	BRCA	Prembolizumab	PD	Lymph node	6
10	COAD	Nivolumab	PD	Primary tumor	2
11	BLCA	Prembolizumab	PD	Peritoneum	2
12	BRCA	Atezolizumab + CD40	PD	Liver	6
13	UC-U	Prembolizumab	SD	Liver	2
14	CESC	Prembolizumab	PR	Lymph node	4
15	CDC-K	Nivolumab	PR	Kidney	1
16	BRCA	Atezolizumab	SD	Subcutaneous / Cutaneous	7
17	BLCA	Atezolizumab + CD40	SD	Other	2
18	BLCA	Ipilimumab + Nivolumab	PR	Lymphnode	2
19	BLCA	Ipilimumab + Nivolumab	CR	Lymphnode	2
20	COAD	Prembolizumab	PD	Subcutaneous / Cutaneous	3
21	LIHC	Nivolumab + LAG3	CR	Liver	2
22	BLCA	Prembolizumab	PR	Lymph node	1
23	READ	Atezolizumab + CD3 (WP)	PD	Liver	3
24	OV	Atezolizumab + BET inhibitor	PD	Subcutaneous	6
25	BRCA	Chemo + Prembolizumab	PR	Lymph node	3
26	BRCA	Chemo + Prembolizumab	PD	Lymph node	1
27	COAD	Atezolizumab + CD3 (WP)	PD	Liver	2
28	OV	Atezolizumab + BET inhibitor	SD	Lymph node	3
29	READ	Atezolizumab + CD40	PD	Liver	4

BLCA; Bladder Urothelial Carcinoma, BRCA; Breast invasive carcinoma, CCA-IG; clear cell adenocarcinoma - intern genitalia, CDC-K; collecting duct carcinoma – kidney, CESC; Cervical squamous cell carcinoma and endocervical adenocarcinoma, COAD; Colon adenocarcinoma, LIHC; Liver hepatocellular carcinoma, OV; Ovarian serous cystadenocarcinoma, PAAD; Pancreatic adenocarcinoma, READ; Rectum adenocarcinoma, SKCN; Skin Cutaneous Melanoma, UC-U; Urothelial carcinoma – Urethra.

In total, 15 patients were treated with PD-1 ICI (Prembolizumab, Ipilimumab and/or Nivolumab) and 14 patients with PD-L1 ICI (Atezolizumab), often in combination with other treatment modalities. Tumor biopsies were primarily taken from metastatic sites, including lymph node and liver. DNA was purified from blood (germline) and tumor to determine the tumor specific somatic mutations through WXS. RNA was sequenced from the tumor sample to determine the transcriptional landscape, including the expression of genes carrying neoepitopes and immune-relevant genes. Furthermore, targeted microarrays were performed on the tumor material and analyzed for a panel of known cancer related genes.

Response to therapy was reported by the response evaluation criteria in solid tumors (RECIST) as best obtained response. The patient cohort included two complete responders (CR), five patients with partial response (PR), four with stable disease (SD), and 18 with progressive disease (PD). In this study, we investigate the differences between patients with PD and non-progressive disease (CR, PR, SD). The cohort included twenty deceased patients due to progressive disease, while ten patients were still under observation by end of the analysis (five of these with no signs of disease progression) (Figure 1A). All patients had previously been treated with at least one and up to seven prior treatment lines, none of which were immunotherapies (Supplementary Table 1). It is noteworthy that two patients obtain CR, and both received a combination of checkpoint inhibition treatments, nivolumab and ipilimumab, and nivolumab and LAG-3, respectively. When looking into the different checkpoint inhibition combinations used and the cancer type, no obvious difference in clinical response related to treatment regimen nor cancer type was observed (Figure 1B). It should be noted, however, that this study is not intended to identify the difference between tumor indications, since each tumor type is scarcely represented.

NEOEPITOPE LOAD IS ASSOCIATED WITH NON-PROGRESSIVE DISEASE

First, we investigated biomarkers known to influence overall survival in uniform cohorts of one cancer type i.e. TMB and neoepitope load (i.e. the number of tumor-specific peptides originating from somatic mutations and predicted to bind to patient-specific HLA, and where the gene of origin is expressed). Based on the analysis of the sequencing data from the 29 patients, when observing the predicted neoepitope load, extracted using the MuPeXI pipeline (Bjerregaard et al., 2017), sorted upon neoepitope load, patients do cluster based on progressive compared to non-progressive disease (Figure 1B). Interestingly, the two patients with a very high neoepitope load (both colon cancer with microsatellite instability (MSI) were found in the PD patient group (patient no. 10 and 20). No significant difference between patients with progressive disease vs. non-progressive disease was observed in the TMB for all mutations identified (Figure 1C), but we did observe a significant difference when evaluating TMB only for non-synonymous mutations (Figure 1D), provided a more significant difference between the two patient groups (Figure 1E).

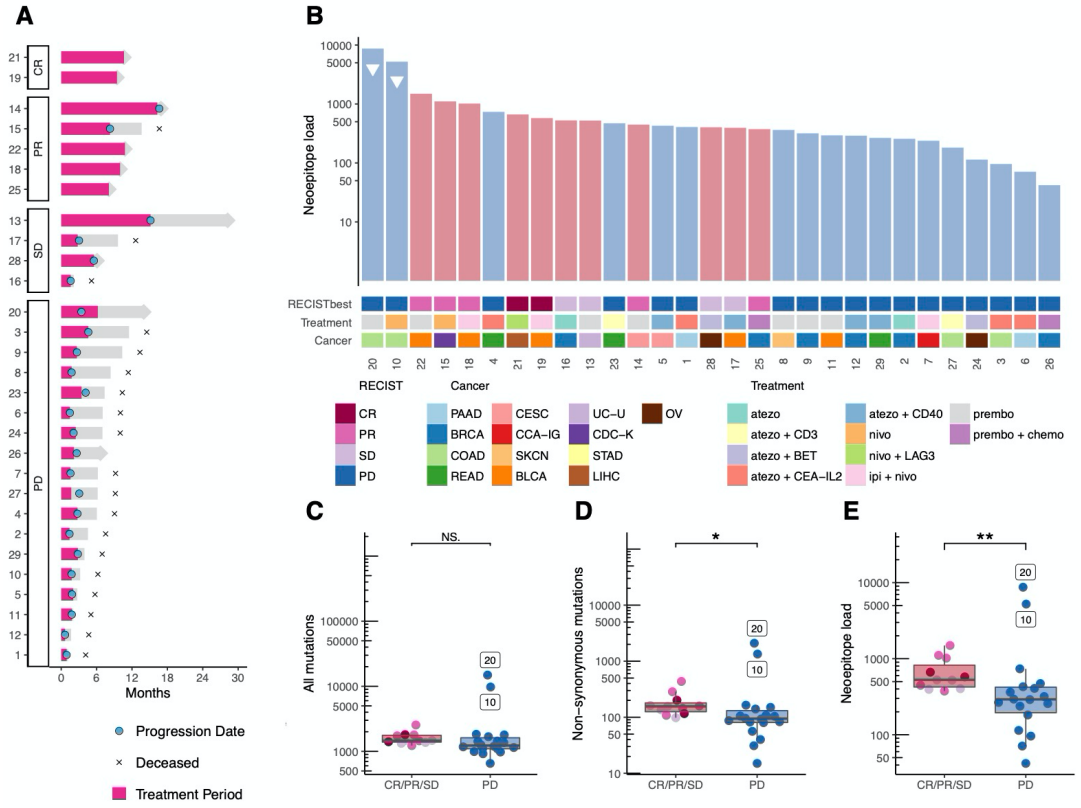


Figure 1. Treatment trajectory, Tumor mutational bourdon and neoepitope load. **A)** The treatment trajectory for each patient is plotted according to response, with annotations of key dates, treatment period is highlighted in pink. **B)** Each patient is represented in a barplot of predicted neoepitope load with annotations according to best obtained RECIST criteria, cancer type, and immunotherapy treatment combination. Patients with high microsatellite instability are annotated with a triangle. BLCA; Bladder Urothelial Carcinoma, BRCA; Breast invasive carcinoma, CCA-IG; clear cell adenocarcinoma - intern genitalia, CDC-K; collecting duct carcinoma – kidney, CESC; Cervical squamous cell carcinoma and endocervical adenocarcinoma, COAD; Colon adenocarcinoma, LIHC; Liver hepatocellular carcinoma, OV; Ovarian serous cystadenocarcinoma, PAAD; Pancreatic adenocarcinoma, READ; Rectum adenocarcinoma, SKCN; Skin Cutaneous Melanoma, STAD; Stomach adenocarcinoma, UC-U; Urothelial carcinoma – Urethra. **C-E)** The mutation and neoepitope load are colored according to best obtained RECIST criteria and grouped by progressive vs non-progressive disease statistic test are med with Wilcox test. **C)** Total mutation burden (p-value = 0.069, Wilcoxon rank sum test). **D)** Non-synonymous mutations (p-value = 0.012, Wilcoxon rank sum test) and **E)** number of predicted neopeptides - referred to as neoepitope load (p-value = 0.0086).

SELECTED T CELL SIGNATURES IDENTIFIES PATIENTS WITH TREATMENT BENEFIT

To investigate whether distinct gene signatures differentiate the patients with progressive disease vs. non-progressive disease, we performed a differential expression analysis of a 770-pan-cancer-immune-related-gene-panel from both the expression array data and the RNAseq data. Due to higher sensitivity, we display the analysis of the microarray data in the main figures.

The differential expression analysis revealed that 322 microarray probes, (Figure 2A) were differentially enriched between the two groups. Following double cluster analysis of the mean of these enriched probes, they condense into 188 genes. This gene signature tends to cluster according to disease outcome (RECIST), with a particular clustering of the patients with non-progressive disease vs. progressive disease (Figure 2B). We note that the two MSI colorectal cancer patients with the highest neoepitope load are clustered together with the other progressive disease patients according to the gene enrichment signature from the TME, indicating that an unfavorable tumor microenvironment, may override the role of the high TMB in promoting tumor foreignness (Figure 2B). To identify biological pathways of interest, a gene set enrichment analysis (GSEA) was performed, revealing lymphocyte differentiation (Figure 2C) and more specifically, T cell differentiation pathways (Figure 2D) to be significantly enriched in the group of patients with non-progressive disease.

To investigate potential biomarkers from the tumor microenvironment (TME) that could be used to identify patients that would benefit from treatment, we investigated the intratumoral T cell presence and associated factors. Based on two different probe sets we find PD-L2 (gene synonym PDCD1L2), highlighted in Figure 2A, to be preferentially expressed in the TME from patients with non-progressive disease. When evaluating the expression level of PD-L2 in the individual patients, we observed a significantly higher expression in patients with non-progressive disease compared to progressive disease patients (Figure 2E) and confirmed in RNAseq (Supplemental Figure 1A). The ICI treatment with anti-PD-Ll blocks the interaction between PD-1 and PD-L1 but have no direct effect on the binding of PD-L2 to PD-1. Hence, we re-evaluated the PD-L2 expression data by splitting patients into two groups depending on the treatment received, anti-PD-1 or anti-PD-L1 respectively. The number of subjects is low as a consequence of this split, and the PD-L1 treated cohort includes only SD patients in the ‘non-progressor’ group. Albeit the data suggest that the predictive effect of PD-L2 expression is stronger in the anti-PD-1 treated patients (Figure 2F).

Cytolytic activity (CYT) of T cells, calculated as the geometric mean of the gene expression of granzyme A (GZMA) and perforin (PRF1) has previously been described as a biomarker for response to immunotherapy (Rooney et al., 2015). We investigated CYT as a biomarker in this cohort and found that CYT has a significantly higher expression in patients with non-progressive disease (Figure 2G) further confirmed in the RNAseq (Supplemental Figure 1B). Further, we examined the T cell infiltrate in the sense of detectable T cell receptor complementarity-determining region 3 (CDR3) sequences from the bulk RNAseq data, and we find that the patients with non-progressive disease have a significantly higher T cell diversity compared to patients with progressive disease (Figure 2H). T cell diversity is defined as the Shannon Entropy, calculated based on the number of unique CDR3 sequences, or T cell clones, detected in the individual patient (Shannon, 1948; Stewart et al., 1997). The T cell diversity correlates with CYT (person correlation = 0.7) as both strategies quantify the T cell infiltrate (Supplementary Figure 1C). We found that both neoepitope load, PD-L2 expression, CYT, and T cell diversity can be used as potential biomarkers to distinguish non-progressive disease patients from progressive disease patients. By combining predicted neoepitopes with PD-L2 expression non-progressive disease patients are clustered in the high-high area split by the median of each value (Supplementary Figure 1D). The same pattern can be observed with CYT

and PD-L2, as well as neoepitope load and CYT (Supplementary Figure 1E+F). Consequently, these may be interesting features for a combined biomarker.

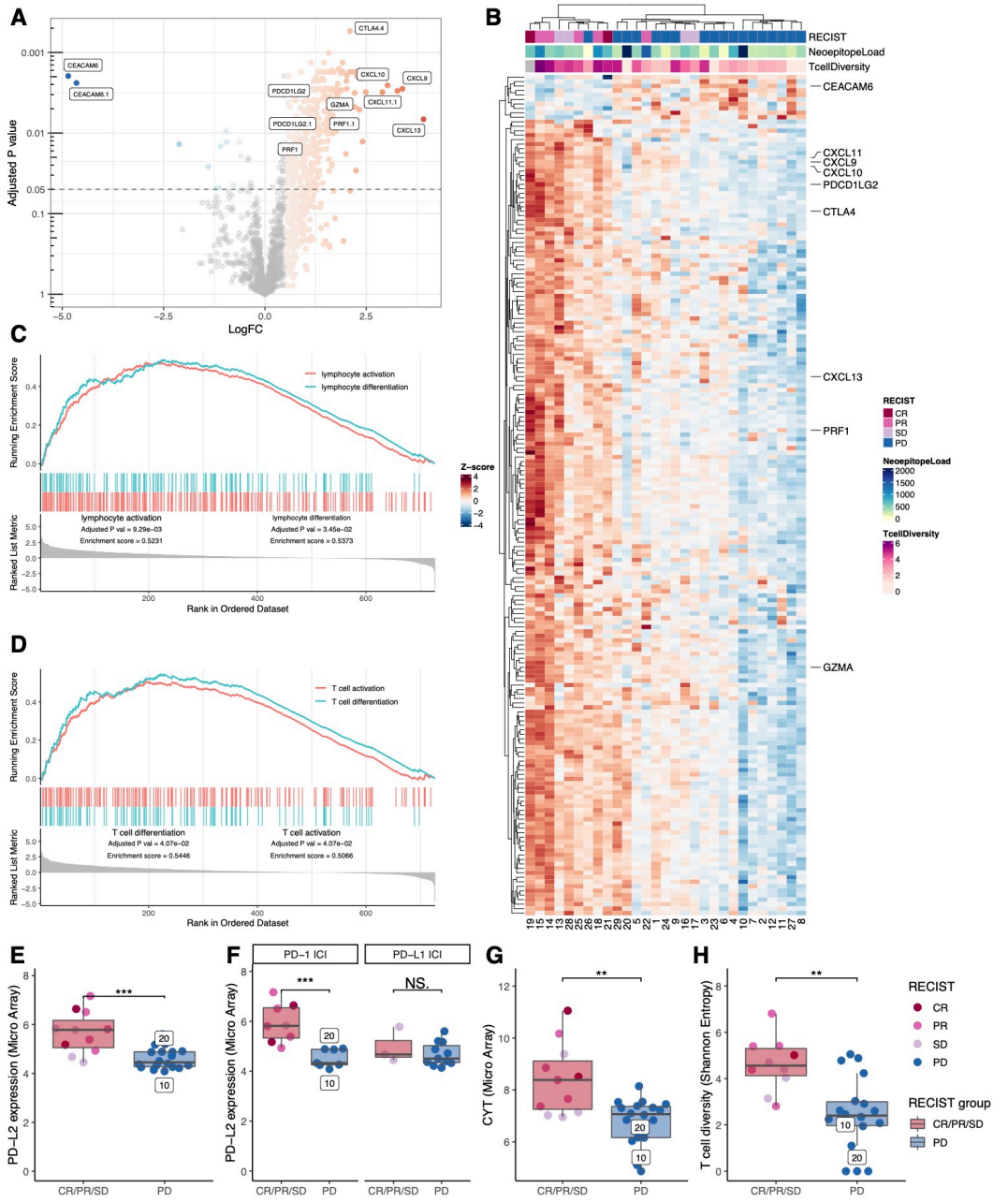


Figure 2. Immunological transcriptional profiling. Differential gene expression analysis between patients with progressive disease vs non-progressive disease displayed as a volcano plot. **A)** Shoving all probes extracted from the 770 gene PanCancer Immune Profiling Panel colored according to log foldchange (LogFC) for probes with an adjusted p-

value below 0.05 and LogFC above 0.5 and below minus 0.5. Probes with the highest variety in LogFC and lowest p-values are highlighted together with probes for PRF1, GZMA, and PDCDLG2. **B)** The mean expression of the significant probes was gathered to reveal 188 significantly differentially expressed genes with z-scores displayed in a double clustered heatmap annotated with best-obtained RECIST criteria, neoepitope load, and T cell diversity (Shannon Entropy). **C+D)** Significantly gene ontology (GO) pathways enriched in the non-progressive group from the gene set enrichment analysis. **C)** Lymphocyte activation and lymphocyte differentiation. **D)** T cell activation and T cell differentiation. **E-H)** Comparing non-progressive with progressive disease patients and the statistic test are made with Wilcoxon rank sum test. **E)** Expression of PD-L2 was found to be significantly higher in patients with the non-progressive disease (p-value 0.00069). **F)** PD-L2 expression was separated into patients receiving PD-1 immune checkpoint inhibitors (ICI) and those who have received PD-L1 ICI where a significant difference in PD-L2 expression can be found for those patients treated with PD-1 ICI (p-value = 0.00031) and not for those who have been treated with PD-L1 ICI (p-value = 0.46). **G)** The same was found for the cytolytic value (CYT), measured as the geometric mean of granzyme A and perforin (p-value = 0.0032). **H)** T cell infiltration analysis identifying CDR3 sequences from bulk tumor RNAseq data showed that the patients benefiting from treatment had a significantly higher T cell diversity measured by the Shannon Entropy (p-value = 0.0017).

COMBINED BIOMARKERS IMPROVE SURVIVAL PROBABILITY

To examine the probability of four different suggested biomarkers, Neoepitope load (NeoLoad), PD-L2 expression (PDL2), CYT, and T cell diversity (Tdiv), and combinations hereof to identify patients with a favorable clinical outcome, we applied cox regression to analyze the association with OS and PFS. We analyzed all biomarkers individually and all possible combinations of the biomarkers. For the analysis, two groups ('high' and 'low') were established for each biomarker split by their median value. For a patient to be included in the "high" category that patient must have values above the median for all biomarkers within the combination while the remaining patients are then placed in the "low" category. Additionally, we included combinations of whether patients had "high" in three or more and two or more of any of the four investigated biomarkers compared to the remaining patients. Hazard ratios with corresponding confidence intervals were calculated for each biomarker and all combinations (Figure 3A). From this analysis, we found that the best combination of biomarkers according to the hazard ratio from PFS was obtained when CYT, NeoLoad and PDL2 were combined. The three best biomarker combinations were illustrated using Kaplan-Meier curves (Figure 3B). All three combinations can significantly separate favorable from unfavorable patient outcome, based on PFS and OS, respectively.

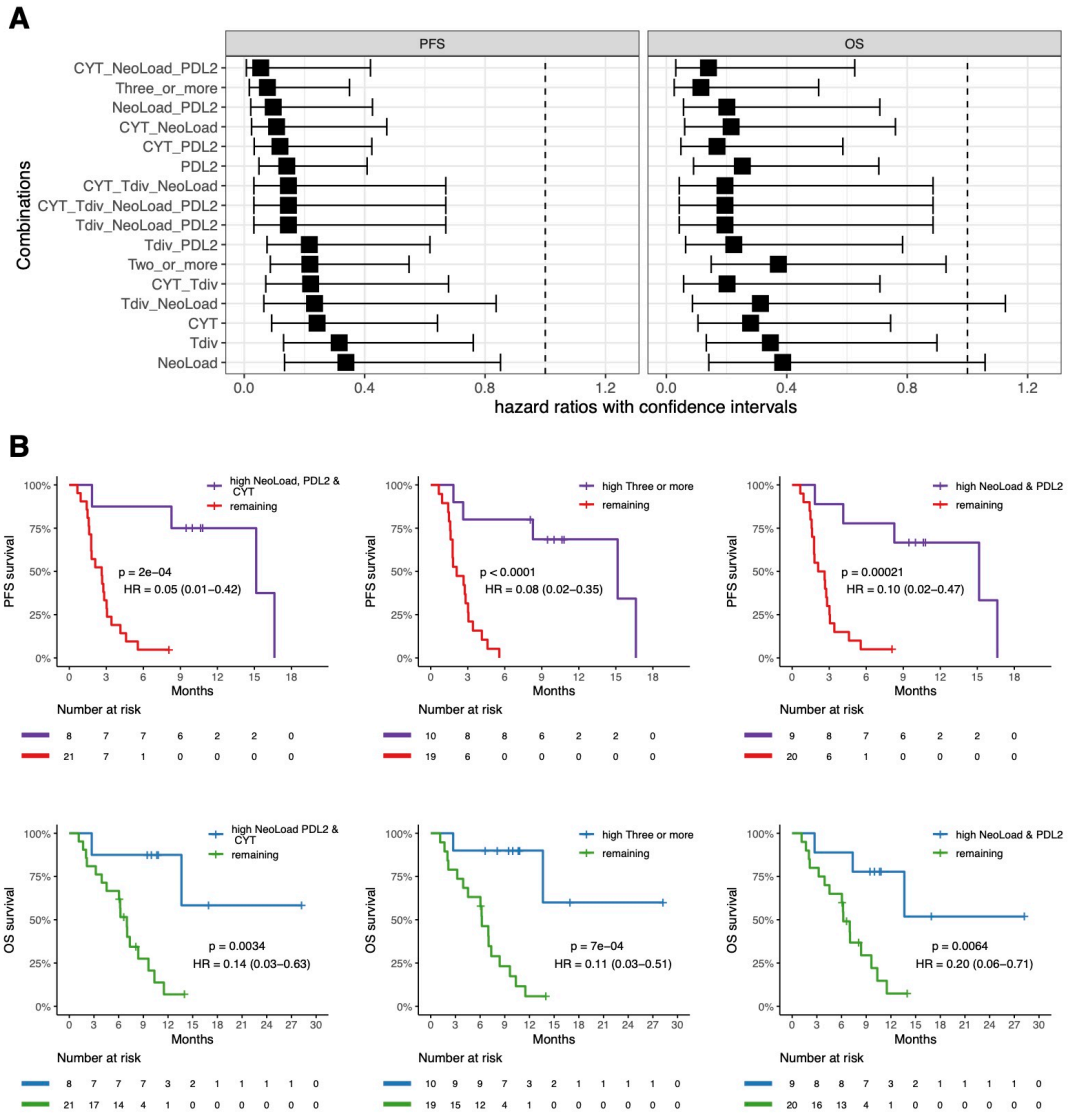


Figure 3. Survival analysis. **A)** Cox-regression analysis for four suggested biomarkers PD-L2 expression (PDL2), Cytolytic Activity (CYT), T cell diversity measured by the Shannon Entropy (Tdiv), and Neoepitope load (NeoLoad), and all combinations of these four biomarkers where "high" were defined by values above the median. On the x-axis, the square indicates the hazard ratio (HR) and the error bars indicate the confidence interval. Left; the analysis for Overall Survival (OS) and right; the analysis for progression-free survival (PFS). **B)** Kaplan-Meier curves for the top three combinations obtained from the analysis made from the cox-regression analysis. Left; a combination of high NeoLoad, high PD-L2 expression, and high CYT showed significantly improved survival probability for both PFS (p=0.0002) and OS (p=0.0034). Middel; patients with three or more high signatures also had increased survival probability both for PFS (p<0.0001) and OS ($p=7 \cdot 10^{-4}$), right; patients with high NeoLoad, and PDL2, respectively, also had increased survival probability both for PFS (0.00021) and OS (p=0.0064).

VALIDATION COHORT

WXS and RNAseq from a cohort of 24 Metastatic Urothelial carcinoma (mUC) patients all treated with anti-PD-L1 ICI (Holm et al., 2022) were used to validate the investigated biomarkers. The first combination with NeoLoad, CYT, and PDL2 was significant in identifying patients with longer PFS and OS (Figure 4A, left). The combination of any three or more biomarkers categorized as high also showed a significant difference in PFS and borderline non-significant separation in OS (Figure 4A, middle). But by combining only PDL2 and NeoLoad a most significant separation was obtained related to both OS and PFS (Figure 4A, right). Neoepitope load, CYT, T cell diversity, and PD-L2 were also individually investigated for their predictive value in the validation cohort, but no significance was observed based on the single parameters (Supplementary Figure 2). Cox-regression analyses were conducted for the validation cohort, as in the primary cohort, using all suggested biomarkers and all combinations (Figure 4B). Again, this demonstrated that NeoLoad and PDL2 was the best combination to predict patient's outcome, related to both PFS and OS.

Overall, both the primary cohort and the validation cohort agree that a combination of biomarkers was better at predicting survival than a single biomarker alone. Summarizing the two cohorts, patients with high neoepitope load and high PD-L2 expression and the same combination including high CYT resulted in a significantly improved survival probability.

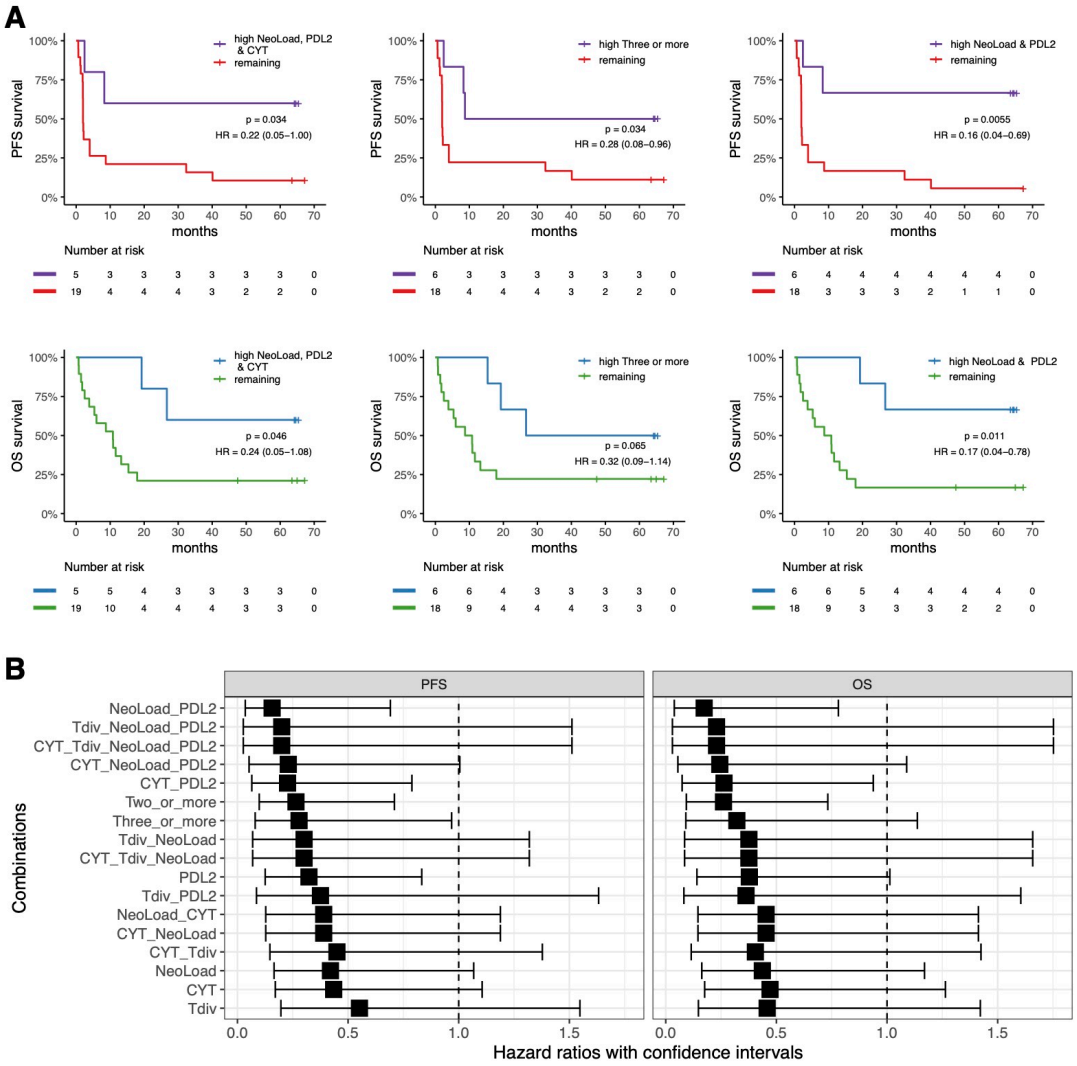


Figure 4. Validation cohort. Cox-regression and survival analysis from the validation cohort. **A)** Kaplan-Meier curves for best three combinations found in the cohort from the basket trail with progression-free survival (PFS) in the top and overall survival at the bottom where high indicates values above median for the cohort. Left; comparing patients with high neoepitope load (NeoLoad), high cytolytic activity (CYT), and high PD-L2 expression (PDL2) to the remaining patients which showed that patients with high NeoLoad, CYT, and PDL2 had a higher PFS and OS probability (p-value = 0.035 and 0.046). Middle; three or more signatures as high compared to the remaining where PFS is significant in PFS (p=0.034) and borderline non-significant in OS (p=0.065). Right; patients with high NeoLoad and PDL2, respectively, compared to the remaining patients resulted in a significantly higher PFS probability for the patients in the "high" group (p= 0.0055) and a significantly higher OS probability (p-value = 0.011). **B)** Cox-regression analysis of the four different suggested biomarkers including PD-L2 expression (PDL2), Cytolytic Activity (CYT), T cell diversity measured by the Shannon Entropy (Tdiv), and Neoepitope load (NeoLoad) and all their combinations illustrated for overall survival (OS) and Progression-free survival (PFS). Black squares indicate the hazard ratio whereas the error bars indicate the confidence interval.

DISCUSSION

Despite the recent success of immunotherapy, the objective response rate rarely reaches >50% (Yarchoan et al., 2017; Chowell et al., 2022). Hence, there is a need to segregate patients likely to respond and understand the biological basis of treatment success and failure.

In this study, we explore the impact of neoepitope load, PD-L2 expression, cytolytic transcriptional signature (CYT), and T cell diversity as biomarkers for predicting the outcome of ICI treatment. Our data suggest that individual parameters can serve as biomarkers to distinguish progressive from non-progressive patients, but a combination of biomarkers is substantially stronger in selecting patients with clinical benefit and can better bridge the heterogeneity observed within patient cohorts and across cohorts. As such, the combination of NeoLoad and PD-L2 expression could significantly identify patients with clinical benefit in both our primary and validation cohorts. Additionally, the selected biomarkers provided a significant separation of patients, measured by their progression free survival and overall survival, which may relate to both a predictive and prognostic value of such immune signatures.

The study was conducted as so-called 'basket' trials, where patients with numerous different cancer types are subjected to the same clinical strategy and suggests that such markers can be used for patient stratification across different tumor types and ICI treatments. This practice relates to the increased understanding that the characteristics associated with clinical responses to immunotherapy, are often tumor-type agnostic, and is defined by the immune and inflammatory signature, and foreignness of both the tumor cells and the TME. Such favorable characteristics can be more or less prevalent in different tumor types, leading to different response rates to ICI, but the mechanism on interaction and influence on treatment is most often tumor agnostic (Taube et al., 2014; Chowell et al., 2022).

Our data demonstrate that combining biomarkers is more robust than using a single biomarker and analysis from the validation cohort supports those findings. As an illustrative case, patient no. 26, initially categorized as a partial responder progressed quickly after the first treatment (therefore not meeting the requirement of sustained a RECIST response for at least 2 months after treatment initiation). This patient turned out to have a gene expression profile corresponding to non-progressors (Figure 2B) but displayed the lowest detected neoepitope load of this cohort. This example suggests that not only is the right immunologic gene expression profile of the tumor tissue important but also a sufficient neoepitope load is needed, most likely for the tumor cells to be 'visible' to the immune system. On the other hand, a high neoepitope load alone is not sufficient, as seen for the non-responding MSI patients no. 10 and 20, who, despite having the highest neoepitope load of the cohort (Figure 1D), displayed an immunological gene signature comparable to the progressive disease patients (Figure 2B). Interestingly, previous data has demonstrated high response rates to checkpoint inhibition therapy in patients with MSI tumors (Le et al., 2015), and based on such data checkpoint inhibition therapy is approved for the treatment of all MSI cancers, despite origin (Lemery et al., 2017). The two MSI-high cases included in our cohort did not respond to therapy, and illustrate that additional biomarkers are needed to identify those patients where checkpoint inhibition is not sufficient even in this category of patients with a high TMB.

Taken together, our study demonstrates the need to combine different markers rather than relying on isolated markers when selecting patients likely to benefit from ICI. The interactions determining how T cells recognize and, ultimately, kill cancer cells are the result of myriad processes and modulating immune response by check-point inhibition is only a single trigger in a larger biological cascade. Attempts to establish a multiparametric system comprising the mechanisms behind these interactions, such as the cancer immunogram (Blank et al., 2016) have been made, but the applicability in a clinical setting requires algorithms capable of managing not only large-scale data but also different types of data. The idea of combining neoantigen and immune signatures as a biomarker has been suggested in melanoma patients treated with adoptive cell transfer (ACT) (Lauss et al., 2017b). Machine learning strategies with multiple biomarkers have also been used to predict patient outcome for treatment with ICI but these machine learning algorithms needs large-scale data to make valuable predictions (Acharjee et al., 2020) and are therefore not used in this study, where the patient cohort is relatively small. The strategy applied here has the limitation that by separating patients into two groups with respect to high and low expression of certain gene signatures, some patients display borderline characteristics, and hence may be false categorized. The strength of machine learning approaches is their capacity to address a continuum of expression and a large number of parameters, thereby avoiding the need for strict and pre-defined cut-off values. But as mentioned, this requires very large datasets to avoid overfitting results and to capture the variability that is observed within and across cancer patient cohorts. Access to biological and clinical data from such large cohorts, where sequencing data from different biological specimens are available at high quality is still a major limitation towards developing such algorithms.

PD-L2 interaction with PD-1 inhibits T cell activation(Latchman et al., 2001). Furthermore, the function and importance of PD-L2 has recently been investigated and suggested as an important target for cancer (Solinas et al., 2020). We see a correlation between high expression of PD-L2 and better survival probability both for patients treated with anti-PD-1 and anti-PD-L1. This trend could be explained by high expression of PD-L2 in TME, being a signature of immune activity in the tumor site. This agrees with a study that showed a positive correlation between high PD-L2 expression and lymphocytic infiltration and improved overall survival (Obeid et al., 2016). Furthermore, for patients treated with anti-PD-1, the PD-L2 molecule is directly involved in the immunosuppressive axis that is being blocked by treatment.

A challenge associated with most of the biomarkers currently identified as relevant for the prediction of response to immunotherapy, including those described here, is that they require the availability of tumor material. Future initiatives are heading towards an understating of how susceptibility to immunotherapy can be evaluated by studying circulating tumor cells (CTCs) and circulating tumor DNA (ctDNA) in peripheral blood. Such material might be useful to determine the neoepitope load if sufficiently representative of the tumor.

In conclusion, this study adds to the potential impact of PD-L2, neoepitope load, CYT, and T cell diversity as potential biomarkers. Data from our study and the validation cohort, suggest that PD-L2 and neoepitope load both with and without CYT significantly predict patient survival. Due to the small sample size, our results need further validation in larger cohorts.

AUTHOR CONTRIBUTION

AB: conceived the concept, performed data analysis, generated figures, discussed the data and wrote the manuscript. AMB: conceived the concept, performed data analysis, generated figures, discussed the data, and wrote the manuscript; VL: Gathered patient material, discussed the data, and wrote the manuscript; OO, CWY and AE: provided technical assistance, discussed data, and revised the manuscript; CB: discussed data, and revised the manuscript; MMS provided patient material and reviewed the manuscript; IMS: discussed the data and revised manuscript; FSN: provided technical platform and revised manuscript; SAF; provided patient material for validation cohort; UL: provided patient material, conceived the concept, discussed the data, and wrote the manuscript; SRH: conceived the concept, discussed the data, and wrote the manuscript.

COMPETING INTERESTS STATEMENT

SRH is the cofounder of Tetramershop and PokeAcell and is the co-inventor of the patents WO2015185067 and WO2015188839 for the barcoded MHC technology which is licensed to Immudex and co-inventor of the licensed patent for Combination encoding of MHC multimers (EP2088/009356), licensed: Sanquin, NL. None of the above is directly related to the data presented in this manuscript. The authors declare that there is no conflict of interest.

ACKNOWLEDGMENTS

We would like to thank all patients who participated in the clinical trial.

FUNDING

Funding was provided by Novo Nordisk Foundation and European Research Council.

DATA AVAILABILITY

Related to our test cohort:
The Capital Region of Denmark do not allow sequencing data being reposited in open repositories, however all data will be available upon request to the authors, based on an individual data sharing agreement.

Related to our validation cohort:
All WES and RNAseq data is available upon application at dbGaP at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001743.v1.p1. GRCh38 reference genome is available at https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.39. All other relevant data are available from the authors upon request. A description of covariate data is previously published (Snyder et al., 2017).

REFERENCES

Acharjee, A., Larkman, J., Xu, Y., Cardoso, V. R., and Gkoutos, G. v. (2020). A random forest based biomarker discovery and power analysis framework for diagnostics research. *BMC Med Genomics* 13. doi: 10.1186/s12920-020-00826-6.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence [online] Data Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bjerregaard, A.-M., Nielsen, M., Hadrup, S. R., Szallasi, Z., and Eklund, A. C. (2017). MuPeXI: prediction of neo-epitopes from tumor sequencing data. *Cancer Immunology, Immunotherapy* 66, 1123–1130. doi: 10.1007/s00262-017-2001-3.

Blank, C. U., Haanen, J. B., Ribas, A., and Schumacher, T. N. (2016). The "cancer immunogram." *Science* 352, 658–60. doi: 10.1126/science.aaf2834.

Blank, C. U., Rozeman, E. A., Fanchi, L. F., Sikorska, K., van de Wiel, B., Kvistborg, P., et al. (2018). Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma. *Nat Med* 24, 1655–1661. doi: 10.1038/s41591-018-0198-0.

Bolotin, D. A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I. Z., Putintseva, E. V, et al. (2015). MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* 12, 380–1. doi: 10.1038/nmeth.3364.

Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 34, 525–7. doi: 10.1038/nbt.3519.

Brown, S. D., Hapgood, G., Steidl, C., Weng, A. P., Savage, K. J., and Holt, R. A. (2016). Defining the clonality of peripheral T cell lymphomas using RNA-seq. *Bioinformatics*, btw810. doi: 10.1093/bioinformatics/btw810.

Burdett, N., and Desai, J. (2020). New biomarkers for checkpoint inhibitor therapy. *ESMO Open* 5. doi: 10.1136/esmoopen-2019-000597.

Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., et al. (2018). Development of Tumor Mutation Burden as an Immunotherapy Biomarker: Utility for the Oncology Clinic. *Annals of Oncology*. doi: 10.1093/annonc/mdy495.

Chowell, D., Yoo, S. K., Valero, C., Pastore, A., Krishna, C., Lee, M., et al. (2022). Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat Biotechnol* 40. doi: 10.1038/s41587-021-01070-8.

Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., et al. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 31, 213–9. doi: 10.1038/nbt.2514.

Doroshow, D. B., Bhalla, S., Beasley, M. B., Sholl, L. M., Kerr, K. M., Gnjatic, S., et al. (2021). PD-L1 as a biomarker of response to immune-checkpoint inhibitors. *Nature Reviews Clinical Oncology 2021 18:6* 18, 345–362. doi: 10.1038/s41571-021-00473-5.

Felix Krueger (2021). TrimGalore https://github.com/FelixKrueger/TrimGalore.

Goodman, A. M., Kato, S., Bazhenova, L., Patel, S. P., Frampton, G. M., Miller, V., et al. (2017). Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. *Mol Cancer Ther* 16, molcanther.0386.2017. doi: 10.1158/1535-7163.MCT-17-0386.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847–2849. doi: 10.1093/bioinformatics/btw313.

Herbst, R. S., Soria, J. C., Kowanetz, M., Fine, G. D., Hamid, O., Gordon, M. S., et al. (2014). Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature 2014 515:7528* 515, 563–567. doi: 10.1038/nature14011.

Holm, J. S., Funt, S. A., Borch, A., Munk, K. K., Bjerregaard, A. M., Reading, J. L., et al. (2022). Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma. *Nature Communications 2022 13:1* 13, 1–17. doi: 10.1038/s41467-022-29342-0.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *The Journal of Immunology*, ji1700893. doi: 10.4049/jimmunol.1700893.

Latchman, Y., Wood, C. R., Chernova, T., Chaudhary, D., Borde, M., Chernova, I., et al. (2001). PD-L2 is a second ligand for PD-1 and inhibits T cell activation. *Nat Immunol* 2, 261–268. doi: 10.1038/85330.

Lauss, M., Donia, M., Harbst, K., Andersen, R., Mitra, S., Rosengren, F., et al. (2017). Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. *Nat Commun* 8, 1–11. doi: 10.1038/s41467-017-01460-0.

Le, D. T., Uram, J. N., Wang, H., Bartlett, B. R., Kemberling, H., Eyring, A. D., et al. (2015). PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 372, 2509–20. doi: 10.1056/NEJMoa1500596.

Lemery, S., Keegan, P., and Pazdur, R. (2017). First FDA Approval Agnostic of Cancer Site — When a Biomarker Defines the Indication. *New England Journal of Medicine* 377, 1409–1412. doi: 10.1056/NEJMp1709968.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–60. doi: 10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17, 10–12. doi: http://dx.doi.org/10.14806/ej.17.1.200.

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., et al. (2016). The Ensembl Variant Effect Predictor. *Genome Biol* 17, 122. doi: 10.1186/s13059-016-0974-4.

Obeid, J. M., Erdag, G., Smolkin, M. E., Deacon, D. H., Patterson, J. W., Chen, L., et al. (2016). PD-L1, PD-L2 and PD-1 expression in metastatic melanoma: Correlation with tumor-infiltrating immune cells and clinical outcome. *Oncoimmunology* 5. doi: 10.1080/2162402X.2016.1235107.

Rooney, M. S., Shukla, S. a, Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. *Cell* 160, 48–61. doi: 10.1016/j.cell.2014.12.033.

Ros-Martínez, S., Navas-Carrillo, D., Alonso-Romero, J. L., and Orenes-Piñero, E. (2020). Immunoscore: a novel prognostic tool. Association with clinical outcome, response to treatment and survival in several malignancies. *https://doi-org.proxy.findit.cvt.dk/10.1080/10408363.2020.1729692* 57, 432–443. doi: 10.1080/10408363.2020.1729692.

Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* 51, 202–206. doi: 10.1038/S41588-018-0312-8.

Sha, D., Jin, Z., Budczies, J., Kluck, K., Stenzinger, A., and Sinicrope, F. A. (2020). Tumor Mutational Burden as a Predictive Biomarker in Solid Tumors. *Cancer Discov* 10, 1808–1833. doi: 10.1158/2159-8290.CD-20-0522.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal* 27, 623–656.

Smyth, G. K. (2005). “limma: Linear Models for Microarray Data,” in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (New York: Springer-Verlag), 397–420. doi: 10.1007/0-387-29362-0_23.

Snyder, A., Nathanson, T., Funt, S. A., Ahuja, A., Buros Novik, J., Hellmann, M. D., et al. (2017). Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: An exploratory multi-omic analysis. *PLoS Med* 14. doi: 10.1371/journal.pmed.1002309.

Solinas, C., Aiello, M., Rozali, E., Lambertini, M., Willard-Gallo, K., and Migliori, E. (2020). Programmed Cell Death-Ligand 2: A Neglected But Important Target in the Immune Response to Cancer? *Transl Oncol* 13. doi: 10.1016/j.tranon.2020.100811.

Stewart, J. J., Lee, C. Y., Ibrahim, S., Watts, P., Shlomchik, M., Weigert, M., et al. (1997). A Shannon entropy analysis of immunoglobulin and T cell receptor. *Mol Immunol* 34, 1067–82. doi: S0161589097001302 [pii].

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* 30, 3310–6. doi: 10.1093/bioinformatics/btu548.

Tang, J., Shalabi, A., and Hubbard-Lucey, V. M. (2018). Comprehensive analysis of the clinical immuno-oncology landscape. *Annals of Oncology* 29, 84–91. doi: 10.1093/annonc/mdx755.

Taube, J. M., Klein, A., Brahmer, J. R., Xu, H., Pan, X., Kim, J. H., et al. (2014). Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clinical Cancer Research* 20. doi: 10.1158/1078-0432.CCR-13-3271.

Tuxen, I. V., Jønson, L., Santoni-Rugiu, E., Hasselby, J. P., Nielsen, F. C., and Lassen, U. (2014). Personalized oncology: genomic screening in phase 1. *APMIS* 122, 723–33. doi: 10.1111/apm.12293.

Tuxen, I. V., Rohrberg, K. S., Oestrup, O., Ahlborn, L. B., Schmidt, A. Y., Spanggaard, I., et al. (2018). Copenhagen Prospective Personalized Oncology (CoPPO)-Clinical Utility of Using Molecular Profiling to Select Patients to Phase I Trials. *Clin Cancer Res*, clincanres.1780.2018. doi: 10.1158/1078-0432.CCR-18-1780.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43, 11.10.1-33. doi: 10.1002/0471250953.bi1110s43.

Weese, D., Holtgrewe, M., and Reinert, K. (2012). RazerS 3: faster, fully sensitive read mapping. *Bioinformatics* 28, 2592–9. doi: 10.1093/bioinformatics/bts505.

Wells, D. K., van Buuren, M. M., Dang, K. K., Hubbard-Lucey, V. M., Sheehan, K. C. F., Campbell, K. M., et al. (2020). Key Parameters of Tumor Epitope Immunogenicity Revealed Through a Consortium Approach Improve Neoantigen Prediction. *Cell* 183, 818-834.e13. doi: 10.1016/j.cell.2020.09.015.

Wood, M. A., Weeder, B. R., David, J. K., Nellore, A., and Thompson, R. F. (2020). Burden of tumor mutations, neoepitopes, and other variants are weak predictors of cancer immunotherapy response and overall survival. *Genome Med* 12. doi: 10.1186/S13073-020-00729-2.

Yang, H., Zhou, X., Sun, L., and Mao, Y. (2019). Correlation between PD-L2 expression and clinical outcome in solid cancer patients: A meta-analysis. *Front Oncol* 9. doi: 10.3389/fonc.2019.00047.

Yarchoan, M., Hopkins, A., and Jaffee, E. M. (2017). Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *New England Journal of Medicine* 377. doi: 10.1056/nejmc1713444.

Yearley, J. H., Gibson, C., Yu, N., Moon, C., Murphy, E., Juco, J., et al. (2017). PD-L2 Expression in Human Tumors: Relevance to Anti-PD-1 Therapy in Cancer. *Clinical Cancer Research* 23, 3158–3167. doi: 10.1158/1078-0432.CCR-16-1761.

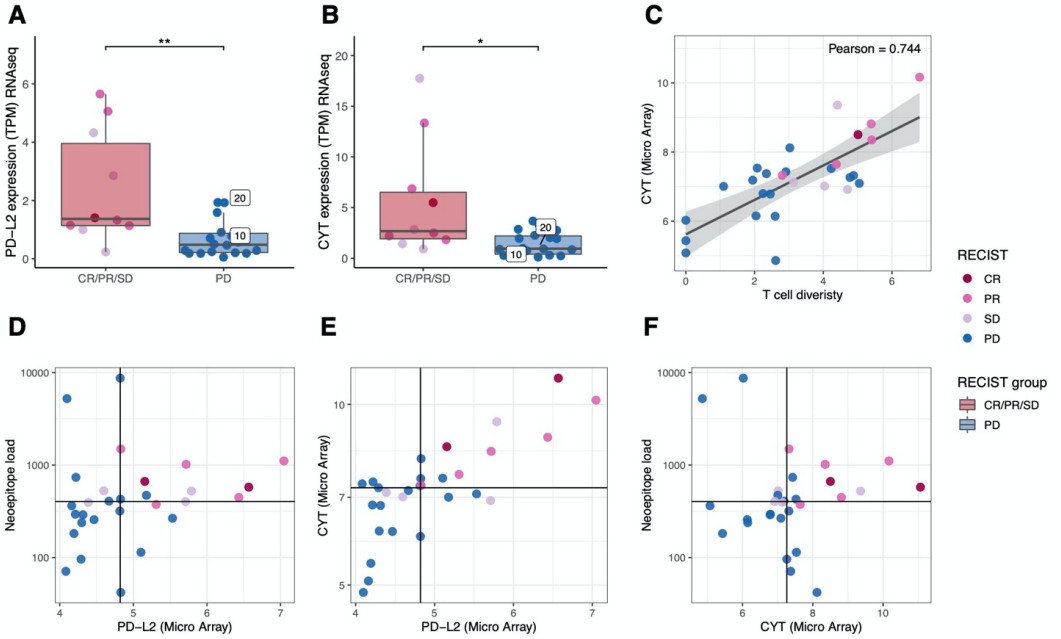
Yu, G. (2021). enrichplot: Visualization of Functional Enrichment Result. *R package version 1.12.3*.

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). ClusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* 16. doi: 10.1089/omi.2011.0118.

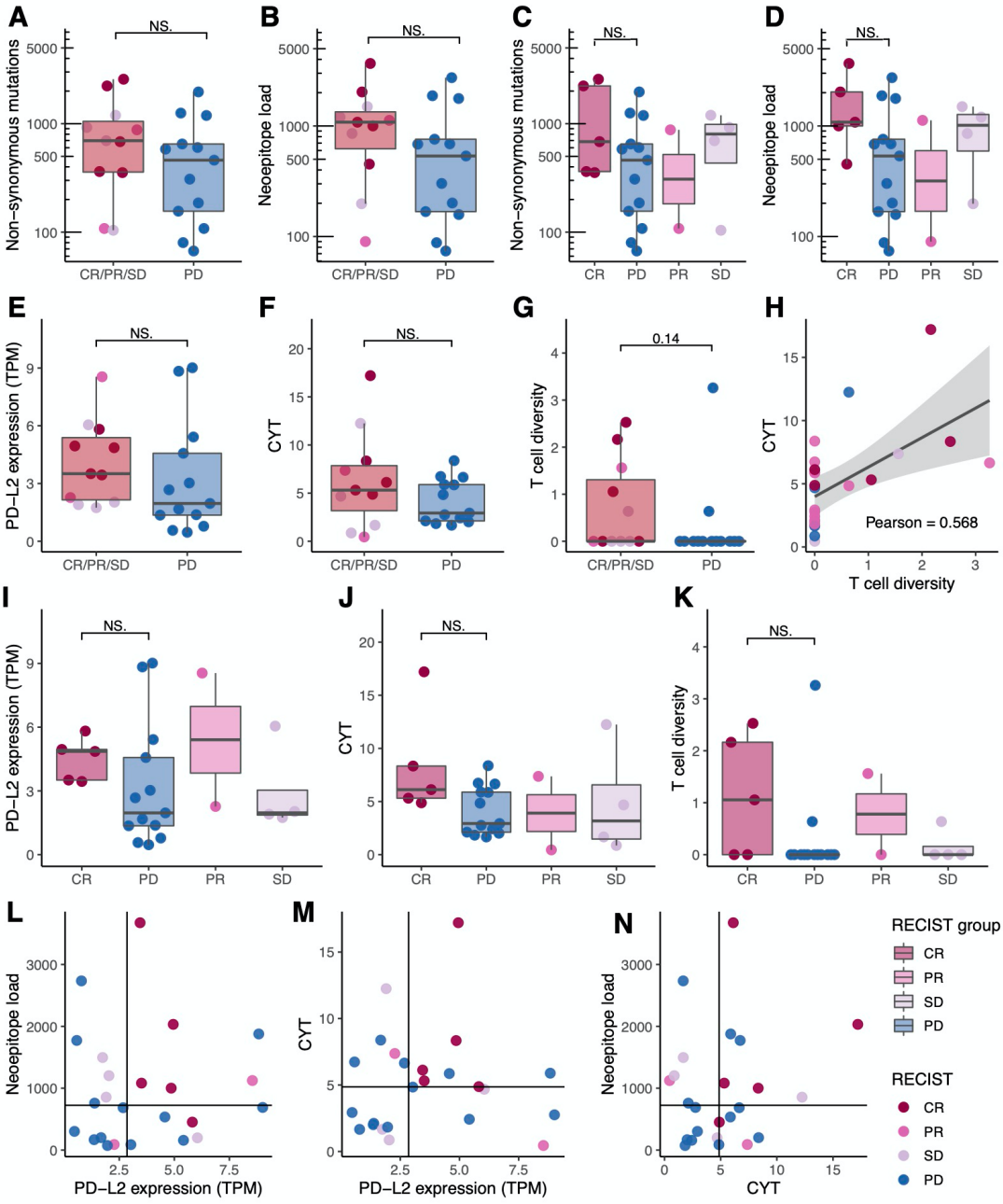
Supplementary Table 1. Clinical characteristics of included patients.

		N	N (%)	Mean	95% CI Lower	95% CI Upper
Gender	Male	10	34.5			
	Female	19	65.5			
Age				56.03 (28-74)	51.3	60.8
Number of treatment lines previous to immune therapy				3.2 (1-7)	2,5	3.9
Death	Alive	10	34.5			
	Dead	19	65.5			
Subtype	BLAC	5	17.2			
	BRAC	6	20.7			
	CCA-IG	1	3.4			
	CDC-K	1	3.4			
	CESC	2	6.9			
	COAD	4	13.8			
	LIHC	1	3.4			
	OV	2	6.9			
	PAAD	2	6.9			
	READ	3	10.3			
	SKCN	1	3.4			
	UC-U	1	3.4			

BLCA; Bladder Urothelial Carcinoma, BRCA; Breast invasive carcinoma, CCA-IG; clear cell adenocarcinoma - intern genitalia, CDC-K; collecting duct carcinoma – kidney, CESC; Cervical squamous cell carcinoma and endocervical adenocarcinoma, COAD; Colon adenocarcinoma, LIHC; Liver hepatocellular carcinoma, OV; Ovarian serous cystadenocarcinoma, PAAD; Pancreatic adenocarcinoma, READ; Rectum adenocarcinoma, SKCN; Skin Cutaneous Melanoma, UC-U; Urothelial carcinoma – Urethra.



Supplementary Figure 1. Biomarker from RNAseq, T cell signature correlation, and biomarker combinations. A-C) the expression from CYT and PD-L2 were validated in RNAseq where the non-progressive group (CR/PR/SD) was compared to the progressive group (PD) group. **A)** PD-L2 expression from RNAseq (TPM) (p-value = 0.0032). **B)** Cytolytic activity (CYT) calculated from RNAseq data (p-value = 0.014). **C)** CYT from RNAseq and T cell diversity with correlation (Pearson correlation = 0.744), colored by RECIST. **D)** Neoepitope load vs PD-L2 expression showed that non-progressive disease patients were gathered in the upper right corner with above median expression in both. **E)** same observation by observing PD-L2 expression combined with CYT. **F)** Neoepitope load vs. CYT also showed clustering of non-progressive patients in the upper right corner.



Supplementary Figure 2. Analysis from validation cohort. The analysis made from the validation cohort compared the suggested biomarker with RECIST criteria all tests comparing are performed with the Wilcox rank sum test and besides the correlations which were performed with person correlation. **A+B**) comparing non-progressive and progressive from the RECIST group CR/PR/SD vs. SD. **A)** no significant differences were obtained by observing Tumor Mutational burden (TMB) (p-value = 0.21). **B)** Neither for the Neoepitope load (p-value = 0.15). **C+D)** Comparing RECIST group individually and statistical differences between CR and PD were observed. **C)** For the TMB no significant difference

were obtained (p-value = 0.14). **D)** Observing the neoepitope load resulted in a better separation but non-significant (p-value = 0.076). **E-G)** Comparing non-progressive and progressive patients. **E)** PD-L2 expression (TPM) (p-value = 0.15). **F)** Cytolytic activity (CYT) (p-value = 0.49). **G)** T cell diversity (p-value = 0.14). **H)** T cell diversity and CYT (person correlation = 0.568). **I-K)** Comparing RECIST criteria individually where statistic test comparing PD with CR. **I)** PD-L2 expression (p-value = 0.11). **J)** CYT (p-value = 0.076). **K)** T cell diversity (p-value = 0.081). **L-N)** Combination of some of the suggested biomarkers where the horizontal and vertical lines indicate the median. **L)** Predicted neoepitopes vs. PD-L2 expression. **M)** CYT vs. PD-L2 expression **N)** Neoepitope load vs. CYT.

4 MANUSCRIPT II

Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma

Jeppe Sejerø Holm*, Samuel A. Funt*, **Annie Borch**, Kamilla Kjærgaard Munk, Anne-Mette Bjerregaard, James L. Reading, Colleen Maher, Ashley Regazzi, Phillip Wong, Hikmat Al-Ahmadie, Gopa Iyer, Tripti Tamhane, Amalie Kai Bentzen, Nana Overgaard Herschend, Susan De Wolf, Alexandra Snyder, Taha Merghoub, Jedd D. Wolchok, Morten Nielsen, Jonathan E. Rosenberg, Dean F. Bajorin* Sine Reker Hadrup*.
* these author are contributed equally.
Nature communications - 11 April 2022

Contribution: This Manuscript is made in collaboration with Jeppe Sejerø Holm and Samuel Funt, who share first authorship. They have collected patient samples and performed all the wet-lab experiments. I have been the main responsible bioinformatician in this study. Mainly represented in Figure 5 and Figure 6.

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-022-29342-0 OPEN

Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma

Jeppe Sejerø Holm[1,9], Samuel A. Funt[2,3,9], Annie Borch[1], Kamilla Kjærgaard Munk[1], Anne-Mette Bjerregaard[1], James L. Reading[4], Colleen Maher[2,3,5], Ashley Regazzi[2,3], Phillip Wong[2,3,5], Hikmat Al-Ahmadie[6], Gopa Iyer[2,3], Tripti Tamhane[1], Amalie Kai Bentzen[1], Nana Overgaard Herschend[1], Susan De Wolf[2], Alexandra Snyder[2,3], Taha Merghoub[2,3,5], Jedd D. Wolchok[2,3,5,7], Morten Nielsen[8], Jonathan E. Rosenberg[2,3], Dean F. Bajorin[2,3,9] & Sine Reker Hadrup[1,9]✉

CD8$^+$ T cell reactivity towards tumor mutation-derived neoantigens is widely believed to facilitate the antitumor immunity induced by immune checkpoint blockade (ICB). Here we show that broadening in the number of neoantigen-reactive CD8$^+$ T cell (NART) populations between pre-treatment to 3-weeks post-treatment distinguishes patients with controlled disease compared to patients with progressive disease in metastatic urothelial carcinoma (mUC) treated with PD-L1-blockade. The longitudinal analysis of peripheral CD8$^+$ T cell recognition of patient-specific neopeptide libraries consisting of DNA barcode-labelled pMHC multimers in a cohort of 24 patients from the clinical trial NCT02108652 also shows that peripheral NARTs derived from patients with disease control are characterised by a PD1$^+$ Ki67$^+$ effector phenotype and increased CD39 levels compared to bystander bulk- and virus-antigen reactive CD8$^+$ T cells. The study provides insights into NART characteristics following ICB and suggests that early-stage NART expansion and activation are associated with response to ICB in patients with mUC.

[1] Experimental and Translational Immunology, Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark. [2] Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [3] Weill Cornell Medical College, New York, NY 10065, USA. [4] Cancer Immunology Unit, Research Department of Hematology and Cancer Research UK, Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK. [5] Parker Institute for Cancer Immunotherapy, San Francisco, CA, USA. [6] Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [7] Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. [8] Section of Bioinformatics, Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark. [9] These authors contributed equally: Jeppe Sejerø Holm, Samuel A. Funt, Dean F. Bajorin, Sine Reker Hadrup ✉email: sirha@dtu.dk

The anti-tumor T cell response induced by ICB of the programmed death 1 (PD-1) / programmed death-ligand 1 (PD-L1) axis can result in deep and durable responses in patients with a variety of metastatic cancers[1–5]. In pre-treatment tumors, a high mutational burden is correlated with a beneficial response to immune checkpoint blockade (ICB) across multiple indications, which is thought to be, at least in part, due to increased presentation of exogenous neoantigens displayed by major histocompatibility complex (MHC) class I to $CD8^+$ T cells reinvigorated by ICB[6–9]. Indeed, a growing body of evidence suggests that $CD8^+$ T cells not only infiltrate responding tumors but also undergo rapid and robust proliferation in the peripheral blood of patients following treatment with ICB[10–17]. With regard to the specificity of this $CD8^+$ T cell response, the interrogation of neoantigen recognizing T cells (NARTs) has mostly been limited to tumor tissue[18]. Importantly, since the relevant neoantigens are unique for the individual patient, detailed NART analysis requires the prediction of potential neoantigens and T cell screening with a unique set of neoantigens for each patient. This effort has been hindered by technological barriers limiting the analytic range required to comprehensively characterize the vast number of neoantigens potentially presented as well as the diversity of potential human leukocyte antigen class I (HLA) genotypes.

To address these issues, we focused on a previously reported cohort of patients who were treated with the anti-PD-L1 antibody atezolizumab for metastatic urothelial carcinoma (mUC)[19]. Clinical outcome of mUC has significantly benefited from the introduction of ICB treatment, with approximately 20% of patients with previously fatal disease experiencing long-term survival[20,21]. Additionally, long-term follow-up was obtained for the patients under study ($n = 24$), with some remaining progression free 5+ years after treatment initiation. Peripheral blood samples collected pre-, during-, and post-treatment were comprehensively screened using patient-specific neopeptide-MHC (pMHC) multimer libraries, labeled with DNA barcodes[22], allowing for high-throughput detection of $CD8^+$ T cell populations recognizing any such neopeptides in one parallel reaction. We interrogated NART dynamics and phenotype using this novel technique and evaluated for associations with clinical outcome.

Results

Neoepitope prediction and T cell screening. Neoepitopes were predicted for each individual patient from Whole Exome Sequencing (WES) and RNAseq by use of the MuPeXI platform[23]. Potential neopeptide candidates were selected based on their respective MHC-I binding affinity and expression level, and the experimental availability of the recombinant MHC-I molecules relevant for the given peptide. From the 24 patients, 56 different HLA ABC-haplotypes are represented, and at the time of analysis 31 of these were available for pMHC multimer generation. On average, four HLA haplotypes were covered for each patient. All HLA-feasible neopeptides with a predicted Eluted Ligand (EL)%Rank score[24] <0.5 and expression level >0.1 transcripts per million (TPM) were selected, yielding between 14-587 HLA-binding neopeptides per patient (Fig. 1a). To allow complete evaluation of potential neoepitopes in patients with low mutational burden, additional neopeptides with higher EL%Rank score, derived from genes with expression level >0.1 TPM were included, based on lowest EL%Rank score, until a minimum of 200 neopeptides was reached for T cell recognition analyses per patient. As a result, each patient was analyzed using pMHC multimer libraries displaying between 200-587 patient-unique neopeptides (Fig. 1b). In total, 6237 HLA-feasible neopeptides across the 24 patients were predicted and included for T cell analyses (Table 1).

In this patient cohort, neither the tumor mutational burden (TMB) nor number of predicted neopeptides with EL%Rank <0.5 was predictive of ICB outcomes (Fig. 1c, d). Hence, we evaluated for the presence of circulating NARTs to gain greater insights into the nature of anti-tumor immunity following initiation of anti-PD-L1 therapy.

Presence of NARTs in patient PBMC samples. Based on the selected neopeptides, barcoded pMHC multimer libraries were generated for each patient matching their HLA-type (Fig. 2a). In addition to neopeptides, one to 17 HLA-matching virus-derived peptides from cytomegalovirus (CMV), Epstein-Barr viral (EBV), and influenza (FLU, together; CEF) were included in each patient's peptide library for internal assay validation and to compare NARTs to virus-antigen reactive T cells (VARTs; Table 1). T cell recognition was examined by peripheral blood mononuclear cell (PBMC) staining and sorting of a pMHC associated barcode in the multimer-binding T cell population. Significant T cell recognition of a given neopepitope (NART response) was defined as a Log_2 fold change (Log_2FC) >2 and false-discovery rate (FDR) < 0.1, based on previous investigations[22].

PBMC samples, collected just prior to administration of atezolizumab, were screened for the presence of NARTs ($n = 85$ samples, median = 3 samples per patient) at the indicated time points (Table 1 + Supplementary Table 1). Of note, PBMCs were available from sampling up to >231 weeks after treatment initiation for six long-term responders.

Representative outputs from patient screenings for NARTs is depicted in Fig. 2b and Supplementary Fig. 2a, b. For patient #2389 (ongoing complete response [CR] per best response evaluation criteria in solid tumors [RECIST] version 1.1 criteria, >280 weeks after treatment initiation), the T cell recognition (Log_2FC) is depicted for each of the pMHC multimers evaluated in this patient ($n = 203$; Fig. 2b). The outcome is listed according to the time points of blood sampling and grouped by the evaluated HLA molecules. An emergence of new T cell populations recognizing neoepitopes presented on HLA-A*01:01 and HLA-B*40:01 is seen from pre-treatment to 3 weeks post-treatment, while a T cell response towards a FLU epitope presented on HLA-A*01:01 (VSDGGPNLY) is detected at the majority of screened timepoints. Based on additional sample availability, 65 NART responses detected at 3- or 9 weeks post-treatment were interrogated using neo-pMHC tetramers (Fig. 2c). 50 NART responses were validated, whereas an additional nine responses were borderline detectable (Supplementary Fig. 3). These borderline detectable responses may represent low-affinity NARTs, likely detected only by using DNA barcode-labelled multimers due to enhanced sensitivity compared to conventional tetramer-based detection for detection of such T cells[22]. In parallel with the screening of patient samples, a selected set of healthy donor (HD) PBMC samples were evaluated to validate all pMHC multimer libraries based on the included CEF-peptides (Supplementary Fig. 2c). A few pMHC multimer complexes demonstrated unspecific binding in all samples and were excluded from further analyses.

At the pre-treatment time point, T cells recognizing neoepitopes were detected in the 18 of 24 of patients (median 2; range 0 to 10 NART responses per patient). After 3 weeks of atezolizumab treatment, T cells recognizing neoepitopes were detected in the 17 of 22 patients (median 3; range 0 to 13 NART responses per patient; Table 1). 18 of 22 patients, including four with no detectable pre-treatment NARTs, developed NARTs post-treatment that were not originally present pre-treatment. 45 of

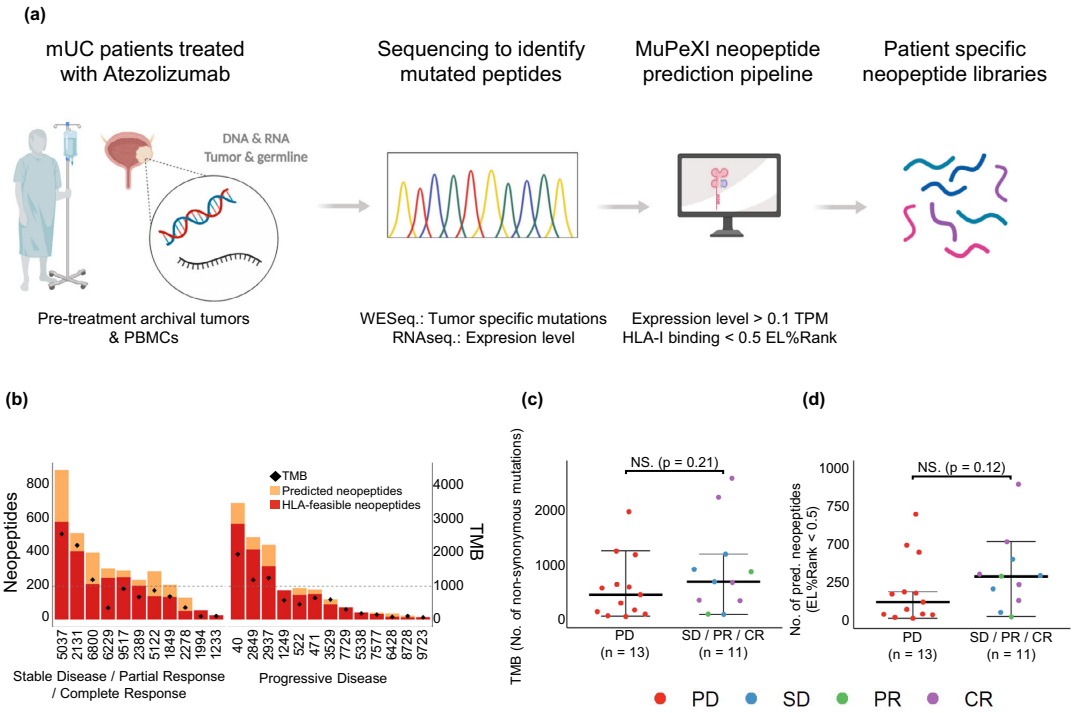


Fig. 1 Prediction of patient-specific neopeptides. a Overview of workflow for identification and prediction of mutation-derived neopeptides in mUC patients. Created with BioRender.com. **b** Number of predicted and HLA-feasible neopeptides with EL%Rank <0.5, and TMB (diamond) in the patient cohort. Patients grouped according to best RECIST 1.1 criteria, PD ($n = 13$ patients; median = 123, range = 16–700 neopeptides with EL%Rank <0.5 predicted) and SD/PR/CR patients ($n = 11$ patients; median = 292, range = 26–898 neopeptides with EL%Rank <0.5 predicted). Dotted line represents minimum 200 neopeptides included in panel for each patient. **c**, **d** TMB and number of prediced neopeptides with EL%Rank <0.5 for patients with PD compared to SD/PR/CR patients. For (**c**, **d**) groups were compared using non-parametric two-sided Mann–Whitney test and data is presented as median values ± largest/smallest value within upper/lower quartile ± 1.5 IQR. NS. Not Significant. Source data are provided as a Source Data file.

unique 148 NART responses were observed at multiple time-points (Supplementary Table 1). There was no immediate association between the number of unique NART responses detected throughout treatment and the number of predicted and evaluated neoepitopes (Fig. 2d). This indicates that other parameters, beyond TMB, influences the capacity to drive such T cell responses upon ICB initiation.

Enhancement of NART responses three weeks post-treatment is associated with improved clinical outcome. At the pre-treatment time point, no association between the number of NART responses and best RECIST response was observed. However, a difference in the kinetics of NART responses was noted over the course of treatment, with NART populations increasing three weeks post-treatment and then contracting in the majority of patients with disease control (defined as patients with SD, PR, and CR) but not in patients with progressive disease (PD; Fig. 3a). Indeed, at the three week post-treatment time point, patients with disease control tended to have a higher number of NART responses compared to patients with PD (Fig. 3a, $p = 0.067$), and also when comparing patients with a CR versus PD (Fig. 3b, c). No significant differences between patient response groups were observed at subsequent time points. The change (*delta*) in number of detected NART responses between pre-treatment and three weeks post-treatment was calculated to better approximate patient-specific NART dynamics (Fig. 3d–e). A significant increase in ΔNART responses was observed for patients with disease control compared to those with PD (Fig. 3d, $p = 0.012$), with CR patients experiencing the largest ΔNART responses compared to PD patients (Fig. 3e, $p = 0.022$). Although a substantial smaller library of CEF-derived epitopes was included in the analyses compared to neoepitopes, no changes in the VART response repertoire were observed during treatment (Supplementary Fig. 5). The frequency of NART responses was estimated based on pMHC multimer staining and the fraction of barcode reads assigned to the given populations (see Materials and Methods). The sum of estimated frequencies (SEFs) across all patient samples ranged between 0.01% and 3.9% ($n = 62$, 0.55% average) (Supplementary Fig. 4a). Individual response estimated frequencies range from 0.01% up to ~2.83% ($n = 221$), with >98% of response frequencies being below 1% (mean = 0.15 %, median = 0.048%). Hence, SEFs in patient samples are not skewed by single, large NART response frequencies. Neither the absolute SEF nor the change in SEF from baseline to 3-weeks post treatment was associated to clinical outcome (Fig. 3f–i). Thus, based on this evaluation, a substantial difference in the number of NART responses was observed between patients with PD and those with CR, and the breadth of the NART repertoire rather than the combined estimated frequencies of such populations was the key parameter associated to favorable clinical outcome in the setting of ICB treatment. The increase *delta* in number of NART responses from pre- to 3 week post-treatment for patients with disease control indicate that these patients tended to rapidly raise a broader T cell neoantigen recognition repertoire post-treatment.

Table 1 Patient clinical data, feasible HLA-I information, neo- and viral peptides library size and no. of T cell responses for all patients. No. of T cell responses indicated for each patient at respective timepoints.

Patient	Clinical data	Feasible HLA-I			HLA-I-feasible neopeptides		Viral peptides	T cell neoepitope responses, per time point						
	Outcome (Best RECIST 1.1)	HLA-A	HLA-B	HLA-C	Predicted (EL%Rank <0.5)	Library size	Library size	Pre-treatment	3 weeks	9 weeks	20-29 weeks	49-73 weeks	156-166 weeks	231+ weeks
#40	PD[2]	A0201/A3201	B1801	C0401/C0701	576	576	9	5	-	-	-	-	-	-
#9723	PD[2]	A0101/A2402	B0702/B0801	C0701/C0702	16	200	16	1	-	-	-	-	-	-
#1249	PD[2]	A0301/A6801	B3501/B4402	C0401/C0501	177	200	12	1	2	-	-	-	-	-
#2937	PD	A3201	B1801/B3503	C0401	322	322	1	9	5	-	-	-	-	-
#1994	SD	A0301	B0702	C0702	56	200	9	3	2	3	-	-	-	-
#522[1]	PD	A2402/A3001	B5701	C0602	150	200	6	2	3	2	-	-	-	-
#6428	PD	A2601	B2705	C0202	18	200	1	0	0	0	-	-	-	-
#7577	PD	A0101/A0201	B5101	C0602/C0701	35	200	14	1	2	0	-	-	-	-
#8728	PD	-	B4403/B5101	-	14	200	4	5	2	0	-	-	-	-
#471	PD	A0201/A6801	B3503/B5101	C0401	155	200	11	0	0	0	-	-	-	-
#3529	PD	A0205/A2402	-	C0602	92	200	5	5	3	7	-	-	-	-
#5338	PD	A0201/A2402	B4001/B4002	C0202/C0304	44	200	16	1	0	0	-	-	-	-
#2849	PD	A0201/A0301	B0702/B2705	C0702	420	420	17	5	1	2	-	-	-	-
#7729	PD	A0101/A0301	B0801/B1302	C0602/C0701	75	200	15	5	3	5	3	-	-	-
#9517	SD	A0101/A0301	B3501/B3801	C0401	254	254	13	2	7	1	4	-	-	-
#1849	SD	A0101/A1101	-	C0401/C0602	136	200	7	10	9	14	-	-	-	-
#6800[1]	SD	A0301	B1501	C0401/C0602	214	214	10	0	5	1	-	-	-	-
#6229	CR	A3301/A3201	B1302	C0401/C0602	250	250	4	5	13	3	-	2	-	-
#5037	CR	A2601	B3801/B4002	C0202	587	587	2	2	6	4	-	2	0	0
#2389	CR	A0101/A2402	B3801/B4001	C0304	203	203	9	2	4	3	-	0	0	1
#2131	CR	A0201	B0801	C0202/C0701	411	411	14	0	4	1	-	6	5	4
#2278	CR	A0201/A2402	-	C0202	52	200	13	1	2	1	-	3	1	1
#1233	PR	A1101/A2601	B2705/B3501	C0202/C0702	26	200	8	0	0	2	-	-	1	0
#5122	PR	A2402/A3201	B3801	-	140	200	4	1	0	0	-	-	0	0

A dash (-) indicates no sample screening at the given time point. Clinical outcome determined from Best RECIST 1.1 criteria during therapy. [1]TCRb analysis for tumor not performed due to failed sequencing quality control. [2]Patient only scanned at baseline. Source data are provided as a Source Data file.

As described, predicted neoepitopes were originally included based on an EL%Rank score <0.5 selection criteria, but for 15 patients the neoepitope library sizes were extended to reach a minimum of 200 neoepitopes. We therefore evaluated if the potential differences in neoepitope characteristics and library size would influence our findings. Comparable analyses were conducted including only neoepitopes with EL%Rank score <0.5 and expression level >2 TPM or all library sizes were set to 200. For both such analyses, we observed similar trends as given for the total neopeptide libraries (Supplementary Fig. 4b±d). Hence, no bias in our findings was introduced by the original selection criteria.

Separating patients based according to Durable Clinical Benefit (DCB; progression-free survival, PFS >6 months) did not yield a significant difference in NART response numbers (Supplementary Fig. 4e). Furthermore, although not statistically significant, a trend for improved PFS and overall survival (OS) was seen for the patients with higher ΔNART response numbers from pre- to 3 weeks post-treatment (>median 0 ΔNART responses; Supplementary Fig. 4f–h).

Peripheral blood TCR metrics display similar kinetics as NARTs. TCR diversity and clonality for PBMCs and TILs have been previously shown to be correlated with response to ICB[14,25,26]. For this cohort, a higher fraction of the T cell clones present in tumor were seen to expand in the blood 3 weeks post-treatment for patients with DCB[19]. We observed a rapid spike in bulk TCR clonality early post-treatment, similar to the observed development in the number of NART responses (Supplementary Fig. 4i), even though changes in bulk TCR clonality or diversity did not differentiate patients with and without response to therapy (data not shown). Although bulk TCR sequencing does not identify the antigen specificity of individual clones, the parallel kinetics between the NART response development and TCR clonality for patients with a favorable clinical outcome is noteworthy. This further supports the contention that clonal expansion and T cell reinvigoration occurs early post-treatment following ICB.

Phenotypic characterization of NARTs indicate increased proliferation of NARTs in patients with disease control. To characterize the phenotypic profile of NARTs, a custom multicolor flow cytometry antibody panel was designed to characterize T cell differentiation, exhaustion, activation, and migration (Supplementary Fig. 6a). Both phycoerythrin (PE)-neoepitope multimers and viral pMHC multimers conjugated to allophycocyanin (APC) were included in the panel to further differentiate the phenotypic profiles of NARTs and VARTs. pMHC multimer-binding T cells were sorted and barcodes sequenced for epitope reactivity, whilst in parallel, the phenotypic profile of NARTs and VARTs were characterized. Patient PBMC samples from pre-treatment and at 3 weeks and 231+ weeks post-treatment were selected based on the initial early NART-response, and patient sample and multimer library availability ($n = 34$).

Data was visualized using Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction plugin. Variations in population distribution were observed when faceting UMAPs by patient, either pre- to post-treatment or disease control versus PD patients (Fig. 4a, b). Populations that were enriched in density post-treatment were characterized by expression of Ki67, PD-1, and in part CD39 (Fig. 4c). In particular, Ki67 and PD-1-expressing NARTs appeared to be enriched post-treatment in disease control patients compared to PD patients. In contrast, NARTs expressing CD57 appeared more frequent in PD patients post-treatment. Hence, guided by the

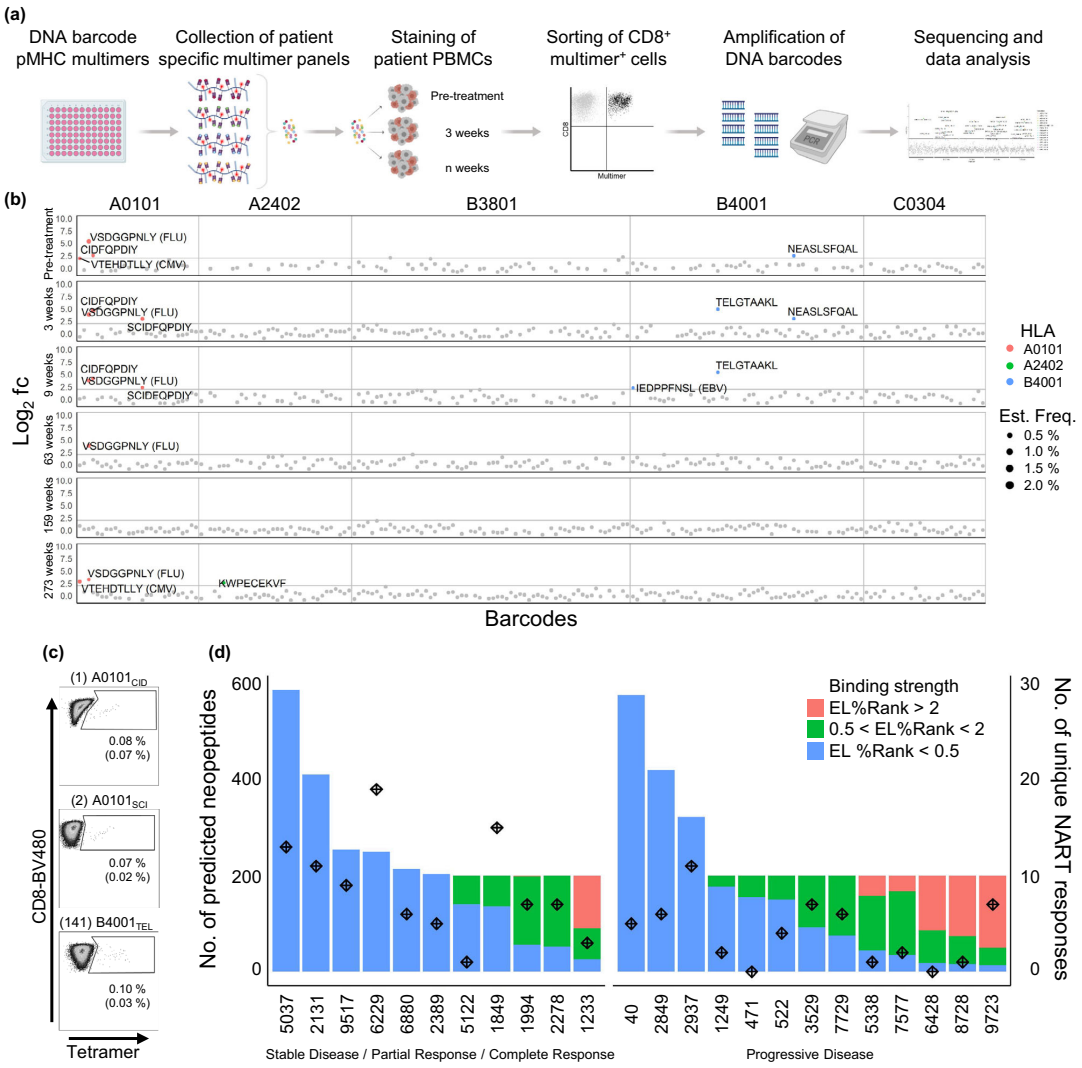


Fig. 2 Detection of neoepitope-reactive T cell responses in mUC cancer patients. a Overview of workflow for detection of neoantigen specific T cells in mUC cancer patients. Experimental steps include building of DNA-barcoded pMHC multimer libraries from predicted neopeptides, collection of multimer panels, staining of patient PBMCs with multimers, sorting of multimer-binding CD8$^+$ T cells, amplification and sequencing of DNA barcodes, and data analysis. Created with BioRender.com. **b** Representative output from patient #2389, screening of detected NART responses. Log_2 fold change (fc) of sequenced pMHC associated barcodes enriched by T cell sorting over the input library at each timepoint: pre-treatment, three weeks, nine weeks, 63 weeks, 159 weeks and at 273 weeks post-treatment. Labelled points represent significantly enriched pMHC co-attached barcodes with Log_2 fc >2, count fraction >0.1% and $p < 0.001$, determined as T cell responses, among the sorted multimer positive CD8$^+$ T cell populations. T cell responses are colored based on peptide-presenting HLA-type, text labelled with peptide sequence, and sized according the estimated frequency of the peptide-recognizing T cell population. If the peptide is derived from a virus, the virus is annotated in brackets, otherwise it is a mutation-derived neopeptide. Grey points represent non-enriched barcodes. Horizontal line at Log_2 fc = 2. Vertical line separating peptide-presenting HLA-types. **c** Tetramer validations of three NART responses in patient #2389 at 9-weeks post-treatment. CD8+ cells shown, gated for tetramer$^+$ populations. Frequency of tetramer pMHC population given in percentage, together with estimated frequency derived from multimer screening given in brackets. **d** Library sizes of predicted HLA-I-feasible neopeptides per patient, colored bars according to predicted HLA-I binding strength, compared to total no. of unique NART responses across screened samples for each patient (stars). Source data are provided as a Source Data file.

signatures from the UMAP, the frequency of the parameters that appeared increased in disease control patients post-treatment, i.e. KI67, CD39 and PD-1, were quantified for 'bulk CD8 T cells', 'NARTs', and 'VARTs' in the individual patients and the evaluated time-points, based on the full dataset. We observed an increase in the frequency of Ki67$^+$ (bulk CD8 $p = 0.00034$, NARTs $p = 0.0054$, VARTs $p = 0.001$) and PD-1$^+$ (bulk CD8 $p = 0.041$, NARTs $p = 0.043$) CD8 T cells from baseline to three

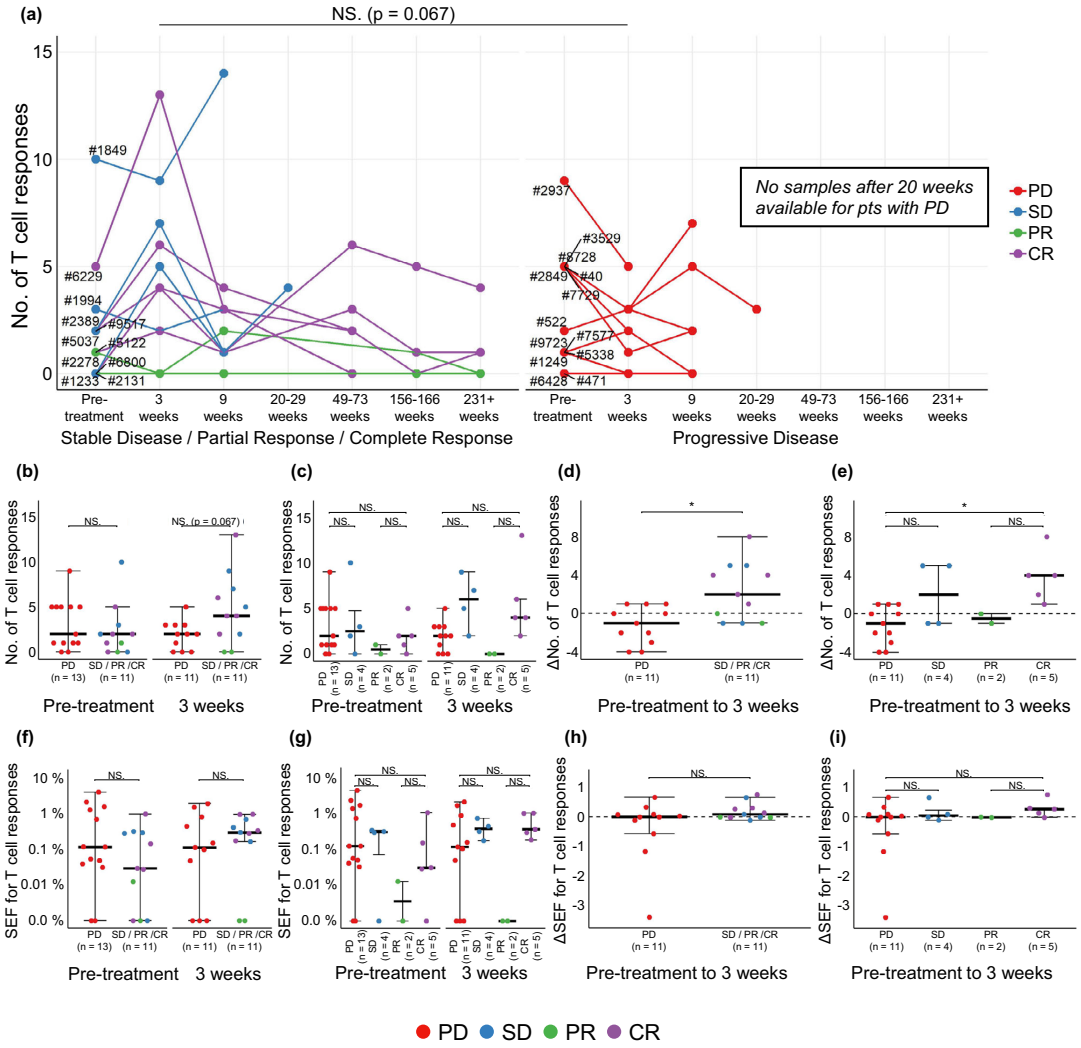
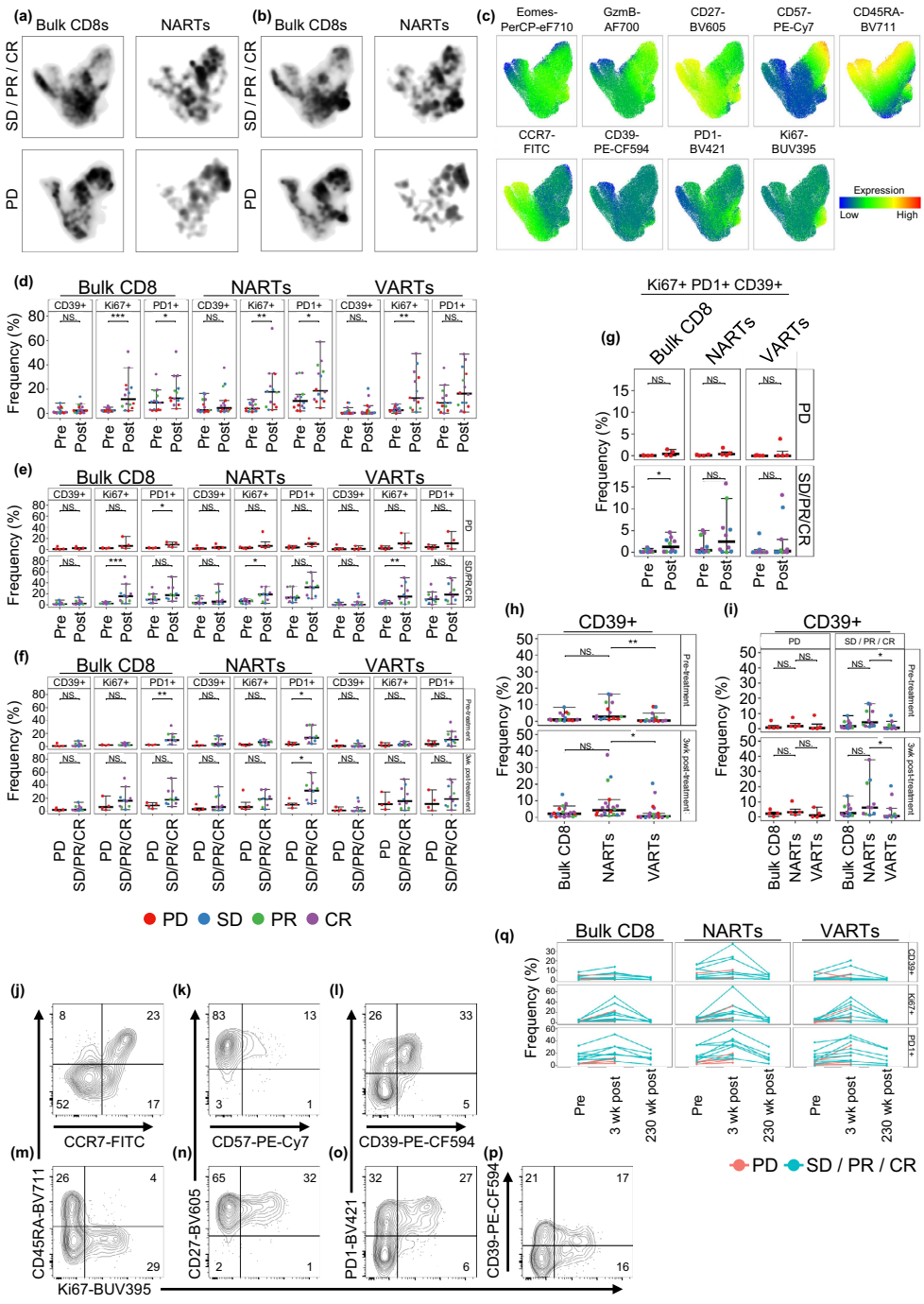


Fig. 3 Impact of NARTs in ICB. a No. of detected NART responses for all patients over course of therapy. Filled circles at respective time points indicate screening of the given sample. Patients are clustered according SD, PR, and CR ($n = 11$ patients) and with PD ($n = 13$ patients). **b**–**c** Number of detected NART responses prior to treatment ($n = 24$ patients) and at three weeks post-treatment ($n = 22$ patients) for each patient, **d**–**e** the change in number of NART responses between pre- and 3 weeks post-treatment for patients with minimum two time points sampled ($n = 22$ patients), **f**–**g** SEF for NART populations at pre- and at 3 weeks post-treatment, and (**h**–**i**) change in SEF for NART populations between pre- and three weeks post-treatment, in all figures evaluated as PD compared to SD/PR/CR patients or individual Best RECIST 1.1 groups. For (**b**), (**d**), (**f**) + (**h**) groups were compared using non-parametric two-sided Mann–Whitney test and Kruskal–Wallis Dunn's multiple comparison test for (**c**), (**e**), (**g**) + (**i**). For (**b**–**i**), data is presented as median values ± largest/smallest value within upper/lower quartile ± 1.5 IQR. NS. Not Significant, $^*p < 0.05$. Source data are provided as a Source Data file.

weeks post-treatment across all evaluated subpopulations (Fig. 4d); indicating a general signature of T cell activation as a consequence of ICB. Importantly, this increase was almost exclusively observed for patients with disease control, with a significance for both Ki67$^+$ bulk CD8 T cells ($p = 0.00088$), Ki67$^+$ NARTs ($p = 0.011$), and Ki67$^+$ VARTs ($p = 0.0091$). PD-1$^+$ bulk CD8 T cells were increased slightly from pre- to post-treatment for PD patients ($p = 0.03$; Fig. 4e), but should be reflected based on a complete absence of PD-1$^+$ bulk CD8 T cells prior to treatment initiation. In the VART population only, we observed a marginal, non-significant increase in T cell activation by Ki67$^+$ and PD-1$^+$ in the PD group (Fig. 4e). It is evident that several patients in the disease control group have elevated levels of PD-1$^+$ CD8 T cells, especially within 'bulk CD8' and 'NART', prior to therapy (bulk CD8 pre-treatment $p = 0.0089$; NARTs pre- $p = 0.02$; Fig. 4f). A similar enhanced level of both Ki67$^+$ and PD-1$^+$ 'bulk CD8' and CD39$^+$, Ki67$^+$ and PD-1$^+$ NARTs is observed at 3wks post-treatment, although only significant for PD-1$^+$ NARTs ($p = 0.028$; Fig. 4f). Interestingly, we observed that triple-positive Ki67$^+$ PD-1$^+$ CD39$^+$ CD8 T cells (Bulk, NARTs



or VARTs) were completely absent in the PD group (Fig. 4g). Lastly, at 3 weeks post-treatment, up to 60% of NARTs are Ki67^{+}CD45RA^{-} cells and tend to constitute a larger subpopulation for patients with disease control (Supplementary Fig. 6b), with the majority of the cells being CD27^{+} rather than CD57^{+}, implying a state of activation rather than terminal differentiation (Supplementary Fig. 6c). No difference in frequencies of terminally differentiated, CD57^{+}CD45RA^{+}GzmB^{+} triple-positive cells were seen between patient groups or during treatment (Supplementary Fig. 6e, f). Detectable NARTs primarily comprise of T_{Em} cells, with

Fig. 4 Phenotypic characterisation of bulk CD8 T cells, NARTs and VARTs. a–b UMAPs of bulk concatenated CD8 + T cells and NARTs at (**a**) pre-treatment and (**b**) 3-week post-treatment for patients with SD/PR/CR ($n = 8$ patients) and PD ($n = 4$ patients). **c** Expression of phenotype markers on UMAP. Cells expressing similar parameters are clustered based on expression patterns. **d–f** Parent population frequencies of selected and (**g**) for triple-positive Ki67$^+$ PD-1$^+$ CD39$^+$ bulk CD8, NART and VART subpopulations for each patient, either at pre- ($n = 14$ patients) and 3 weeks post-treatment ($n = 14$ patients) and between patient with SD/PR/CR ($n = 10$ patients) and PD ($n = 4$ patients). **h–i** CD39$^+$ frequencies for bulk CD8 T cells, NART, and VART subpopulations at pre- and post-treatment for patients with SD/PR/CR ($n = 10$ patients) and PD ($n = 4$ patients). **j–p** Representative example of flow contour plots (5%, outliers shown) of key parameters for the NART subpopulation at three-week post-treatment for patient #2389 (DCB; Best RECIST 1.1 CR). **q** Ki67$^+$, CD39$^+$, and PD-1$^+$ subpopulation frequencies for bulk CD8 T cells, NARTs, and VARTs at pre-treatment ($n = 14$ patients) and at three weeks ($n = 14$ patients) and 230+ weeks post-treatment ($n = 6$ patients). For (**d–g**) groups were compared using non-parametric two-sided Mann–Whitney test and Kruskal–Wallis Dunn's multiple comparison test for (**h**) + (**i**). For (**d–i**), data is presented as median values ± largest/smallest value within upper/lower quartile ± 1.5 IQR. NS. Not Significant, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. wk: week, Pre: Pre-treatment, Post: Post-treatment. In (**d**), (**e**) and (**g**) Post: three weeks post-treatment. Source data are provided as a Source Data file.

a smaller fraction of Naïve NARTs (Supplementary Fig. 6g). This indicates that T cell recognition based on NARTs from the naïve repertoire are also captured to some extent.

Importantly, CD39 seems to be the parameter that best differentiates the PBMC NART population from the VART population (pre- $p = 0.007$, post-treatment $p = 0.018$; Fig. 4h). The CD39 expression is particularly evident for NARTs, both pre and post-treatment, in the patients with disease control (pre-$p = 0.013$, post-treatment $p = 0.041$; Fig. 4i), indicating recent antigen exposure for this group of T cells. Previously, CD39 has been seen to differentiate tumor specific CD8$^+$ tumor infiltrating lymphocytes (TILs) from bystander TILs[27]. Interestingly, PD patient #7577 expresses higher frequencies of Ki67$^+$ cells than the remaining PD patients, but with few CD39$^+$ NARTs detected both pre- and post-treatment, suggesting lack of antigen recognition for such NARTs, despite proliferation following ICB (Supplementary Fig. 6c, d).

Examples of the above NART subpopulations at three weeks post-treatment is shown in Fig. 4j–p (patient #2389, Best RECIST 1.1 CR). For patients with long-term clinical response, we further evaluated PBMCs at a late time-point post-treatment (231+ weeks post-treatment, $n = 6$). The majority of these patients experienced an initial burst in the frequencies of Ki67$^+$, PD-1$^+$ and CD39$^+$ subpopulations, which declined to pre-treatment levels at the late time-point evaluated (Fig. 4q).

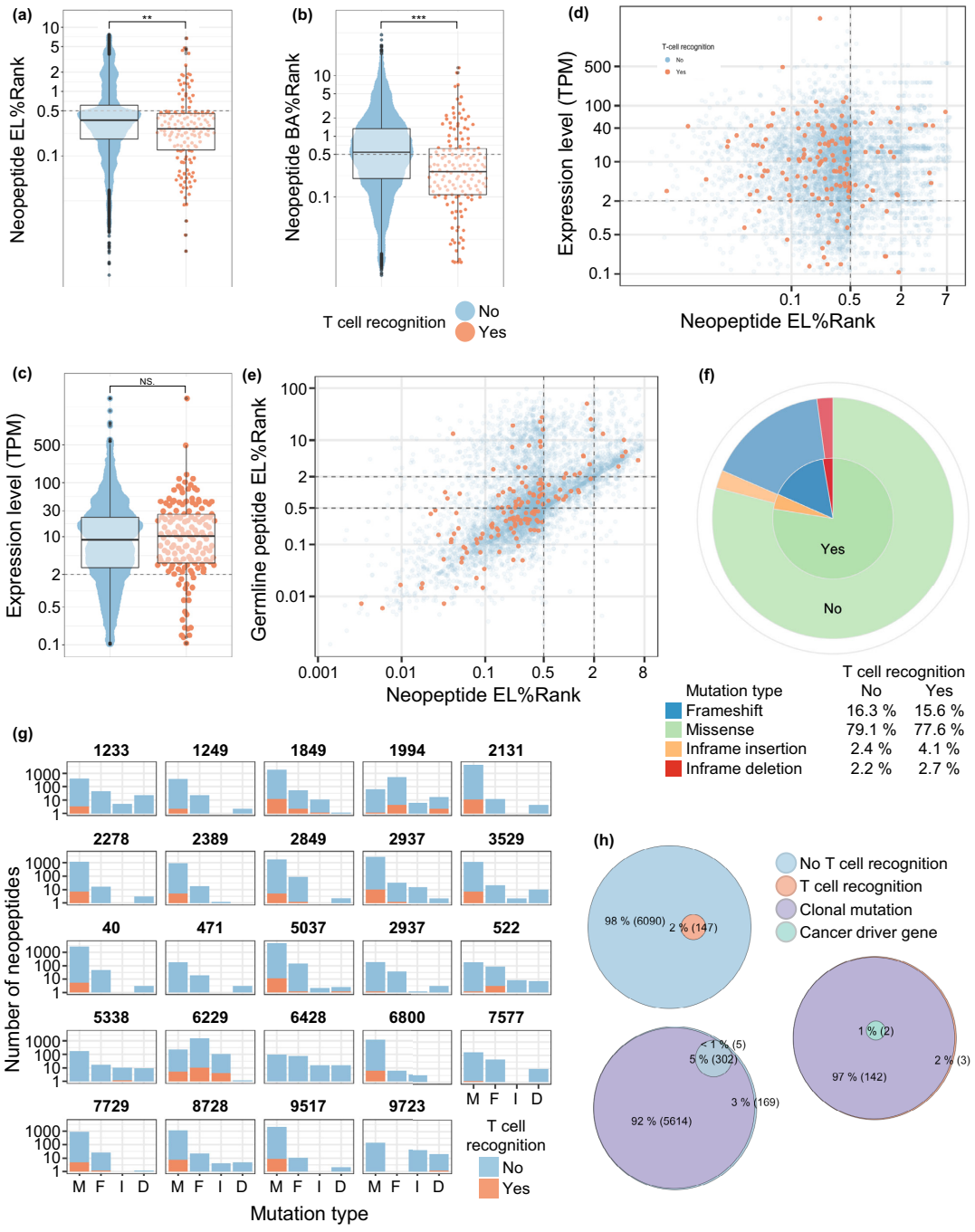
Taken together, a proliferative burst of NARTs is observed following a single dose of PD-L1 blockade, which has been noted previously for bulk CD8 T cells[13]. During this burst, NARTs in patients with disease control tend to be in a Ki67$^+$ state, and mostly of a CD45RA$^-$, PD-1$^+$ phenotype, favoring CD27 expression over CD57. Furthermore, this NART subpopulation can be identified in PBMCs based on CD39 expression.

Eluted ligand rank-score is a key correlate for neoepitope recognition by CD8$^+$ T cells. It is of key interest to precisely predict which tumor neoepitopes are recognized by T cells. Hence, we evaluated a number of features that may impact the likelihood for T cell recognition. T cell recognized neoepitopes had lower percentile rank both related to EL%Rank ($p = 0.0016$) and binding affinity prediction (BA%Rank, $p < 0.0001$), whereas neopeptide-related gene expression level did not differ between T cell recognized and non-recognized neopeptides ($p = 0.73$; Fig. 5a–c). An enrichment of T cell recognized neoepitopes was observed for predicted neopeptides with EL%Rank < 0.5 and expression level >2 TPM ($p < 0.001$; Fig. 5d), indicating that gene-expression level in combination with EL is relevant in predicting immunogenic neoepitopes in this cohort. Furthermore, there were no differences in improved- or conserved HLA-binders, as both where equally represented in the T cell recognized fraction ($p = 0.19$, Fig. 5e)[28]. Interestingly, we did not observe improved T cell recognition of neopeptides derived from certain classes of mutations ($p =$ NS Fig. 5f). However, a substantial pool of predicted neopeptides derived from non-missense mutations elicited a T cell response, which was also seen in patients with renal cell carcinoma (RCC)[29]. Such non-missense mutations where unevenly distributed, but present in the majority of evaluated patients (Fig. 5g)

Recently, clonality of predicted neopeptides has been associated with T cell immune reactivity following ICB, with peptides derived from clonal mutations dominating the elicited T cell responses following PD-1- and CTLA-4 blockade[30,31]. To define favorable characteristics of CD8$^+$ T cell-reactive neopeptides, we investigated their clonality, gene origin, expression level, and HLA-binding affinity. Peptides derived from clonal mutations make up the majority of neopeptides included in the libraries (5,756 of 6,237; 92%), similar to what has previously been seen in NSCLC[30]. Of 6,237 neopeptides screened in this cohort, 148 unique neopeptides were observed to elicit a T cell response. Of these, 143 neopeptides were derived from clonal mutations, with two originating from cancer driver genes (GPC3 and MAML2). No difference in T cell recognition was observed towards clonal or non-clonal neoepitopes ($p = 0.4$), but that may be due to the large fraction of clonal mutations observed in this cohort. Also, no preference for cancer driver genes amongst recognized peptides was observed ($p = 0.066$) and NARTs recognized peptides from a multitude of non-classical cancer driver genes in this cohort ($n = 120$ genes; Fig. 5h). Together, these results point to the importance of HLA-peptide binding affinity in successfully predicting immunogenic neoepitopes as well as a potential influence of neoantigen expression.

Pre-treatment TME mRNA gene signatures are associated with post-treatment NART repertoire and phenotypic characteristics. Having established the characteristics of NART responses, we subsequently analyzed mRNA expression patterns in pre-treatment tumors to evaluate the potential determinants in the tumor microenvironment (TME) driving post-treatment NART response. To better understand the composition of the TME, we applied differential expression analysis (DEA) to determine overly expressed genes and used Microenvironment Population Counter (MCP) to estimate immune cell population abundancy in the TME.

From the pre-treatment RNA-seq data, differentially expressed genes were assessed based on the number of detected NART responses at three weeks post-treatment (high >= median no. of responses of three; Fig. 6a). Interestingly we observed a strong clustering of patients based on the TME gene expression patterns when evaluated based on the differentially expressed genes - patients with high NART response numbers tended to have a higher mRNA expression of genes such as CD3D, PPARG, and TNFSF15, involved in T cell activation and differentiation[32–36], indicating a T cell stimulating pre-treatment TME in patients with high number of post-treatment NART responses. In



contrast, for patients with lower number of NART responses, we observed high expression of FN1, ITGA5 and COL3A1, which is correlated with poor survival for cancer patients[37–39]. Additionally, CXCL5, involved in angiogenesis[40,41], was increasingly expressed both across patients with low number of NART responses and PD patients. Furthermore, for patients with high NART responses post-treatment, the gene ontology (GO) enrichment analysis of the DEA results showed pre-treatment enrichment of gene sets associated with the TCR complex and antigen presentation, potentially facilitating an improved NART response post-treatment (Fig. 6b, c).

Applying the MCP-counter method, patients with a high number of detected NART responses post-treatment had a higher expression of genes that together define various subtypes of immune cells,

Fig. 5 Molecular characteristics of T cell neoepitopes. a–c EL%Rank score, BA%Rank score and expression level of predicted neopeptides, grouped according to T cell recognition (Yes, $n = 147$ neopeptides/No, $n = 6,090$ neopeptides). Dotted lines indicate split for groups in z-test, applied in Fig. 5d. **d** Expression levels of neopeptide-origin genes and predicted EL%Rank score of neopeptides. Colored according to T cell recognition of neopeptides. **e** Peptide EL%Rank scores of wild type (normal) and mutant neopeptides, colored according to T cell recognition. **f** Proportions of neoepitopes eliciting T cell recognition or not, grouped according to neopeptide-induced type of mutation. **g** Distribution of predicted neopeptides and neoepitope recognition, grouped according the type of peptide-inducing mutation across patients (M = Missense, F = Frameshift, I = Inframe insertion, D = Inframe deletion). **h** Venn-diagram of non- and T cell recognized predicted neopeptides and if they are derived from clonal mutations or cancer-driver genes. For (**a–c**) groups were compared using non-paired *t*-test and data is presented as boxplot ranged by 25th and 75th quartiles with median values, and whiskers as ±largest/smallest value within upper/lower quartile ± 1.5 IQR. NS. Not Significant, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Source data are provided as a Source Data file.

including bulk T cells ($p = 0.042$) and with a trend for CD8 T cells ($p = 0.066$), while exhibiting lower levels of fibroblasts, indicating a less inhibitory TME[42] ($p = 0.0005$; Fig. 6d–h). Together, pre-treatment TME patterns are associated with the post-treatment NART response repertoire, with a pre-treatment T cell activating mRNA signature in the tumor within patients that tended to raise a broader NART response post-treatment.

Discussion

This study utilized a high-throughput screening approach to serially interrogate CD8 T cell recognition of patient-specific neopeptides predicted from the pre-treatment tumor mutagenome in the peripheral blood of 24 patients with mUC treated with anti-PD-L1-therapy. Several findings are important. First, we observed an increase in NART responses from pre-treatment to three weeks post-treatment in patients with disease control. At this time point, the overall neoepitope recognition breadth, not the estimated frequency of such CD8 T cell populations, was associated with clinical radiographic response. This may reflect the finding that the majority of NART responses are low frequent. Second, phenotypic characterization of NARTs revealed an association between Ki-67$^+$ PD-1$^+$ NARTs three weeks post-treatment and clinical outcome. Third, CD39 was expressed on a higher fraction of the NARTs compared with VARTs in the blood, suggesting this marker could be used to identify anti-tumor T cells. The same difference was not observed in bulk CD8$^+$ T cells, which may include additional tumor-antigen specific T cells not captured in our screening for neoepitope recognition. Fourth, TME mRNA expression patterns pre-treatment were associated with increased NART responses three weeks post-treatment. Finally, in silico modelling of neoepitope prediction partially recapitulated the T cell recognition of NARTs, demonstrating peptide HLA-binding, and mutation gene-expression to affect neoepitope T cell recognition.

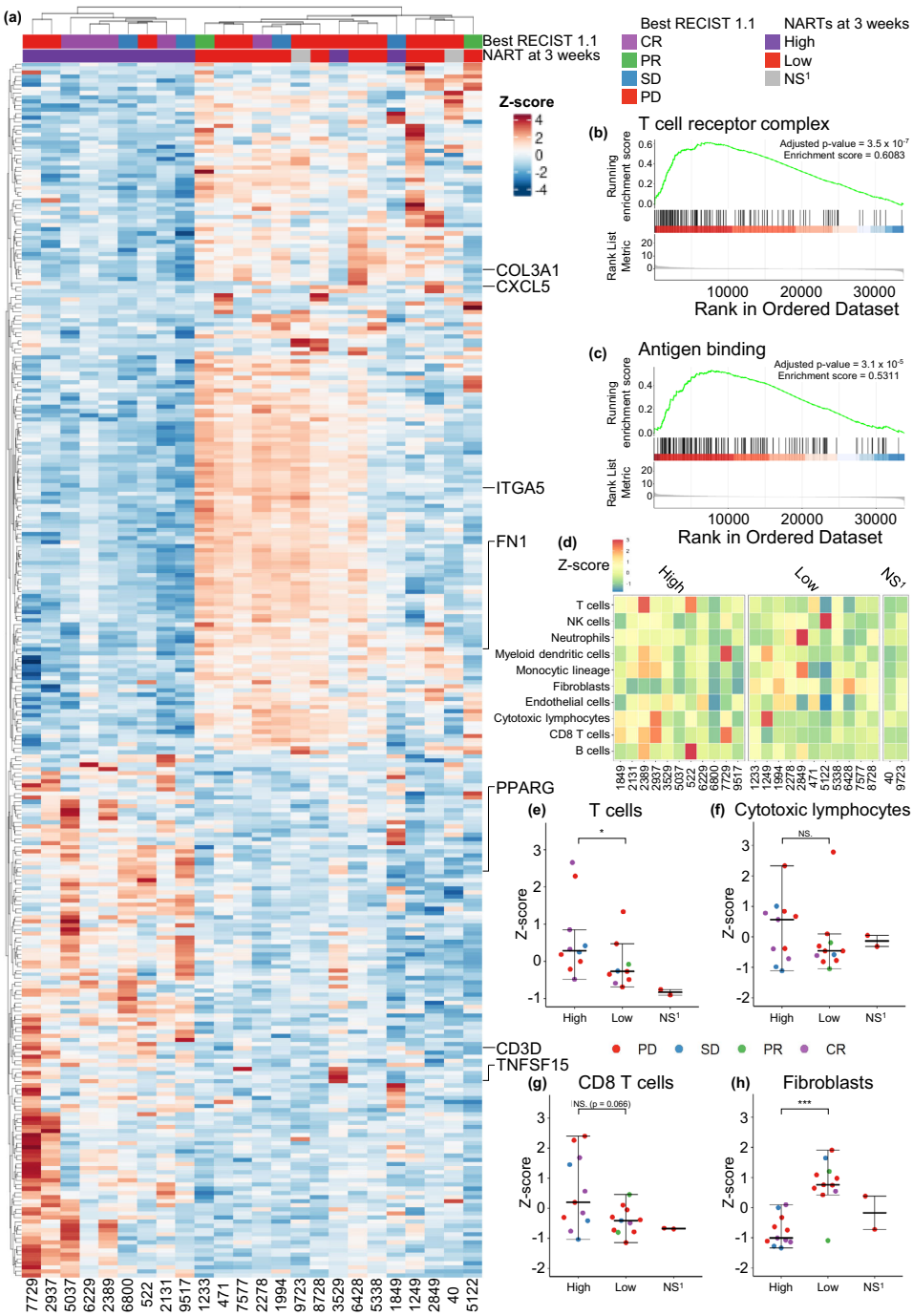
The apparent kinetics of NART responses detected in our study are consistent with a growing body of literature indicating that ICB rapidly induces an immunological T cell response in the peripheral blood, where multiple reports describe early peripheral T cell turnover, expansion and activation after ICB initiation within 7–21 days[12–17,19,43]. Both effector T cell expansion[14] and Ki67+ PD-1+ CD8 T cell increase[13] in the periphery three weeks post-treatment have been linked to clinical outcome to ICB. Our findings provide novel insight, as they shed light on the specificities and temporal dynamics of neoepitope-recognizing CD8 T cell responses through the detection and quantification of circulating NARTs following ICB. Yet whether emerging NARTs are truly de-novo primed or present pre-treatment at sub-detection levels in periphery cannot be deduced, although recent studies have suggested ICB-induced de-novo recruitment of naïve tumor-specific T cells from the lymph nodes[44].

It should also be noted that the median time to response for patients included in this cohort is 2.1 months (95% CI 2.0–2.2)[5] consistent with the time of the first radiographical disease assessment, comparable to other ICB clinical trials in mUC patients (median time to response 1.4–2.1 months)[45–50]. Consequently, any earlier tumor reduction prior to first scan is not measured, but the rapid time to response indicates that clinical response to therapy occurs early. Hence, these observations are consistent with our findings of early proliferation, as most response to therapy are captured at 9 weeks post-treatment.

Furthermore, interrogating NARTs revealed kinetics that were not visible from investigating bulk TCR-seq. These NART responses were personal to each patient under study and persisted, albeit at lower levels than the initial 3 weeks peak, for over 5 years post-treatment initiation in some long-term responders. NART responses shared between patients were not identified in this study.

The ability to in silico model the T cell neoantigen responses induced in vivo, as well as the clinical relevance of in silico modeling of antigenic diversity beyond association of TMB and neopeptide prediction to clinical outcome, are areas of active investigation. The importance of gene expression level for neoantigen quality is previously described, yet thresholds remain undefined[51], in contrast to the established importance of HLA-I binding affinity[28]. We observed that the majority of NART populations recognized neopeptides with EL%Rank <0.5 and expression level >2 TPM, and that analyses based on only these responses also were associated with clinical outcome (Supplementary Fig. 4b–d). However, a multitude of NART responses towards neopeptides outside these specifications was detected, and could likewise serve as important targets for anti-tumor immunity. Although the combined EL%Rank scores and expression level provide the best parameters for neopeptide T cell reactivity here, these results also suggest that future neoepitope prediction and selection should include additional parameters, such as tumor immunogenicity, immune priming, and peptide sequence similarity to known self- and infection-derived antigens, as incorporated in recently described neoantigen fitness models[52–56]. At the initiation of study, neopeptides were selected only based on HLA-binding affinity and expression level >0.1 TPM, limiting potential bias in prediction and selection of immunogenic neoepitopes. Recently, selection based on such key characteristics were supported, but also revealed the persistent challenge in fully defining the parameters critical for high accuracy neoepitope prediction[57]. Lastly, the results propose that predicted neopeptides outside these current thresholds should be included in future NART screenings.

Interestingly, we observed that mRNA-expression patterns in pre-treatment tumors differ between patients mounting a wider post-treatment NART response and those that do not. In particular, high T cell infiltration seems to be important for the generation of NARTs. Although needing validation, our results also suggest that pre-treatment TME mRNA gene expression patterns may also be useful when predicting NART responses. Also of interest, clinical outcome may potentially be driven by a combined favorable pre-treatment TME and induced post-treatment NART repertoire, requiring further interrogation.



Phenotypic characterization of NARTs allowed for important observations in the peripheral blood that have heretofore been made mostly in the tumor microenvironment. Recently, in NSCLC patients, activated progenitor-like (TCF-1$^+$ PD-1$^+$) T cells with proliferative capacities have been observed, highlighting the likely importance of this cell type to the anti-tumor response[58]. Our analyses revealed that a similar PD-1$^+$ Ki67$^+$ NART population was detected post-treatment primarily in patients who derived benefit from therapy. Although phenotypic changes to some extent seemed to be antigen-independent, our results still suggest that the reactivity and proliferative tendencies of NARTs in the peripheral blood contribute to tumor clearance

Fig. 6 Transcriptomic analysis of TME related to the level of NARTs post therapy. a Differentially expressed genes ($n = 295$) from DEA of all patient genes, related to high versus low NART responses at three weeks post-treatment. **b**, **c** Two significant gene sets from the GO enrichment analysis, antigen binding and T cell receptor complex, respectively. **d** Heatmap of immunological cell signatures across patients, grouped by high or low number of NART responses at 3-week post-treatment, and (**e**–**h**) associations towards high ($n = 11$ patients) or low number ($n = 11$ patients) of NART responses at 3-week post-treatment for (**e**) T cells, **f** CTLs, **g** CD8 T cells, and (**h**) Fibroblasts. No three weeks post-treatment PBMCs were available from patients #40 and #9723, hence these patients were not included in the analyses or comparisons. For (**a**) and (**d**–**h**) NS[1] denotes patients not screened at three weeks post-treatment due to sample unavailability. For e)-h) groups were compared using non-parametric two-sided Mann–Whitney test and data is presented as median values ± largest/smallest value within upper/lower quartile ± 1.5 IQR. NS. Not Significant, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Source data are provided as a Source Data file.

and that an activated T cell profile is associated to favorable clinical outcome. Further, in lung and colorectal tumors, CD39 expression has been utilized to distinguish between bystander and NARTs in the microenvironment[27]. We demonstrate here that CD39 expression may also helpful identify NARTs from bystander T cells in the blood without the need to procure and immediately process fresh tumor tissue.

There are notable limitations to our study. First, the sample size was small. However, the screening of the 24 patients under study was unselected and comprehensive, with 200-587 neopeptides included per patient yielding a total of 6237 neopeptides. Patient samples were also longitudinally collected and screened, which allowed for detection of NART responses for up to 5+ years post-treatment in some long-term responders. Second, we did not evaluate the contribution of MHC-II restricted peptides and the role of CD4 T cells, which were recently shown to mediate and drive anti-tumor cytotoxicity and immunogenicity following ICB, also in UC[59–61]. Finally, we did not perform on-treatment biopsies and were therefore unable to evaluate NART trafficking and tumor immunoediting during treatment.

Of additional interest, the T cell neoepitope recognition profile of two patients with partial response to therapy differed significantly from the other patients with disease control. Patient #5122 (PFS/OS = 1932, best RECIST 1.1 PR) differed from other patients with long-term progression-free survival in that the number of NART responses decreased from baseline and maintained no detectable NART responses throughout treatment. However, the patient was observed to harbor a *PDL1* gene amplification, which has been associated with positive outcomes following ICB treatment[62]. Patient #1233 displayed NART dynamics and phenotype as well as a pre-treatment TME mRNA expression pattern associated with PD in the larger cohort; thus, no such characteristics explain the apparent clinical response. These cases indicate the complexity of clinical response and resistance to ICB.

Despite these limitations, our observations add to the current understanding of anti-tumor immunity induced by ICB and suggest that important insights into NART dynamics can be made from pre-treatment archival tumors and post-treatment peripheral blood interrogation alone; NARTs are indeed detectable in the periphery and to a higher degree among patients that benefit from therapy. These findings warrant further investigation, to both improve ICB clinical outcome prediction and to investigate the mechanistic underpinnings of ICB, whether pre-existing NARTs rise to a detectable levels[63] or if NARTs are recruited and activated as a result of ICB[44,64,65].

Methods and Materials

Study design and participant samples. Patients had mUC and were treated with atezolizumab 1200 mg intravenously (IV) every 21 days at Memorial Sloan Kettering Cancer Center as part of the IMVigor210 trial[5]. Cross sectional imaging was performed every 9 weeks for the first 12 months following cycle 1 day followed by every 12 weeks thereafter. Best overall response was determined by radiologic assessment of response, using RECIST version 1.1. In patients in whom clinical progression was determined based on symptoms and decline and functional status, this determination superseded radiologic categorization. Patients that were only screened at baseline were classified as with progressive disease. All patients provided written informed consent to both the IMvigor 210 trial and an Institution Review Board-approved biospecimen protocol permitting tissue and blood collection, sequencing, and correlative studies.

Twenty-four patients included in a previously published multi-omic analysis[19] were the subject of this study (Supplementary Fig. 1). PFS and OS were updated for this cohort. Patients with PFS > 6 months ($n = 9$; PFS < 6 months $n = 15$) were stated as having DCB from treatment, with other outcomes data in the supplement (Supplementary Table 2). Patient formalin-fixed paraffin-embedded (FFPE) tumor and PBMC samples were obtained and prepared as previously described[19] . Blood samples were drawn from patients prior to IV infusion on the day of treatment, pre-treatment and during treatment, and PBMCs were isolated and cryopreserved at −150 °C in Human Serum Albumin (HSA)/10% DMSO until analysis.

HD samples were collected by approval of the local Scientific Ethics Committee, with donor written informed consent obtained according to the Declaration of Helsinki. HD blood samples were obtained from the blood bank at Rigshospitalet, Copenhagen, Denmark. All samples were obtained anonymously. PBMCs from HDs were obtained from whole blood by density centrifugation on Lymphoprep (Axis-Shield PoC, cat# 1114544) in Leucosep tubes (Greiner Bio-One, cat# 227288) and cryopreserved at −150 °C in fetal calf serum (FCS, Gibco, cat#10500064) + 10% dimethyl sulfoxide (DMSO, Sigma-Aldrich, cat#C6164).

WES, RNA- and TCR-seq, HLA typing and Next-Generation Sequencing data processing. Patient WES and RNAseq data, HLA typing and TCRβ CDR3 region amplification was acquired as described as part of a previous study on the patient cohort[19]. In total, 22 patients had tumor and PBMC material of sufficient quality for both WES, RNA-seq, TCRβ-seq and minimum one pretreatment PBMC sample, as tumor TCR analysis was not performed on patient tumor samples from patients #522 and #6800 due to failed sequencing quality control.

Novel for this study, raw FASTQ files from WES and RNAseq were analyzed in the following manner. First, both data sets were pre-processed for quality using Trim Galore version 0.4.0[66], which combines the functions of Cutadapt[67] and FastQC 0.11.2:[68] trimming the reads below an average Phred score of 20 (default value), cutting out standard adaptors such as those from Illumina, and running FastQC to evaluate data quality. Variant calling was performed following the Genome Analysis Toolkit (GATK) best practice guidelines for somatic variant detection[69]. Reads were aligned to the human genome (GRCh38) using the Burrows-Wheeler Aligner[70] version 0.7.15:q with default mem options and with a reading group provided for each sample for compatibility with the following steps. Duplicate reads were marked using Picard-tools version 2.9.1 MarkDuplicates. Base recalibration was performed with GATK version 3.7 to reduce false-positive variant

calls. SNV and indel calls were made using GATK version 3.8's build in a version of MuTect2[71] designed to call variants, both SNVs and indels, from matched tumor and normal samples. Kallisto 0.42.1[72] was used to determine the gene expression in transcript per million (TPM) from RNAseq data.

Neopeptide prediction and selection. The VCF output files from GATK's MuTect2 was given as input to the neopeptide predictor MuPeXI version 1.1.3[23] together with RNAseq expression values obtained from Kallisto. HLA alleles of each patient were inferred from the WES data using OptiType version 1.2[73] with default settings after filtering the reads aligning to the HLA region with RazerS version 3.4.0[74]. Identified mutations were used to predict 9, 10, and 11 amino acid peptides, sorted according to the EL% Rank score of the mutated neopeptides using NetMHCpan 4.0[24]. All HLA-I-feasible neopeptides with an EL%Rank score <0.5 with expression level >0.1 TPM were selected for peptide synthesis. For patients where the number of HLA-I-feasible predicted neopeptides did not exceed 200, the highest-ranking 200 peptides were selected to constitute the patient neopeptide library.

Peptides. All selected mutation derived and virus control peptides were ordered and purchased from Pepscan (Pepscan Presto BV, Lelystad, Netherlands). Peptides were dissolved to 10 mM in DMSO following arrival and stored at −18 °C prior to use.

MHC monomer production and generation of peptide-MHC complexes. The production of MHC monomers was performed as previously described[75,76]. In brief, MHC class I heavy chains and human β2m were expressed in E.coli strain BL21(DE3) pLysS (Novagen, cat#69451). Inclusion bodies containing expressed proteins were harvested by washing in detergent buffer and wash buffer and solubilized in 8 M urea buffer (8 M Urea, 50 mM K-HEPES pH 6.5, and 100 µM β-mercaptoethanol). Final purified inclusion bodies were stored at −80 °C until used. MHC class I molecules were obtained by in vitro folding of heavy chain and β2m light chain with respective UV-sensitive ligand in folding buffer (0.1 M Tris pH 8.0, 500 mM L-Arginine-HCl, 2 mM EDTA, 0.5 mM oxidized glutathione and 5 mM reduced glutathione) at 4 °C[77,78] or by using disulfide-stabilized empty MHC I complexes as previously reported[79]. After folding for 3–5 days, folded protein was biotinylated using BirA biotin-protein ligase standard reaction kit (Avidity, LLC- Aurora, Colorado). Finally, biotinylated monomer complexes were purified with size-exclusion column (Waters, BioSuite SEC Column, 125 Å, 13 µm SEC, 21.5 mm × 300 mm) with HPLC (Waters Corporation, USA)) and stored at −80 °C until further use. Specific pMHC complexes were generated by UV-induced peptide exchange[75,77].

Detection of peptide-MHC specific T cells by DNA barcode-labelled multimers. Patient-specific libraries of predicted neopeptides and virus control peptides (size 201-589 peptides per patient) were generated as DNA barcode-labelled pMHC multimers as previously described[22]. In short, patient specific neoepitope MHC multimer libraries were generated by multimerizing exchanged pMHC molecules on a PE-labeled polysaccharide-backbone (for neopeptides), and APC-labeled polysaccharide-backbone (for virus-derived epitopes) coupled to DNA barcoded-labeled dextran backbone Thus, a specific peptide is linked to a unique DNA barcode together with a fluorescent label, serving as a tag for the given pMHC. Patient and HD PBMCs were stained with an up-concentrated pool of multimers together with 50 nM dasatinib. Samples screened only for T cell multimer recognition were stained with an antibody mix consisting of CD8-BV480 (BD, cat. #566121, clone RPA-T8, 2 µl), dump channel antibodies (CD4-FITC (BD, cat. #345768, 1.25 µl), CD14-FITC (BD, cat. #345784, 3.125 µl), CD19-FITC (BD, cat. #345776, 6.25 µl), CD40-FITC (Serotech, cat. #MCA1590F, 2.5 µl), and CD16-FITC (BD, cat. #335035, 1.56 µl)) and a dead cell marker (LIVE/DEAD Fixable Near-IR; Invitrogen, cat. #L10119, 0.1 µl). Multimer binding T cells were sorted as lymphocytes, single, live, $CD8^+$, $FITC^-$ and PE^+. Samples screened for T cell multimer recognition and phenotypic characterization were stained with an antibody mix composed of T cell lineage markers (CD3-BV786 (BD, cat. #563799, clone SK7, 5 µl), CD4-BV650 (BD, cat. #563876, 2.5 µl), and CD8-BV480 (BD, cat. #566121, clone RPA-T8, 2 µl)), characterization markers (Ki67-BUV395 (BD, cat. #564071, clone B56, 2.5 µl), 4-1BB-BUV737 (BD, cat. #741867, clone 4B4-1, 2.5 µl), PD1-BV421 (BioLegend, cat. #329920, clone EH12.2H7, 3 µl), CD27-BV605 (BioLegend, cat. #302829, clone O323, 2.5 µl), CD45RA-BV711 (BD, cat. #563733, clone HI100, 2.5 µl), CCR7-FITC (BioLegend, cat. #353215, clone G043H7, 5 µl), Eomes-PerCP-eFlour710 (eBioscience, Thermo Fisher Scientific cat. #46-4877-41, clone WD1928, 2.5 µl), CD39-PE-CF594 (BD, cat. #563678, clone Tu66, 2.5 µl), CD57-PECy7 (BioLegend, cat. #393309, clone QA17A04, 2.5 µl), and GranzymeB-AlexaFlour700 (BioLegend, cat. #372221, clone QA16A02, 1.25 µl)), and a dead cell marker (LIVE/DEAD Fixable Near-IR; Invitrogen, cat. #L10119, 0.1 µl). Multimer binding CD8 + T cells were sorted as lymphocytes, single, live, $CD3^+$, $CD8^+$, $CD4^-$ and either PE^+ or APC^+ on either FACSAria™ Fusion or FACSMelody™ instruments (BD Biosciences). All sorted T cells were pelleted by centrifugation. From isolated cells and from a stored baseline aliquot of multimer pool (diluted 10,000x in final PCR reaction), the specificities of $multimer^+$ $CD8^+$ T cells were decoded by amplification, subsequent purification using QIAquick PCR Purification kit (Qiagen, cat. #28104), and ultimate sequencing of DNA barcodes at PrimBio Research Institute (PA, USA) using an Ion Torrent PGM 316 or 318 chip (Life Technologies). Sequencing data were processed by the software package Barracoda, available online at (https://services.healthtech.dtu.dk/service.php?Barracoda-1.8). The tool identifies the DNA barcodes annotated for a given experiment, assigns a sample ID and pMHC specificity to each DNA barcode, and counts the total number of reads and clonally reduced reads for each peptide-MHC-associated DNA barcode. Log_2FC in read counts mapped to a given sample relative to the mean read counts mapped to triplicate baseline samples are estimated using normalization factors determined by the trimmed mean of M-values method. FDRs were estimated using the Benjamini–Hochberg method. A minimum read count fraction of 0.1% for a given DNA barcode of the total DNA barcode number in that given sample was set as threshold to avoid false-positive detection of T cell responses due to low number of reads in the baseline samples. DNA barcodes with FDR <0.1% (corresponding to $p < 0.001$), read count fraction > (total read count for full barcode library/barcode library size), $Log_2FC > 2$ over the input values for the total pMHC library, and CD8+ T cell estimated frequency of >0.01 % were considered to be true T cell responses, based on previous studies[22,29,80].

Detection of peptide-MHC specific T cells by fluorescently-labelled tetramers. For selected neopeptides, pMHC tetramers were generated for staining of neoepitope-specific T cells. Neopeptides were selected based on the observed NART responses from the DNA barcode-labelled multimer screening. Following the observed increase in NART responses at 3 weeks post-treatment, NART responses in 3-week post-treatment PBMC samples were interrogated wherever sufficient patient material remained, otherwise 9-week post-treatment samples were analyzed. Single-fluorochrome pMHC specificity tetramers using were generated as described in detail previously[81,82], using a library of streptavidin (SA)-conjugated flourochromes consisting of PE-SA (BioLegend, cat. #405204), APC-SA (BioLegend,

cat. #405207), BV421-SA (BD, cat. #563259), PE-Cy7-SA (BD, cat. #557598), BV605-SA (BD, cat. #563260), PE-CF594-SA (BD, cat. #562284), BV650-SA (BD, cat. #563855), BUV395-SA (BD, cat. #564176). Up to eight patient-specific pMHC tetramers per sample were investigated. PBMC samples were stained with respective library of pMHC tetramers and with an antibody mix consisting of CD8-BV480, dump channel antibodies and a dead cell marker, as above. Tetramer-specific T cells analyzed as lymphocytes, single, live, $CD8^+$, $FITC^-$ and $tetramer^+$ cells. Due to staining strategy, $tetramer^+$ cells were gated by being $CD8^+$.

Analytical processing of detected T cell responses. For each patient, patient PBMC samples at timepoints pre-, during- and post-treatment were stained with the respective patient pMHC multimer library and relevant fluorescent antibodies. Concurrent with patient PBMC samples, PBMCs from HDs, HLA-matched with the respective patient as best possible, were also stained with the patient's pMHC multimer library and CD8 T cell subset identification antibodies (one to three HD samples per staining, median = 2). Presence of NARTs was determined based on the enrichment in barcode reads for a given neoepitope, visualized as the Log_2FC for each pMHC specificity in each patient sample, longitudinally evaluated over the course of treatment (Fig. 2a). For all patient and HD sample screenings, samples were stained with the entire multimer library. Predicted for and included in the screenings, the library of neopeptides presented on HLAs C0202 and C0501 ($n = 515$) were subsequently excluded from downstream analysis due to observations of substantial unspecific binding of these HLAs to Killer-cell immunoglobulin-like receptor (KIR). Furthermore, to avoid signals from potential pMHC-elements with unspecific binding, all HLA-matching patient-derived NART responses detected in HD samples were excluded from the pool of patient T cell responses, if the given neopeptide response was detected across all samples for a given patient. Of 6237 screened neopeptides in all patient and HD PBMC samples, 28 neopeptides generated background signals, and thus was excluded from the final library of NART responses.

For any pMHC-coupled DNA barcode in a sample, an estimated frequency of each pMHC-recognizing $CD8^+$ T cell population was estimated based on the read count fraction of the given DNA barcode out of the total fraction of T cells binding the pMHC multimer pool. Estimated frequencies for all NART responses were summed-up for each patient and timepoint to determine the total frequency of NARTs, i.e. sum of estimated frequencies.

Analytical processing of phenotyping data. Flow cytometry results were analyzed using the FlowJo v10 software (TreeStar, Inc.)[83]. For UMAP dimensionality reduction, 3000 representative live, $CD3^+$, $CD4^-$, bulk $CD8^+$ T cells from patient samples pre- and 3 weeks post-treatment were concatenated ($n = 28$) and projected using the UMAP plugin in FlowJo[84]. UMAP was run by selecting the parameters for Eomes, GzmB, CD27, CD57, CD45RA, CCR7, CD39, PD1, and Ki67. 41BB-BUV737 was excluded from further analysis due to significant spectral overlap from CD45RA-BV711. For the FlowSOM algorithm for unsupervised clustering[85], 15 clusters were selected with otherwise default settings.

Differential expression analysis and microenvironment cell populations-counter. Differential expression analysis is performed with all genes where the output from Kallisto version 0.42.1 was used as input to DeSeq2[86] version 1.26.0 from BiocManger in R version 3.6.1 with default option. The median number of detected NART responses at the given time point was used to split the cohort in high versus low number of NART responses. Log-fold change >1 and <−1 together with an adjusted p-value < 0.05 was used as threshold for over- and under-expressed genes for the analysis. The heatmap illustrations were generated with ComplexHeatmap from Bioconductor[87]. The GO enrichment analysis is developed using R version 4.1.1 with the built in packages; enrichplot version 1.13.2[88], clusterProfier version 4.0.5[89] with Benjamin Hochberg at p value adjustment. Cell populations abundancy was estimated from bulk RNA sequencing data using Microenvironment Cell Populations-counter (MCP-counter)[90]. The expression matrix obtained from Kallisto was fed as input to ebecht/MCPcounter from GitHub in R version 4.0.2 with Hugo-symbols as feature Type.

Neopeptide clonality. Allele copy number, purity and ploidity were found using Sequenza version 3.0[91]. As input, bam files from normal and tumor were given to Sequenza-utils version 3.0 bam2seqz with CRCh38 as reference followed by Sequenza seqz_binding. To run the Sequenza copynumber call with CRCh38, the R packages copynumber[92] with minor modifications from Shixiang/copynumber[93] was applied. Sequenza results were generated with the Sequenza packages in R version (3.6.1) and copynumber information from Sequenza were merged with mutations file from Mutect2 and used as input to PyClone. To locate clonal mutations, PyClone version (0.13.0)[94] was applied with the best estimated cellularity given from Sequenza, and max cluster of 10 and minimum size of 0 to yield all possible mutations. Afterwards, clonal mutations were filtered with a cluster size of minimum 5 and cellularity of minimum 90.

Figures and statistical analysis. Figure 1a and 2a were created in Biorender. Graphs in Fig. 1b–d, 2b, c, 3a–i, 4d–i + q, 5a–h, 6d–h, Supplementary Fig. 2a–c, 4a–e +m, 5a–e and 6j–p were generated using the *ggplot2* package in R v3.6.1 and v4.0.2. For Fig. 3f–g and Supplementary Fig. 4a, to facilitate plotting on logarithmic scale, in cases where SEF was 0 due to no detected T cell responses, 0.001% was added to SEF. Groups in Fig. 1c–d, 3a, b, d, f, h, 4d–g, 5g and 6e–h, and Supplementary Fig. 4a–c + e, 5a–c +e and 6b +e–g were compared using non-parametric two-sided Mann–Whitney test, in Fig. 3c, e, g, i and 4h–i, and Supplementary Fig. 4d and 5d using Kruskal–Wallis Dunn's multiple comparison test, and in Fig. 5a–c using non-paired t-test. Mann–Whitney and *t*-test conducted using the *ggsignif* package[95] and Dunn's test using the *rstatix* package[96]. Proportion test (z-test)[97] was applied for Fig. 4b and 5d–f + h. The *eulerr* package[98] was applied for Fig. 5h. Figure 6a was generated from the *ComplexHeatmap* and Fig. 6b–c from *clusterProfiler* and *enrichplot* packages in R. All survival plots (Supplementary Fig. 4f-h) were made using Kaplan Meyer-curves and hazard ratios with the *survival* and *survivminer* packages in R[99]. Figure 4a–c + j–p and Supplementary Fig. 3 and 6a were generated from FlowJo v10. For all figures; NS. Not Significant, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All WES and RNAseq data is available upon application at dbGaP at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001743.v1.p1. GRCh38 reference genome is available at https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.39. The source data presented in figures are provided as Source Data file. All other relevant data are available from the authors upon request. Source data are provided with this paper.

Code availability

The applied codes are embedded into MuPeXI and Barracoda, both publicly available tools, and can be reached and accessed at https://github.com/ambj/MuPeXI and https://services.healthtech.dtu.dk/service.php?Barracoda-1.8, respectively. All other data and R

scripts to reproduce figures can be obtained from the corresponding author upon request.

Received: 19 January 2021; Accepted: 7 March 2022;
Published online: 11 April 2022

References

1. Bellmunt, J. et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N. Engl. J. Med* **376**, 1015–1026 (2017).
2. Motzer, R. J. et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N. Engl. J. Med* **373**, 1803–1813 (2015).
3. Garon, E. B. et al. Pembrolizumab for the treatment of non–small-cell lung cancer. *N. Engl. J. Med.* **372**, 2018–2028 (2015).
4. Ribas, A. Tumor Immunotherapy Directed at PD-1. *N. Engl. J. Med* **366**, 2517–2519 (2012).
5. Rosenberg, J. E. et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: A single-arm, multicentre, phase 2 trial. *Lancet* **387**, 1909–1920 (2016).
6. Hellmann, M. D. et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N. Engl. J. Med.* **378**, 2093–2104 (2018).
7. Samstein, R. M. et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* **51**, 202–206 (2019).
8. Cristescu, R. et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade–based immunotherapy. *Sci. (80-.).* **362**, eaar3593 (2018).
9. Rizvi, N. A. et al. Mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer. *Science (80-.).* aaa1348 (2015).
10. Tumeh, P. C. et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).
11. Gubin, M. M. et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* **515**, 577–581 (2014).
12. Kamphorst, A. O. et al. Proliferation of PD-1+ CD8 T cells in peripheral blood after PD-1-targeted therapy in lung cancer patients. *Proc. Natl Acad. Sci. USA.* **114**, 4993–4998 (2017).
13. Huang, A. C. et al. T-cell invigoration to tumour burden ratio associated with anti-PD-1 response. *Nature* **545**, 60–65 (2017).
14. Valpione, S. et al. Immune-awakening revealed by peripheral T cell dynamics after one cycle of immunotherapy. *Nat. Cancer* **1**, 210–221 (2020).
15. Im, S. J. et al. Defining CD8+ T cells that provide the proliferative burst after PD-1 therapy. *Nature* **537**, 417–421 (2016).
16. Wu, T. D. et al. Peripheral T cell expansion predicts tumour infiltration and clinical response. *Nature* 1–6. https://doi.org/10.1038/s41586-020-2056-8 (2020).
17. Kim, K. H. et al. The first-week proliferative response of peripheral blood PD-1+CD8+ T cells predicts the response to anti-PD-1 therapy in solid tumors. *Clin. Cancer Res.* **25**, 2144–2154 (2019).
18. Heeke, C. et al. The landscape of neoantigen-reactive CD8+ T cells impacts the clinical outcome in patients receiving adoptive transfer of tumor-infiltrating lymphocytes. Prep.
19. Snyder, A. et al. Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: An exploratory multi-omic analysis. *PLoS Med* **14**, 1–24 (2017).
20. Maher, V. E. et al. analysis of the association between adverse events and outcome in patients receiving a programmed death protein 1 or programmed death ligand 1 antibody. J. Clin. Oncol. JCO.19.00318 https://doi.org/10.1200/JCO.19.00318 (2019).
21. Fradet, Y. et al. Randomized phase III KEYNOTE-045 trial of pembrolizumab versus paclitaxel, docetaxel, or vinflunine in recurrent advanced urothelial cancer: results of >2 years of follow-up. *Ann. Oncol.* **30**, 970–976 (2019).
22. Bentzen, A. K. et al. Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat. Biotechnol.* **34**, 1037–1045 (2016).
23. Bjerregaard, A. M., Nielsen, M., Hadrup, S. R., Szallasi, Z. & Eklund, A. C. MuPeXI: prediction of neo-epitopes from tumor sequencing data. *Cancer Immunol. Immunother.* **66**, 1123–1130 (2017).
24. Jurtz, V. et al. NetMHCpan-4.0: Improved peptide–MHC class i interaction predictions integrating eluted ligand and peptide binding affinity data. *J. Immunol.* ji1700893 https://doi.org/10.1101/149518 (2017).
25. Postow, M. A. et al. Peripheral T cell receptor diversity is associated with clinical outcomes following ipilimumab treatment in metastatic melanoma. *J. Immunother. Cancer* **3**, 3–7 (2015).
26. Han, J. et al. TCR repertoire diversity of peripheral PD-1þCD8þ T cells predicts clinical outcomes after immunotherapy in patients with non–small cell lung cancer. *Cancer Immunol. Res.* **8**, 146–154 (2020).
27. Simoni, Y. et al. Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).
28. Bjerregaard, A. M. et al. An analysis of natural T cell responses to predicted tumor neoepitopes. *Front. Immunol.* **8**, 1–9 (2017).
29. Hansen, U. K. et al. Tumor-Infiltrating T cells from clear cell renal cell carcinoma patients recognize neoepitopes derived from point and frameshift mutations. *Front. Immunol.* **11**, 1–12 (2020).
30. McGranahan, N. et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Sci. (80-.)* **351**, 1463–1469 (2016).
31. Wolf, Y. et al. UVB-induced tumor heterogeneity diminishes immune response in melanoma. *Cell* **179**, 219–235.e21 (2019).
32. Zhang, Z. & Li, L. Y. TNFSF15 modulates neovascularization and inflammation. *Cancer Microenviron.* https://doi.org/10.1007/s12307-012-0117-8 (2012).
33. Kadiyska, T., Tourtourikov, I., Popmihaylova, A.-M., Kadian, H. & Chavoushian, A. Role of TNFSF15 in the intestinal inflammatory response. *World J. Gastrointest. Pathophysiol.* **9**, 73–78 (2018).
34. Wang, M., Windgassen, D. & Papoutsakis, E. T. Comparative analysis of transcriptional profiling of CD3+, CD4+ and CD8+ T cells identifies novel immune response players in T-Cell activation. *BMC Genomics* https://doi.org/10.1186/1471-2164-9-225 (2008).
35. Chowdhury, P. S., Chamoto, K., Kumar, A. & Honjo, T. PPAR-induced fatty acid oxidation in T cells increases the number of tumor-reactive CD8 + T cells and facilitates anti–PD-1 therapy. *Cancer Immunol. Res.* https://doi.org/10.1158/2326-6066.CIR-18-0095 (2018).
36. Yang, Y. et al. CD3D is associated with immune checkpoints and predicts favorable clinical outcome in colon cancer. *Immunotherapy.* https://doi.org/10.2217/imt-2019-0145 (2020).
37. Dannenmann, S. R. et al. Tumor-associated macrophages subvert T-cell function and correlate with reduced survival in clear cell renal cell carcinoma. *Oncoimmunology.* https://doi.org/10.4161/onci.23562 (2013).
38. Zheng, W., Jiang, C. & Li, R. Integrin and gene network analysis reveals that ITGA5 and ITGB1 are prognostic in non-small-cell lung cancer. *Onco. Targets. Ther.* https://doi.org/10.2147/OTT.S91796 (2016).
39. Engqvist, H. et al. Immunohistochemical validation of COL3A1, GPR158 and PITHD1 as prognostic biomarkers in early-stage ovarian carcinomas. *BMC Cancer.* https://doi.org/10.1186/s12885-019-6084-4 (2019).
40. Romero-Moreno, R. et al. The CXCL5/CXCR2 axis is sufficient to promote breast cancer colonization during bone metastasis. *Nat. Commun.* **10**, (2019).
41. Chen, C. et al. CXCL5 induces tumor angiogenesis via enhancing the expression of FOXD1 mediated by the AKT/NF-κB pathway in colorectal cancer. *Cell Death Dis.* **10**, 178 (2019).
42. De Jaeghere, E. A., Denys, H. G. & De Wever, O. Fibroblasts fuel immune escape in the tumor microenvironment. *Trends in Cancer.* https://doi.org/10.1016/j.trecan.2019.09.009 (2019)
43. Huang, A. C. et al. A single dose of neoadjuvant PD-1 blockade predicts clinical outcomes in resectable melanoma. *Nat. Med.* https://doi.org/10.1038/s41591-019-0357-y (2019).
44. Dammeijer, F. et al. The PD-1/PD-L1-checkpoint restrains T cell immunity in tumor-draining lymph nodes. *Cancer Cell* **38**, 1–16 (2020).
45. Balar, A. V. et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet.* https://doi.org/10.1016/S0140-6736(16)32455-2 (2017)
46. Bellmunt, J. et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N. Engl. J. Med.* https://doi.org/10.1056/nejmoa1613683 (2017).
47. Balar, A. V. et al. First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. *Lancet Oncol.* https://doi.org/10.1016/S1470-2045(17)30616-2 (2017).
48. Sharma, P. et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. *Lancet Oncol.* https://doi.org/10.1016/S1470-2045(17)30065-7 (2017).
49. Powles, T. et al. Efficacy and safety of durvalumab in locally advanced or metastatic urothelial carcinoma: Updated results from a phase 1/2 open-label study. *JAMA Oncol.* https://doi.org/10.1001/jamaoncol.2017.2411. (2017).
50. Apolo, A. B. et al. Avelumab, an anti-programmed death-ligand 1 antibody, in patients with refractory metastatic urothelial carcinoma: Results from a multicenter, Phase Ib study. *J. Clin. Oncol.* https://doi.org/10.1200/JCO.2016.71.6795 (2017).
51. Besser, H., Yunger, S., Merhavi-Shoham, E., Cohen, C. J. & Louzoun, Y. Level of neo-epitope predecessor and mutation type determine T cell activation of MHC binding peptides. *J. Immunother. Cancer* **7**, 1–9 (2020).
52. Łuksza, M. et al. A neoantigen fitness model predicts tumor response to checkpoint blockade immunotherapy. *Nature* **551**, 517–520 (2017).
53. Balachandran, V. P. et al. Identification of unique neoantigen qualities in long term pancreatic cancer survivors. *Nature* **551**, 512–516 (2017).

54. Koşaloğlu-Yalçın, Z. et al. Predicting T cell recognition of MHC class I restricted neoepitopes. *Oncoimmunology* **7**, e1492508 (2018).
55. Wickström, S. L. et al. Cancer neoepitopes for immunotherapy: discordance between tumor-infiltrating T cell reactivity and tumor MHC peptidome display. *Front. Immunol.* **10**, 2766 (2019).
56. Leng, Q., Tarbe, M., Long, Q. & Wang, F. Pre-existing heterologous T-cell immunity and neoantigen immunogenicity. *Clin. Transl. Immunol.* **9**, 1–9 (2020).
57. Wells, D. K. et al. Key parameters of tumor epitope immunogenicity revealed through a consortium approach improve neoantigen prediction. *Cell* **183**, 818–834.e13 (2020).
58. Ghorani, E. et al. The T cell differentiation landscape is shaped by tumour mutations in lung cancer. *Nat. Cancer* **1**, 546–561 (2020).
59. Oh, D. Y. et al. Intratumoral CD4+ T cells mediate anti-tumor cytotoxicity in human bladder cancer. *Cell* **181**, 1612–1625.e13 (2020).
60. Alspach, E. et al. MHC-II neoantigens shape tumour immunity and response to immunotherapy. *Nature* **574**, 696–701 (2019).
61. Zander, R. et al. CD4+ T cell help is required for the formation of a cytolytic CD8+ T cell subset that protects against chronic infection and cancer. *Immunity* **51**, 1028–1042.e4 (2019).
62. Goodman, A. M. et al. Prevalence of PDL1 amplification and preliminary response to immune checkpoint blockade in solid tumors. *JAMA Oncol.* **4**, 1237–1244 (2018).
63. Callahan, M. K. & Wolchok, J. D. Recruit or reboot? how does anti-PD-1 therapy change tumor-infiltrating lymphocytes? *Cancer Cell* **36**, 215–217 (2019).
64. Yost, K. E. et al. Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat. Med.* **25**, 1251–1259 (2019).
65. Postow, M. A., Yuan, J., Kitano, S., Lesokhin, A. M. & Wolchok, J. D. Markers for Anti-cytotoxic T-lymphocyte Antigen 4 (CTLA-4) Therapy in Melanoma. in *Molecular Diagnostics for Melanoma: Methods and Protocols* (eds. Thurin, M. & Marincola, F. M.) 83–95 https://doi.org/10.1007/978-1-62703-727-3_6 (Humana Press, 2014).
66. Krueger, F. Trim Galore. *Babraham Bioinformatics* on line (2016) http://www.bioinformatics.babraham.ac.uk/projects/trimgalore/ (2016).
67. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10 (2011).
68. Andrews S. FastQC A. Quality control tool for high throughput sequence data. *Babraham Bioinfo* 3–5 (2018).
69. Van der Auwera, G. A. et al. From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* https://doi.org/10.1002/0471250953.bi1110s43 (2013).
70. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* **25**, 1754–1760 (2009).
71. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
72. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
73. Szolek, A. et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316 (2014).
74. Weese, D., Holtgrewe, M. & Reinert, K. RazerS 3: Faster, fully sensitive read mapping. *Bioinformatics* **28**, 2592–2599 (2012).
75. Rodenko, B. et al. Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nat. Protoc.* **1**, 1120–1132 (2006).
76. Hadrup, S. R. et al. High-throughput t-cell epitope discovery through mhc peptide exchange. *Methods Mol. Biol.* **524**, 383–405 (2009).
77. Toebes, M. et al. Design and use of conditional MHC class I ligands. *Nat. Med.* **12**, 246–251 (2006).
78. Bakker, A. H. et al. Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc. Natl Acad. Sci. U. S. A.* **105**, 3825–3830 (2008).
79. Saini, S. K. et al. Empty peptide-receptive MHC class I molecules for efficient detection of antigen-specific T cells. *Sci. Immunol.* **4**, (2019).
80. Petersen, N. V. et al. T cell recognition of novel shared breast cancer antigens is frequently observed in peripheral blood of breast cancer patients. *Oncoimmunology* **8**, e1663107 (2019).
81. Hadrup, S. R. et al. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat. Methods* **6**, 520–526 (2009).
82. Andersen, R. S. et al. Parallel detection of antigen-specific t cell responses by combinatorial encoding of MHC multimers. *Nat. Protoc.* **7**, 891–902 (2012).
83. Becton Dickinson and Company. FlowJo Software Version 10.6.2. (2019).
84. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform manifold approximation and projection for dimension reduction. (2018).
85. Van Gassen, S. et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytom. Part A* **87**, 636–645 (2015).
86. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
87. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
88. Yu, G. enrichplot: Visualization of functional enrichment result. R package version 1.10.1. (2020).
89. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* https://doi.org/10.1089/omi.2011.0118 (2012).
90. Becht, E. et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218 (2016).
91. Favero, F. et al. Sequenza: Allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol.* **26**, 64–70 (2015).
92. Nilsen, G. et al. Copynumber: Efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics* **13**, 591 (2012).
93. Wang, S. Copynumber. *GitHub Repository* www.github.com/ShixiangWang/copynumber (2020).
94. Roth, A. et al. PyClone: Statistical inference of clonal population structure in cancer. *Nat. Methods* **11**, 396–398 (2014).
95. Ahlmann-Eltze, C. ggsignif: Significance Brackets for 'ggplot2'. R package, version 0.6.0. (2019).
96. Kassambara, A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. (2021).
97. R Core Team. R: A language and environment for statistical computing. (2020).
98. Larsson, J. et al. eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses, version 6.1.0. (2020).
99. Therneau, T. M. A Package for Survival Analysis in R, version 3.2-7. (2020).

Acknowledgements

We thank all donors and patients for participating in the study; B. Rotbøl and A.F. Løye for technical assistance handling the flow cytometry instruments and PBMC samples; C. Heeke for assistance in pMHC multimer gathering; and R. Zappasodi, B. Greenbaum, D. Aggen, M. Hellmann, and V. Balachandran for fruitful discussions of the data. The research was funded in part by the Ludwig Center for Cancer Research; the NIH/NCI Cancer Center Support Grant P30 CA008748, the K12CA184746-01A1 (S.A.F.); Ludwig Collaborative and Swim Across America Laboratory (MSKCC, New York, NY 10065, USA); Parker Institute for Cancer Immunotherapy (MSKCC, New York, NY 10065, USA); Department of Medicine (MSKCC, New York, NY 10065, USA); Weill Cornel Medicine (MSKCC, New York, NY 10065, USA); and the Independent Research Fund Denmark (J.S.H.).

Author contributions

J.S.H. designed and performed experiments, analyzed the data, generated figures, and wrote the manuscript; S.A.F. collected patient material, analyzed the data, provided clinical evaluation, and wrote the manuscript; A.B. and K.K.M. conducted all bioinformatics analyses and generated figures; A.M.B. predicted neoepitopes; J.R. designed phenotypic panels and experiments; C.M. and A.R. managed patient material; P.W., H.A.A., and G.I. collected patient material and provided clinical evaluation; T.T. produced HLA-I complexes; A.K.B. provided technical guidance; N.O.H. selected neopeptides; S.D.W. discussed the data, A.S. provided WES-, RNA-seq- and TCRb-seq-data; T.M. and J.D.W. supervised the study and discussed the data; M.N. designed the in silico platforms and supervised neopeptide prediction; J.E.R. supervised the patient material collection and clinical evaluation, and discussed the data; D.F.B. conceived the concept, supervised the patient material collection and clinical evaluation, discussed the data, and supported funding; S.R.H. conceived the concept, supervised the study, discussed the data, supported funding, and wrote the manuscript. All authors reviewed and approved the manuscript.

Competing interests

SAF has received research support from AstraZeneca, Genentech/Roche, is a consultant/advisory board member for Merck, and owns stock in Urogen, Allogene Therapeutics, Neogene Therapeutics, Kronos Bio, and Inconovir. He is also supported by the National Cancer Institute K12CA184746-01A1 grant and the Bochner-Fleisher Research Scholar in Urologic Oncology Award. HAA is a consultant/advisory board member for Bristol Myers Squibb, EMD Serono, AstraZeneca/MedImmune, and Janssen Biotech. GI has a consulting or advisory role for Bayer, Janssen, and Mirati Therapeutics; and has received research funding from Mirati Therapeutics, Novartis, Debiopharm Group, and Bayer. PW is a consultant for Sellas Life Sciences and Leap Therapeutics. AKB and SRH are co-inventors of the licensed patents for DNA Barcoded MHC-multimers (WO2015185067 and WO2015188839), licensee: Immudex, DK. SDW is supported by the MSK Clinical Scholars T32 (5T32CA009512) and Young Investigator Award from the American Society of Clinical Oncology. AS is an employee of and owns stock in Merck and Co. TM is a consultant for Leap Therapeutics, Immunos Therapeutics and Pfizer, and co-founder of Imvaq therapeutics; has equity in Imvaq therapeutics; has received research support from Bristol Myers Squibb, Surface Oncology, Kyn Therapeutics, Infinity Pharmaceuticals, Peregrine Pharmeceuticals, Adaptive Biotechnologies, Leap Therapeutics and

Aprea; and is inventor on patent applications related to work on oncolytic viral therapy, alphavirus-based vaccines, neo-antigen modeling, CD40, GITR, OX40, PD-1 and CTLA-4. JDW is a consultant for Amgen, Apricity, Arsenal IO, Ascentage Pharma, AstraZeneca, Astellas, Boehringer Ingelheim, Bristol Myers Squibb, Chugai, Dragonfly, F Star, Eli Lilly, Georgiamune, Imvaq, Merck, Polynoma, Psioxus, Recepta, Trieza, Truvax, Sellas, and Werewolf Therapeutics; has equity in Tizona Pharmaceuticals, Imvaq, Beigene, Linneaus, Apricity, Arsenal IO, and Georgiamune; and has received research support from Bristol Myers Squibb and Sephora. JDW is also the co-inventor of the following licensed patents: Xenogeneic DNA Vaccines (USPTO, US7556805), licensee: Merial; Myeloid-derived suppressor cell (MDSC) assay (EPO, PCT/US2013/027475), licensee: Serametrix; Anti-PD1 Antibody (USPTO, US10323091), licensee: Agensus; Anti-CTLA4 antibodies (USPTO, US10144779), licensee: Agensus; Anti-GITR antibodies and methods of use thereof (USPTO, US10155818/US10280226), licensee: Agenus/Incyte. JER holds stock and other ownership interests in Illumina; has received honoraria from AstraZeneca, Bristol-Myers Squibb, Chugai Pharma, Clinical Care Options, Clinical Mind, Intellisphere, Medscape, Peerview, Research To Practice, UpToDate, and Vindico; has a consulting or advisory role for Adicet Bio, Agensys, Astellas Pharma, AstraZeneca/MedImmune, Bayer, BioClin Therapeutics, Bristol-Myers Squibb, EMD Serono, Fortress Biotech, GlaxoSmithKline, Inovio Pharmaceuticals, Janssen Oncology, Lilly, Merck, Pharmacyclics, QED Therapeutics, Roche/Genentech, Seattle Genetics, Sensei Biotherapeutics, and Western Oncolytics; has received institutional research funding from Astellas Pharma, AstraZeneca, Bayer, Genentech/Roche, Incyte, Jounce Therapeutics, Mirati Therapeutics, Novartis, QED Therapeutics, Seattle Genetics, and Viralytics; holds an institutional interest in a patent for a predictor of platinum sensitivity; and has received travel and accommodation expenses from Bristol-Myers Squibb and Genentech/Roche. DFB is a consultant for Bristol Myers Squibb, Merck, Genentech-Roche, AstraZeneca, and Pfizer; and has received research support from Merck, Genentech-Roche, Astra-Zeneca, Novartis, and Bristol-Myers Squibb. SRH is the co-founder of Tetramershop and PokeACell and co-inventor of the licensed patent for Combination encoding of MHC multimers (EP2088/009356), licensee: Sanquin, NL. The remaining authors declare no conflicts of interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-29342-0.

Correspondence and requests for materials should be addressed to Sine Reker Hadrup.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

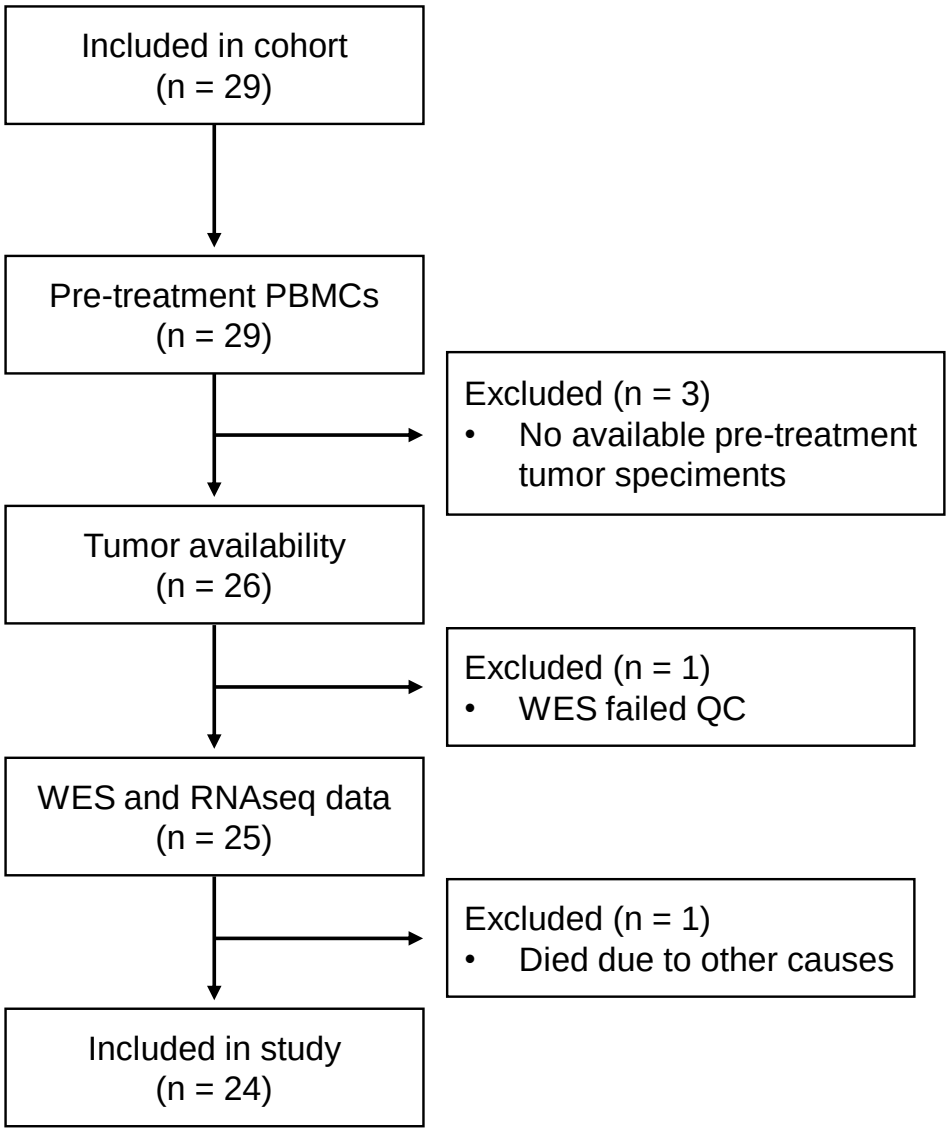
Supplementary Information

Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma

Jeppe Sejerø Holm[1*], Samuel A. Funt[2,3*], Annie Borch[1], Kamilla Kjærgaard Munk[1], Anne-Mette Bjerregaard[1], James L. Reading[5], Colleen Maher[2,3,7], Ashley Regazzi[2,3], Phillip Wong[2,3,7], Hikmat Al-Ahmadie[6], Gopa Iyer[2,3], Tripti Tamhane[1], Amalie Kai Bentzen[1], Nana Overgaard Herschend[1], Susan De Wolf[2], Alexandra Snyder[2,3], Taha Merghoub[2,3,7], Jedd D. Wolchok[2,3,7,8], Morten Nielsen[4], Jonathan E. Rosenberg[2,3], Dean F. Bajorin[2,3*], Sine Reker Hadrup[1*§]

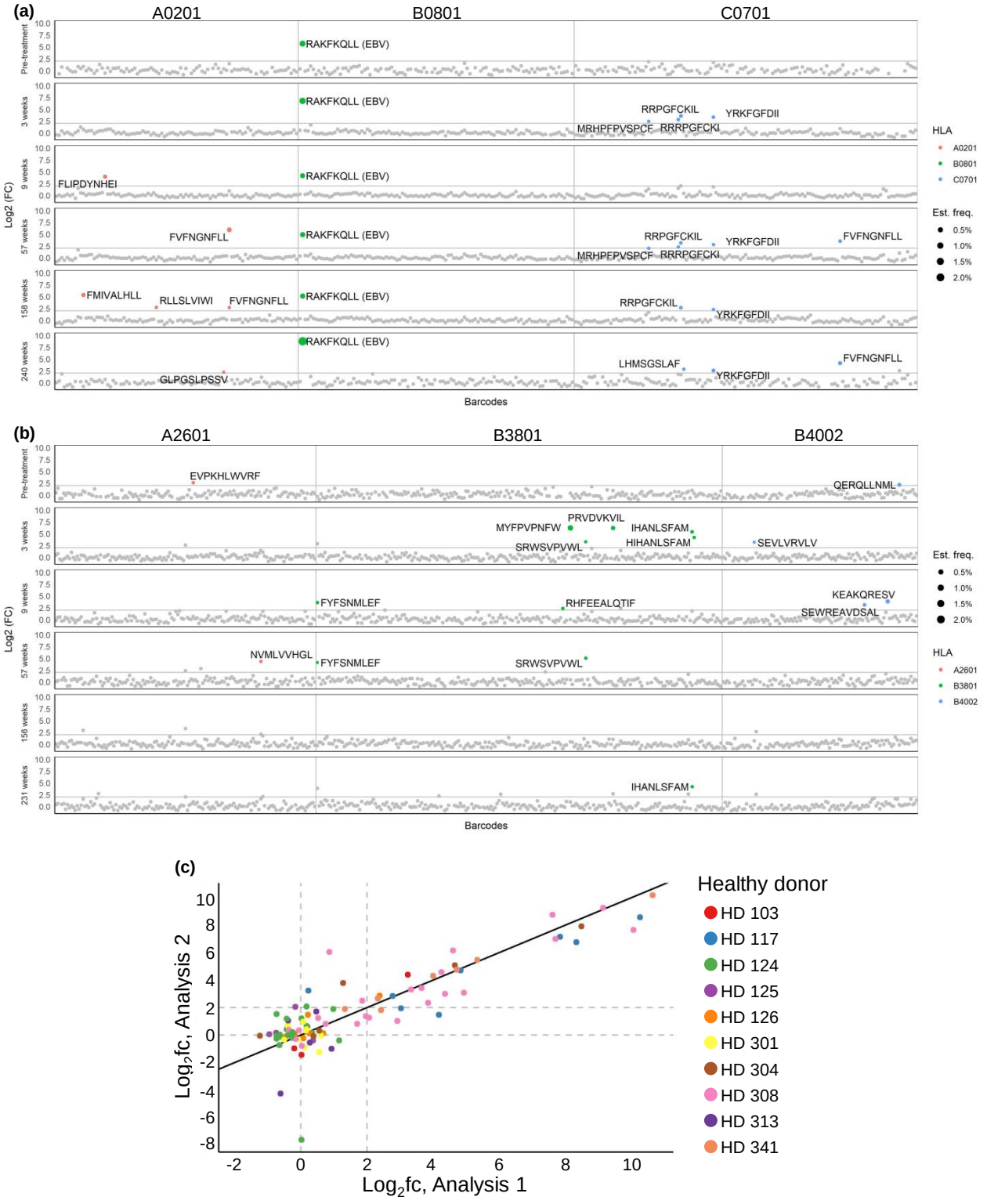
Supplementary Figures 1-6

Supplementary Tables 1-3



Supplementary Figure 1

Supplementary Fig. 1 Consort diagram of patients in cohort included in previous study (Snyder et al., 2017) and inclusion criteria for patient samples analyzed here. Patients #0979, #7592, and #8214 had no available pre-treatment archival tumor speciments, QC for WES of patient #4072 failed, and patient #9881 died of other causes. These patients were excluded from this study.

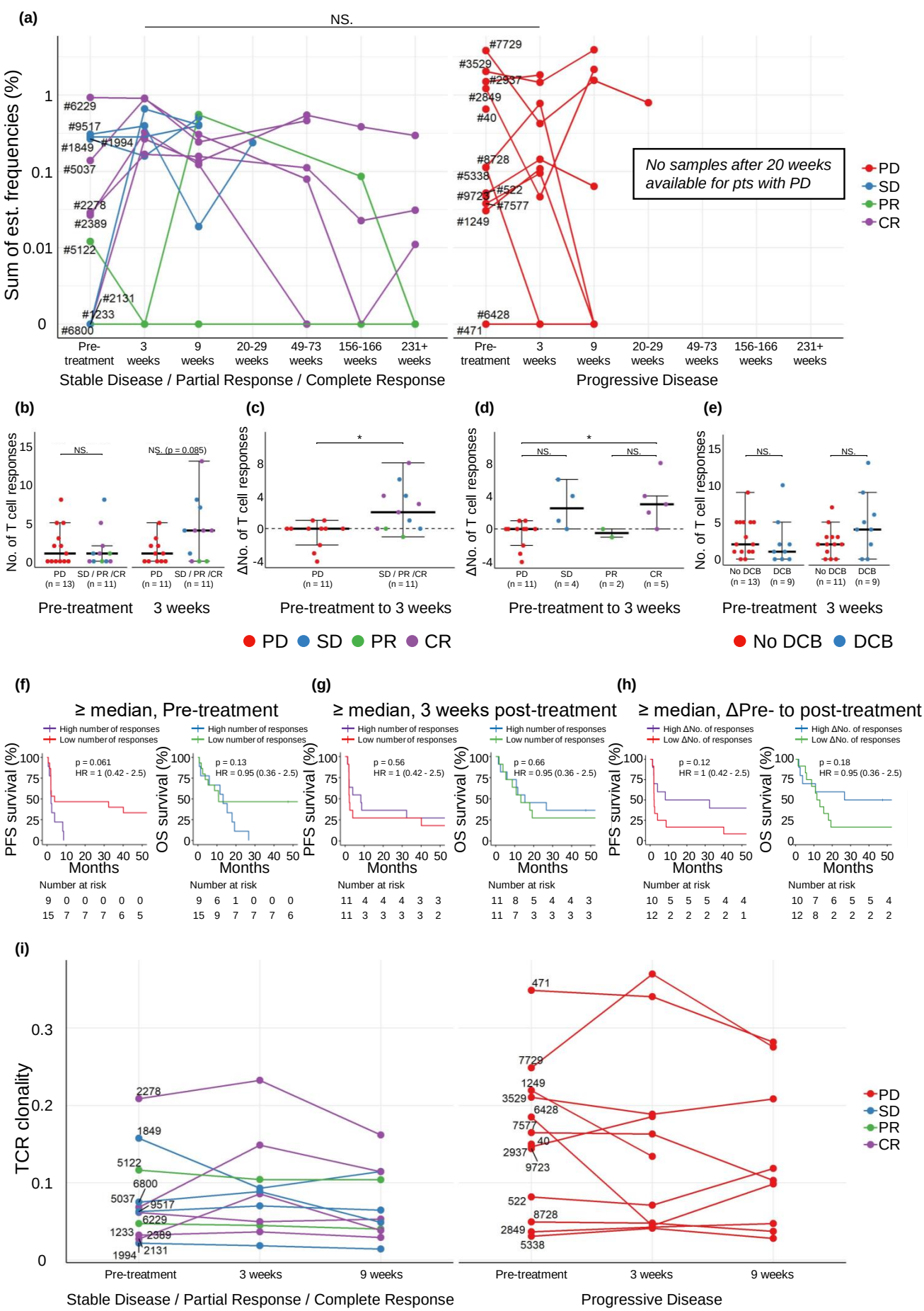


Supplementary Figure 2

Supplementary Fig. 2 Detection of neoepitope-reactive T cell responses in mUC cancer patients. Additional representative output from patients **(a)** #2131 and **(b)** #5037, screening of detected NART responses. Log_2FC of sequenced pMHC associated barcodes enriched by T cell sorting over the input library at stated timepoints. Labelled points; Log_2 FC > 2, count fraction > 0.1 % and $p < 0.001$, determined as T cell responses. Colored based on peptide-presenting HLA-type, text labelled with peptide sequence, and sized according the estimated frequency of the peptide-recognizing T cell population. For virus-derived peptides the virus is stated in brackets. Grey points represent non-enriched barcodes. Horizontal line at Log_2 fc = 2. Vertical line separating peptide-presenting HLA-types. **(c)** Log_2FC of viral CMV, EBV, and FLU-derived pMHC-associated barcodes (n = 72) from duplicate stainings of HD PBMC controls (n = 10). Source data are provided as a Source Data file.

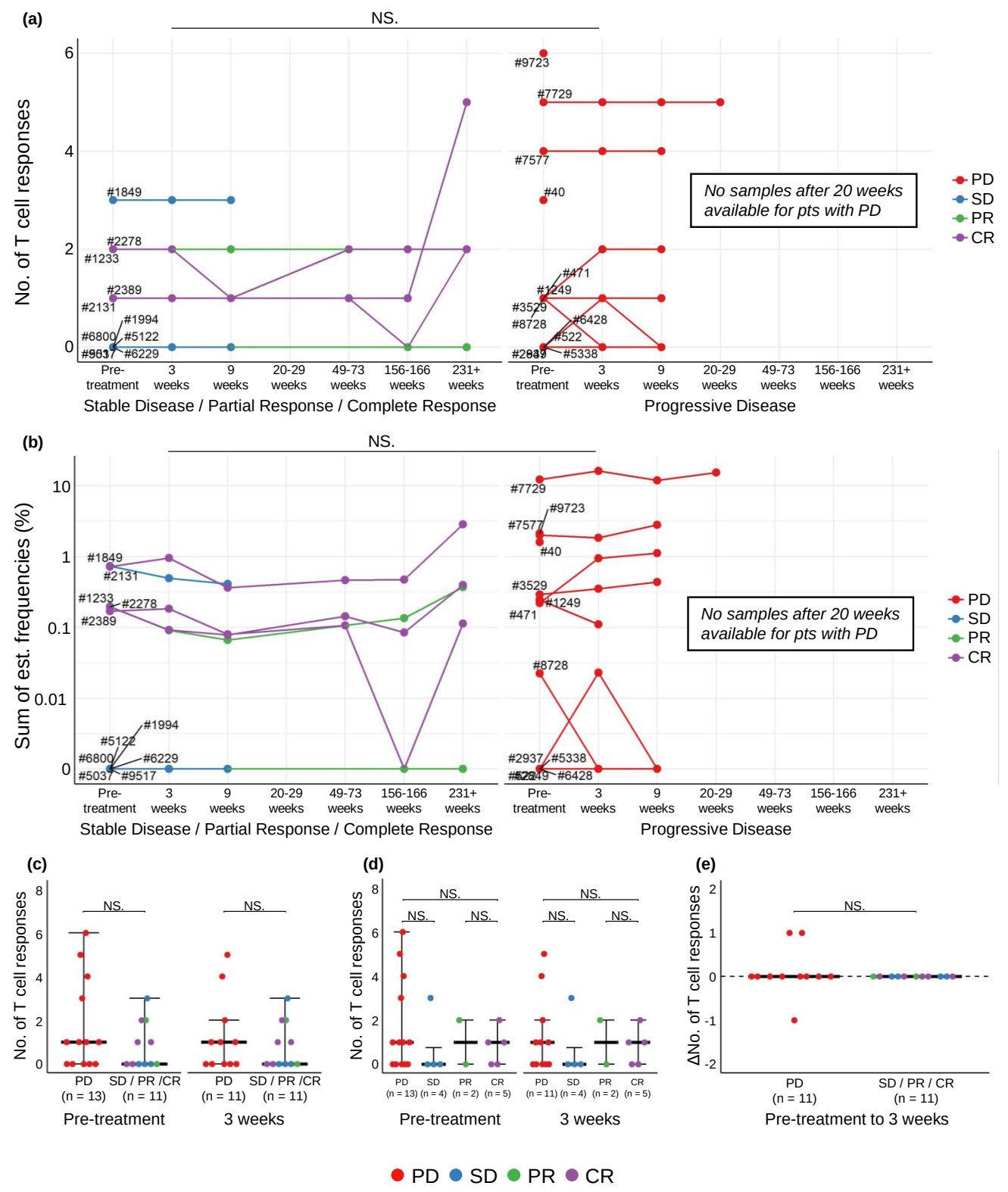
Supplementary Figure 3

Supplementary Fig. 3 Tetramer pMHC stainings for NART response validations. CD8+ cells shown, gated for tetramer+ populations. Selected patient PBMC samples, either at 3- or 9-weeks post-treatment depending on sample availabily at time of study, were stained with tetramers generated for recognized neoepitopes seen in multimer screening at given timepoint. Neoepitope number and sequence, presenting HLA-type, estimated frequency from multimer screening, and tetramer+ NART frequency listed for each pMHC. Green markings indicate positive validation of NART response, with yellow markings indicating NART responses at border of detection level using tetramer stainings. PBMCs from 3-weeks post-treatment; patients #1249, #2131, #2278, #2849, #2937, #3529, #5037, #522, #6229, #6800, #7729, #8728, and #9517. PBMCs from 9-week post-treatment; patients #1233 and #1849.



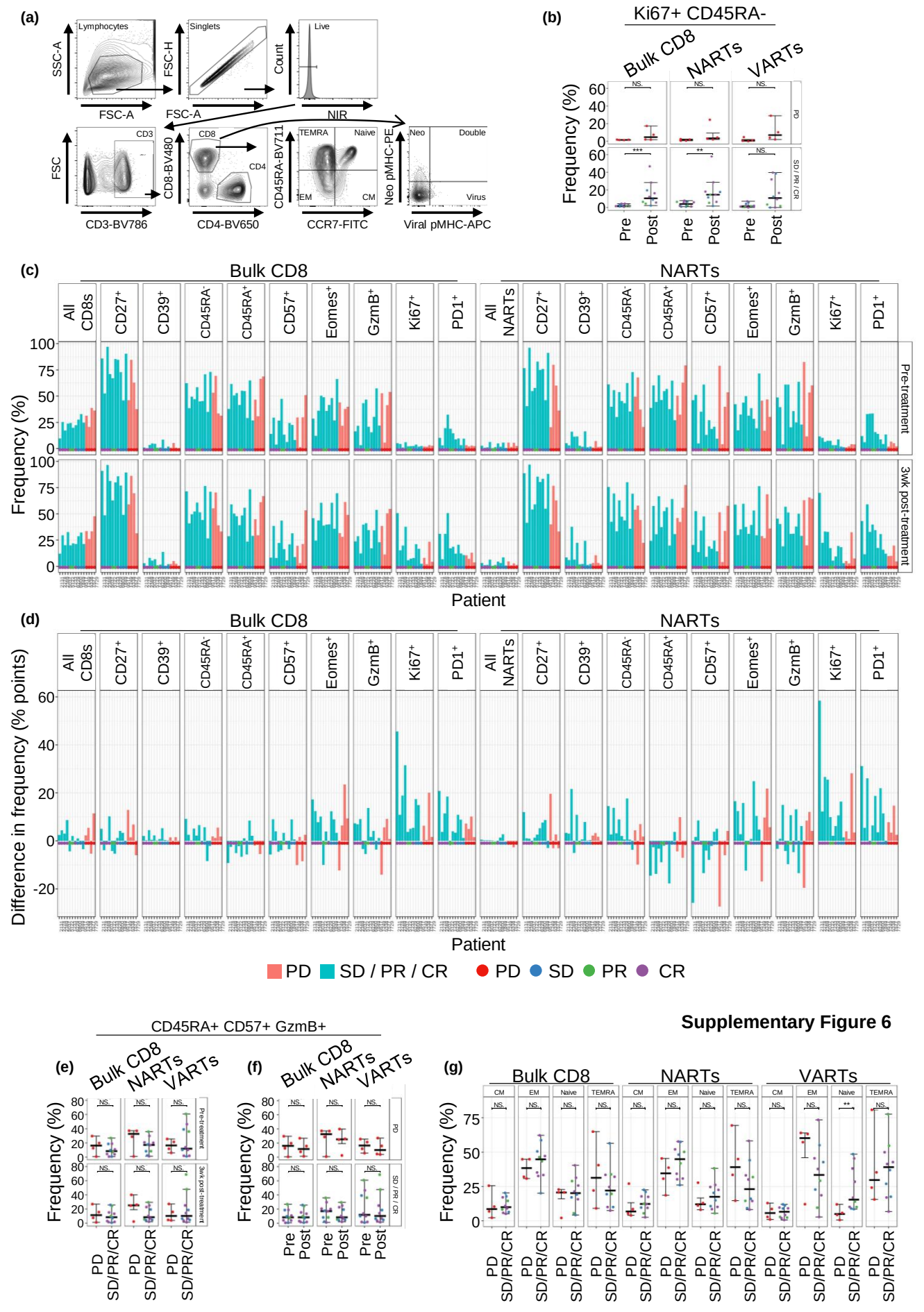
Supplementary Figure 4

Supplementary Fig. 4 Impact of NARTs in anti-PD-L1 treatment. **(a)** SEFs for NART populations over course of therapy. Points at respective time point indicate screening of the given sample. Patients clustered according to patients with SD, PR, and CR (n = 11 patients) and with PD (n = 13 patients). **(b)** Number of NART responses for patients and **(c-d)** the change from pre- to 3-week post-treatment, between SD/PR/CR and PD patients, only for NART responses towards neopeptides with EL%Rank < 0.5 and expression level > 2 TPM. **(e)** Number of NART responses for patients with or without DCB. **(f-h)** Kaplan-Meier estimation curves for patient PFS and OS grouped according to High or Low number of NART responses at **(f)** pre-treatment (> median = 2 responses) or at **(g)** 3-week post-treatment (≥ median = 3 responses), or **(h)** change in number of NART responses (> median = 0 response increase). **(i)** Bulk PBMC TCRb cloanlity at pre-treatment, 3 weeks- and 9 weeks post-treatment. Points at respective time point indicate PBMC sample and TCRb sequencing data availability. For b)-e) groups were compared using non-parametric two-sided Mann-Whitney test and Kruskal-Wallis Dunn's multiple comparison test for d). For a)-e), data is presented as median values +/- largest/smallest value within upper/lower quartile +/- 1.5 IQR. NS.: Not Significant, *: $p < 0.05$. Source data are provided as a Source Data file.



Supplementary Figure 5

Supplementary Fig. 5 VARTs in anti-PD-L1 treatment. Timeline of **(a)** number of VART responses and **(b)** SEF for VART populations over course of therapy. Points at respective time point indicate screening of the given sample. Patients clustered according to patients with SD, PR, and CR (n = 11 patients) and with PD (n = 13 patients). **(c)** Number of VART responses for patients, clustered according to disease control and **(d)** according to RECIST criteria, at pre- and 3-week post-treatment. **(e)** The change in number of VART responses for patients, clustered according to disease control. For c) + e) groups were compared using non-parametric two-sided Mann-Whitney test and Kruskal-Wallis Dunn's multiple comparison test for d). For c)-e), data is presented as median values +/- largest/smallest value within upper/lower quartile +/- 1.5 IQR. NS.: Not Significant. Source data are provided as a Source Data file.



Supplementary Figure 6

Supplementary Fig. 6 Phenotypic characterisation of bulk CD8 T cells, NARTs and VARTs. **(a)** Gating strategy for sorting and characterisation of PE-neo- and APC-viral pMHC multimer-binding singlet, Live (NIR-), CD3+, CD8+/CD4- lymphocytes. **(b)** Parent population frequencies of Ki67+CD45RA- bulk CD8 T cells, NARTs, and VARTs for SD/PR/CR (n = 10 patients) and PD patients (n = 4 patients) at pre- and 3 weeks post-treatment. **(c-d)** Parent population frequencies of bulk CD8 T cells and NARTs for all non-lineage parameters at pre- and 3 weeks post-treatment, and the change in frequency between time points. **(e-f)** Parent population frequencies of triple-positive CD45RA+ CD57+ GzmB+ bulk CD8 T cells, NARTs, and VARTs for PD (n = 4 patients) and SD/PR/CR patients (n = 10 patients) at pre- and 3 weeks post-treatment. Mann-Whitney test for l)-p). **(g)** Parent population frequencies of Central- (CM) and Effector Memory (EM), Naive and TEMRA bulk CD8 T cells, NARTs, and VARTs for PD (n = 4 patients) and SD/PR/CR patients (n = 10 patients) at 3 weeks post-treatment. For b) + e)-g) groups were compared using non-parametric two-sided Mann-Whitney test and data is presented as median values +/- largest/smallest value within upper/lower quartile +/- 1.5 IQR NS.: Not Significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$. wk: week, Pre: Pre-treatment, Post: 3 weeks post-treatment. Source data are provided as a Source Data file.

Pt #	T cell neoepitope responses, per time point						
	Pre-treatment	3 weeks	9 weeks	20-29 weeks	49-73 weeks	156-166 weeks	231+ weeks
#40	#128 HSSEVTMTL (A3201) #138 TLIFARLTI (A3201) #388 SYDTALQARI (C0401) #414 VFMEENSKL (C0401) #463 IRLRPRHVL (C0701)			*Subsequent samples not available*			
#9723	#96 QQQQQAQTL (B0801)			*Subsequent samples not available*			
#1249	- #60 EASVKVIHR (A6801)	#45 WAVFPSIVGR (A6801) #60 EASVKVIHR (A6801)		*Subsequent samples not available*			
#2937	#42 ATIKVNGTI (A3201) #60 SLLEHVEEY (A3201) #91 MVEAGAGTSW (B1801) #95 SEFDNLRTL (B1801) - #194 EPWVSIKKF (B3503) - #234 FPLPWWDLM (B3503) #307 LYDQGYTSL (C0401) #308 ILYDQGYTSL (C0401) #309 LYDQGYTSLG (C0401)	- - - - #134 TEKPPYIEV (B1801) - #219 SATEAFGEL (B3503) - #307 LYDQGYTSL (C0401) #308 ILYDQGYTSL (C0401) #309 LYDQGYTSLG (C0401)		*Subsequent samples not available*			
#1994	- - #89 YTYAAMLRI (C0702) - #147 LLLAAVVSH (A0301) - #186 NSVPPPPPL (B0702)	- #23 KPMTLFQIQF (B0702) - - #147 LLLAAVVSH (A0301) - -	#21 RPRLLLLVL (B0702) - - #115 FYILSSGLI (C0702) - #164 GLPNVFGLGR (A0301) -	*Subsequent samples not available*			
#522	#19 SFLVHPYGF (A2402) - #128 FQRFHKLHYL (C0602) -	- #118 RRVAMRRWI (C0602) #128 FQRFHKLHYL (C0602) #149 FYSSRKRLL (C0602)	- #118 RRVAMRRWI (C0602) #128 FQRFHKLHYL (C0602) -	*Subsequent samples not available*			
#6428	No responses	No responses	No responses	*Subsequent samples not available*			
#7577	- #9 CMDFNSNGKY (A0101)	#1 ANDNSPFMLY (A0101) #9 CMDFNSNGKY (A0101)	No responses	*Subsequent samples not available*			
#8728	#118 TEKPIQRNPG (B4403) #141 SPLPQSPQV (B5101) #144 LPLKLQSEV (B5101) #153 HPQDFRDHPV (B5101) #170 QLPLKLQSEV (B5101) - -	- - - - - #174 VGIFILCTV (B5101) #176 QSPQVLQQL (B5101)	No responses	*Subsequent samples not available*			
#471	No responses	No responses	No responses	*Subsequent samples not available*			
#3529	#66 MRCLVQHIL (C0602) #76 YARFLQSNAY (C0602) #78 YRAQVYVPV (C0602) - #149 KGYEGYYVL (C0602) #177 YRSGHQLHC (C0602) -	#66 MRCLVQHIL (C0602) #76 YARFLQSNAY (C0602) #78 YRAQVYVPV (C0602) - - - -	#66 MRCLVQHIL (C0602) #76 YARFLQSNAY (C0602) #78 YRAQVYVPV (C0602) #130 RTFSMQVAL (A0205) #149 KGYEGYYVL (C0602) #177 YRSGHQLHC (C0602) #179 AFQIAMKLL (A2402)	*Subsequent samples not available*			
#5338	#161 HPHPHPHAF (C0304)	No responses	No responses	*Subsequent samples not available*			
#2849	#249 LPEARRPRL (B0702) - #298 GRHILVAWK (B2705) #371 YFSQEQWGL (C0702) #385 YFKSDELQF (C0702) #401 VHARVINFF (C0702)	- - - - - #401 VHARVINFF (C0702)	- #289 TRVKCVVSM (B2705) - - - #401 VHARVINFF (C0702)	*Subsequent samples not available*			
#7729	- #149 QTIDKAKYI (C0602) #150 SQRETTWTF (C0602) #161 LLPPYKQSI (C0602) #162 CSYSEPHYM (C0602) #198 LLPPYKQSI (C0701)	- #149 QTIDKAKYI (C0602) - #161 LLPPYKQSI (C0602) #162 CSYSEPHYM (C0602) -	#132 IRTDSVLIL (C0602) #149 QTIDKAKYI (C0602) #150 SQRETTWTF (C0602) #161 LLPPYKQSI (C0602) #162 CSYSEPHYM (C0602) -	- #149 QTIDKAKYI (C0602) - #161 LLPPYKQSI (C0602) #162 CSYSEPHYM (C0602) -	*Subsequent samples not available*		
#9517	- - #21 DADLLRPHAY (A0101) #212 TRLFLFHLL (B3801) - - - - -	- #9 LVTEDTTICY (A0101) #21 DADLLRPHAY (A0101) - #217 YVDYPIYDML (C0401) #218 SFEEYLKLL (C0401) #220 QWPPFVVTL (C0401) #229 YLQTYGAEL (C0401) #242 MFSGVAVYL (C0401)	- #9 LVTEDTTICY (A0101) - - - - - - -	#8 VTEDTTICY (A0101) #9 LVTEDTTICY (A0101) #21 DADLLRPHAY (A0101) - - - - #229 YLQTYGAEL (C0401) -	*Subsequent samples not available*		

Supplementary Table 1

Pt #	T cell neoepitope responses, per time point						
	Pre-treatment	3 weeks	9 weeks	20-29 weeks	49-73 weeks	156-166 weeks	231+ weeks
#1849	#1 QLEQLMQLY (A0101) #17 YTDQISKYA (A0101) - - #63 VTSEVSNLK (A1101) #132 LRSRHSTRI (C0602) - #146 LRSSQRMVI (C0602) - #152 SRYGGGLAV (C0602) #154 FSFSKSRRI (C0602) #155 FNYKPLHTI (C0602) #164 LRSVSPWTY (C0602) - #185 FSKSRRILL (C0602)	- #17 YTDQISKYA (A0101) - - - #132 LRSRHSTRI (C0602) #133 LRACTRSSM (C0602) #146 LRSSQRMVI (C0602) #150 FRNSANATSL (C0602) #152 SRYGGGLAV (C0602) #154 FSFSKSRRI (C0602) #155 FNYKPLHTI (C0602) - - #185 FSKSRRILL (C0602)	#1 QLEQLMQLY (A0101) #17 YTDQISKYA (A0101) #18 ILEYTDQISKY (A0101) #53 SSEPPFGPK (A1101) #63 VTSEVSNLK (A1101) #132 LRSRHSTRI (C0602) - #146 LRSSQRMVI (C0602) #150 FRNSANATSL (C0602) #152 SRYGGGLAV (C0602) #154 FSFSKSRRI (C0602) #155 FNYKPLHTI (C0602) #164 LRSVSPWTY (C0602) #167 MSMPKGRVV (C0602) #185 FSKSRRILL (C0602)			*Subsequent samples not available*	
#6800	No responses	#84 SLKRLRNTK (A0301) #109 YLAKHTILY (B3501) #118 IPQDSIAQIY (B3501) #123 FQISSGISF (B3501) - #196 YSLELGMTL (C0401)	- - - - #190 KYMQMNYAL (C0401) -			*Subsequent samples not available*	
#6229	#14 RMHYLPQLK (A3001) - - - - #36 HQRPGHLLA (A3001) - - - - - - #205 IYHQPTHLW (C0602) - - - #235 SRPLPVAAV (C0602) #243 WRSERRSWV (C0602) -	- #18 ASKPIFMDR (A3001) #19 GTRMPTSTY (A3001) #20 KVFSHHAYI (A3001) #21 KVFSHHAYF (A3001) - #41 VFHTPTVIK (A3001) #46 KVSFSGCMVK (A3001) #74 RLYTGKKPY (A3201) #80 SQWRKAPGW (A3201) - #201 FQLDQITAL (C0602) #205 IYHQPTHLW (C0602) #226 RHRDPVPEL (C0602) - #234 SRWNMSRRL (C0602) - #243 WRSERRSWV (C0602) -	- - - - - - - - - #94 IYHQPTHLW (A3201) - - - - - - - #243 WRSERRSWV (C0602) #250 SRRVREASL (C0602)	*Sample not available*	- - - - - - - - - - - - - #230 FRSRWKEQY (C0602) - - #243 WRSERRSWV (C0602) -	*Subsequent samples not available*	
#5037	#41 EVPKHLWVRF (A2601) - - - - - - - - - - - - #437 QERQLLNML (B4002)	- - - - - #267 MYFPVPNFW (B3801) #275 SRWSVPVWL (B3801) #289 PRVDVKVIL (B3801) #331 IHANLSFAM (B3801) #332 HIHANLSFAM (B3801) #362 SEVLVRVLV (B4002) - - -	- - #136 FYFSNMLEF (B3801) #263 RHFEEALQTIF (B3801) - - - - #419 SEWREAVDSAL (B4002) #431 KEAKQRESV (B4002) -	*Sample not available*	- #73 NVMLVVHGL (A2601) - - - - #275 SRWSVPVWL (B3801) - - - - - - - -	No responses	No responses
#2389	#1 CIDFQPDIY (A0101) - - #159 NEASLSFQAL (B4001)	#1 CIDFQPDIY (A0101) #2 SCIDFQPDIY (A0101) - #141 TELGTAAKL (B4001) #159 NEASLSFQAL (B4001)	#1 CIDFQPDIY (A0101) #2 SCIDFQPDIY (A0101) - #141 TELGTAAKL (B4001) -	*Sample not available*	No responses	No responses	- - #26 KWPECEKVF (A2402) - -
#2131	No responses	- - - - - #312 MRHPFPVSPCF (C0701) #323 RRRPGFCKI (C0701) #324 RRPGFCKIL (C0701) - #336 YRKFGFDII (C0701) -	- #23 FLIPDYNHEI (A0201) - - - - - - - - -	*Sample not available*	- - - - #65 FVFNGNFLL (A0201) #312 MRHPFPVSPCF (C0701) #323 RRRPGFCKI (C0701) #324 RRPGFCKIL (C0701) - #336 YRKFGFDII (C0701) #383 FVFNGNFLL (C0701)	#15 FMIVALHLL (A0201) - #40 RLLSLVIWI (A0201) - #65 FVFNGNFLL (A0201) - - - #336 YRKFGFDII (C0701) -	- - - #63 GLPGSLPSSV (A0201) - - - - #325 LHMSGSLAF (C0701) #336 YRKFGFDII (C0701) #383 FVFNGNFLL (C0701)
#2278	- - - - #75 RFDAFVLFL (A2402) - -	- - - - #75 RFDAFVLFL (A2402) #96 GIFGLNLALF (A2402) -	- - - - - - #118 FYRALMSNTY (A2402)	*Sample not available*	- - #71 RWYAICHPFMF (A2402) #72 WYAICHPFMF (A2402) #75 RFDAFVLFL (A2402) - -	#13 LLAKGLVLL (A0201) - - - - - -	- #44 RFDAFVLFL (A0201) - - - - -
#1233	No responses	No responses	#37 VLIFAVVGM (A2601) - #58 ARTVRPVSL (B2705)	*Samples not available*		#56 EDFLLHINF (A2601) -	No responses
#5122	#154 GRGEGPIWL (B3801)	No responses	No responses	*Samples not available*		No responses	No responses

Supplementary Table 1 cont.

Supplementary Table 1 Detailed patient T cell neoepitope response data. The patient-specific peptide no., peptide sequence, and peptide-presenting HLA is stated for each T cell neoepitope response. '-' or 'No responses' indicate no neoepitope T cell responses were detected towards the specific neoepitope or at all in the PBMC sample at the given timepoint. Source data are provided as a Source Data file.

Clinical characteristics					
Patient	DCB *(PFS > 6 months)*	Outcome *(Best RECIST 1.1)*	PFS *(Days)*	OS *(Days)*	Ultimate time point *(Weeks)*
#40	No DCB	PD[1]	20	24	0
#9723	No DCB	PD[1]	19	22	0
#1249	No DCB	PD[1]	41	54	3
#2937	No DCB	PD	37	44	3
#1994	No DCB	SD	121	403	9
#522	No DCB	PD	58	76	9
#6428	No DCB	PD	61	163	9
#7577	No DCB	PD	60	117	9
#8728	No DCB	PD	61	546	9
#471	No DCB	PD	61	329	9
#3529	No DCB	PD	64	182	9.4
#5338	No DCB	PD	67	266	9.7
#2849	No DCB	PD	75	586	11
#7729	No DCB	PD	60	354	20
#9517	No DCB	SD	121	331	29
#1849	DCB	SD	265	468	9
#6800	DCB	SD	985	1445	9
#6229	DCB	CR	252	812	49.3
#5037	DCB	CR	1954	1954	231.1
#2389	DCB	CR	1990	1990	237.1
#2131	DCB	CR	1965	1965	240
#2278	DCB	CR	2046	2046	244.3
#1233	DCB	PR	1221	1977	247.4
#5122	DCB	PR	1932	1932	256.3

Supplementary Table 2

Supplementary Table 2 Detailed patient clinical data. Durable Clinical Benefit (DCB) based on Progression-free survival (PFS) > 6 months and clinical outcome determined from Best RECIST 1.1 criteria during therapy. [1]Patient only scanned at baseline. PFS and OS data updated as per March 10th 2020. Source data are provided as a Source Data file.

5 MANUSCRIPT III

Neoantigen-reactive CD8+ T cells affect clinical outcome of adoptive cell therapy with tumor-infiltrating lymphocytes in melanoma

Nikolaj Pagh Kristensen*, Christina Heeke*, Siri A. Tvingsholm*, **Annie Borch**, Arianna Draghi, Michael D. Crowther, Ibel Carri, Kamilla K. Munk, Jeppe Sejerø Holm, Anne-Mette Bjerregaard, Amalie Kai Bentzen, Andrea M. Marquard, Zoltan Szallasi, Nicholas McGranahan, Rikke Andersen, Morten Nielsen, Göran B. Jönsson, Marco Donia, Inge Marie Svane, Sine Reker Hadrup.
* these author are contributed equally.
The Journal of clinical investigation - (JAN 2022), Vol. 132, no. 2

Contribution: This Manuscript is conducted in collaboration with Nikolaj Pagh Kristensen, Christina Heeke, and Siri A. Tvingsholm, who have done all the experimental experiments. I have contributed to all the bioinformatic analyses, mainly presented in Figure 6 and Supplementary Figure 10.

Neoantigen-reactive CD8+ T cells affect clinical outcome of adoptive cell therapy with tumor-infiltrating lymphocytes in melanoma

Nikolaj Pagh Kristensen,[1] Christina Heeke,[1] Siri A. Tvingsholm,[1] Annie Borch,[1] Arianna Draghi,[2] Michael D. Crowther,[2] Ibel Carri,[3] Kamilla K. Munk,[1] Jeppe Sejerø Holm,[1] Anne-Mette Bjerregaard,[1] Amalie Kai Bentzen,[1] Andrea M. Marquard,[1] Zoltan Szallasi,[4] Nicholas McGranahan,[5] Rikke Andersen,[2] Morten Nielsen,[3,6] Göran B. Jönsson,[7] Marco Donia,[2] Inge Marie Svane,[2] and Sine Reker Hadrup[1]

[1]Section for Experimental and Translational Immunology, Department of Health Technology, Technical University of Denmark (DTU), Kongens Lyngby, Denmark. [2]National Center for Cancer Immune Therapy (CCIT-DK), Department of Oncology, Copenhagen University Hospital, Herlev, Denmark. [3]Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín, Buenos Aires, Argentina. [4]Danish Cancer Society Research Center, Copenhagen, Denmark. [5]Cancer Genome Evolution Research Group, University College London Cancer Institute, London, United Kingdom. [6]Section for Bioinformatics, Department of Health Technology, DTU, Kongens Lyngby, Denmark. [7]Division of Oncology and Pathology, Department of Clinical Sciences Lund, Faculty of Medicine, Lund University, Lund, Sweden.

BACKGROUND. Neoantigen-driven recognition and T cell–mediated killing contribute to tumor clearance following adoptive cell therapy (ACT) with tumor-infiltrating lymphocytes (TILs). Yet how diversity, frequency, and persistence of expanded neoepitope-specific CD8+ T cells derived from TIL infusion products affect patient outcome is not fully determined.

METHODS. Using barcoded pMHC multimers, we provide a comprehensive mapping of CD8+ T cells recognizing neoepitopes in TIL infusion products and blood samples from 26 metastatic melanoma patients who received ACT.

RESULTS. We identified 106 neoepitopes within TIL infusion products corresponding to 1.8% of all predicted neoepitopes. We observed neoepitope-specific recognition to be virtually devoid in TIL infusion products given to patients with progressive disease outcome. Moreover, we found that the frequency of neoepitope-specific CD8+ T cells in TIL infusion products correlated with increased survival and that neoepitope-specific CD8+ T cells shared with the infusion product in posttreatment blood samples were unique to responders of TIL-ACT. Finally, we found that a transcriptional signature for lymphocyte activity within the tumor microenvironment was associated with a higher frequency of neoepitope-specific CD8+ T cells in the infusion product.

CONCLUSIONS. These data support previous case studies of neoepitope-specific CD8+ T cells in melanoma and indicate that successful TIL-ACT is associated with an expansion of neoepitope-specific CD8+ T cells.

FUNDING. NEYE Foundation; European Research Council; Lundbeck Foundation Fellowship; Carlsberg Foundation.

Introduction

Adoptive cell therapy with expanded tumor-infiltrating lymphocytes (TIL-ACT) can mediate durable tumor regression in patients with metastatic melanoma (1, 2). Furthermore, TIL-ACT has a high objective response rate even after the failure of checkpoint inhibitor therapy (1–4). TIL-ACT therefore represents an attractive treatment option for metastatic melanoma patients with high unmet medical needs. Current predictors of tumor regression and long-term survival after ACT include tumor-mutational burden (TMB) and neoantigen load (5), which have recently emerged as independent predictors of outcome across multiple immunotherapies (6, 7). Moreover, transcriptomic evidence implicates antigen presentation within the tumor microenvironment before TIL-ACT (5) as an important additional factor, suggesting that antigen presentation and immune recognition of mutation-derived neoantigens contribute to therapeutic benefit in TIL-ACT. While immune recognition and tumor cell killing are generally associated with a positive outcome (8, 9), evaluation of T cell recognition of mutation-derived neoantigens within TIL infusion (TIL Inf) products and peripheral blood after infusion have only been reported in case studies of complete responders (CRs) (10–15). We aimed to systematically assess T cell recognition toward

Authorship note: NPK, CH, and SAT are co–first authors.
Conflict of interest: SRH is the cofounder of Immumap, Tetramer-shop, and PokeAcell and is the coinventor of patents WO2015185067 (Determining antigen recognition through barcoding of MHC multimers) and WO2015188839 (General detection and isolation of specific cells by binding of labeled molecules) for the barcoded MHC technology that is licensed to Immudex. MD has received honoraria for lectures from Novartis and Roche. IMS is a cofounder, shareholder, and advisor for IO Biotech. IMS has an advisory board relationship with or has lectured for Roche, Novartis, MSD, Celgene, Incyte, TILT Bio, Pfizer, and BMS AstraZeneca.

Submitted: May 4, 2021; **Accepted:** November 18, 2021; **Published:** January 18, 2022.
Reference information: *J Clin Invest*. 2022;132(2):e150535.
https://doi.org/10.1172/JCI150535.

neoantigens in TIL-ACT and the influence of such recognition on therapeutic outcome. Recent advances in T cell technologies have led to the possibility of comprehensive screening of T cell recognition against large libraries of patient-derived neoepitopes (12, 16–18). Here, we used DNA barcode–labeled pMHC multimers to screen for $CD8^+$ T cell recognition, using 151 to 585 predicted neoepitopes per patient, presented in a multimeric form in the context of patient-matched HLA-I molecules (19, 20). Using this strategy, we determined the presence of $CD8^+$ T cells recognizing mutation-derived neoepitopes, here denoted as neoantigen-reactive T cells (NARTs), in TIL Inf products from 26 patients with metastatic melanoma. Furthermore, we examined the persistence of such T cells in samples of peripheral blood collected at multiple time points after therapy. This comprehensive mapping of NARTs demonstrates a substantial T cell reactivity level toward patient-derived neoepitopes and a positive influence on clinical outcome following TIL-ACT. This highlights the importance of detecting and enhancing the levels of such T cells in TIL-ACT.

Moreover, this study provides essential data to support efforts to identify the few immunogenic neoepitopes that give rise to T cell recognition out of the large number of predicted neopeptides. Recent efforts have been made to identify the parameters that determine the immunogenicity of a given neoepitope (21) and facilitate more accurate prediction of such sequences for therapeutic measures. In the current study, we evaluated a total of 5921 predicted neopeptides and identified T cell recognition toward 106 (1.8 %) of these in TIL Inf products. Using this large data set, we further assessed the influence of HLA binding, antigen expression level, clonality, TMB, and type of mutation on immunogenicity (i.e., recognition of a given neopeptide).

Results

Identification of neoepitope-reactive $CD8^+$ T cells. In a cohort of metastatic melanoma patients treated with TIL-ACT (Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/JCI150535DS1), prediction of patient-specific mutated HLA-I epitopes was performed using whole exome sequencing (WES) and RNA-Seq on tumor material and normal tissue PBMCs. The in silico neopeptide prediction platform MuPeXI (https://services.healthtech.dtu.dk/service.php?MuPeXI-1.1) was employed to identify single nucleotide variants and indels/frameshifts from the sequencing data specific to the cancer material (5, 20). Mutation-derived peptides were subsequently ranked using netMHCpan (20, 22) and transcription of the corresponding gene (transcripts per million [TPM]) (see Methods) with the aim of including at least 200 neopeptides per patient.

We covered 30 different HLA alleles ranging from 2 to 6 HLA alleles per patient (average, 4.4 HLAs) (Supplemental Figure 1, A and C); however, HLA-C*02:02 and C*05:01 were excluded from data analyses due to technical concerns. Thus, the final neopeptide library ranged from 151 to 585 peptides per patient (Supplemental Figure 1B), with the most frequent alleles in our cohort being HLA-A*01:01 and C*03:04 (Supplemental Figure 1C). In addition to neopeptides, we also included a small set of known $CD8^+$ T cell epitopes derived from common human viruses EBV, CMV, and influenza virus (FLU). These represent "bystander" T cells in the TIL Inf product and also serve as positive controls for the technical process.

DNA barcode–labeled neopeptide libraries were constructed as described previously (19) using UV-mediated peptide-MHC exchange (23, 24) and fluorescent streptavidin-labeled dextrans (see Methods). PBMCs and TIL Inf products were stained with patient-specific multimer libraries followed by sorting of multimer-binding $CD8^+$ T cells. The coattached DNA barcodes were amplified from the sorted T cell population to reveal antigen specificity (ref. 19 and Figure 1A). We defined biologically relevant NARTs as NARTs with an estimated frequency of at least 0.01% and without presence in partially HLA-matching healthy donor PBMCs. To assess the reproducibility of our pMHC multimer library screens, we screened TIL samples of 9 patients twice with the same library, demonstrating a correlation between technical replicates ($R = 0.55$; $P < 2.2 \times 10^{-10}$; Supplemental Figure 1D).

An example of the analysis of enriched DNA barcodes and their corresponding pMHC in a TIL Inf product from patient M22 (partial responder [PR]) is depicted in Figure 1B and for patients M14 (progressive disease [PD]) and M26 (CR) in Supplemental Figure 2. In patient M22, NARTs were detected for 3 of 4 HLA molecules included, although most reactivity was seen against HLA-A*01:01–restricted peptides. Of interest, 7 HLA-A*01:01–restricted neoepitopes recognized by the M22 TIL Inf product comprised the C-terminal amino acid sequence SILSY ($AKAP9^{P1796L}$), and $CD8^+$ T cells specific for each of these peptides were confirmed in TIL Inf products with single-tetramer staining (Supplemental Figure 3A).

From in silico structural models of the interaction between the different $AKAP9^{P1796L}$ peptide variants and the HLA-A*01:01 molecule, we observed that leucine (L), introduced by the mutation, protruded from the HLA-binding groove for potential interaction with a TCR. Furthermore, the four 8 to 10 mer epitope variants shared this conformation when bound to HLA-A*01:01 (Supplemental Figure 3B). This suggests that the $AKAP9^{P1796L}$ amino acid substitution has given rise to multiple neoepitopes that may be recognized by the same population of $CD8^+$ T cells, but with different affinities. The binding affinity hierarchy can be assessed both by the estimated frequency (Supplemental Figure 3C) and the MFI of the tetramer populations (Supplemental Figure 3D) and indicates favorable interaction with the 9 mer and 10 mer neoepitopes holding the SILSY motif.

Screening of TIL Inf products from 26 melanoma patients with personalized multimer libraries resulted in the detection of 106 different NART populations across the cohort. NARTs were detected in 18 out of 26 TIL Inf products, ranging from 0 to 13 NART populations per sample. To avoid any potential bias based on differences in HLA coverage, the number and frequency of detected NARTs were normalized to the average HLA coverage of the cohort (4.4 HLAs per patient). Following HLA normalization, the median number of NARTs per TIL Inf product was 3.7 (range 0–12.1, Figure 1C). Additionally, we detected the presence of virus-specific $CD8^+$ T cells toward a selected list of virus-derived epitopes in half of the TIL Inf products (13 out of 26 patients, Figure 1C), which is in line with previous analyses of TIL Inf products (25, 26). Across all TIL Inf products, we observed an estimated NART frequency of 0%–38.6% (median = 0.63%) of total $CD8^+$ T cells (Figure 1C).

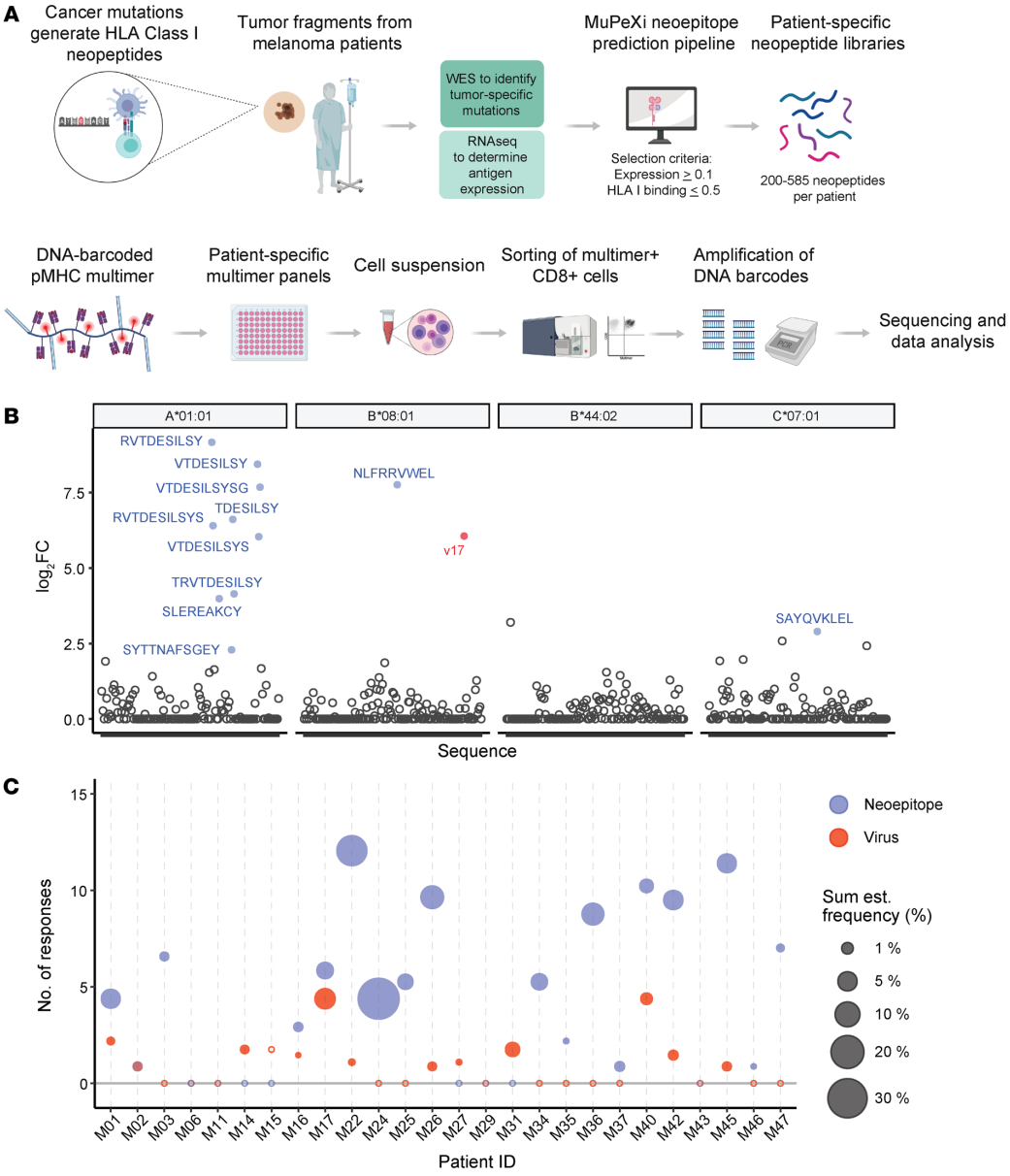


Figure 1. Detection of neoepitope-specific CD8⁺ T cells in expanded TILs of melanoma. (**A**) Melanoma-specific mutation-derived peptides were predicted to bind patient's HLA molecules using the prediction platform MuPeXI. DNA barcode–labeled MHC multimers with either neopeptides or virus-derived peptides were assembled on a PE-labeled streptavidin-conjugated dextran backbone. Multimer-binding NARTs were fluorescence sorted and T cell specificities decoded by barcode sequencing. (**B**) Examples of neoepitope- and virus-specific CD8⁺ T cells detected in expanded TILs of melanoma patient M22 (PR) across available HLAs. Significant barcode enrichment is defined based on a $\log_2$ FC of the number of barcode reads compared with triplicate baseline samples. $P \leq 0.001$ (egdeR) after correction for multiple hypothesis testing (see Methods). Blue, NARTs; red, virus-specific CD8⁺ T cells; black, multimers with nonenriched barcodes. V17 annotate EBV peptide RAKFKQLL. (**C**) Number and frequency of neoepitope- and virus-specific CD8⁺ T cells in TIL samples across cohort of 26 melanoma patients. Blue, NARTs; red, virus-specific CD8⁺ T cells. Number of and frequency of NARTs were normalized to absolute HLA coverage (see Methods). Sum est. frequency, sum of estimate frequency.

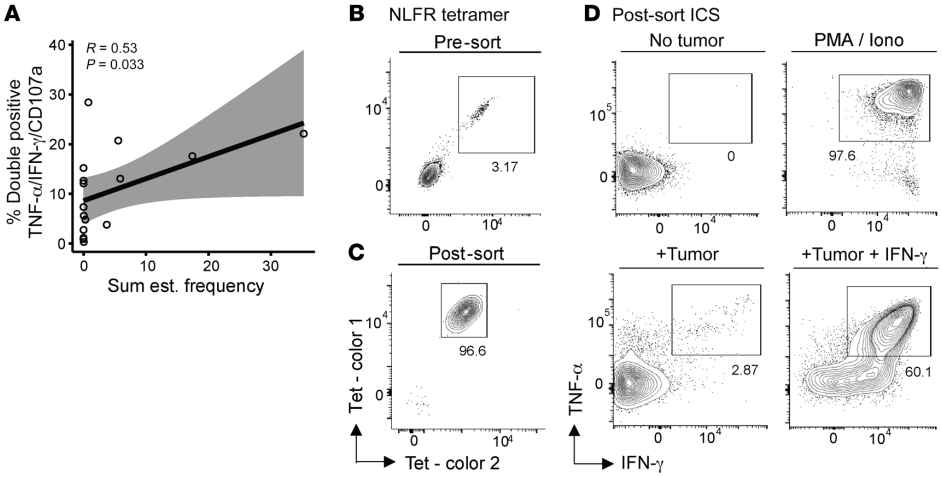


Figure 2. Autologous tumor recognition by enriched NARTs. (**A**) Correlation of TIL reactivity to autologous tumor (measured by intracellular cytokine staining) and sum of estimated NART frequency. TIL reactivity toward an autologous tumor cell line was defined as positive for 2 out of the 3 proteins TNF-α, IFN-γ, and CD107a. Sixteen patients with available tumor reactivity data were included from both responder (n = 6) and nonresponders (n = 10). R and P values from Spearman's correlation with 95% CIs in gray. NART frequency was normalized to absolute HLA coverage (see Methods). (**B** and **C**) HLA-B*08:01–restricted, NLFR-specific CD8$^+$ T cells from M22 TIL Inf product were sorted based on 2-color tetramer binding (**B**) and expanded in vitro followed by NLFR-tetramer staining (**C**). (**D**) Tumor reactivity as measured by TNF-α/IFN-γ release after coculture of expanded, NART-specific cell products with or without autologous tumor cell lines, with PMA/ionomycin or with autologous tumor cell line and IFN-γ. NLFR, NLFRRVWEL from USP34^{S1391F}. TIL reactivity data shown in **A** originate from previous study (4), and the assay was performed as described previously (66).

Recognition of melanoma tumor cells by NARTs in vitro. The TIL Inf product from most patients (16 of 26) was previously analyzed for tumor recognition properties in terms of cytokine secretion toward an autologous tumor cell line, generated from the same tumor biopsy as the TIL Inf product (4). The estimated frequency of NARTs identified in this study correlated with the capacity of the TIL Inf product to recognize the tumor, indicating that detected NARTs may indeed contribute to tumor cell recognition (Figure 2A). While a significant association was observed, the effect on cytokine secretion from other immune subsets, tumor antigen classes, or NARTs restricted to HLA alleles not included in our study cannot be excluded.

We additionally investigated the direct tumor-recognition capacity of sorted and expanded neoepitope-specific T cell populations. From the patient M22 TIL Inf product, we sorted USP34^{S1391F}–derived NLFR-HLA-B*08:01–specific T cells using tetramers. The presence of such T cells was verified (3.2%, Figure 2B), and postsort expansion resulted in purity of greater than 96% (Figure 2C). The expanded NLFR-HLA-B*08:01–specific T cells displayed tumor recognition determined by cytokine secretion upon coculture with an autologous tumor cell line with (60.1%) and without (2.87%) pretreatment with IFN-γ (Figure 2D). Thus, tumor recognition was specific and greatly enhanced by IFN-γ pretreatment of the autologous tumor cell line. It has previously been demonstrated that IFN-γ pretreatment enhances MHC-I expression and antigen presentation in both autologous (8) and established tumor cell lines (27). We also sorted CD8$^+$ T cells specific to 2 AKAP9^{P1796L} peptide variants followed by rapid expansion (Supplemental Figure 4A), which recognized their respective AKAP9^{P1796L} variants (Supplemental Figure 4B). These sorted T cell populations both recognized autologous tumor cell lines with and without pretreatment with IFN-γ (Supplemental Figure 4C). This indicates that some multimer-detected NARTs are capable of further expansion and can specifically recognize autologous tumor cell lines.

The number and frequency of NARTs are associated with the clinical outcome of TIL-ACT. Next, we investigated whether higher diversity (number of responses) and frequency of NARTs in TIL Inf products correlate with improved clinical efficacy of TIL-ACT. NARTs were detectable across all Response Evaluation Criteria in Solid Tumors (RECIST), version 1, groups (28), although they were severely depleted from TIL Inf products given to patients that developed PD (n = 6) (Figure 3A). Overall, NARTs tended to demonstrate greater diversity in products from responders compared with nonresponders (Figure 3B).

The estimated NART frequency within TIL Inf products was significantly higher in responders compared with nonresponders (Figure 3, C and D, and Supplemental Figure 5, A and B), suggesting that NART frequency affects clinical outcome. Tumor mutational burden and number of predicted neoepitopes were uniformly distributed across RECIST groups (Supplemental Figure 5, D and H), and no difference was observed between responders and nonresponders (Supplemental Figure 5, E and I). Tumor mutational burden was, however, associated with longer progression-free survival (PFS) (Supplemental Figure 5F), as previously indicated (5), although we did not observe a strong influence of the number of predicted neoepitopes on PFS (Supplemental Figure 5J).

Next, we investigated whether the diversity and frequency of NARTs within TIL Inf products affected PFS and overall survival (OS). Patients in whom the number of NARTs was above the medi-

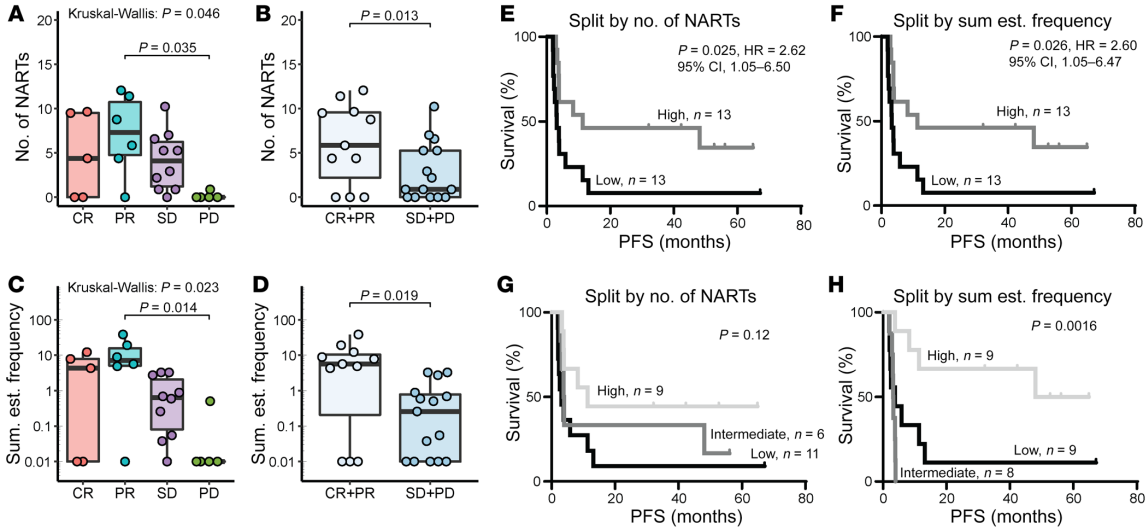


Figure 3. Frequency of NARTs correlates with increased survival after TIL-ACT. (**A** and **B**) NART diversity represented as the number of NARTs detected in TIL Inf products for each patient according to RECIST (**A**) and clinical response (**B**). (**C** and **D**) NART frequency represented as the sum of estimated frequency of NARTs detected in TIL Inf products for each patient according to RECIST (**C**) and clinical response (**D**). (**E** and **F**) PFS for the cohort split by median NART diversity (median = 3.65 NARTs) (**E**) and median NART frequency (median = 0.63 %) (**F**). (**G** and **H**) PFS for the cohort splits by high (>66th percentile), intermediate (> 33rd percentile), and low groups (≤33rd percentile). (**G**) NART diversity. 66th percentile = 5.65 NARTs. 33rd percentile = 0.88 NARTs. (**H**) NART frequency. 66th percentile = 3.26%. 33rd percentile = 0.03%. *P* values were calculated using Kruskal-Wallis test followed by Dunn's multiple comparison test in **A** and **C**; only significant comparisons are shown. Nonparametric Mann-Whitney *U* test was used for **B** and **D**. Box plot whiskers represent IQR. *P* values and HRs were calculated using the Mantel-Cox test and log-rank approach, respectively (**F**). *P* values for **G** and **H** were calculated using log-rank test. Both number of and frequency of NARTs were normalized to absolute HLA coverage (see Methods). *n* = 26 for all plots. All values displayed on a logarithmic scales were increased by 0.01 to account for 0 values.

an of 3.7 (high, $n = 13$) had increased PFS ($P = 0.025$, HR 2.62; 95% CI = 1.05–6.50) compared with patients below the median (low, $n = 13$; Figure 3E). Likewise, patients with a high NART frequency within TIL Inf products (median = 0.7%) (high, $n = 13$) demonstrated significantly improved PFS ($P = 0.026$, HR 2.60; 95% CI = 1.05–6.47) compared with patients with low NART frequency (low, $n = 13$; Figure 3F). High NART frequency also showed a positive effect on OS (Supplemental Figure 6B); however, no such correlation was found with NART diversity (Supplemental Figure 6A). Note that OS might also be affected by subsequent treatment given after TIL-ACT.

Interestingly, the clinical impact of NART frequency was most prominent for patients above the 66th percentile. For NART frequency, the high patient group (above the 66th percentile, $n = 9$) showed significantly longer PFS ($P = 0.0016$; Figure 3H) and OS ($P = 0.021$; Supplemental Figure 6D) compared with the intermediate patients (equal to or below the 66th percentile and greater than the 33rd percentile, $n = 8$) or low patients (equal to or below the 33rd percentile, $n = 9$). In contrast, NART diversity did not significantly affect survival (PFS and OS) when comparing groups split by the 66th and 33rd percentiles (Figure 3G and Supplemental Figure 6C). The 66th and 33rd percentiles corresponded to a frequency of 3.26% and 0.03%, respectively, while the same percentiles for NART diversity were 5.65 and 0.88 NARTs, respectively.

In our analysis, T cells that recognized different overlapping peptides originating from the same mutation were defined as multiple individual NART populations. However, T cell recognition of multiple neopeptides could also arise from crossreactivity of a single NART population toward several similar epitopes. To avoid any bias in our data analyses based on such potential crossreactive T cell populations, we reduced the number of detected NART responses to the number of unique immunogenic somatic mutations recognized by NARTs (median = 2.6) and redid our survival analysis using the most frequent NART as a proxy for recognition of all overlapping epitopes from the same nonsynonymous mutation. The result showed a similar association: both NART diversity and frequency correlated with increased PFS, whereas only frequency correlated with increased OS (Supplemental Figure 6, E–H), ensuring that contribution from T cell recognition of overlapping epitopes did not bias our overall observation. In summary, these data suggest that high frequency of NARTs positively affects therapeutic outcome following TIL-ACT.

NARTs are detected in peripheral blood after TIL-ACT and decline over time. As indicated by others (10, 29), an essential factor for TIL-ACT efficacy is the capacity of transferred T cells to persist in patients following therapy. This can be measured based on their presence in peripheral blood over time after transfer. For 19 patients, available blood samples were taken 8 days before TIL-ACT and at different time points after TIL Inf, i.e., less than 1 month after TIL-ACT, less than 4 months after TIL-ACT, less than 12 months after TIL-ACT, less than 24 months after TIL-ACT, and less than 48 months after TIL-ACT (Supplemental

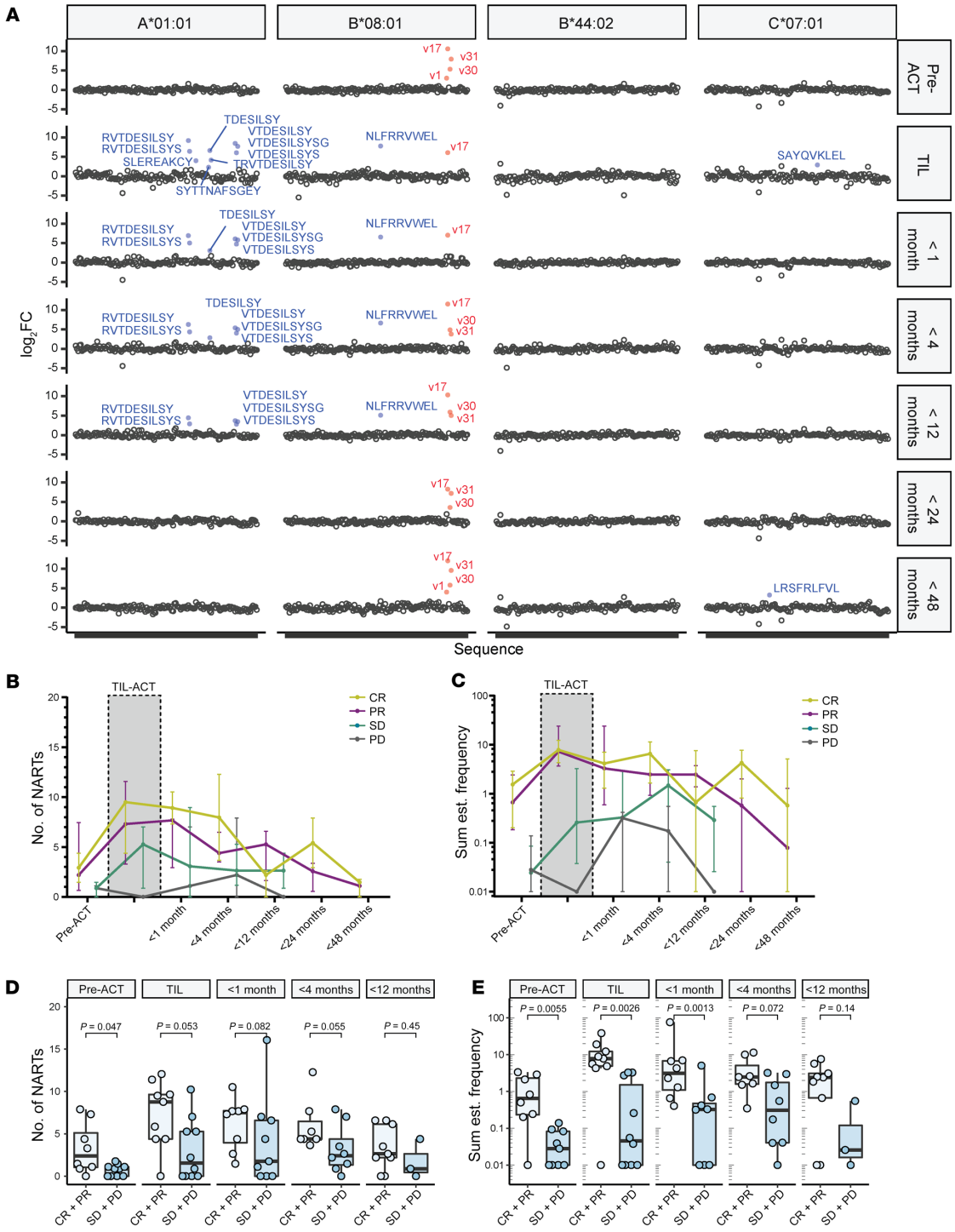

A
A*01:01
B*08:01
B*44:02
C*07:01
Pre-ACT
TIL
< 1 month
< 4 months
< 12 months
< 24 months
< 48 months
log2FC
Sequence
TDESILSY
RVTDESILSY
RVTDESILSYS
VTDESILSY
VTDESILSYSG
VTDESILSYS
SLEREAKCY
TRVTDESILSY
SYTTNAFSGEY
NLFRRVWEL
SAYQVKLEL
LRSFRLFVL
v17
v31
v30
v1
B
TIL-ACT
No. of NARTs
CR
PR
SD
PD
Pre-ACT
<1 month
<4 months
<12 months
<24 months
<48 months
C
TIL-ACT
Sum est. frequency
D
Pre-ACT
TIL
<1 month
<4 months
<12 months
P = 0.047
P = 0.053
P = 0.082
P = 0.055
P = 0.45
CR + PR
SD + PD
E
P = 0.0055
P = 0.0026
P = 0.0013
P = 0.072
P = 0.14

Figure 4. NARTs appear in peripheral blood and decline in frequency following TIL-ACT. (**A**) Output example from screening paired PBMCs from 19 patients. Virus- and neoepitope-specific CD8+ T cells in patient M22 (PR) in pre-ACT PBMCs, TIL Inf product, and PBMCs following TIL-ACT. Blue, NARTs; red, virus-specific CD8+ T cells; black, multimers associated with nonenriched barcodes. Significant barcode enrichment is defined based on a $\log_2$ FC of the number of barcode reads compared with triplicate baseline samples. $P < 0.001$ (egdeR) (see Methods). V1 annotate FLU peptide ELRSRYWAI, v17 annotate EBV peptide RAKFKQLL, v30 annotate EBV peptide QAKWRLQTL, and v31 annotated EBV peptide FLRGRAYGL. (**B** and **C**) Median number of NARTs. Error bars indicate IQR. Points were displaced for visual purposes. (**B**) Number of NART responses and sum of estimated NART frequency (**C**) over time in TIL Inf product and available PBMC samples. Patients were divided according to RECIST groups. (**D** and **E**) Box plots representing diversity (**D**) and frequency (**E**) of NARTs for each patient according to RECIST groups. *P* values were calculated using Mann-Whitney *U* test. Nineteen patients had both TIL Inf products and PBMCs available, but the number of samples at each time point varied according to sample and data availability (Supplemental Table 1 and Supplemental Figure 8). NART frequency could not be calculated for M40 PBMCs before ACT and for M40 PBMCs less than 1 month after treatment (see Methods) and are therefore excluded in **C** and **E**. Whiskers represent IQR. NART frequencies were normalized to HLA coverage of the given patient (see Methods). All values displayed on logarithmic scales were increased by 0.01 to account for 0 values.

Table 1). NARTs present in the first or later PBMC samples after ACT were defined as engrafted. Furthermore, if a given NART was detected in multiple later PBMC samples, that NART was regarded as persisting. Each sample was screened for T cell recognition toward neopeptides included in the full patient-specific neoepitope-MHC library, exemplified by patient M22 (PR) (Figure 4A). In M22, only virus-specific T cells could be detected in the pre-ACT PBMC sample, namely, B*08:01-restricted CD8+ cells capable of binding epitopes FLU-ELR (v1), EBV-RAK (v17), EBV-QAK (v30), and EBV-FLR (v31). These virus-specific CD8+ T cells were detectable throughout most time points, while NARTs engrafted (PBMC <1 month) and persisted in the following PBMC samples up to 1 year after treatment.

Similar NART kinetics were observed in patient M45 (PR), with NARTs recognizing overlapping neoepitope containing the mutated sequence SAGA ($SORC2^{A1093S}$) (Supplemental Figure 7). $SORC2^{A1093S}$ was first recognized in the M45 TIL Inf product, and immune recognition persisted in PBMCs until the last recorded time point (<12 months). Furthermore, M45 showed immune recognition toward the same neoepitope DIHF ($ZNF786^{M87I}$) bound to multiple HLA alleles (HLA-A*01:01, A*24:02, and B*13:02). Recognition of $ZNF786^{M87I}$ was initially discovered in the TIL Inf product, and while it persisted on HLA-A*24:02 until the last time point for M45, it appeared to incompletely persist on HLA-A*01:01 and B*13:02. Overall, this suggests that $ZNF786^{M87I}$ produces a promiscuous neoepitope capable of binding multiple HLAs, with a preference for HLA-A*24:02. HLA promiscuity is otherwise known to occur for viral epitopes (30).

The median NART diversity and frequency across RECIST categories were followed to assess the overall kinetics of NARTs after ACT. Note that most nonresponders did not have PBMC samples for less than 12 months and thereafter (7 of 10). NART diversity increased markedly when comparing pre-ACT PBMCs and the TIL Inf product and declined over time after TIL-ACT in the CR, PR, and stable disease (SD) patient groups, displaying the expansion of NART populations in the TIL Inf product and their persistence after therapy (Figure 4B). NART frequency demonstrated kinetics similar to those of NART diversity. However, only responders appeared to have substantial frequencies of NARTs within TIL Inf products (Figure 4C). Unlike those in the other groups, patients with PD did not display any NARTs within TIL Inf products ($n = 3$); however, they did appear to have ongoing NART recognition in peripheral blood before and after therapy, although at lower frequencies (Figure 4, B and C).

Finally, we compared responders and nonresponders in relation to NART diversity across all time points and found that responders had a higher level of NART diversity in PBMCs collected before TIL-ACT (Figure 4D). Similarly, we found increased NART frequency in responders before TIL-ACT, within TIL Inf products, and at early time points following infusion (>1 month; Figure 4E).

In conclusion, we observed a broad repertoire of NARTs recognizing single neoepitopes, overlapping neoepitopes, and HLA promiscuous neoepitopes in TIL Inf products of metastatic melanoma patients treated with TIL-ACT. These NARTs showed signs of engraftment and could persist in peripheral blood after TIL-ACT. Furthermore, we observed that responders had a higher estimated NART frequency before and following TIL-ACT in peripheral blood, supporting prior prospective efforts (31).

Engrafted neoepitope-specific CD8+ T cells dominate immune recognition in responders of TIL-ACT. To better understand the dynamic relationship among preexisting, ongoing, and TIL-derived immune recognition, we annotated each detected NART according to its first appearance from 8 days prior to therapy (pre-ACT PBMCs) to the last available time point. Thus, if a NART population appeared exclusively in pre-ACT samples, it was annotated pre-ACT. If a given NART was detected in both pre-ACT PBMCs and in the given TIL Inf product, it was denoted pre/TIL, while if it first appeared in the infusion product, it was denoted TIL. Finally, if a NART population first appeared in a later PBMC sample it was regarded as novel, annotated with its first time of appearance and followed from there on out (see patient overview in Supplemental Figure 8).

Using this categorization, we observed that persisting NARTs derived from the TIL Inf product (Pre/TIL plus TIL) were present across responders and patients with SD at multiple time points after infusion, but absent in patients with PD (Figure 5, A and B, and Supplemental Figure 8). Additionally, we observed that 7 of 8 responders and 5 of 10 nonresponders with available pre-ACT material had preexisting NARTs (pre-ACT plus pre/TIL). Preexisting NARTs are likely clinically relevant, as TIL Inf products from responders were overall dominated by preexisting immune recognition that was further expanded to high frequencies within the TIL Inf product (pre/TIL) (Figure 5B and Supplemental Figure 8). Note, however, that the presence of preexisting NARTs that were further expanded did not appear sufficient to generate a clinical response, as we also observed pre/TIL NARTs in 3 patients with SD (Supplemental Figure 8). The perceived therapeutic benefit of preexisting NARTs that were further expanded may therefore relate more to the high frequency and persistence after expansion in selected patients than to their presence alone.

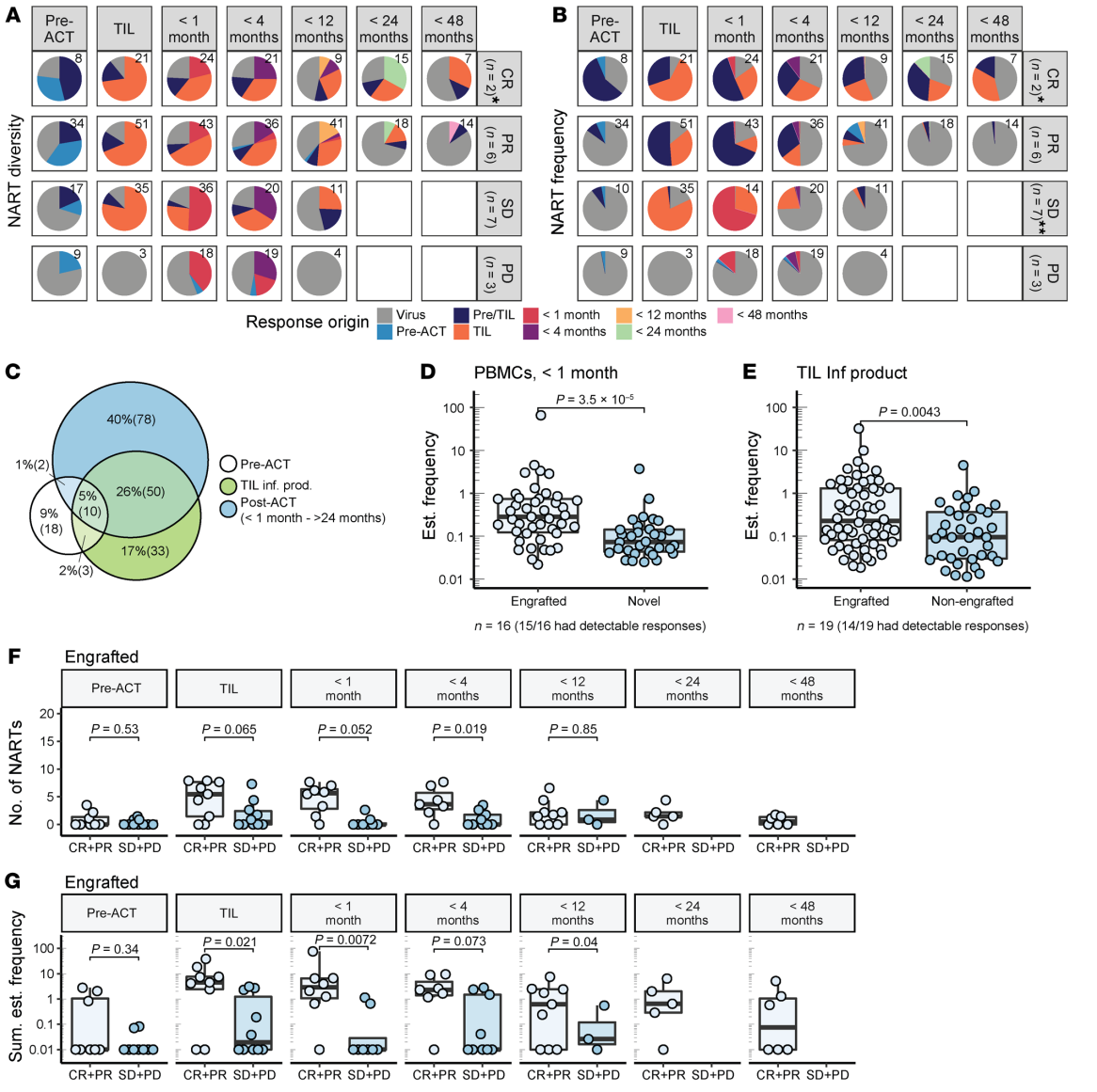


Figure 5. Responding patients have high-frequency engrafting NARTs in their TIL Inf product. (**A** and **B**) Each NART population was annotated and colored according to first appearance in pre-ACT PBMCs, TIL Inf products, and post-ACT PBMCs (<1 month to <48 months). Black numbers specify the total number of NARTs detected for the specific time and RECIST group. (**A**) Distribution of NARTs within RECIST groups according to first appearance. (**B**) Distribution of NART frequency within RECIST groups according to first appearance. *M01 (CR) did not have pre-ACT and <1 month PBMCs available and was excluded from analysis to avoid a biased distribution. **Frequency data could not be calculated for M40 pre-ACT and M40 post-ACT <1 month, which were excluded (see Methods). (**C**) Venn diagram showing the overlap of detected NARTs among pre-ACT PBMCs, TIL Inf products, and all post-ACT PBMC samples. *n* = 19. (**D**) The estimated frequency of each NART population detected less than 1 month after infusion. Responses were either regarded as engrafted (i.e., also detected in TIL Inf) or novel. *n* = 16. M01 and M40 were excluded as stated for **A** and **B**; M29 did not have detectable antigen-specific $CD8^+$ T cells before the second time point after ACT. (**E**) The estimated frequency of each NART population observed in TIL Inf products. Nonengrafted versus engrafted (i.e., detected at least once at a later time points). *n* = 19. (**F** and **G**) Number and frequency of engrafted NARTs, defined by presence in both TIL Inf product and after TIL-ACT. *n* varied according to sample availability (Supplemental Table 1 and Supplemental Figure 8). M40 before ACT and <1 month PBMCs were excluded from **G** (see Methods). Sum of estimated frequency in **G** was increased by 0.01 to account for 0 values. *P* values from Mann-Whitney *U* test. Whiskers represent IQR.

We observed that 62.5% (60 of 96) of NARTs observed in TIL Inf products were also detectable after ACT (Figure 5C). Furthermore, 57% of NARTs detected after ACT were novel and did not originate from the TIL Inf product (80 of 140), whereas 43% originated from the TIL Inf product (60 of 140; Figure 5C). These novel NARTs were transiently appearing and could represent epitope spreading. However, their appearances may not necessarily have therapeutic benefit, as they were observed across all RECIST groups (Figure 5, A and B) and present at lower frequency than newly engrafted NARTs (TIL NARTs present in post-ACT PBMCs) (Figure 5D). Finally, we observed that engrafted NARTs derived from the TIL Inf product (TIL plus pre/TIL) had a higher estimated frequency compared with their nonengrafted counterparts in the TIL Inf product (Figure 5E), suggesting engraftment to be associated with prior frequency.

To evaluate the impact of engrafted NART populations separately from that of nonengrafted and novel NARTs, appearing only in TIL Inf products and post-ACT PBMCs, respectively, we compared the diversity and frequency of engrafted NARTs (pre/TIL and TIL) in responders and nonresponders with available PBMCs throughout all time points (Figure 5, F and G). Interestingly, we observed that nonresponders had a markedly lower diversity (Figure 5F) and frequency (Figure 5G) of engrafted NARTs compared with responders in the first 2 sampling time points after ACT (<1 month and <4 months). These data suggest that responders were treated with TIL Inf products characterized by high-frequency, engrafting NARTs, whereas nonresponders were treated with TIL Inf products containing a relatively lower frequency of NARTs that were unable to engraft and persist after ACT. This is in line with prior TCR-sequencing efforts (29).

The characteristics of immunogenic neoepitopes. Based on the large screen presented here, we evaluated T cell recognition against 5921 predicted neopeptides that were selected based on their HLA-binding characteristics and gene transcriptional levels in tumor next-generation sequencing (NGS) data. Of these predicted neopeptides, we detected specific $CD8^+$ T cell recognition toward 204 neoepitopes in either TIL Inf products or PBMC samples from melanoma patients, while the remaining 5717 were not recognized by T cells in the evaluated patients (Figure 6A). The pool of immunogenic neoepitopes displayed characteristics related to both clonality and C/T mutations similar to that of the total library of evaluated neopeptides (Figure 6A). Hence, we did not observe a specific enrichment of T cell recognition toward clonal mutations, as has previously been suggested for non–small cell lung cancer (NSCLC) (32). Interestingly, cancer-driver genes (33) are significantly overrepresented in the fraction of immunogenic neoepitopes compared with the fraction of nonimmunogenic neopeptides (Figure 6A; 6.5% versus 3.3%, $P = 0.0043$). However, we did not find any immunogenic neoepitopes to be shared among patients, as has previously been observed in TILs isolated from colorectal cancer (34).

Our neopeptide library was preselected for predicted HLA binding. Within this pool, neoepitopes can be classified as either conserved binders (CBs), i.e., neopeptides with HLA binding similar to that of the mutated peptide versus the germ-line sequence, or improved binders (IBs), where the mutation affects HLA-binding capabilities, resulting in a neopeptide with improved HLA affinity compared with the germline sequence, as defined in Bjerregaard et al. (ref. 35 and Figure 6B). Immunogenic neoepitopes were represented in both categories, and we did not observe a significantly different distribution of immunogenic versus nonimmunogenic neopeptides among CBs versus IBs (3.4% CB versus. 3.5% IB, $P = 0.99$; Figure 6B). Furthermore, within the selected HLA affinity range evaluated here, we did not observe any further impact of HLA percentage rank score on neopeptide immunogenicity, evaluated as the potential enrichment of immunogenic neoepitopes below a percentage rank score of 0.5 ($P = 0.71$, z test; Figure 6C). In line with previous findings (34), we observed an enrichment of genes with RNA expression greater than 2 TPM among immunogenic neopeptides (Figure 6D; $P = 0.001$, z test).

TMB and predicted neoepitope load have previously been demonstrated as predictive for TIL-ACT outcome (5). We found a strong correlation between TMB and number of predicted neoepitopes (Supplemental Figure 9A). However, we did not find a correlation between TMB or the number of predicted neoepitopes and NART diversity and NART frequency (Supplemental Figure 9, B–E). This indicates that the presence of NARTs in TIL Inf products is an independent marker of therapeutic outcome in patients with metastatic melanoma. Since the interpatient variation in neopeptide library size may affect the correlation, we also correlated the number of NARTs detected and estimated frequency within the top 151 predicted neoepitopes so that it showed an equal representation of all patients (Supplemental Figure 9G). Again, no strong correlation was evident (Supplemental Figure 9, H–K), as multiple patients with low TMB showed medium-sized populations of neoepitope-specific $CD8^+$ T cells in their respective TIL Inf products (Supplemental Figure 9G). This emphasizes the need to improve our predictive capacity for identification of those neoepitopes that give rise to functional T cell recognition and tumor cell killing and furthermore highlights that other parameters, beyond TMB, affect immune recognition.

The tumor microenvironment has a substantial influence on the capacity of the immune system to mount a T cell response toward the tumor and for such T cells to exert their function. Although the generation of TIL Inf products is conducted in vitro, the tumor microenvironment may still affect the capacity for T cell expansion and function. We used the available transcriptomic data from our neoepitope prediction pipeline as input for a differential gene expression analysis, grouping patients according to higher or lower than median sum of estimated NART frequency within TIL Inf products. From this, we observed 226 differentially expressed genes (Figure 6E), that were associated with 383 enriched Gene Ontology (GO) gene sets (36). The top 20 enriched GO gene sets were a collection of humoral and B cell–mediated mechanisms and several pathways pertaining to the immune cell signal transduction (Supplemental Figure 10). These gene sets are highly relevant in light of the recently revealed relationship among intratumoral lymphoid structures, antigen presentation, and therapeutic benefit following immunotherapy (37). Of further interest, we observed enriched presence of GO terms relating to lymphocyte-mediated immunity (Figure 6F) and increased T cell proliferation (Figure 6G).

Discussion

Immune recognition and tumor killing by cytotoxic T cells are associated with a positive outcome across multiple immunother-

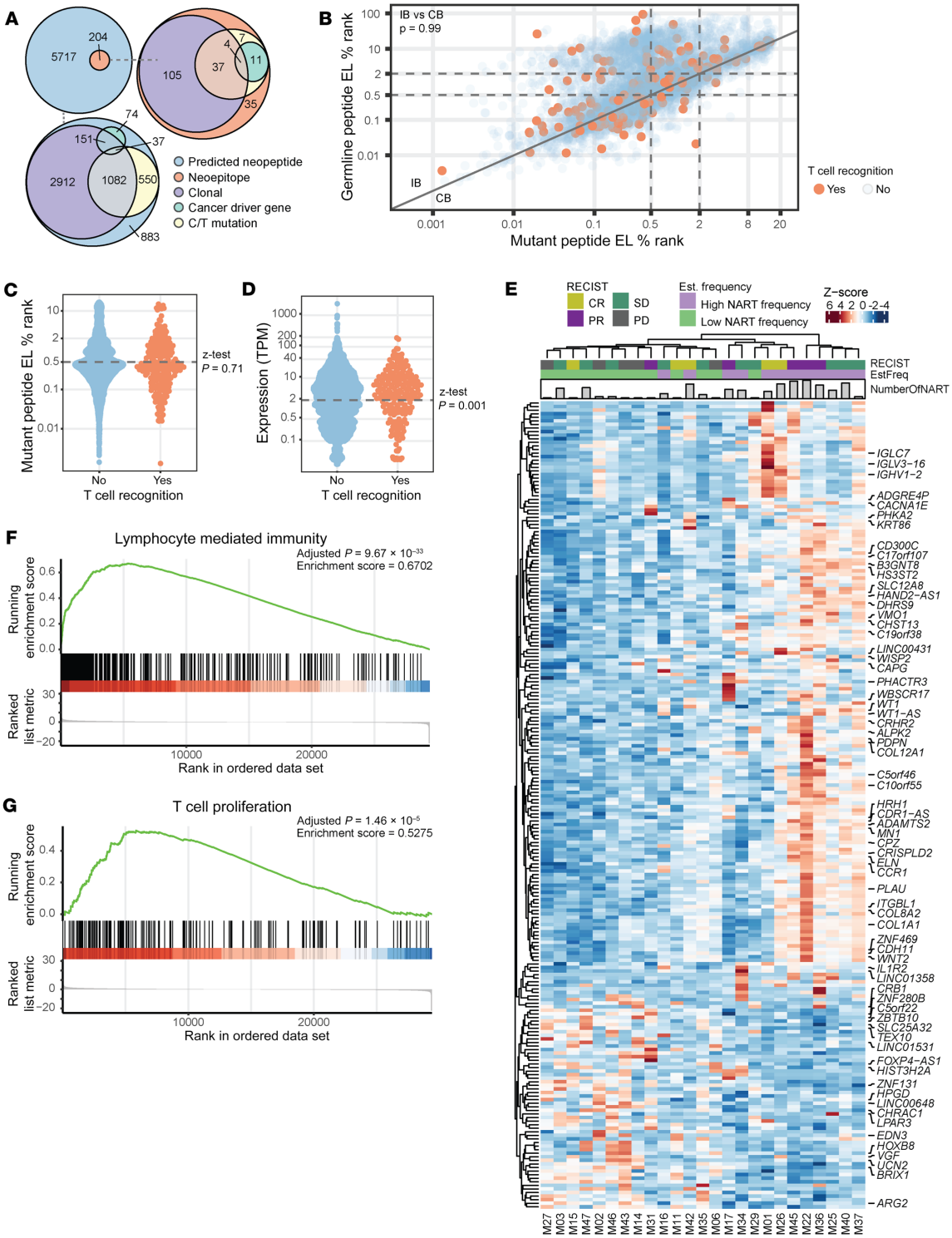

A
204
5717
105
4
7
37
11
35
74
151
37
2912
1082
550
883
Predicted neopeptide
Neoepitope
Clonal
Cancer driver gene
C/T mutation
B
IB vs CB
p = 0.99
Germline peptide EL % rank
Mutant peptide EL % rank
IB
CB
T cell recognition
Yes
No
C
Mutant peptide EL % rank
z-test
P = 0.71
No
Yes
T cell recognition
D
Expression (TPM)
z-test
P = 0.001
No
Yes
T cell recognition
E
RECIST
CR
SD
PR
PD
Est. frequency
High NART frequency
Low NART frequency
Z−score
6 4 2 0 -2-4
RECIST
EstFreq
NumberOfNART
IGLC7
IGLV3−16
IGHV1−2
ADGRE4P
CACNA1E
PHKA2
KRT86
CD300C
C17orf107
B3GNT8
HS3ST2
SLC12A8
HAND2−AS1
DHRS9
VMO1
CHST13
C19orf38
LINC00431
WISP2
CAPG
PHACTR3
WBSCR17
WT1
WT1−AS
CRHR2
ALPK2
PDPN
COL12A1
C5orf46
C10orf55
HRH1
CDR1−AS
ADAMTS2
MN1
CPZ
CRISPLD2
ELN
CCR1
PLAU
ITGBL1
COL8A2
COL1A1
ZNF469
CDH11
WNT2
IL1R2
LINC01358
CRB1
ZNF280B
C5orf22
ZBTB10
SLC25A32
TEX10
LINC01531
FOXP4−AS1
HIST3H2A
ZNF131
HPGD
LINC00648
CHRAC1
LPAR3
EDN3
HOXB8
VGF
UCN2
BRIX1
ARG2
M27 M03 M15 M47 M02 M46 M43 M14 M31 M16 M11 M42 M35 M06 M17 M34 M29 M01 M26 M45 M22 M36 M25 M40 M37
F
Lymphocyte mediated immunity
Adjusted P = 9.67 × 10−33
Enrichment score = 0.6702
Running enrichment score
Ranked list metric
Rank in ordered data set
G
T cell proliferation
Adjusted P = 1.46 × 10−5
Enrichment score = 0.5275
Running enrichment score
Ranked list metric
Rank in ordered data set

Figure 6. Characteristics of immunogenic neoepitopes. (**A**) Venn diagram of 5921 unique pMHC; 204 immunogenic and 5717 nonimmunogenic as determined by the presence of neoepitope-specific CD8$^+$ T cells in patients at any time. The distribution and overlap of immunogenic versus nonimmunogenic neoepitopes deriving from either cancer-driver genes (6.5% versus. 3.3%, $P = 0.0048$, *z* test), C/T mutations (3.4% versus. 3.5%, $P = 0.78$, *z* test), or clonal mutations (80.1% versus 86.0% $P = 0.03$, *z* test). Clonality could not be determined for 913 neopeptides, as WES was performed on autologous tumor cell lines (M22, M24, and a subset of M15). These were excluded from the *z* test, but included in the Venn diagram as subclonal mutations for visualization. (**B**) Eluted ligand (EL) percentage rank score of mutated peptide compared with percentage rank score of the corresponding germline peptide without mutation or nearest germline peptide. Red, immunogenic peptides. 3.4% CB versus 3.5 % IB, $P = 0.99$, *z* test. (**C**) Mutant EL percentage rank score comparing proportion of immunogenic neoepitopes above and below 0.5 percentage rank score (3.3 % versus 3.5, $P = 0.71$, *z* test). (**D**) RNA expression (TPM) comparing proportion of immunogenic peptides with expression above and below 2 TPM (4.2 % versus. 2.6%, $P = 0.001$, *z* test). (**E**) Unsupervised clustering of the 226 differentially expressed gene according to high and low sum of estimated frequency within TIL Inf products split by the median frequency (0.63%). Denoted names were prioritized according to GO terms and known function. (**F**) Enriched GO gene set representing lymphocyte-mediated immunity. (**G**) Enriched GO gene set representing T cell proliferation. Significance threshold or GSEA was set at FDR ≤ 0.01. M24 was excluded from **D**–**G**, as RNA-Seq data were obtained from an autologous tumor cell line. $n = 25$. M22 was included in **D**–**G** using data from the tumor biopsy used for manufacturing of the infusion product.

apies (9, 32, 38); however, the presence of neoepitope-specific CD8$^+$ T cells in TIL-ACT remains incompletely documented outside case responders (10–15). In the present study, we investigated the capacity of TIL Inf products to recognize predicted, HLA-binding neoepitopes originating from expressed, nonsynonymous mutations from 26 patients with metastatic melanoma. To this end, we utilized DNA barcode–labeled pMHC multimers from which we quantified NART diversity and frequency in TIL Inf products and patient PBMCs. We report recognition of a total of 106 neoepitopes within TIL Inf products from this cohort across all 4 RECIST groups. Supporting that the presence of NARTs affects the clinical response to TIL-ACT, we found that NART diversity and frequency were substantially lower in patients with PD when compared with patients with SD and PR and that NART frequency correlated with PFS and was higher in patients with clinical response to TIL-ACT (CR+PR).

We found that both NART diversity and frequency were highly variable across RECIST groups, especially within responding patients: 3 out of 11 CR/PR patients had 0 detectable NART populations. This variability could be due to limitations in neoepitope selection, contribution from other antigen types, insufficient HLA coverage, sampling bias, NART response frequencies below the threshold for detection (i.e., resulting in false-negative detection), or other NART-independent and/or HLA-I–independent pathways such as the MR1-dependent immune-recognition pathway (39).

Following each NART population from first appearance to last available PBMC time point further uncovered that responders were characterized by circulating NARTs of higher diversity and frequency in pretreatment PBMCs. This is interesting because pretreatment circulating NARTs could represent a biomarker for ongoing tumor recognition by CD8$^+$ T cells, which, in extension, could provide a noninvasive way to measure immune activity of the tumor. However, identification of NARTs is a laborious and patient-specific process, and for biomarker purposes, a simpler measurement of NART reactivity should be developed. Responders were furthermore predominantly treated with TIL Inf products of high NART frequency capable of engrafting and persisting after TIL-ACT at an estimated frequency higher than 0.01%. Additionally, we observed that engrafted NARTs initially appeared with an overall higher estimated frequency in the TIL Inf product compared with nonengrafted NARTs, which indicates that successful NART expansion precedes successful engraftment. As mentioned, the persistence of tumor antigen–specific TCRs has been hypothesized to drive therapeutic benefit following TIL-ACT (29). Interestingly, this hypothesis has recently been supported in the metastatic melanoma setting (40), where the persistence of neoantigen-specific TCRs after TIL-ACT correlated with CD39$^-$ CD69$^-$ stem-like T cells capable of self-renewal, differentiation, and further expansion upon stimulation. Future efforts to discover and quantify the presence of NARTs may benefit from a simultaneous characterization of stem-like phenotypes to increase our understanding of why certain NARTs are superior in their capacity for expansion and persistence. Together with our current report, this identifies an unmet need to improve the manufacturing of TIL Inf products to increase the frequency of tumor-specific CD8$^+$ T cells that are able to engraft and persist in patients after ACT.

Interestingly, we observed that 2 out of 3 patients with PD and multiple patients with SD appeared to have NARTs in peripheral blood despite the lack of persisting NART populations in the TIL Inf product. This suggests that selected nonresponders had ongoing tumor recognition that was not expanded by the TIL-manufacturing process (i.e., failure to expand meaningful NARTs), perhaps due to poor tumor immune infiltration (i.e., immunologically "cold"). Thus, development of technologies to expand tumor-specific CD8$^+$ T cells from peripheral blood may be beneficial for the future treatment of patients that do not benefit from conventional TIL-ACT. Given information on the antigen recognized in peripheral blood, other strategies, such as therapeutic vaccination (41, 42), could furthermore be combined to increase the likelihood of generating long-lasting CD8$^+$ and CD4$^+$ memory T cells from TIL-ACT.

We additionally observed novel NARTs at multiple time points after infusion in both responders and nonresponders. This might illustrate epitope spreading as a result of tumor-cell killing in responders. However, these late-emerging NART populations were present at a lower frequency and appeared to be more transient than those transferred in the TIL Inf product. Thus, epitope spreading, with T cell recognition of preexisting mutations and their derived peptide products, does not appear to play a major role following TIL-ACT. However, this does not exclude a potential therapeutic role for epitope spreading based on T cell recognition toward novel mutations occurring after immunotherapy.

Finally, we observed that lymphocyte activity and proliferation within the tumor microenvironment were associated with higher NART frequency in TIL Inf products, suggesting that ongoing immune activity within the tumor supports the manufacturing of TIL Inf products containing a high frequency of NARTs. Superior T cell proliferation and response to checkpoint inhibition is associated with intratumoral tertiary lymphoid structures, which

maintain a niche of professional antigen-presenting cells and proliferating T cells (37, 43). Tertiary lymphoid structures could, therefore, possibly support the successful expansion of TILs prior to successful TIL-ACT. However, the relationship among ongoing T cell proliferation, successful TIL expansion, and therapeutic response remains undetermined.

Both TIL expansion and posttransfer persistence of $CD8^+$ NARTs may additionally be affected by supporting $CD4^+$ T cells (44). So far, no differences have been observed between CD8-enriched TIL products and TIL products containing different lymphocytes (although the majority are CD8; ref. 45). Furthermore, epitope spreading as evaluated here for $CD8^+$ T cells may likewise occur for $CD4^+$ T cells, and further insight into the relationship between $CD4^+$ and $CD8^+$ tumor-reactive T cells and the relevance for shared antigen recognition are critical aspects for addressing future improvements in immunotherapy. However, technical limitations still prohibit detailed epitope mapping of $CD4^+$ NARTs, as conducted here for $CD8^+$ NARTs (46).

In this study, we screened for recognition among 5921 predicted neopeptides arising from nonsynonymous mutations, of which we found recognition of 1.8% (106 neoepitopes) in TIL Inf products and additionally 98 neoepitopes in peripheral blood before or after TIL Inf, making a T cell recognition percentage of 3.4%. This illustrates that neoepitope prediction is feasible, but it remains a cumbersome approach to identifying neoepitope-specific $CD8^+$ T cells in metastatic melanoma. While recent efforts have led to significant improvements in the prediction of antigen processing and HLA binding (47), a gap remains in our ability to predict which of the presented neoepitopes are able to give rise to T cell recognition (21). Among the neoepitopes recognized by T cells in this study, we observed an enrichment of neoepitopes derived from cancer-driver genes and genes expressed above 2 TPM. However, despite these characteristics, the majority of the neoepitopes detected were derived from passenger mutations, and no stringent criteria could be assigned to determine the neoepitopes driving T cell recognition.

In conclusion, our study describes the critical contribution of NARTs to the clinical outcome in TIL-ACT therapy and provides a thorough characterization of neoantigens recognized by T cells in this therapeutic context. To this end, our study highlights a critical need for improving TIL-ACT manufacturing and the capacity to predict immunogenic neoepitopes. Strategies to improve the expansion and engraftment of NARTs in TIL Inf products should further improve clinical outcome.

Methods

Patient material. To study the role of NARTs in TIL-ACT in melanoma, we evaluated 26 patients with unresectable or metastatic melanoma enrolled in a phase I/II clinical study of ACT (ClinicalTrials.gov NCT00937625). Demographic and clinical information for each patient ID are available in previous reports (4, 5, 26). TIL Inf products were generated by expanding TILs in vitro from tumor lesions following a rapid expansion protocol (REP) with high-dose IL-2, as described previously (48). All patients were included at the time of progression from previous treatment or treatments with either IL-2/IFN-α and/or anti–CTLA-4 treatment and/or DC vaccination and/or temozolomide and/or vemurafenib (26). Furthermore, as specified previously, 4 patients received vemurafenib between surgical resection and TIL-ACT (M27, M29, M35, M36; ref. 4). Clinical response was assessed according to RECIST 1.0. Among the 26 patients, 5 were CRs, 6 were PRs, 10 were SD, and 5 were PD patients (4), with a median PFS and OS of 3.85 and 23.25 months, respectively. Using DNA barcode–labeled pMHC multimers, we analyzed the TIL Inf products from all 26 patients for neoepitope-specific $CD8^+$ T cells. From 19 of these patients, we additionally analyzed corresponding PBMC samples before and at multiple time points after TIL-ACT (Supplemental Table 1). Tumor sequencing data (RNA and WES) were available from 26 of the 27 patients enrolled in the trial. PBMCs from healthy donors were obtained from whole blood by density centrifugation on Lymphoprep in Leucosep tubes and cryopreserved at –150°C in FCS (Gibco, Thermo Fisher Scientific) plus 10% DMSO.

TIL sorting and expansion. Young TILs were thawed and cultured overnight at 37°C in complete medium (CM) (RPMI-1640 supplemented with 10% heat-inactivated human serum), 100 U/mL penicillin, 100 μg/mL streptomycin, 1.25 μg/mL fungizone, and 6000 IU/mL IL-2. Cells were washed twice in R0 (RPMI 1640, 100 U/mL penicillin, 100 μg/mL streptomycin) and stained with 0.2 μg of pMHC tetramers for 10 minutes at 37°C. Tetramers were assembled from fluorescent-streptavidin conjugates (PE, catalog 405204, BioLegend; APC, catalog 405243, BioLegend; BV421, catalog 563259, BD) and biotinylated, recombinant UV-cleavable pMHC-1 (23, 24). An empty disulfide-stabilized monomer was used for A*02:01-Y84C (49). Anti–CD4-FITC (clone SK3, catalog 345768, BD) and anti–CD8-PerCP (clone SK1, catalog 345774, BD) antibodies were added for a further 20 minutes at 37°C. Cells were washed with R0, resuspended in R0 plus 10% heat-inactivated human serum, and sorted by flow cytometry using the BD FACSAria cell sorter (BD Biosciences) into a 96-well plate. Sorted $CD8^+$ $tetramer^+$ cells were expanded in 2 consecutive minirapid expansions 9 days apart, on day 0 and day 9. The second minirapid expansion was omitted in cases with abundant proliferation. In brief, 5×10^5 allogeneic feeder cells from healthy donors, 30 ng/mL anti-CD3 antibody (clone OKT3, Janssen-Cilag), master mix made of 50% CM and 50% rapid expansion medium (RM) consisting of AIM-V medium (Gibco, Thermo Fisher Scientific), and 1.25 μg/mL fungizone supplemented with 6000 IU/mL IL-2 with 10% heat-inactivated human serum (HS) were added to sorted cells and cultured at 37°C; 50% of the media (without OKT-3) was replaced after 5 days and subsequently every 2 days.

Intracellular cytokine assay. Tumor cells were either pretreated with IFN-γ (100 IU/mL, Peprotech) or left untreated for 3 days. TILs were then added in a 1:1 ratio, with protein transport inhibitors brefeldin A (1:1000 dilution, GolgiPlug, catalog 555029, BD), Monensin (1:1000 dilution, GolgiStop, catalog 554724, BD), and anti–CD107a-BV421 antibodies (clone H4A3, BD 562623). Tumor cells and TILs were cocultured for 5 hours, after which all cells were stained with Near-IR LIVE/DEAD (Life Technologies) and for surface markers CD3-FITC (clone SK7, BD 345764), CD8-QDot605 (clone 3B5, Thermo Fisher Q10009), and CD4-BV711 (clone SK3, BD, catalog 563028). Subsequently, the cells were fixed and permeabilized (eBioscience) overnight and stained for intracellular cytokines TNF-APC (clone MAb11, BD catalog, 554514) and IFN-γ–PE-Cy7 (clone B27, BD, catalog 557643). Cells were analyzed on a Novocyte Quanteon (ACEA Biosciences). See details related to antibodies used in Supplemental Table 2.

Neoepitope prediction. WES and RNA-Seq data were obtained from digested tumor fragments, except for M22 and M24, for which autologous tumor cell lines were used. Two WES files from M15 were utilized and their results combined, one from an autologous tumor digest and another from an autologous tumor cell line. All WES data were obtained from tumor material from the same biopsy as was used for manufacturing of the corresponding TIL Inf products, expect for M22, for which the tumor cell line was derived from an earlier time point. FASTQ files from WES and RNA-Seq were preprocessed using Trim Galore (50), version 0.4.0. WES reads were aligned to the human genome (GRCh38) using Burrows-Wheeler Aligner (51), version 0.7.15, with default mem parameters, and duplicate reads were marked using MarkDuplicates from Picard Tools (52), version 2.9.1. Peptides were extracted and prioritized using MuPeXI (20), version 1.1.3, and netMHCpan, version 4.0, (22), providing as input the somatic variants obtained following GATK, version 3.8.0, best practices, the RNA-Seq expression values calculated using Kallisto, version 0.42.1 (53), and the HLA alleles inferred from normal WES samples using OptiType, version 1.2 (54). For patients with high neoantigen load, all predicted neoepitopes with a percentage rank score of 0.5 or less and TPM of 0.1 or more were included. For patients with lower neoantigen load, we lowered the expression threshold to 0.01 TPM or more and selected the top 200 predicted neopeptides according to percentage rank score. All predicted neopeptides and virus control peptides were synthesized and purchased from Pepscan (Pepscan Presto) and dissolved to 10 mM in DMSO.

For each cancer-specific nonsynonymous mutation, the HLA-I–binding potential of mutation-derived peptides was predicted using netMHCpan, version 4.0 (20, 22). For each patient, a minimum of 200 top-ranking neopeptides were included. The ranking was based on the predicted HLA-I binding (percentage rank score) and the transcription of the corresponding gene as RNA TPM.

Clonality. Copy number, purity, and ploidity were found using Sequenza, version 3.0 (55). As input, printed reads from normal and tumor were used as input to Sequenza. Sequenza-utils, version 3.0, bam2seqz with GRCh38 was used as a reference. To run the Sequenza copy number call with GRCh38, the R packages Shixiang/copynumber, version 1.26.0 (56), was applied. The created seqz files were used as input to sequenza-utils seqz_binding, and the outputs were used to Sequenza utils snp2seqz. To reduce the amount of false negatives according to the built-in mutations called from Sequenza, copy number files from the mutect2 output were merged with the copy number call from the bam files. Sequenza results and PyClone inputs were generated with the Sequenza packages in R, version 3.6.1. To find clonal mutations, PyClone, version 0.13.0 (57), was applied with the cellularity given from Sequenza and max cluster of 30 and minimum size of 0 to get all possible mutations given. Clonal mutations were filtered with a cluster size of minimum 80 and cellularity of minimum 90. Clonality could not be computed for M22, M24, and part of M15, as the underlying WES data came from autologous tumor cell lines.

Generation of DNA barcode–labeled pMHC multimers. Oligonucleotides containing distinct 25 mer nucleotide sequences (58) were purchased from LGC Biosearch Technologies. All oligos carry a 6 nt unique molecular identifier (59). Oligonucleotides modified with a 5′ biotin tag (oligo A) were joined to unmodified, partially complementary oligonucleotides (oligo B) to generate more than 1000 unique double-stranded AxBy DNA barcodes. Combinations of A and B oligos (one of each) were mixed with 5× Sequenase Reaction Buffer Mix (PN 70702, Affymetrix) to final concentrations of 26 μM (oligo A) and 52 μM (oligo B), respectively, heated to 65°C for 2 minutes, and allowed to anneal by cooling slowly to less than 35°C over 15 to 30 minutes. The annealed oligo As and Bs were elongated to create double-stranded AxBy DNA barcodes by adding Sequenase polymerase (70775Y, Affymetrix), 20 μM DTT, and 800 μM or 72 μM dNTPs, followed by incubation for 5 to 10 minutes at room temperature. Elongated AxBy barcodes were diluted in nuclease-free water plus 0.1% Tween to 2.17 μM (with respect to the A oligo) and stored at −20°C. Attachment of 5′ biotinylated AxBy DNA barcodes to PE- and streptavidin-conjugated dextran (Fina Biosolutions) was performed by mixing the 2 components at final concentrations of 154 nM dextran backbone and 77 nM barcode in order to obtain 0.5 barcodes for each dextran backbone and subsequent incubation for 30 minutes at 4°C.

Refolded, biotinylated pMHC-I was subsequently added at a stoichiometry of approximately 16.5 pMHC molecules per dextran; these were generated through UV-mediated exchange of cleavable ligands as described previously (23, 24). In brief, MHC monomers bound to UV-sensitive ligands were mixed with HLA-matching peptides at a final concentration of 50 μg/mL monomer and 100 mM peptide and exposed to UV light for 60 minutes (366 nm). Afterwards, pMHC monomers were centrifuged for 5 minutes at 3300*g* and then coupled to DNA barcode– and PE-labeled dextran backbones to a final concentration of 35 μg/mL monomer and 4.2×10^{-8} M barcode- and PE-labeled dextran backbone and incubated for 20 minutes on ice. Then a freezing buffer was added to reach PBS plus 0.5% BSA plus 100 μg/mL herring DNA plus 2 mM EDTA plus 5% glycerol and 909 nM D-biotin, and after 20 minutes on ice, the pMHC multimers were stored at −20°C until use.

T cell staining with barcode-labeled pMHC multimers. Cryopreserved cells were thawed, washed twice in RPMI plus 10% FCS, and then washed in barcode-cytometry buffer (PBS plus 0.5% BSA plus 100 μg/mL herring DNA plus 2 mM EDTA). Before staining, MHC multimers were thawed on ice, centrifuged for 5 minutes at 3300*g*, and 1.5 μL (0.043 μg) of each distinct pMHC was taken from each well, avoiding potential aggregates in the bottom, and pooled. The volume of the reagent pool was reduced by ultrafiltration to obtain a final volume of approximately 80 μL of pooled MHC multimers per staining. Centrifugal concentrators (Vivaspin 6, 100,000 Da, Sartorius) were saturated with BSA before use. Following ultrafiltration, the pool of multimers was spun at 10,000*g* for 2 minutes to sediment potential aggregates. An aliquot of approximately 5 μL of the MHC multimer reagent pool was stored at −20°C for later baseline analysis. Up to 10 million cells were stained in 80 μL with 50 nM dasatinib and multimer pools in a 15-minute incubation at 37°C. Following incubation, the cells were stained with an antibody mix containing CD8-BV480 (clone RPA-T8, BD, catalog 566121), dump channel antibodies (CD4-FITC (clone SK3, BD, catalog 345768), CD14-FITC (clone MϕP9, BD, catalog 345784), CD19-FITC (clone 4G7, BD, catalog 345776), CD40-FITC (clone LOB7/6, Serotech, catalog MCA1590F), and CD16-FITC (clone NLP15, BD, catalog 335035), and a dead cell marker (LIVE/DEAD Fixable Near-IR; Invitrogen L10119) and incubated for 30 minutes at 4°C. Samples were stained with antibodies in a total volume of 100 μL. See staining concentrations in Supplemental Table 2. Cells were washed 3 times in barcode cytometry buffer and fixed in 1% paraformaldehyde (PFA) for 0.5 to 24 hours before they were washed twice

and resuspended in barcode-cytometry buffer. Cells were acquired within a week after multimer staining.

Sorting of pMHC multimer$^+$ T cells. Multimer-binding CD8$^+$ T cells were sorted on a FACSAria Fusion or FACSMelody Cell Sorter (BD) into BSA-saturated tubes containing 100 μl of barcode/cytometry buffer. We gated on single, live, CD8$^+$, and dump channel–negative (CD4, CD14, CD16, CD19, and CD40) lymphocytes and sorted all multimer-positive PE cells within this population. As tested and described in Bentzen et al. (19), inclusion of CD8$^+$ multimer negative cells in the sorting gate does not have an impact on the final results because the fluorescence signal is used only for sorting out the relevant cells. Determination of antigen specificity is done solely based on the DNA barcode. The sorted cells were centrifuged for 10 minutes at 5000*g*, and the buffer was removed. The cell pellet was stored at −80°C. The percentage of multimer$^+$ CD8$^+$ T cells was used as input for estimation of epitope-specific CD8$^+$ T cells (see *Processing of sequencing data from DNA barcodes*). Three samples were run without exported flow cytometry files, precluding adequate estimation of frequency after sequencing of DNA barcodes (M15, TIL Inf product; M40, pre-ACT PBMCs; and M40, PBMCs <1 month). TIL Inf product from M47 was stained again to estimated percentage of multimer$^+$ CD8$^+$ T cells. M15 had no significant hits among barcoded multimers (i.e., sum of estimated frequency was set to 0%). See antibody assay details in Supplemental Table 2.

DNA barcode amplification. DNA barcode amplification was performed using Taq PCR Master Mix Kit (QIAGEN, 201443) and 3 μM of forward and reverse primers (LGC Biosearch Technologies). PCR amplification was conducted on sorted multimer-binding T cells (in <19 μL of buffer) and on a triplicate of the stored aliquot of the MHC multimer reagent pool (diluted 10.000× in the final PCR) under the following conditions: 95°C for 10 minutes; 36 cycles: 95°C for 30 seconds, 60°C for 45 seconds, 72°C 30 for seconds, and 72°C for 4 minutes. The multimer reagent pool was used as the baseline to determine the number of DNA barcode reads within a nonprocessed MHC multimer reagent library. PCR products were purified with a QIAquick PCR Purification Kit (QIAGEN)m and the amplified DNA barcodes were sequenced at PrimBio using an Ion Torrent PGM 316 or 318 chip (Life Technologies).

Processing of sequencing data from DNA barcodes. Sequencing data were processed by the software package Barracoda, available online (https://services.healthtech.dtu.dk/service.php?Barracoda-1.8). This tool identifies the barcodes used in a given experiment, assigns PCR used sample IDs and pMHC specificity to each barcode, and counts the total number (clonally reduced) of reads for each DNA barcode. Furthermore, it accounts for barcode enrichment based on methods designed for the analysis of RNA-Seq data, implemented in the R package edgeR; specifically, $\log_2$ fold changes (FCs) in read counts mapped to a given sample relative to the mean read counts mapped to triplicate baseline samples are estimated using normalization factors determined by the trimmed mean of M values method. Enriched barcodes were regarded as significant when the adjusted *P* value was below 0.001, which equals an FDR< 0.1 (estimated using the Benjamini–Hochberg method). Barracoda outputs were further processed and annotated using R 4.0.2 — adding relevant clinical information and excluding signals arising from insufficient read depth (percentage of read count < 0.1). Furthermore, biologically relevant barcode enrichment was defined as an estimated frequency of 0.01% or more and without presence in partially HLA-matching healthy donor PBMCs. 227 Multimers were excluded due to technical concerns regarding HLA-C*05:01 (M22, 140 multimers; M27, 46 multimers) and HLA-C*02:02 (M43, 41 multimers). Peptide missannotations, which originated from pipetting errors discovered through cross referencing of ordering and annotation tables (M27, 1 multimer; M35, 40 multimers; M46, 1 multimer), were also excluded. Frequency of a pMHC-specific CD8$^+$ T cell population was estimated based on the percentage of read count of the associated barcode out of the total percentage of the multimer-positive CD8$^+$ T cell population. Sum of estimated frequency represents the pooled frequencies of all T cell populations in a given sample. Due to differences in number of producible HLA molecules, the number and frequency of neoepitope-specific CD8$^+$ T cells were normalized to the mean absolute HLA coverage in the cohort: (average HLA covered [across all panels]/HLA covered [patient panel]).

Structural analysis of overlapping mutated peptides and HLA binding. Structural pMHC models were generated using the method described in TCRpMHCmodels (60). All peptides were bound to HLA-A*01:01 and the sequence for this MHC molecule was downloaded from the IMGT database (61). To get the electrostatic potential for each of the pMHC models, hydrogen atoms were added using the phenix.reduce protocol previously described (62), after which Delphi (63) was used to calculate the electrostatic potential with the following parameters: scale = 1.0, perfil = 70.0, indi = 4.0, exdi = 80.0, prbrad = 1.4, salt = 0.15, ionrad = 2.0, bndcon = 2, linit = 800, maxc = 0.0001, sigma = 2.0, srfcut = 20.0 and gaussian = 1. The electrostatic potential from Delphi was finally virtualized using PyMOL (https://pymol.org/).

Differential expression analysis. RNA-Seq data for differential gene expression analysis exclusively came from tumor digests, i.e., no autologous tumor cell lines were used. Output files from Kallisto were used as input to DESeq2, version 1.26.0, from R/bioconductor with default options (64) to find differential expressed genes (adjusted $P < 0.05$, related to high and low sum of estimated frequency split by the median and PFS split by equal or below the median). GO enrichment analysis was performed using R, version 4.0.2, with the packages enrichplot, version 1.11.0.991 (65), and clusterProfiler, version 3.16.1, with Benjamini-Hochberg at *P* value adjustment (36).

Code availability. MuPeXi used for neoepitope prediction is available for all users at https://services.healthtech.dtu.dk/service.php?MuPeXI-1.1 and has been published (20). Visualization of pMHCs was generated as described in Methods. Analysis of DNA barcodes was performed as described in Methods, and the bioinformatics pipeline is available (https://services.healthtech.dtu.dk/service.php?Barracoda-1.8). Code used for further analysis and visualization was written in R as performed as described in methods.

Statistics. Statistical analysis of DNA barcoding data was performed using the software package Barracoda as described above. Survival analysis used percentiles and medians (number of NARTS or frequency) to define thresholds to split the cohort. Any values matching the threshold were treated as belonging to the lower group. Mantel-Cox test was used to evaluate the effect of NARTs on survival, and HRs were calculated using the log-rank approach with GraphPad Prism 8. Correlations were tested using nonparametric, 2-sided Spearman's correlation test, except for Supplemental Figure 1D, where we used a 2-sided Pearson's correlation. Two-sided *z* tests (prop.test) were applied where specified for Figure 6, A, C, and D. All 2-group comparisons were performed using nonparametric Mann-Whitney *U* test with a significance threshold of 0.05. Multigroup comparisons

were performed by an initial nonparametric Kruskal-Wallis test followed by post hoc Dunn's multiple comparison test.

Study approval. This study was conducted using TILs and PBMCs from patients enrolled in a clinical study (ClinicalTrials.gov NCT00937625). All patients signed a written consent form according to the Declaration of Helsinki. The study was approved by the local ethics committee for the capital region of Denmark (Region H). Likewise, healthy donor samples were collected by approval of the local Scientific Ethics Committee for the capital region of Denmark (Region H), with donor written, informed consent obtained according to the Declaration of Helsinki. Healthy donor blood samples were obtained from the blood bank at Rigshospitalet, Copenhagen, Denmark. All samples were obtained anonymously

Author contributions

NPK, CH, and SAT performed experiments, analyzed data, generated figures, and wrote the manuscript. AB conducted all bioinformatics analyses and generated figures. AD and MDC performed experiments and analyzed data. IC predicted neoepitopes. KKM conducted structural analyses of pMHC and generated figures. JSH assisted neopeptide selection and pMHC multimer production. AMB assisted neopeptide prediction. AKB provided technical guidance. AMM supported data analyses. ZS codesigned in silico platforms and supported funding. NM assisted in bioinformatics guidance. RA provided patient material, diagnosed and characterized patients, and generated tumor cell lines. MN designed in silico platforms and supervised neoepitope prediction. GBJ conducted sequencing analysis and discussed data. MD provided patient material, cosupervised the study, and discussed data. IMS provided patient material, cosupervised the study, discussed data, and revised the manuscript. SRH conceived the concept, supervised the study, discussed data, and wrote the manuscript. NPK, CH, and SAT are listed as co–first authors. NPK led the effort through the revision phase and is therefore listed first. CH and SAT are listed in alphabetical order.

Acknowledgments

We thank all donors and patients for participating in the study and B. Rotbøl, A.F. Løye, and A.D. Burkal for excellent technical assistance with handling flow cytometry instruments and blood samples. This research was funded primarily through the NEYE foundation grant "A personalized approach for increased clinical efficacy of cancer immunotherapy" and in part supported by the European Research Council, StG 677268 NextDART, Lundbeck Foundation Fellowship R190–2014–4178, and the Carlsberg Foundation.

Address correspondence to: Sine Reker Hadrup, Kemitorvet Bld. 204, 2800 Kongens, Lyngby, Denmark. Phone: 45.27.12.52.21; Email: sirha@dtu.dk.

CH's present address is: CBio A/S, Copenhagen, Denmark. JSH's present address is: PokeAcell, Copenhagen, Denmark. AMB's present address is: Novo Nordisk A/S, Måløv, Denmark. AMM's present address is: Clinical Microbiomics, Copenhagen, Denmark.

1. Rosenberg SA, et al. Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy. *Clin Cancer Res*. 2011;17(13):4550–4557.
2. Besser MJ, et al. Adoptive transfer of tumor-infiltrating lymphocytes in patients with metastatic melanoma: intent-to-treat analysis and efficacy after failure to prior immunotherapies. *Clin Cancer Res*. 2013;19(17):4792–4800.
3. Itzhaki O, et al. Establishment and large-scale expansion of minimally cultured "young" tumor infiltrating lymphocytes for adoptive transfer therapy. *J Immunother*. 2011;34(2):212–220.
4. Andersen R, et al. Long-lasting complete responses in patients with metastatic melanoma after adoptive cell therapy with tumor-infiltrating lymphocytes and an attenuated IL-2 regimen. *Clin Cancer Res*. 2016;22(15):3734–3745.
5. Lauss M, et al. Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. *Nat Commun*. 2017;8(1):1–10.
6. Goodman AM, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther*. 2017;16(11):2598–2608.
7. Samstein RM, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet*. 2019;51(2):202–206.
8. Donia M, et al. Methods to improve adoptive T-cell therapy for melanoma: IFN--γ enhances anticancer responses of cell products for infusion. *J Invest Dermatol*. 2013;133(2):545–552.
9. Rosenberg SA, et al. Treatment of patients with metastatic melanoma with autologous tumor-infiltrating lymphocytes and interleukin 2. *J Natl Cancer Inst*. 1994;86(15):1159–1166.
10. Zhou J, et al. Persistence of multiple tumor-specific T-cell clones is associated with complete tumor regression in a melanoma patient receiving adoptive cell transfer therapy. *J Immunother*. 2005;28(1):53–62.
11. Lu YC, et al. Mutated PPP1R3B is recognized by T cells used to treat a melanoma patient who experienced a durable complete tumor regression. *J Immunol*. 2013;190(12):6034–6042.
12. Lu YC, et al. Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions. *Clin Cancer Res*. 2014;20(13):3401–3410.
13. Huang J, et al. T cells associated with tumor regression recognize frameshifted products of the CDKN2A tumor suppressor gene locus and a mutated HLA class I gene product. *J Immunol*. 2004;172(10):6057–6064.
14. Robbins PF, et al. Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nat Med*. 2013;19(6):747–752.
15. Prickett TD, et al. Durable complete response from metastatic melanoma after transfer of autologous T cells recognizing 10 mutated tumor antigens. *Cancer Immunol Res*. 2016;4(8):669–678.
16. Newell EW, et al. Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. *Nat Biotechnol*. 2013;31(7):623–629.
17. Joglekar AV, et al. T cell antigen discovery via signaling and antigen-presenting bifunctional receptors. *Nat Methods*. 2019;16(2):191–198.
18. Kula T, et al. T-scan: a genome-wide method for the systematic discovery of T cell epitopes. *Cell*. 2019;178(4):1016–1028.
19. Bentzen AK, et al. Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat Biotechnol*. 2016;34(10):1037–1045.
20. Bjerregaard AM, et al. MuPeXI: prediction of neo-epitopes from tumor sequencing data. *Cancer Immunol Immunother*. 2017;66(9):1123–1130.
21. Wells DK, et al. Key parameters of tumor epitope immunogenicity revealed through a consortium approach improve neoantigen prediction. *Cell*. 2020;183(3):818–834.
22. Jurtz V, et al. NetMHCpan-4.0: improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol*. 2017;199(9):3360–3368.
23. Rodenko B, et al. Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nat Protoc*. 2006;1(3):1120–1132.
24. Toebes M, et al. Design and use of conditional MHC class I ligands. *Nat Med*. 2006;12(2):246–251.
25. Kvistborg P, et al. TIL therapy broadens the tumor-reactive CD8(+) T cell compartment in melanoma patients. *Oncoimmunology*. 2012;1(4):409–418.
26. Ellebaek E, et al. Adoptive cell therapy with autologous tumor infiltrating lymphocytes and low-dose Interleukin-2 in metastatic melanoma patients. *J Transl Med*. 2012;10(1):1–12.
27. Effern M, et al. Adoptive T cell therapy targeting different gene products reveals diverse and con-

text-dependent immune evasion in melanoma. *Immunity*. 2020;53(3):564–580.
28. Villaruz LC, Socinski MA. The clinical viewpoint: definitions, limitations of RECIST, practical considerations of measurement. *Clin Cancer Res*. 2013;19(10):2629–2636.
29. Robbins PF, et al. Cutting edge: persistence of transferred lymphocyte clonotypes correlates with cancer regression in patients receiving cell transfer therapy. *J Immunol*. 2004;173(12):7125–7130.
30. Frahm N, et al. Extensive HLA class I allele promiscuity among viral CTL epitopes. *Eur J Immunol*. 2007;37(9):2419–2433.
31. Gros A, et al. Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nat Med*. 2016;22(4):433–438.
32. Mcgranahan N, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*. 2016;351(6280):1463–1469.
33. Hodis E, et al. A landscape of driver mutations in melanoma. *Cell*. 2012;150(2):251–263.
34. Parkhurst MR, et al. Unique neoantigens arise from somatic mutations in patients with gastrointestinal cancers. *Cancer Discov*. 2019;9(8):1022–1035.
35. Bjerregaard AM, et al. An analysis of natural T cell responses to predicted tumor neoepitopes. *Front Immunol*. 2017;8:1566.
36. Yu G, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287.
37. Cabrita R, et al. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature*. 2020;577(7791):561–565.
38. Liu S, et al. Efficient identification of neoantigen-specific T-cell responses in advanced human ovarian cancer. *J Immunother Cancer*. 2019;7(1):1–17.
39. Crowther MD, et al. Genome-wide CRISPR–Cas9 screening reveals ubiquitous T cell cancer targeting via the monomorphic MHC class I-related protein MR1. *Nat Immunol*. 2020;21(2):178–185.
40. Krishna S, et al. Stem-like CD8 T cells mediate response of adoptive cell immunotherapy against human cancer. *Science*. 2020;370(6522):1328–1334.
41. Carreno BM, et al. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science*. 2015;348(6236):803–808.
42. Lövgren T, et al. Complete and long-lasting clinical responses in immune checkpoint inhibitor-resistant, metastasized melanoma treated with adoptive T cell transfer combined with DC vaccination. *Oncoimmunology*. 2020;9(1):1792058.
43. Jansen CS, et al. An intra-tumoral niche maintains and differentiates stem-like CD8 T cells. *Nature*. 2019;576(7787):465–470.
44. Borst J, et al. $CD4^+$ T cell help in cancer immunology and immunotherapy. *Nat Rev Immunol*. 2018;18(10):635–647.
45. Dudley ME, et al. CD8+ enriched "young" tumor infiltrating lymphocytes can mediate regression of metastatic melanoma. *Clin Cancer Res*. 2010;16(24):6122–6131.
46. Hadrup SR, Newell EW. Determining T-cell specificity to understand and treat disease. *Nat Biomed Eng*. 2017;1(10):784–795.
47. Sarkizova S, et al. A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat Biotechnol*. 2020;38(2):199–209.
48. Donia M, et al. Simplified protocol for clinical-grade tumor-infiltrating lymphocyte manufacturing with use of the Wave bioreactor. *Cytotherapy*. 2014;16(8):1117–1120.
49. Saini SK, et al. Empty peptide-receptive MHC class I molecules for efficient detection of antigen-specific T cells. *Sci Immunol*. 2019;4(37):eaau9039.
50. *TrimGalore*. Version.06.7. Felix Krueger; 2021. https://github.com/FelixKrueger/TrimGalore.
51. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–1760.
52. *Picard*. Version 2.26.6. BroadInstitue; 2021. https://github.com/broadinstitute/picard.
53. Bray NL, et al. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34(5):525–527.
54. Szolek A, et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*. 2014;30(23):3310–3316.
55. Favero F, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol*. 2015;26(1):64–70.
56. Wang S. *Copynumber*. Wang Shixiang; 2020. https://github.com/ShixiangWang/copynumber.
57. Roth A, et al. PyClone: statistical inference of clonal population structure in cancer. *Nat Methods*. 2014;11(4):396–398.
58. Xu Q, et al. Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc Natl Acad Sci U S A*. 2009;106(7):2289–2294.
59. Kivioja T, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods*. 2012;9(1):72–74.
60. Jensen KK, et al. TCRpMHCmodels: Structural modelling of TCR-pMHC class I complexes. *Sci Rep*. 2019;9(1):1–12.
61. Lefranc MP, et al. IMGT®, the international ImMunoGeneTics information system R 25 years on. *Nucleic Acids Res*. 2015;43(d1):D413–D422.
62. Liebschner D, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. 2019;75(pt 10):861–877.
63. Li L, et al. DelPhi: a comprehensive suite for DelPhi software and associated resources. *BMC Biophys*. 2012;5(1):9.
64. Love MI, et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):1–21.
65. Yu G, Hu E. enrichplot: Visualization of Functional Enrichment Result. https://yulab-smu.top/biomedical-knowledge-mining-book/. Accessed December 1, 2021.
66. Donia M, et al. Aberrant expression of MHC class II in melanoma attracts inflammatory tumor-specific CD4+ T- cells, which dampen CD8+ T-cell antitumor reactivity. *Cancer Res*. 2015;75(18):3747–3759.

Supplementary Materials

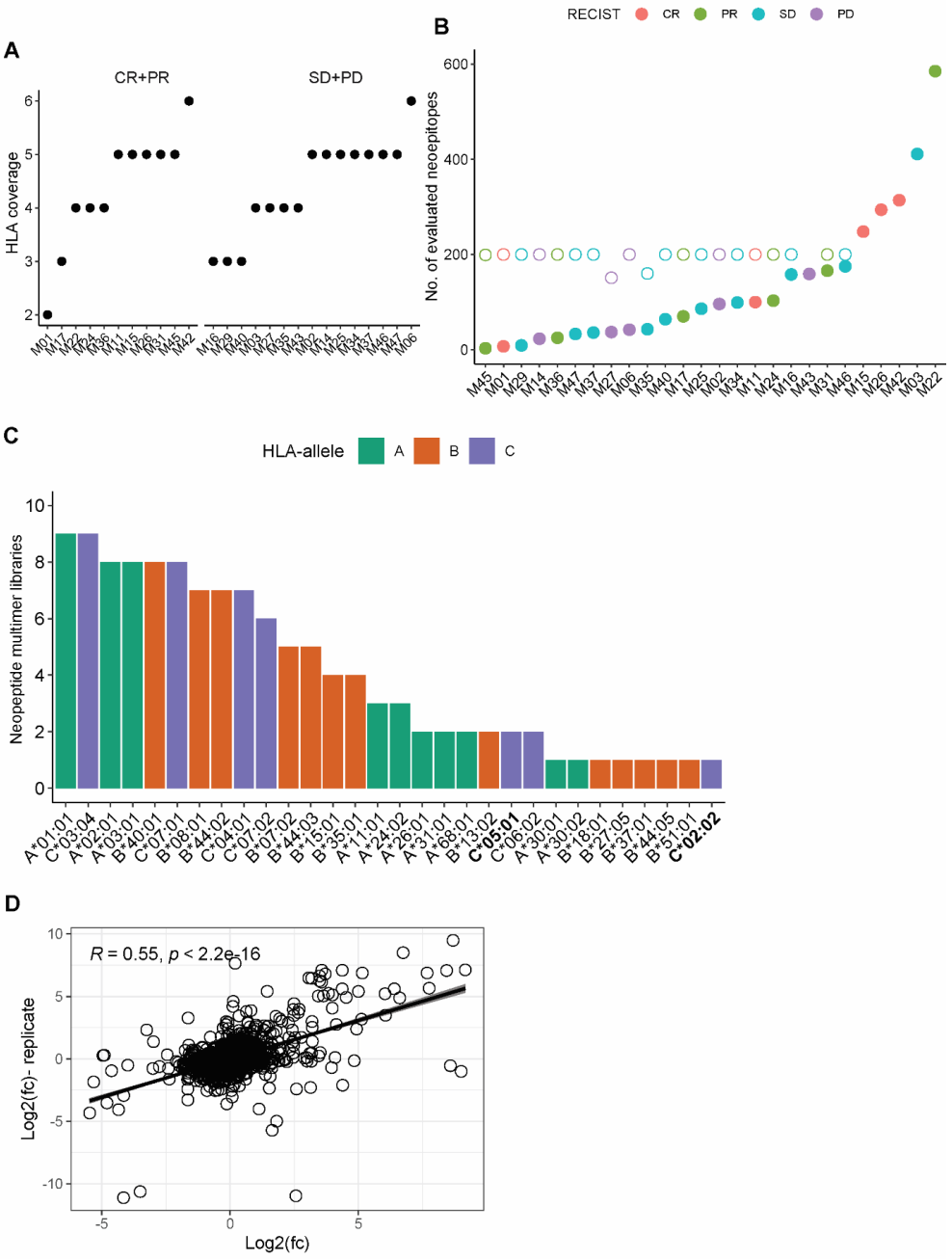
- **Supplemental Table 1:** Availability of peripheral blood samples pre- and post-ACT in melanoma patient cohort.
- **Supplemental Table 2:** Data reporting for fluorochrome-conjugated antibodies used in flow cytometry.
- **Supplemental Figure 1:** Supporting information regarding neopeptide multimer libraries.
- **Supplemental Figure 2:** Detection of NARTs in TIL Inf products from two melanoma patients.
- **Supplemental Figure 3:** Detection of NARTs specific for AKAP^{P1796L} peptide variants ('SILSY' variants).
- **Supplemental Figure 4:** Tumor recognition by expanded, AKAP^{P1796L}-specifc NARTs.
- **Supplemental Figure 5:** Tumor-mutational burden, estimated frequency and diversity of neoepitope-specific CD8 T cells in TIL-ACT.
- **Supplemental Figure 6:** Impact of NART diversity and frequency, as well as presence of immunogenic mutations on overall and progression-free survival.
- **Supplemental Figure 7:** NARTs in TIL samples and peripheral blood over time.
- **Supplemental Figure 8:** Distribution and temporal appearance of NARTs in TIL-ACT treated patients.
- **Supplemental Figure 9:** NART diversity and frequency within TIL Inf products does not correlate with tumor mutational burden (TMB) or number of predicted neoepitopes.
- **Supplemental Figure 10:** Exploratory analysis of differentially expressed genes.

Patient ID	PBMC pre-ACT 8 days prior	TIL Inf Prod	PBMC post-ACT < 1 month	PBMC post-ACT < 4 months	PBMC post-ACT < 12 months	PBMC post-ACT < 24 months	PBMC post-ACT < 48 months
M01		√			√		√
M14	√	√	√	√			
M17	√	√	√		√	√	√
M22	√	√	√	√	√	√	√
M24	√	√	√	√	√	√	√
M25	√	√	√	√			
M26	√	√	√	√	√	√	√
M27	√	√	√	√			
M29	√	√	√	√			
M31	√	√	√	√	√		
M34	√	√	√				
M35	√	√	√				
M36	√	√	√	√	√		
M40	√	√	√		√		
M42	√	√	√	√	√	√	√
M43	√	√	√	√	√		
M45	√	√	√	√	√		
M46	√	√	√	√	√		
M47	√	√	√	√			

Supplemental Table 1. Availability of peripheral blood samples pre- and post-ACT in melanoma patient cohort.

Marker	Assay	Fluorochrome	Clone	Supplier	Cat. Number	Staining conc.
CD8	Multimer screening	BV480	RPA-T8	BD	566121	2:100
CD4	Multimer screening	FITC	SK3	BD	345768	1.25:100
CD14	Multimer screening	FITC	$M\varphi$P9,	BD	345784	3.13:100
CD16	Multimer screening	FITC	NLP15	BD	335035	1.56:100
CD19	Multimer screening	FITC	4G7	BD	345776	6.25:100
CD40	Multimer screening	FITC	LOB7/6	Serotech	MCA1590F	2.5:100
CD4	TIL sorting	FITC	SK3	BD	345768	240ng/ul
CD8	TIL sorting	PerCP	SK1	BD	345774	500ng/ul
CD107a	ICS	BV421	H4A3	BD	562623	0.3:50
CD3	ICS	FITC	SK7	BD	345764	5ng/ul
CD8	ICS	QDOT605	3B5	Thermo Fischer	Q10009	0.2:50
CD4	ICS	BV711	SK3	BD	563028	1:50
TNFα	ICS	APC	Mab11	BD	554514	4ng/ul
IFNγ	ICS	PE-Cy7	B27	BD	557643	1.5:50

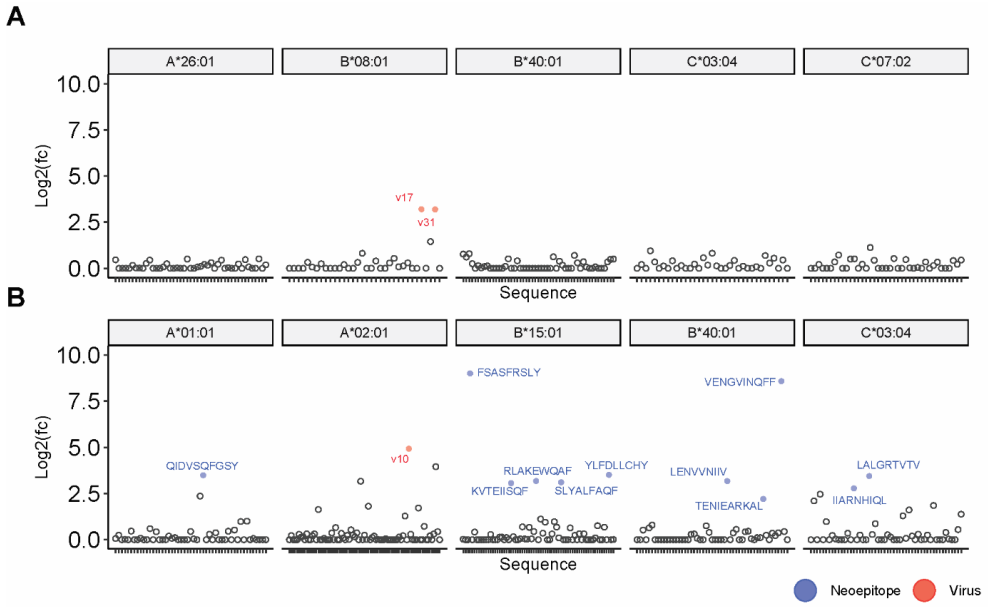
Supplemental Table 2: Data reporting for fluorochrome-conjugated antibodies used in flow cytometry. ICS, Intracellular Cytokine Staining.


A
CR+PR
SD+PD
HLA coverage
B
RECIST CR PR SD PD
No. of evaluated neoepitopes
C
HLA-allele A B C
Neopeptide multimer libraries
A*01:01
C*03:04
A*02:01
A*03:01
B*40:01
C*07:01
B*08:01
B*44:02
C*04:01
C*07:02
B*07:02
B*44:03
B*15:01
B*35:01
A*11:01
A*24:02
A*26:01
A*31:01
A*68:01
B*13:02
C*05:01
C*06:02
A*30:01
A*30:02
B*18:01
B*27:05
B*37:01
B*44:05
B*51:01
C*02:02
D
R = 0.55, p < 2.2e-16
Log2(fc)- replicate
Log2(fc)

41

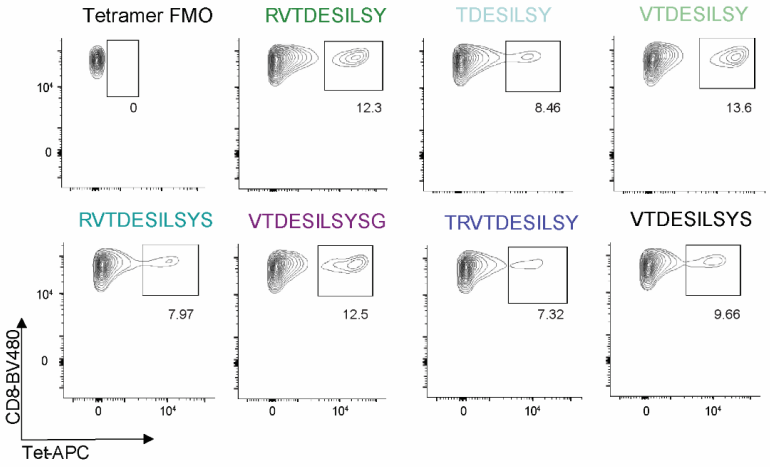
42

Supplemental Figure 1. Supporting information to neopeptide multimer libraries. (**A**) HLA coverage per patient. Number of HLAs evaluated for each patient-specific multimer library. (**B**) Neopeptide library size. Closed circles represent the number of predicted neoepitopes that can bind producible HLAs with %rank $\leq$ 0.5, and expression (TPM) $\geq$ 0.1. Open circles represent the number of evaluated multimers (see methods). Note, that hollow circles are overlaid by filled circles for M43, M15, M26, M42, M03, and M22. (**C**) HLA alleles and their prevalence in assembled multimer libraries. Bold: excluded HLA-alleles due to technical issues. **(D**) Correlation of replicate multimer screens in TIL Inf samples of nine patients. Shown is the log2(fc) change of barcode read counts compared to triplicate panel baseline. Normality was tested using Shapiro-Wilk's method followed by Pearson correlation. Grey shading represent the 95% confidence intervals.

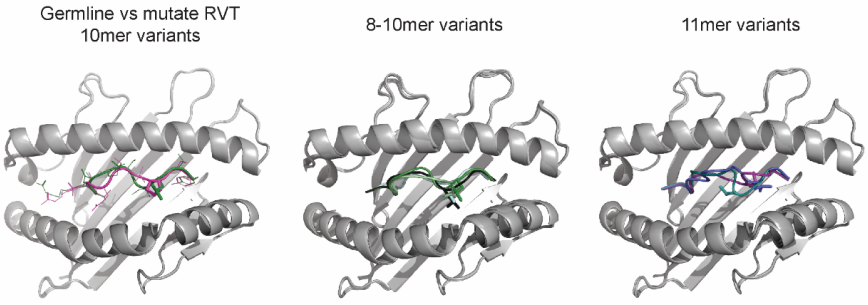


Supplemental Figure 2. Detection of NARTs in TIL Inf products from two melanoma patients. Example of a full screening using barcoded pMHC multimers for detection of both neoepitope-and virus-epitopes specific CD8+ T cells in TIL Inf products for melanoma patient M14 (PD) (**A**) and M26 (CR) (**B**). The data is separated according to the HLA alleles included in the screen. Blue: Virus-specific CD8 T cells. Red: NARTs. Black: Non-enriched barcodes. V10 annotate FLU peptide FLYALALLL, v17 annotate EBV peptide RAKFKQLL, and v31 annotate EBV virus peptide FLRGRAYGL.

A



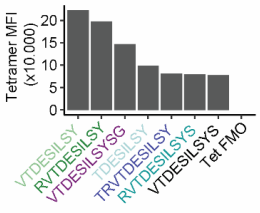
B



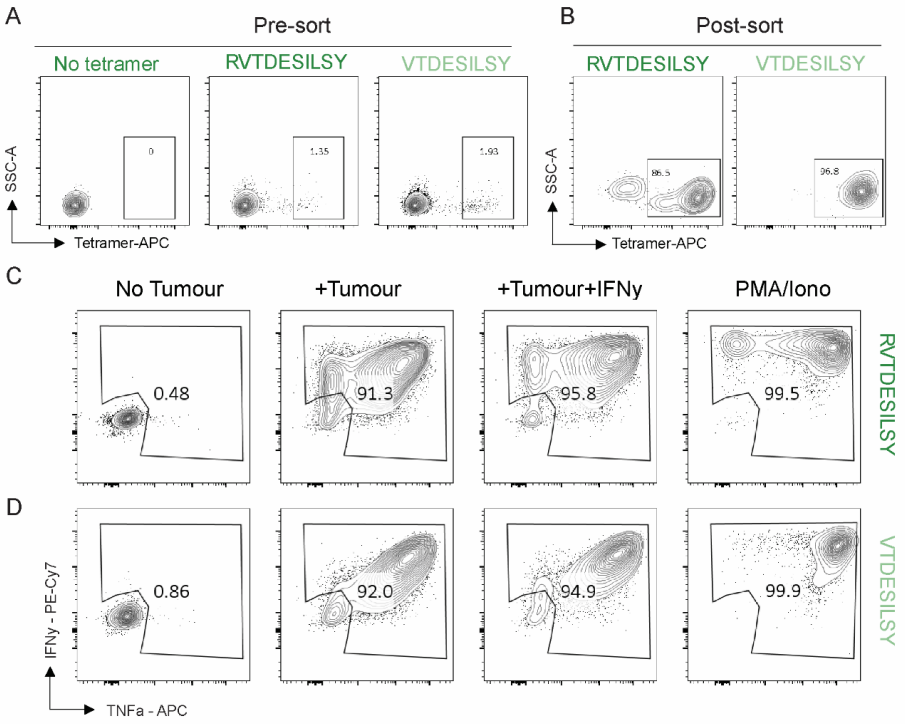
C

Mutated peptide	Germline peptide	Length	Est. freq (%)
RVTDESLSY	RVTDESPSY	10	9.88
TDESLSY	TDESPSY	8	2.06
VTDESLSY	VTDESPSY	9	1.92
RVTDESLSYS	RVTDESPSYS	11	1.10
VTDESLSYSG	VTDESPSYSG	11	0.91
VTDESLSYS	VTDESPSYS	10	0.47
TRVTDESLSY	TRVTDESPSY	11	0.36

D



Supplemental Figure 3. Detection of NARTs specific for AKAP9^{P1796L} peptide variants ('SILSY' variants). (**A**) Dump channel negative (CD4-CD14-CD16-CD19-CD40-), neoepitope-specific CD8+ T cells from M22 TIL Inf product. All 7 predicted AKAP9^{P1796L} neopeptides tested with APC tetramers; all 7 were restricted to HLA-A*01:01. (**B**) Prediction of MHC binding confirmation to HLA-A*01:01 for all AKAP9^{P1796L} peptide variants using TCRpMHCmodels as described in the methods. (**C**) Summary table with peptide lengths and estimated frequencies of the NART population to each of the peptide variants. (**D**) MFI for the corresponding AKAP9^{P1796L} tetramer+ CD8 T cell populations in (A). MFI, median fluorescence intensity. Tet, tetramer. FMO, fluorescence minus one.



93

94 **Supplemental Figure 4. Tumor recognition by expanded, AKAP9^{P1796L}-specifc NARTs.** (**A**)
95 HLA-A*01:01-restriced specific CD3+CD8+ T cells were sorted based on tetramer binding. (**B**)
96 REP expanded cells were tested for neoepitope-recognition using tetramers. (**C-D)** IFNγ and TNFα
97 release following co-cultures with autologous tumor cell lines and tetramer-specific clones recog-
98 nizing RVTDESILSY (C) and VTEDILSY (D). All plots represent gated CD3+CD8+ T cells.

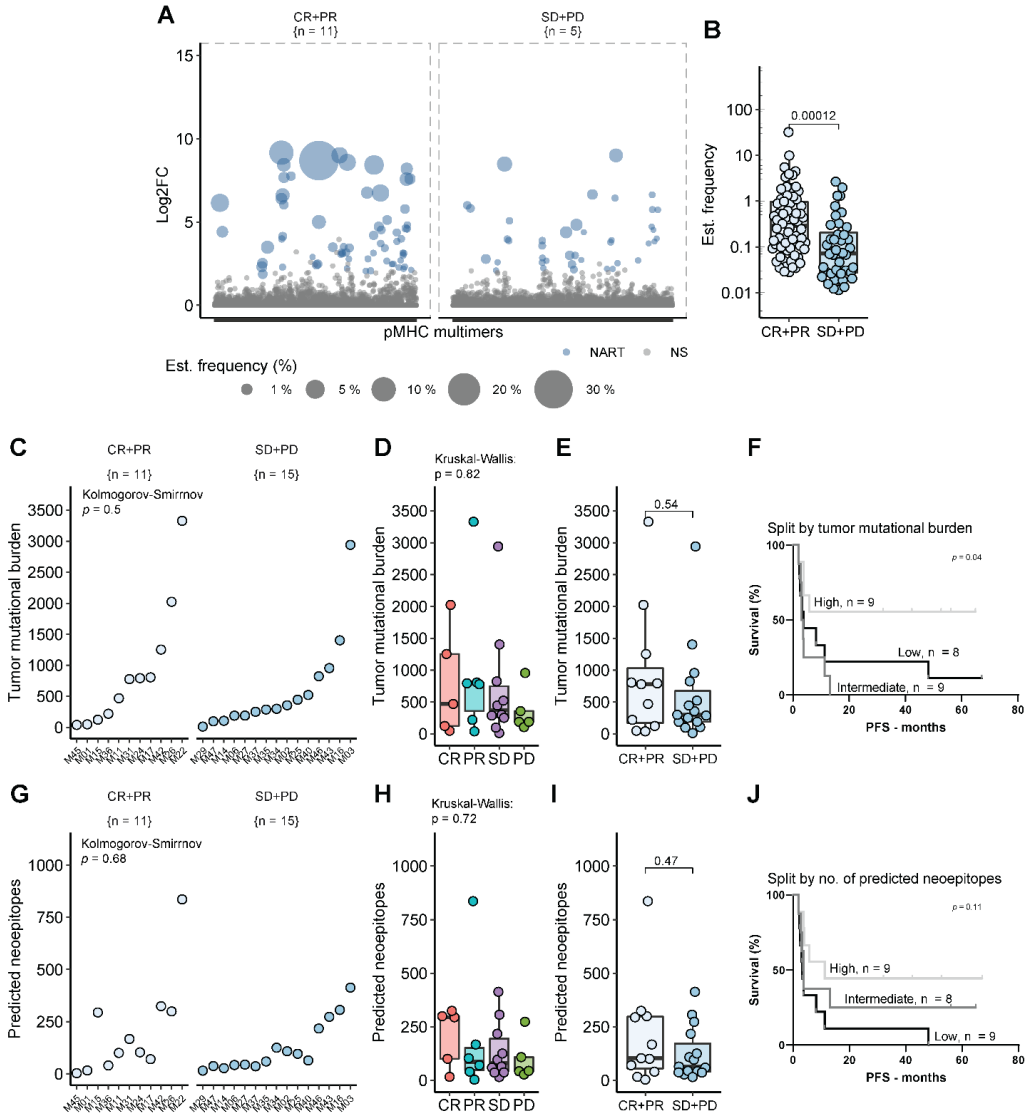
99

100

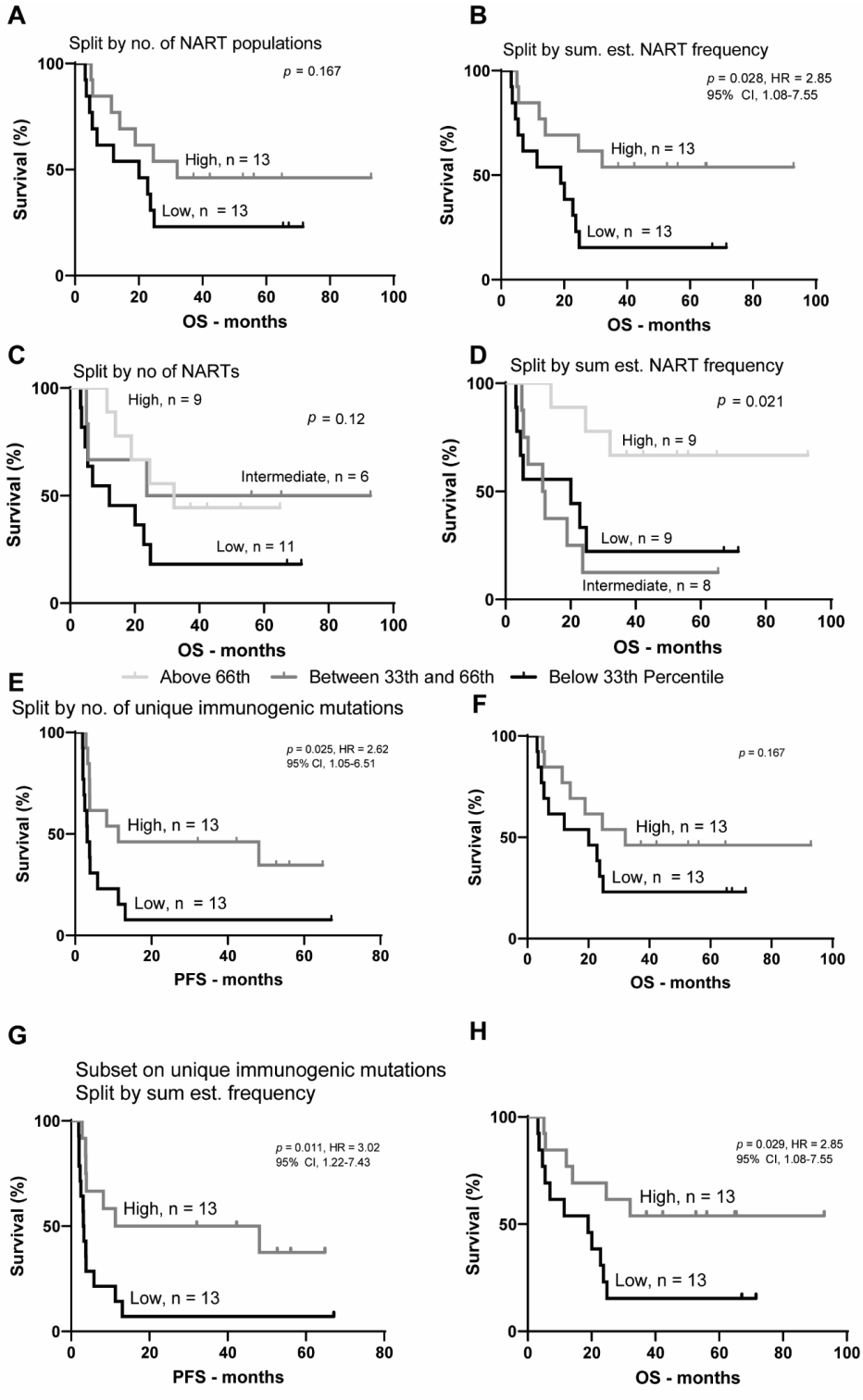
101

102

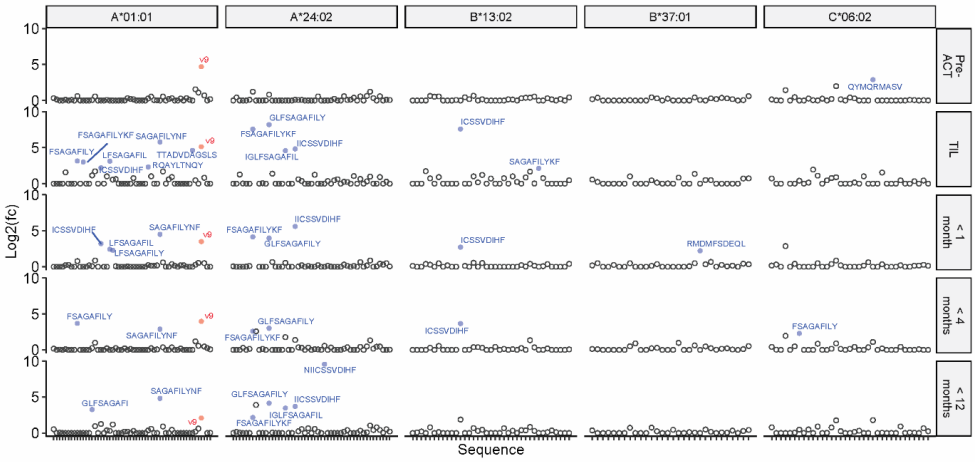
103


A
CR+PR
{n = 11}
SD+PD
{n = 5}
Log2FC
pMHC multimers
Est. frequency (%)
1 %
5 %
10 %
20 %
30 %
NART
NS
B
0.00012
Est. frequency
CR+PR SD+PD
C
CR+PR
{n = 11}
SD+PD
{n = 15}
Kolmogorov-Smirnov
p = 0.5
Tumor mutational burden
D
Kruskal-Wallis:
p = 0.82
CR PR SD PD
E
0.54
F
Split by tumor mutational burden
High, n = 9
Low, n = 8
Intermediate, n = 9
Survival (%)
PFS - months
G
Kolmogorov-Smirnov
p = 0.68
Predicted neoepitopes
H
Kruskal-Wallis:
p = 0.72
I
0.47
J
Split by no. of predicted neoepitopes
High, n = 9
Intermediate, n = 8
Low, n = 9

Supplemental Figure 5. Tumor-mutational burden, estimated frequency and diversity of neoepitope-specific CD8 T cells in TIL-ACT. (**A**) All evaluated multimers and TIL Inf products plotted according to clinical response and the log2(fc) enrichment of the given barcode. Estimated frequency is depicted as the dot size. (**B**) Estimated frequency of all 106 NARTs detected in TIL Inf products. (**C-E**) Tumor mutational burden in the cohort; (**C**) according to patient sorted by highest tumor mutational burden; (**D**) according to RECIST; (**E**) Responders (CR+PR) vs non-responders (SD+PD). **(F)** Progression-free survival for the cohort split by tumor mutational burden. The 66^{th} percentile = 787 mutations, 33^{rd} percentile = 228.67 mutations. (**G-I**) Number of predicted neoepitopes; (**G**) according to patient sorted by highest tumor mutational burden; (**H**) according to RECIST; (**I**) Responders (CR+PR) vs non-responders (SD+PD). (**J**) Progression-free survival for the cohort split by number of predicted neoepitopes. The 66^{th} percentile = 153 predicted neoepitopes, and the 33^{rd} percentile = 49 predicted neoepitopes. Whiskers represent IQR. p-values were calculated using nonparametric Mann-Whitney U test in B, E and I. Kolmogorov-Smirnov was used in C and G to test equality of distributions. Kruskal-Wallis test was used for D and H. Finally, Log-rank and Mantel-Cox was used to calculate p-values and hazard ratios (HR) respectively for F and J.


A
Split by no. of NART populations
p = 0.167
High, n = 13
Low, n = 13
Survival (%)
OS - months
B
Split by sum. est. NART frequency
p = 0.028, HR = 2.85
95% CI, 1.08-7.55
High, n = 13
Low, n = 13
Survival (%)
OS - months
C
Split by no of NARTs
High, n = 9
p = 0.12
Intermediate, n = 6
Low, n = 11
Survival (%)
OS - months
D
Split by sum est. NART frequency
p = 0.021
High, n = 9
Low, n = 9
Intermediate, n = 8
Survival (%)
OS - months
Above 66th
Between 33th and 66th
Below 33th Percentile
E
Split by no. of unique immunogenic mutations
p = 0.025, HR = 2.62
95% CI, 1.05-6.51
High, n = 13
Low, n = 13
Survival (%)
PFS - months
F
p = 0.167
High, n = 13
Low, n = 13
Survival (%)
OS - months
G
Subset on unique immunogenic mutations
Split by sum est. frequency
p = 0.011, HR = 3.02
95% CI, 1.22-7.43
High, n = 13
Low, n = 13
Survival (%)
PFS - months
H
p = 0.029, HR = 2.85
95% CI, 1.08-7.55
High, n = 13
Low, n = 13
Survival (%)
OS - months

Supplemental Figure 6. Impact of NART diversity and frequency, as well as presence of immunogenic mutations on overall and progression-free survival. (**A**-**B**), Overall survival split by median number of NARTs (3.65, A) or by median NART frequency (median = 0.64%, B) within TIL Inf product. (**C**) Overall survival split by 66^{th} and 33^{rd} percentile of NART diversity; 66^{th} percentile = 5.65 NARTs. 33^{rd} percentile = 0.88 NARTs. (**D**) Overall survival split by 66^{th} and 33^{rd} percentile of NART frequency; 66^{th} percentile = 3.26%. 33^{rd} percentile = 0.03%. (**E**-**F**), Overall **(E)** and Progression-free survival **(F)** split by median number of unique immunogenic mutations (3.22 uniquely recognized mutations). (**G**-**H**), Overall **(E)** and Progression-free survival **(F)** split by median NART frequency recognizing unique immunogenic mutations (0.63 %)**.** The highest estimated frequency among a group of NARTs recognizing the same mutation was taken as a proxy for the overall NART frequency of the group. p-values and hazard ratios (HR) from Mantel-Cox test and log-rank approach, respectively. Number of NARTs and NART frequency were normalized to HLA coverage as described in materials and methods. OS, Overall survival. PFS, Progression-free survival. n = 26 for all plots.

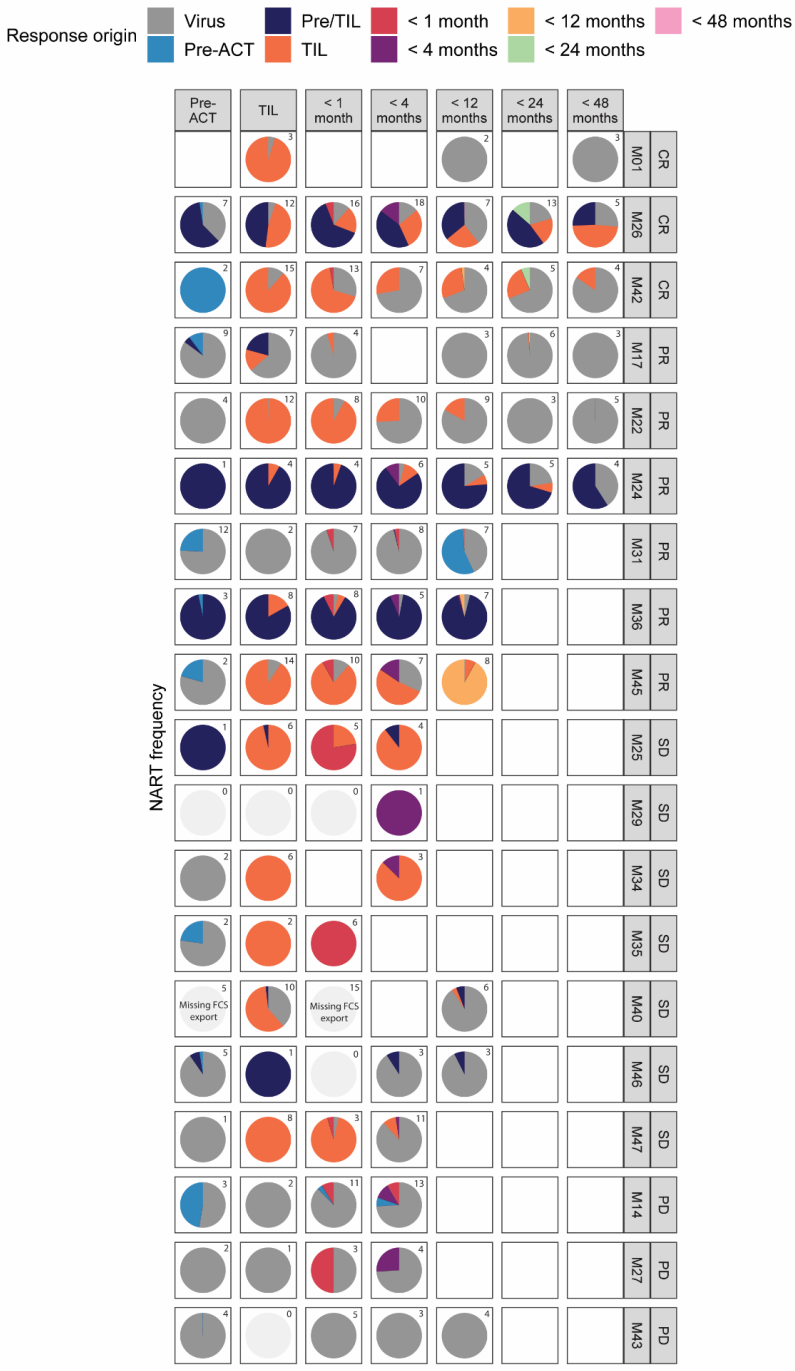


154

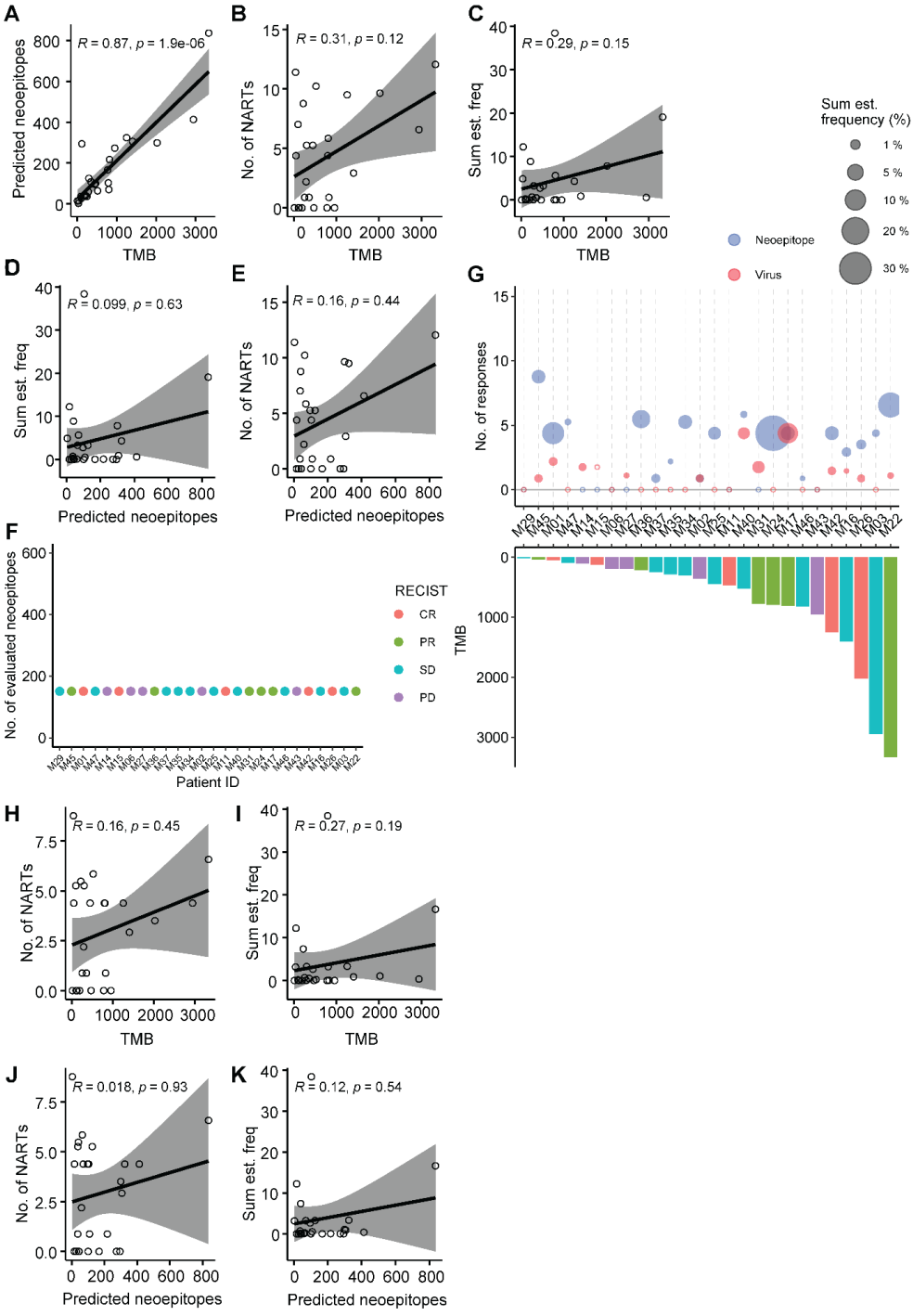
155 **Supplemental Figure 7. NARTs in TIL samples and peripheral blood over time.** Example of
156 a full screen for CD8+ T cell populations in patient M45 (PR), in PBMC before and after therapy,
157 and in the TIL Inf product. Separated according to HLA. Blue: NARTs. Red: responses to virus
158 peptides. Grey dots were considered non-enriched barcodes. V9 annotates CMV peptide
159 YSEHPTFTSQY.

160

161

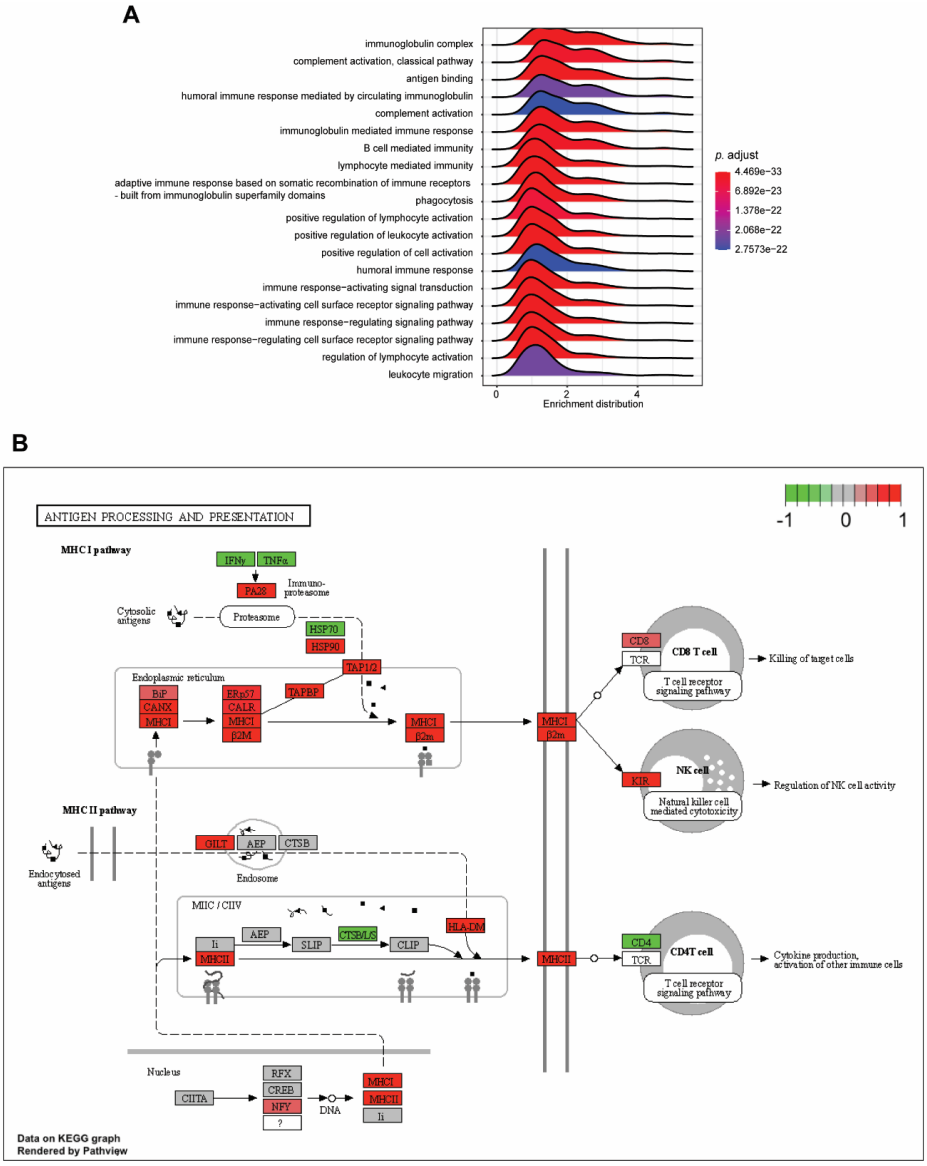

Response origin
Virus
Pre-ACT
Pre/TIL
TIL
< 1 month
< 4 months
< 12 months
< 24 months
< 48 months
Pre-ACT
TIL
< 1 month
< 4 months
< 12 months
< 24 months
< 48 months
NART frequency
M01 CR
M26 CR
M42 CR
M17 PR
M22 PR
M24 PR
M31 PR
M36 PR
M45 PR
M25 SD
M29 SD
M34 SD
M35 SD
M40 SD
M46 SD
M47 SD
M14 PD
M27 PD
M43 PD
Missing FCS export
Missing FCS export

Supplemental Figure 8. Distribution and temporal appearance of NARTs in TIL-ACT treated patients. Pie charts represent the frequency distribution of CD8+T cells specific towards neo- and viral epitopes followed over time from pre-ACT to < 48 months after therapy. Individual colors represents the group of NARTs appearing at a given time point. Virus responses are colored in grey. The total number of NART and virus responses within each circle is given in upper left corner for each time point. Missing FCS files from flow cytometry precludes frequency estimation in M40 Pre-ACT and <1 month PBMC samples.


A
R = 0.87, p = 1.9e-06
Predicted neoepitopes
TMB
B
R = 0.31, p = 0.12
No. of NARTs
TMB
C
R = 0.29, p = 0.15
Sum est. freq
TMB
Sum est. frequency (%)
1 %
5 %
10 %
20 %
30 %
Neoepitope
Virus
D
R = 0.099, p = 0.63
Sum est. freq
Predicted neoepitopes
E
R = 0.16, p = 0.44
No. of NARTs
Predicted neoepitopes
G
No. of responses
TMB
F
No. of evaluated neoepitopes
Patient ID
RECIST
CR
PR
SD
PD
H
R = 0.16, p = 0.45
No. of NARTs
TMB
I
R = 0.27, p = 0.19
Sum est. freq
TMB
J
R = 0.018, p = 0.93
No. of NARTs
Predicted neoepitopes
K
R = 0.12, p = 0.54
Sum est. freq
Predicted neoepitopes

187

Supplemental Figure 9. NART diversity and frequency within TIL Inf products does not correlate with tumor mutational burden or number of predicted neoepitopes. (**A**) TMB vs. number of predicted neoepitopes. **(B-E)** All evaluated multimers (151-585 multimers per patient). (**B**) TMB vs NART diversity. (**C**) TMB vs NART frequency. (**D**) Number of predicted neoepitopes vs NART diversity. (**E**) Number of predicted neoepitopes vs NART frequency. (**F**) Alternative selection strategy used for G through K selecting top 151 neoepitopes with the highest binding potential according to predicted %rank score. (**G**) NART diversity and frequency following alternative selection. Patients were arranged according to TMB. (**H**) TMB vs NART diversity. (**I**) TMB vs NART frequency. (**J**) Number of predicted neoepitopes vs NART diversity. (**K**) Number of predicted neoepitopes vs NART frequency. Diversity and frequency values were normalized to HLA coverage (see materials and methods). R and p-values from Spearman correlation with 95% confidence intervals in grey. All patients were evaluated (n = 26).



204

205 **Supplemental Figure 10. Exploratory analysis of differentially expressed genes.** (**A**) Top 20
206 enriched gene sets according to GO terms. (**B**) KEGG-pathway analysis showing the Antigen pro-
207 cessing and presentations pathway colered by enriched genes according to the GSEA for enriched
208 GO terms. Significance threshold were set with an FDR $\leq$ 0.01.

6 MANUSCRIPT IV

IMPROVE, A feature model to predict neoepitope immunogenicity through broad-scale validation of T cell recognition.

Annie Borch, Ibel Carri*, Birkir Reynisson*, Heli Magali García Álvarez, Kamilla K. Munk, Alessandro Montemurro, Nikolaj Pagh Kristensen, Siri A. Tvingsholm, Jeppe Sejerø Holm, Christina Heeke, Keith Henry Moss, Ulla Kring Hansen,Anna-Lisa Schaap-Johansen, Frederik Otzen Bagger, Carolina Barra, Morten Nielsen,* and Sine Reker Hadrup*.
* these author are contributed equally.
Manuscript in preparation

Contribution: In this Manuscript, I have been main responsible for data collection, feature calculation and model development. Additionally, I have made all figures and written the manuscript.

IMPROVE, A FEATURE MODEL TO PREDICT NEOEPITOPE IMMUNOGENICITY THROUGH BROAD-SCALE VALIDATION OF T CELL RECOGNITION

Annie Borch[1], Ibel Carri[2,*], Birkir Reynisson[3,*], Heli Magali García Álvarez[2], Kamilla K. Munk[1], Alessandro Montemurro[4], Nikolaj Pagh Kristensen[1], Siri A. Tvingsholm[1], Jeppe Sejerø Holm[1], Christina Heeke[1], Keith Henry Moss[1], Ulla Kring Hansen[1], Anna-Lisa Schaap-Johansen[4], Frederik Otzen Bagger[5], Carolina Barra[4], Morten Nielsen[2,4,*] and Sine Reker Hadrup[1*].

[1] Department of Health Technology, Technical University of Denmark, Lyngby, Denmark
[2] Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín, Buenos Aires, Argentina.
[3] PokeAcell, Copenhagen, Denmark
[4] Department of Bioinformatics, Technical University of Denmark, Lyngby, Denmark
[5] Center for Genomic Medicine, Copenhagen University Hospital, Denmark
* These authors contributed equally

Abstract

Mutation-derived neoantigens are critical targets for tumor rejection in immunotherapy, and consequently, the scientific community has been exploring new therapeutic approaches to increase the number of neoantigen-responsive T cells directly. Bioinformatic tools have enabled identifying patient-specific neopeptides from sequencing data. However, such strategies suffer from our limited capacity to precisely select those few neoepitopes that can give rise to T cell mediated recognition of cancer cells among that vast majority of immunologically irrelevant neopeptides. We have extensively screened cancer patients undergoing immunotherapy for T cell recognition of neopeptides using MHC multimers. As such, we have analyzed more than 19,000 neoepitope candidates and identified 519 of these to give rise to T cell responses. Based on this large novel dataset, the characteristics of immunogenic neoantigens were explored to derive a feature set optimized to predict neoepitope immunogenicity. Through optimized feature selection, we can significantly improve the selection of immunogenic neopeptides from individual patients, which substantially enhances our capacity to target such neoantigens through targeted therapeutic approaches.

Keywords

Neoantigen; Neoepitope Prediction; Machine Learning; Immunotherapy; Immunoinformatics

Introduction

The development of effective antitumor immunotherapeutic strategies has valued the identification of tumor antigens. Among these antigens, there is a group called neoantigens, which are tumor-specific and private. To be immunogenic, neoantigens must be sufficiently expressed, cross-presented, and recognized by immune cells. Specifically, CD8+ T cells play an essential role by recognizing fragments of the neoantigens, called neoepitopes, and eliciting an immune response that may lead to tumor regression (1).

In this sense, it has been demonstrated that a patient's neoantigen load is predictive of response to immunotherapies, such as checkpoint inhibition (CPI) (1–3). Besides, personalized immunotherapies based on vaccination with neoantigens are under clinical development and have demonstrated neoepitope-directed T cell responses and good tolerability (4). However, the development of therapeutic strategies targeting neoepitopes depends entirely on our capacity to predict which of the many mutational alterations accumulating in tumors that give rise to T cell recognition. With current tools, only 2-6% of predicted neopeptides are demonstrated to give rise to T cell recognition (5–7). This number needs to be significantly improved to facilitate successful clinical implementation for neoepitope targeting strategies.

Various methods have been developed to predict patient-specific neoepitopes from DNA and RNA sequencing (RNAseq). These methods rely on detecting somatic mutations that generate neopeptides and predicting their binding to the patient's MHC to generate a list of neoepitope candidates (neopeptides that have chances to be immunogenic but need to be experimentally validated) (8–12)Even though these methods have been useful to make great advances in personalized cancer immunotherapy, the predicted candidates often contain a high proportion of false positives that will not elicit an immune response (5,6). To improve the specificity of neoepitope detection, multiple mutations and neoepitope characteristics have been investigated as predictors of neoantigen immunogenicity, and machine learning methods have also been developed to rank neoepitope candidates. However, their predictive performance is still limited (9). One of the main challenges for developing accurate neoepitope immunogenicity predictors is the limited available data that experimentally distinguish the immunogenic neoepitopes from the non-immunogenic neopeptides. In addition, experimental evaluation of neopeptide-specific T cell responses is laborious and expensive, requires a patient specific peptide selection, and the breadth of peptides that can be evaluated is limited by scarce patient samples (13)).

In this study, we explore the characteristics of immunogenic neoepitopes in order to improve their prediction. To this aim, we arranged a large dataset of more than 19000 neoepitope candidates screened for neoantigen reactive CD8+ T cells (NARTs) that were derived from three studies covering in total 70 patients with different tumor types. All three studies used the same technology to determine T cell recognition of neoepitopes, based on barcoded MHC multimers (13), hence distinguishing the immunogenic neoepitopes from the non-immunogenic neopeptides.

Another challenging aspect of neoantigen prediction is the particular characteristics of each patient's tumor and immune system. It is known that tumors evolve to be less immunogenic by the process of immunoediting (1,7,14). The downregulation of MHC transcription, the induction of T cell exhaustion, and the modulation of the immune infiltrate by the production of different suppressor cytokines are some of the mechanisms that favor cancer progression. The introduction of immunotherapies in immunocompetent patients may alter this natural course by shifting the tumor microenvironment (TME) profile, stimulating neoantigen cross-presentation, and enhancing the T cells activity to finally develop an effective antitumoral immune response (1). However, not all patients respond to immunotherapy, and even those that do, may not take full advantage of the immunogenic potential of the neoantigens.

As a consequence, when screening patients for T cell recognition toward neopeptides, such data will comprise a substantial number of false-negative data, i.e. neopeptides that have the potential to be immunogenic but are not recognized by T cells in the given cancer patients, due to intrinsic immune and tumor characteristics. Such event forms a false negative data sink that challenge our predictive capacity. To compensate for this potential bias and hence improve the predictive capacity, we took into account also the association between neoantigen immunogenicity and the characteristics of the TME (15,16).

Based on the described data, we developed machine learning models that integrate i) the neopeptide sequence, ii) neopeptide-derived features such as their physicochemical properties, the source mutation qualities, the likelihood of antigen presentation and T cell propensity, and iii) patient-specific derived features including MHC expression in tumor cells, the cytolytic activity (CYT) and the different cell populations that constitute the TME. We found that the combination of these features increases the performance for the selection of immunogenic neoepitopes. This result suggests that the challenges in neoantigen prediction can be addressed by integrating multiple factors from the complex antitumor immune response.

Results

Selection of neopeptides and experimental evaluation of their immunogenicity

In this study, we assessed the features associated with the immunogenicity of predicted neoepitope candidates derived from the tumors of 70 cancer patients, each experimental evaluated for T cell recognition covering a range of 100 to 1360 neopeptides per patient. Here, immunogenicity is defined as the capacity of a given neopeptide to raise a detectable CD8 T cell response in the given patient. This study includes a total of 19541 neopeptides, among which 519 (2.6%) were recognized by a T cell population in a given patient. This data was generated through the screening of three patient cohorts, including different tumor entities. One being a cohort of metastatic melanoma patients receiving adoptive cell transfer (ACT) with tumor infiltrating lymphocyte (TIL) (TIL-ACT) (5), second, a cohort of metastatic urothelial carcinoma (mUC) patients who received PD-L1 checkpoint inhibition CPI) (7), and third a basket trial cohort with different cancer types and types of CPI treatment ((17), not published) (Figure 1A). The percentage of immunogenic neoepitopes in the different cohorts was 3.45%, 2.35%, and 2.28%, respectively (Supplementary Figure 1A).

All neopeptides included for T cell screening across the three studies were extracted based on the patient's individual tumor mutational landscape, derived from paired tumor/normal whole-exome sequencing (WES) and RNAseq from each patient. Genome analysis tool kit (GATK) best practices (18) were applied to obtain somatic variants, followed by peptide extraction with MuPeXI (9) (Figure 1B). From this pool of neopeptides, we evaluated, on average 279 (range 100-1360) for T cell recognition per patient. The majority of the neopeptides included for T cell screening were classified as binders (NetMHCpan 4.0, Eluted Ligand % Rank (RankEL) < 2) to the patient's HLA class I molecules. Peptides from patients with a high number of candidate mutations were selected using a strong binding (RankEL < 0.5) threshold, whereas the RankEL threshold was relaxed for patients with fewer neopeptide candidates to meet the inclusion of 100 neopeptides per patient. The selection criteria resulted in a binding score distribution with two peaks (of strong and weak binders), as seen in Supplementary Figure 1B. Only neopeptides from transcribed regions (transcripts per million (TPM) >0.1) were included. (Supplementary Figure 1C). In total, 31 different HLA class I molecules were included for T cell screening across all patients, and T cell responses were found restricted towards 27 of these. Some HLA molecules obtained a significantly higher proportion of immunogenic neopeptides, including HLA-A0101, A0301, B0702, B1501, B4001, B4402, C0202, C0602. (Supplementary Figure 1D). In general, no strong association was observed between the number of neopeptides included for T cell screening and the number of immunogenic neoepitopes identified (Figure 1C). This is probably reflected by the large number of neopeptides evaluated for all patients and the fact that patient-specific features determine tumor immunogenicity.

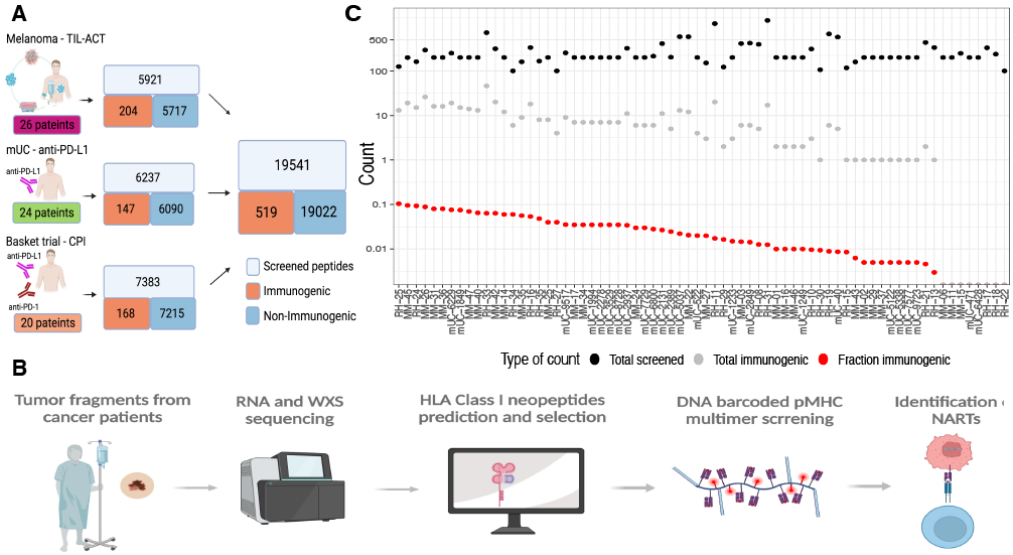


Figure 1: Data overview. A) Data overview illustrating the number of validated peptides for each cohort and the number of patients screened together with a summary of the total amount validated with the number of immunogenic and non-immunogenic neopeptides. B) General workflow of the data generation, including the patient samples being sequenced and patients' specific libraries with neoepitope candidates being generated and screened with patients' samples to find immunogenic neoepitopes. C) Patient overview according to the number of neoepitopes screened (black dots) and immunogenic neoepitopes (grey dots) and fraction immunogenic (red dots). "MM" stands for the melanoma cohort, "mUC" for the mUC cohort, and "RH" for the Basket cohort.

Identification of features associated to neoepitope immunogenicity

Next, we interrogated which characteristics influence the immunogenicity of neoepitopes. To broadly elucidate how different neopeptide characteristics would impact the chance of raising a T cell response against such peptides, we investigated 27 features and their association with neoepitope immunogenicity (Supplementary Table 1). First, we explored if the type of mutation influences its immunogenicity. The mutation types are categorized as missense, frameshift, in-frame insertion, and in-frame deletions. More immunogenic neopeptides were observed in the missense and frameshift category, but these also constituted a larger fraction of the evaluated neopeptides, hence no enrichment was observed ($p = 0.1$, proportion z-test) (Figure 2A).

A previous study has observed the mutation position to be important for the immunogenicity of neoepitopes (19). To investigate neopeptides of different lengths, we applied the predicted 9mer binding core derived from NetMHCpan 4.1 of all the neopeptides with missense mutations (20). Applying this, we could uniformly investigate the role of the mutation position despite their original length. In immunogenic neoepitopes, mutations were predominantly located around the anchor positions of MHC I motifs, while there is a depletion of mutations near position 4 in 8-,9-,10-mers and near position 5 in 11-mers (Figure 2B, top). This dynamic was not observed in the non-immunogenic neopeptides (Figure 2B, bottom). We also observed a significantly increased frequency of mutations in the gap position (outside the core) of 10-mer neoepitopes ($p= 0.01$, proportion test) (Figure 2B) compared to the non-immunogenic ones, suggesting that mutations in longer peptides, in general, are facing out toward the T cell in immunogenic peptides, as gap positions are generally characterized by protruding residues. To differentiate the neopeptides with improved binding compared to wild-type (wt), i.e. neopeptides where the wt version is expected to be poorly presented in the given HLA; and conserved binding (CB), i.e. neopeptides where both the wt and the mutated version

is expected to be well presented on HLA, we calculated differences between mutant and wt RankEL as described in (21), revealing 291 conserved binders and 228 improved binders (IB) among the immunogenic neoepitopes. As expected, the IB mostly have mutations in anchor positions, affecting the HLA binding capacity of such peptides, whereas the CB has mutations in the 'T cell receptor (TCR)-binding' region, pos. 4-8 (Figure 2C).

Considering that tumor neoepitopes derive from normal self-peptides, it is hypothesized that the mutations should generate drastic changes to break the immune tolerance. To determine this, we evaluated the self-similarity of the neopeptide either in the CB or IB category. By comparing CB with IB, we find lower self-similarity in the CB, both within the non-immunogenic and immunogenic neoepitopes (Figure 2D). Comparing non-immunogenic neopepitdes with the immunogenic in each category, no significant difference was observed in self-similarity (Supplementary Figure 2A). These results indicate that the self-similarity on its own was unable to distinguish immunogenic neoepitopes from non-immunogenic neopeptides. Additionally, we compare wt to the mutant peptide with the foreignness score (Foreignness) and the Differential Agretopicity Index (DAI), as they both previously independently and combined have shown to be important in defining immunogenicity for pathogenic and cancer-derived epitopes (22–27). However, we do not observe any significant difference in either of them (Supplementary Figure 2A).

As mutation calling can give false-positive mutation assignments (28–30), we investigated the presence of the WES-called mutation in the transcriptome. We were able to validate whether the mutation was present in the RNA in 85% of the neopeptides. Among these, 47% of neopeptides originated from mutations that could be detected in at least one transcript of the RNAseq, 47% were not found in RNAseq, and 6% had insufficient RNA coverage in the region of the mutation (see method). No significant difference was found with this validation in separating immunogenic neoepitopes from non-immunogenic ones (Figure 2E). Neither when observing the cohort independently (Supplementary Figure 2C). Additionally, the coefficient of the identified transcript (ValMutRNACoef) was not associated with neopeptide immunogenicity (Supplementary Figure 2A). Furthermore, we investigated if the level of expression of a mutated gene was associated with immunogenicity. RNAseq was used to calculate the abundance of the source transcripts. No significant difference was found in the RNA expression level of immunogenic and non-immunogenic neopeptides (Wilcoxon rank-sum test;$p = 0.56$) (Figure 2F). Of note, in this dataset, finding detectable RNA transcripts for the given mutation was election criterion for all neopeptides evaluated for immunogenicity (shown in Figure 1D) and hence could affect the result. However, from this analysis, we found no further enrichment of immunogenicity associated with level of gene expression. Cellular prevalence (CelPrev) and the variant allele fraction (VarAlFrac) are mutation specific features that describe the prevalence of the mutated alleles. We find CelPrev to significantly influence the identification of immunogenic neoepitopes (Supplementary Figure 2B).

Strong binding affinity of the neopeptide to MHC and high stability of the peptide-MHC (pMHC) complex have been shown to characterize immunogenic neoepitopes in previous studies (5, 9, 18). We demonstrate that the predicted binding affinity % Rank (NetMHCpan 4.0, RankBA), RankEL, and the predicted pMHC stability significantly differentiate immunogenic and non-immunogenic neopeptides (Figure 2F, Supplementary Figure 2B). Of note, in this study, the investigated neoepitopes were pre-selected according to RankEL, potentially generating a sampling bias, which may reflect why RankEL displays less separation within two groups compared to RankBA.

Next, we characterize the physicochemical properties of the neopeptides by calculating the molecular weight, aromaticity, instability, isoelectric point, molar extinction coefficient, mean hydrophobicity, and proportion of hydrophobic and aromatic residues. Anchor residues interact with MHC molecules, and non-anchor residues are more exposed to encounter with a

TCR. Considering that the non-anchor region of the peptide may be more relevant for immunogenicity, we extracted this subsequence (from the 4th to the penultimate amino acid) and calculated the mean hydrophobicity and the proportion of small, aromatic, basic, and acidic amino acids. Mean hydrophobicity in non-anchor subsequence of the neopeptide (HydroCore) (Figure 2H) and 'PropHydroAro', a parameter that describes the proportion of hydrophobic and aromatic residues in the peptide (Figure 2I), both significantly differ in non-immunogenic and immunogenic neopeptides. This observation aligns with other findings demonstrating hydrophobicity as a hallmark of the immunogenicity of T cell epitopes in general (31) and specifically for neoepitopes (25).

Finally, we included the prediction of T cell propensity with the PRIME model (PrimeScore). This model pondered the physicochemical properties in non-anchor residues and showed a significant difference in separating the immunogenic from the non-immunogenic neopeptides (Supplementary Figure 2B).

In summary, more than half of the features (17/27) showed a significant difference between immunogenic and non-immunogenic neoepitopes. To assess the performance of these features to independently drive an improved identification of immunogenic neoepitopes, we calculated the area under the receiver operating characteristic curve (AUC) and the partial area under the receiver operating characteristic (ROC) curve at 0.1 (AUC 0.1). This latter metric was included as a mean to focus on the high specificity part of the ROC curve. Each feature independently reaches a max performance of AUC 0.1 = 0.013 and AUC = 0.612. (Figure 2J and Supplementary Figure 2D). Indicating that even though a significant difference was obtained for a feature in differentiating immunogenic and non-immunogenic neoepeptide (Figure 2F-I, Supplementary Figure 2B) a poor performance was obtained when evaluating AUC. In addition, some features, that did not reveal significant differences between immunogenic and non-immunogenic neopeptides, may still play a role in a feature combination matrix. This is illustrated by the combination of RNA expression level > 2 TPM and RankEL < 0.5 selects a higher proportion of immunogenic peptides (proportion test, $p = 0.0001$) compared to each feature independently, where the expression level has a $p = 0.073$ and RankEL has $p = 0.008$ (Figure 2K). Similar results are observed for HydroCore > 0.3 ($p = 8.04 \times 10^{-16}$) and PropHydroAro > 0.4($p = 9.67 \times 10^{-15}$), where the combination has a p-value of 2.2×10^{-16} (Figure 2L). Together these findings encouraged us to use machine-learning strategies to build a multiparametric model for predicting the immunogenicity of neoepitopes.

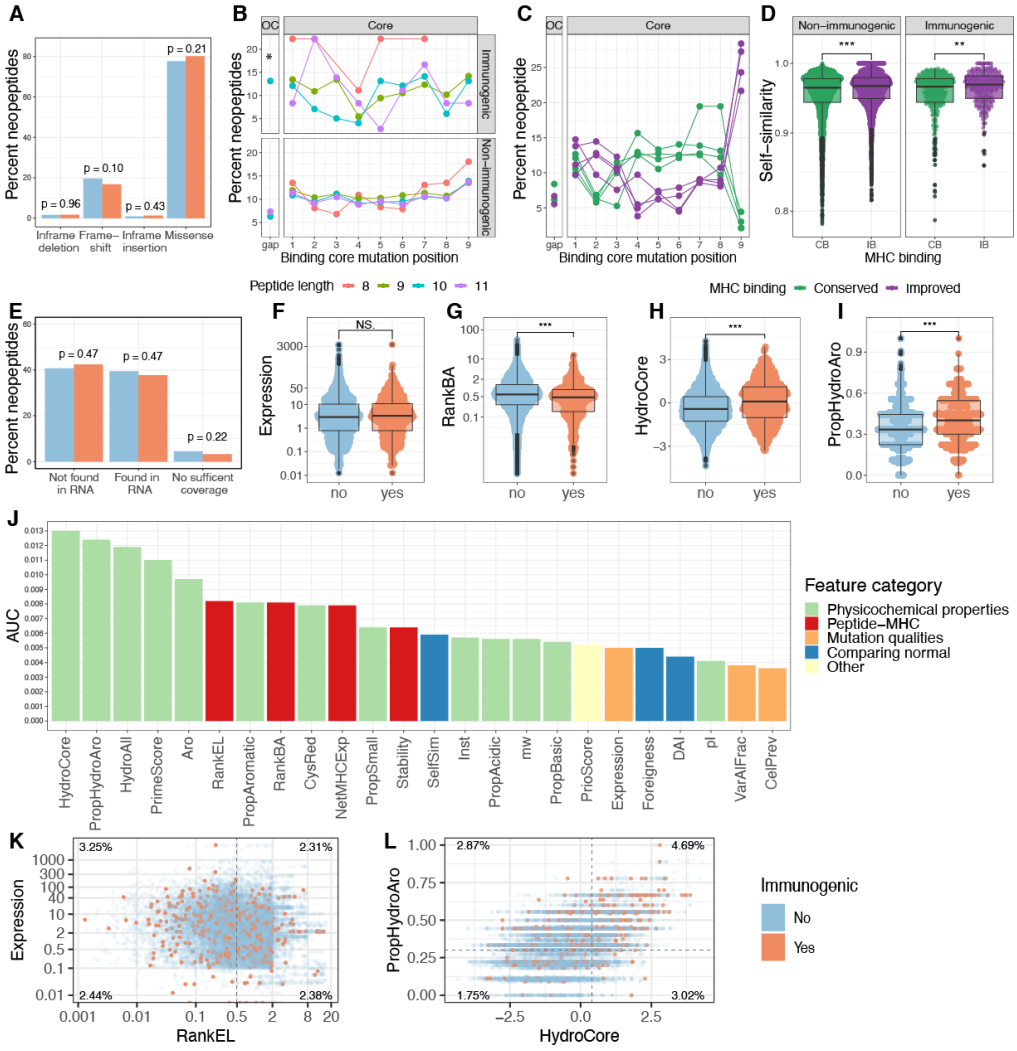


Figure 2: Features and immunogenicity. A) Percent of immunogenic neoepitopes according to the mutation consequence. The p-values are according to the proportion test, testing if the number of immunogenic neoepitopes for each mutation type is present in a higher fraction compared to the non-immunogenic ones. B). Immunogenic neoepitope fraction for all missense according to peptide position and peptide length where the gap position is outside the core (OC) and is significantly enriched for neopeptides with a length of 10 within the immunogenic neoepitopes (p=0.01, prop.test). The peptide is separated into immunogenic and non-immunogenic neopeptides. C) Same as (B) but only according to the immunogenic peptide and then separated into the improved binder (IB) and conserved binders (CB). D) Self-similarity comparing the IB and CB for the neoeptiopes. E) Percent of immunogenic and non-immunogenic neopeptides where the mutation is validated in RNA. The statistic is performed from a proportion test testing if the proportion of found RNA has more immunogenic neoepitopes than those not found. F-I) boxplot comparing the non-immunogenic form immunogenic neopeptides for four selected features, statistic by Wilcox test. F) Expression p = 0.56. G) RankBA $p = 4.2 \cdot 10^{-9}$. H) HydroCore $p = 5.8 \cdot 10^{-14}$. I) PropHydroAro $p = 2.2 \cdot 10^{-16}$. J) AUC01 performance for each feature independently colored by feature type. K+L) Dot-plot combining two features with the fraction of immunogenic vs. non-immunogenic for each quadrant according to the selected threshold. The percent of immunogenic neopeptide for each quadrant is noted. J) Expression with a threshold of 2 vs. RankEL with a threshold of 0.5. K) PropHydroAro threshold of 0.3 vs. HydroCore with a threshold of 0.4.

Improved prediction of immunogenicity by random forest modeling

RF was applied to develop a model that predicts immunogenicity in our dataset, based on the features previously described. The overall workflow is illustrated in Figure 3A. Considering that previous studies demonstrate that highly correlated features reduce the trainability of RF models (32,33), the feature space was reduced to only include features with a Spearman's correlation coefficient lower than 0.7 (higher than -0.7 if negative). This resulted in discarding HydroAll (correlated with HydroCore) and VarAlFrac (correlated with PriorScore) (Supplementary Figure 3A). Furthermore, the binary one-hot-encoded features (i.e. yes/no features), including mutation consequence and mutation position, did not add predictive power (based on a backward feature selection) and therefore were not included in the model. Based on these selection criteria, 22 features were incorporated into the feature-based RF model.

As an alternative strategy to the feature-driven model, we also developed a sequence-based model using the NNAlign method (34)The type and frequency of amino acids in the neopeptide sequences may contain information related to their immunogenicity. This has previously been explored in sequence-based models for the detection of both pathogen-derived epitopes and cancer neoepitopes (35,36). A 5-fold cross-validation scheme was used to train and evaluate both models. To avoid data redundancy, a modified common motif clustering was applied (described in Materials and Methods). This ensures that all the data from the same patient is partitioned together to avoid overfitting on patient-level features. The result was a dataset split into 5 partitions, separated by neoepitope candidate, common motifs, and patients.

The performance of the models was evaluated in terms of AUC and AUC01. Both the RF model and the sequence-based NNAlign model performed significantly better than the selection based on RankEL alone (as previously done, for the immunogenicity testing). The RF model display significantly better prediction (AUC=0.638 and AUC01=0.0147) than NNAlign (AUC=0.605 and AUC01= 0.0145), both outperforming RankEL (AUC=0.537 and AUC01=0.0082). Investigating the performance of an ensemble model, based on the average score of both methods, demonstrated no major improvement with respect to the RF model (Figure 3B). We analyzed the separation of immunogenic and non-immunogenic neopeptides based on the prediction scores. Looking at each cohort independently, a significant separation was found in all cohorts according to the RF model. The NNalign model obtained a significant separation in Basket and Melanoma but not in the mUC cohort (Figure 3C). When considering the performance of each patient, again, the RF model demonstrates a significant improvement both for AUC01 and AUC, compared to the RankEL, with the RF model displaying the best performance (Figure 3D, Supplementary Figure 3B).

Both RF and NNAlign allow interpretation of the rules learned by trained models. Investigating the sequence logos produced by NNAlign, immunogenic neopeptides were found to be enriched in hydrophobic and aromatic residues (Supplementary Figure 3C). Also, when analyzing the feature importance for the RF model, the mean hydrophobicity and PRIME score (36) were found to be the most relevant features (Figure 3E).

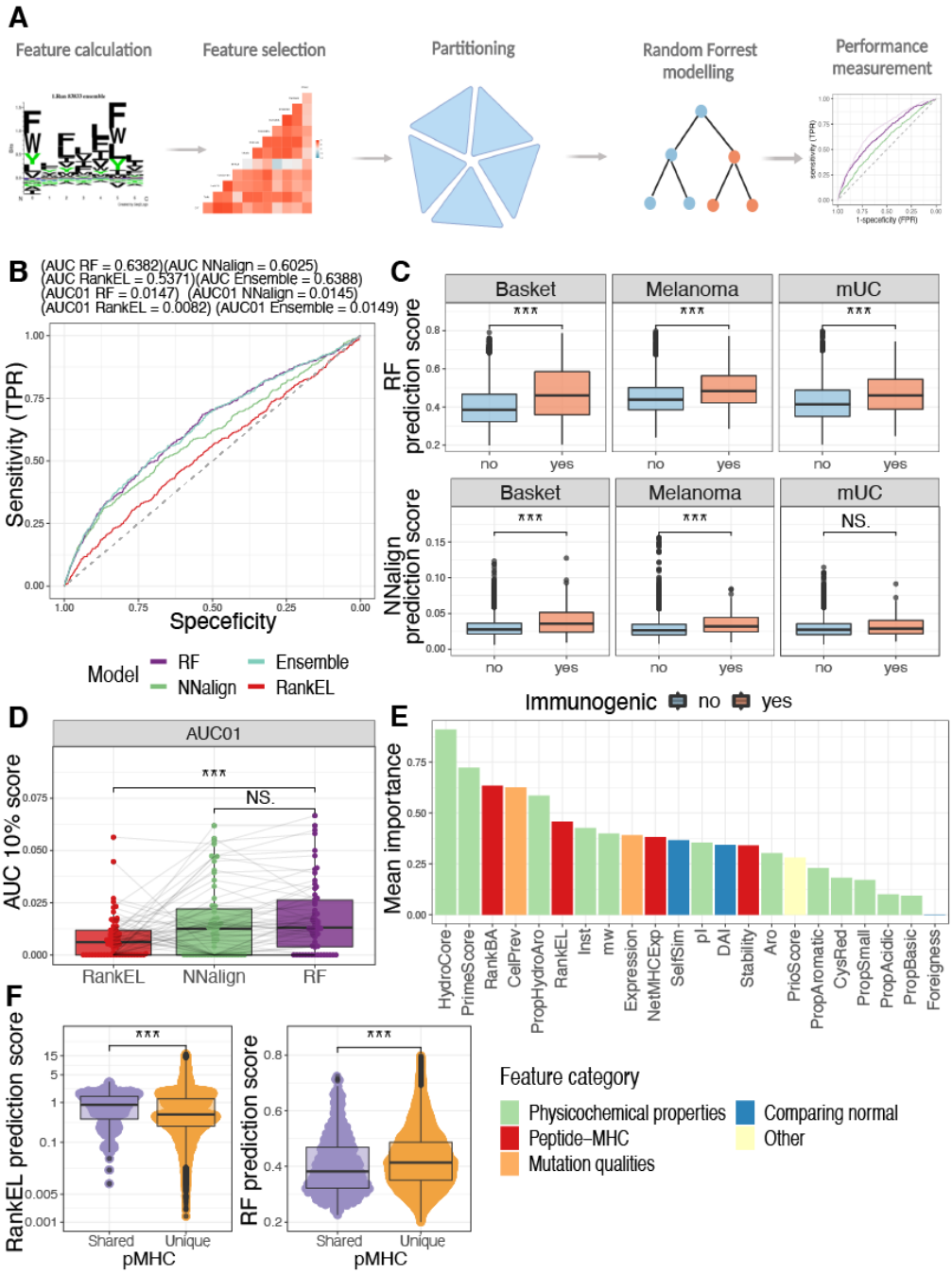


Figure 3: Random Forrest modeling. A) Strategy of the machine learning approach with feature selection, partitioning, and modeling. B) Roc curve with the random forest (RF) model in purple, which performs significantly better than the NNalign in green (p = 0.001, roc.test) and RankEL (p= 7.7^{-8}. C) Prediction score from the NNalign model in top and RF model in bottom according to the immunogenic and non-immunogenic peptide splitted by cohort. The RF model had significant separation in all three cohorts. The basket obtained p-value at = 1.1^{-10}, mUC p-value at 1^{-7}, and Melanoma obtained p-value = 2.6^{-8}, all with non-paired Wilcoxon test. The NNalign model obtained significant separation in Basket (p= 2.5^{-10}, Wilcox test) and melanoma

($p= 6.9^{-09}$, Wilcox test), but non-significant for the mUC cohort (p=0.77). D) AUC01 per patient, RF model is significantly better than RankEL (p-value = 0.00015) but non-significant but slightly improved compared to NNalign (p=0.066). E) Mean feature importance for the RF model colored by the feature category. F) RF prediction score and RankEL prediction score for non-immunogenic neopeptides comparing shared pMHC between patients to unique pMHC. F) Comparing unique and shared non-immunogenic neopeptides for the RF and RankEL prediction scores. The RF has a significantly lower prediction score for the shared compared to the unique ones (p-value = 3.2^{-6}, wilcox test), and the RankEL shows the opposite; the shared have a significantly higher score compared to the unique ones ($p = 6.9^{-09}$, wilcox test).

In general, the prediction scores of the RF model were able to separate the immunogenic neoepitopes and non-immunogenic neopeptides and performed slightly better than the NNAlign model, with both models significantly improving prediction compared to RankEL. Although neoepitopes most often are found as private antigens since most mutations are unique to the given patient's tumor. However, given the size of the data set in this study, we have been able to detect several identical neopeptides present in the tumors of multiple patients. 4% of the dataset corresponds to neopeptides whose sequence is observed in more than one patient and 2% to pMHC present in more than one patient. Following T cell screening we found T cell recognition to 4 pMHC to be 'shared' neoepitopes. Interestingly, these neoepitopes were found to be immunogenic in only one patient and negative in the rest, indicating that patient-specific characteristics may influence the neoantigen-directed immune response. This observation leads us to question whether the negative neopeptides that have been tested in a single patient could be immunogenic in another patient or environment and therefore act as false negatives for the RF model. To investigate this aspect, we hypothesize that pMHCs that have been evaluated as negatives in multiple patients are true negatives with more confidence. We observed that the RF model assigns lower values to these negatives than to private neopeptides ($p = 3.2^{-6}$) (Figure 3F), suggesting that at least a proportion of the private negatives share features comparable to that of the positives, and hence could be false negatives.

Encountering tumor microenvironment improves prediction but not on patient-level

The TME comprises an essential factor in the antitumoral immune response. Earlier studies have suggested that a combination of tumor mutational burden (TMB) and TME can be used as a biomarker to predict the patient's response to immunotherapy (37–39). Also, it has been widely discussed that a hot TME is essential for having TILs recognizing neoantigen (5). The presence of immunogenic neoepitopes will only lead to tumor elimination if active NARTs are present. Cytolytic activity (CYT, geometric mean of GZMA, and PRF1 expression) can be used as an estimate for T cell cytolytic activity and has been found to correlate with the neoantigen load (40). HLA presentation is an essential factor for neoantigen-directed T cell mediated killing of cancer cells. Downregulation of MHC molecules is a known escape mechanism of tumor cells (41,42), and the HLA expression in tumors correlates to higher immune cell infiltration and prolonged survival (1,43).

Based on RNAseq data, we derived information regarding the proportion of the different cell populations that compose the TME, the CYT, and the HLA expression (HLAexp) of each tumor sample included in this study. CYT and HLAexp correlated with the number of immunogenic neoepitopes per patient (Figure 4A-B). Furthermore, HLA expression showed a significant separation between the immunogenic and non-immunogenic neoepitopes (Supplementary Figure 4A), and it was relevant for all HLA loci (Supplementary Figure 4B). Based on these results, we integrated the patient-specific immune-related features in an RF model (RF TME) in addition to the above-defined neopeptide-derived features. A correlation was found between the TME features (Supplementary Figure 4C), and due to the inability of the RF model to detect the importance of highly correlated features, a selection scheme was

therefore implemented. Features were defined by a forward selection based on the highest performance and excluding features with a high correlation with the before-selected features. The AUC was found for each feature and train set, and the mean of each feature was calculated (Supplementary Figure 4D). For example, CYT had the best mean performance, followed by CD8 T cells (TcellsCD8), but since both features were highly correlated (Spearman's correlation coefficient = 0.7, the TcellsCD8 was not included in the model. Using this approach, the final included TME features consist of: CYT, HLA-expression (HLAexp), and monocyte linage (monocyte). Similar to HLAexp and CYT, monocyte also showed a correlation with the number of immunogenic neoepitopes (Supplementary Figure 4E). To validate that this feature selection did not result in overfitting for the data, the feature selection was also made for each train set and used for training and testing. This procedure did not change the performance of the model. From this, we conclude that the potential overfitting imposed by feature selection prior to the model training is minimal, and therefore, for further analyses, we moved forward with the first simpler model using an equal feature space between the different folds in the cross-validation.

When including these four TME features (CYT, HLAexp, and monocytes) with the previously described neopeptide features, the performance of neoepitope detection increases significantly ($p = 0.002$) with a global AUC of 0.66 (Figure 4C). However, when the performance of each patient was investigated, only a significant improvement was achieved in AUC ($p = 0.025$) and none in AUC01 (Figure 4D+E). To elaborate on this, we calculated the difference or delta between the prediction scores derived from the RF TME and RF models (RF TME - RF). It could be observed that peptides from patients with high CYT had a positive delta, while peptides from patients with low CYT had a negative delta (Spearman's correlation coefficient = 0.75 and 0.82) (Supplementary Figure 4F). In other words, the model with TME features favored peptides from patients with high CYT and decreased the prediction score of peptides from patients with low CYT; therefore, TME features themselves were unable to distinguish the immunogenicity of peptides within patients but favored the patients with a favorable TME. Hence, the improvement in AUC likely stems from enrichment of T cell responses in patients with favorable TME, which will also affect the immunogenicity classification in our dataset. Thus, to only determine peptide features associated with immunogenicity, the RF model can be used, but to include the likelihood of finding T cell reactivity towards such neopeptide in a given patient, the RF TME model should be applied.

Finally, we assessed the relative feature importance based on the RF TME model learned rules. Also here, HydroCore was the feature that constituted the highest importance for the model. From the novel features, CYT represented the second relevant feature, whereas HLAexp was the third (Figure 4F). In conclusion, TME and HLA expression are useful features to assess the proclivity of a patient to have a detectable T cell response against immunogenic neoepitopes.

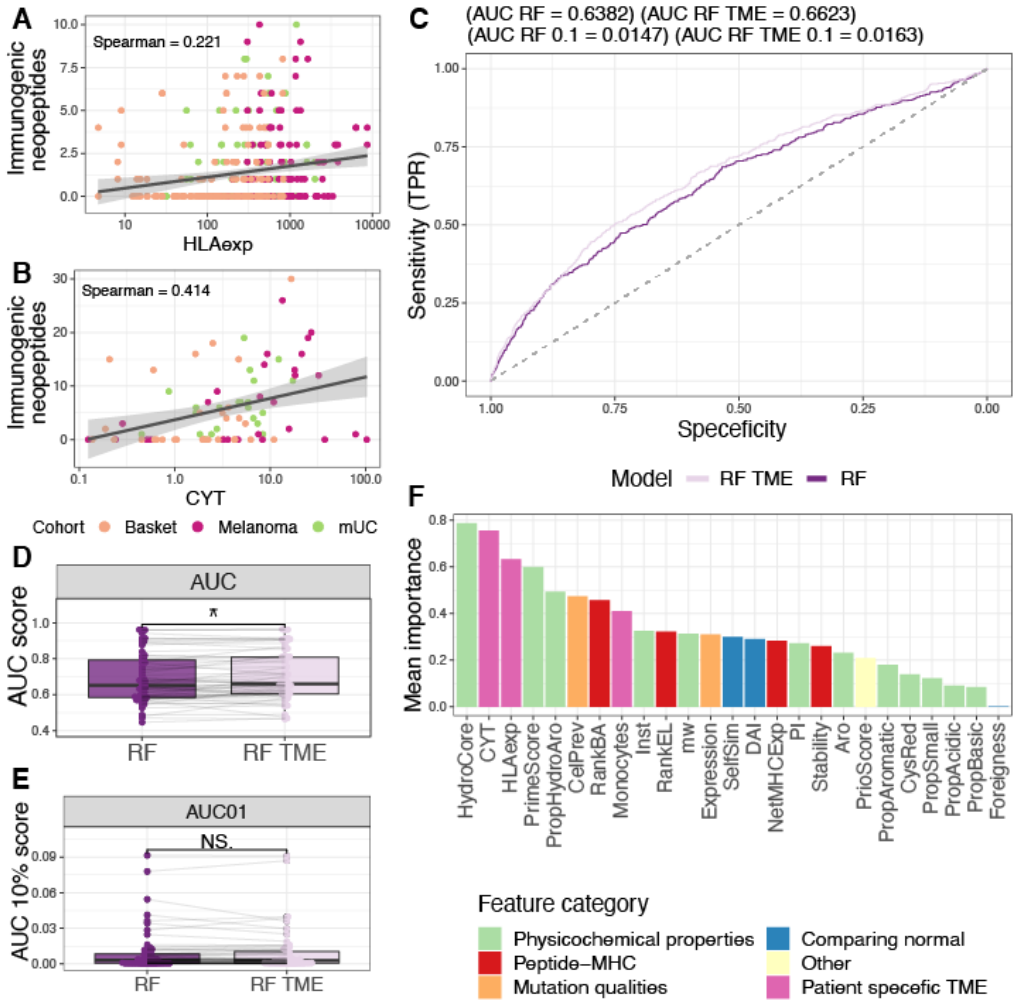


Figure 4, Random Forest with tumor microenvironment (TME) parameters. A) CYT correlation with the number of responses. B) Correlation of HLA expression compared to the number of responses. C) Roc curve illustrating the three models. NNalign in green, The RF model without TME features in dark purple, and RF with TME features in light purple. RF TME is significantly better than RF ($p = 0.002$, roc.test). D+E) AUC per patient for the two models and statistics made with paired Wilcox test D) AUC01 $p = 0.85$, and E) AUC $p = 0.0025$. F) Mean feature importance for the RF with TME features colored by the feature type.

Capturing of true positive neoepitope and association with patent survival

To develop therapeutic interventions targeting neoepitopes, its crucial to precisely determine immunogenic neoepitopes of relevance. Hence, based on the IMPROVE prediction model, we investigated our capacity to catch true positive immunogenic neoepitopes out of the total possible pool. Here, the proportion of observed immunogenic neoepitopes located in the top 20 and 50 ranked neopeptides per patient was estimated according to the predicted value from both models (with and without TME) and compared with the fraction of immunogenic neoepitopes found in the top according to the RankEL and with random sampling. Both RF models showed a higher proportion of identified immunogenic neoepitopes compared to RankEL and random sampling. However, the implementation of the TME did not demonstrate any improvement compared to the model without TME (Figure 5A+B).

We further evaluated the fraction of true-positive (i.e. T cell evaluated immunogenic neoepitopes that could be identified by the different methods. First, we defined a cut-off for selection of predicted 'true' neopeptides using our IMPROVE model. This cut-off was defined as the point where the sensitivity and specificity crosses. This was calculated for the model both with and without TME features included (Figure 5C+D). Next, the models were compared to RankEL with a cutoff at 0.5, defining the strong binders. When implementing the defined cutoffs, the RankEL identified 9737 positive neopeptides, of which 289 were TP with a Matthew correlation coefficient (MCC) at 0.02 and an accuracy of 0.50. The IMPROVE model without TME resulted in 7853 positive neopeptides, of which 308 were TP (MCC at 0.06 and accuracy of 0.60), and the IMPROVE model with TME resulted in 7684 positive neopeptides, of which 320 were TP (MCC at 0.08, and accuracy of 0.61) (Figure 5E). The latter corresponds to the identification of >60% of the TP immunogenic neoepitopes.

Earlier studies have suggested neoantigens as a biomarker for patient outcome and survival to checkpoint inhibition (1). To investigate this in the context of our prediction model, we used the models to predict the immunogenicity score for all possible neoepitopes in the individual patients included in the studies addressed here. Using the previously defined cutoffs for the IMPROVE model (Figure 5C+D), a 0.5 cutoff for the RankEL model, patients were separated into four groups by the quantile related to the number of predicted neoepitopes using the 3 different models. When evaluating the patients' Overall Survival (OS) and Progression Free Survival (PFS), related to their level of neoepitopes, based on this quantile separation Focusing on the two ultimatums (high vs. low), a better overall survival probability was obtained based on the IMPROVE model compared to Rank EL. This was evident both without and with the TME factors included; focusing on the two ultimatums (high vs. low), we find HR: 0.45<0.36>0.35 related to OS and HR: 0.40<0.42>0.34 related to PFS (Figure 5F). The included cohorts were very heterogeneous, given different cancer types and treatments, and hence somewhat differently influenced by the level of neoepitopes when evaluated individually (Supplementary Figure 5). Although the prediction of patient outcome is not the key accomplishment of the IMPROVE prediction strategy, the influence observed on patients outcome suggests that those predicted immunogenic neoepitopes do influence treatment outcome, and survival, and supports that understanding that we select neoepitopes of relevance for active T cell recognition and tumor elimination.

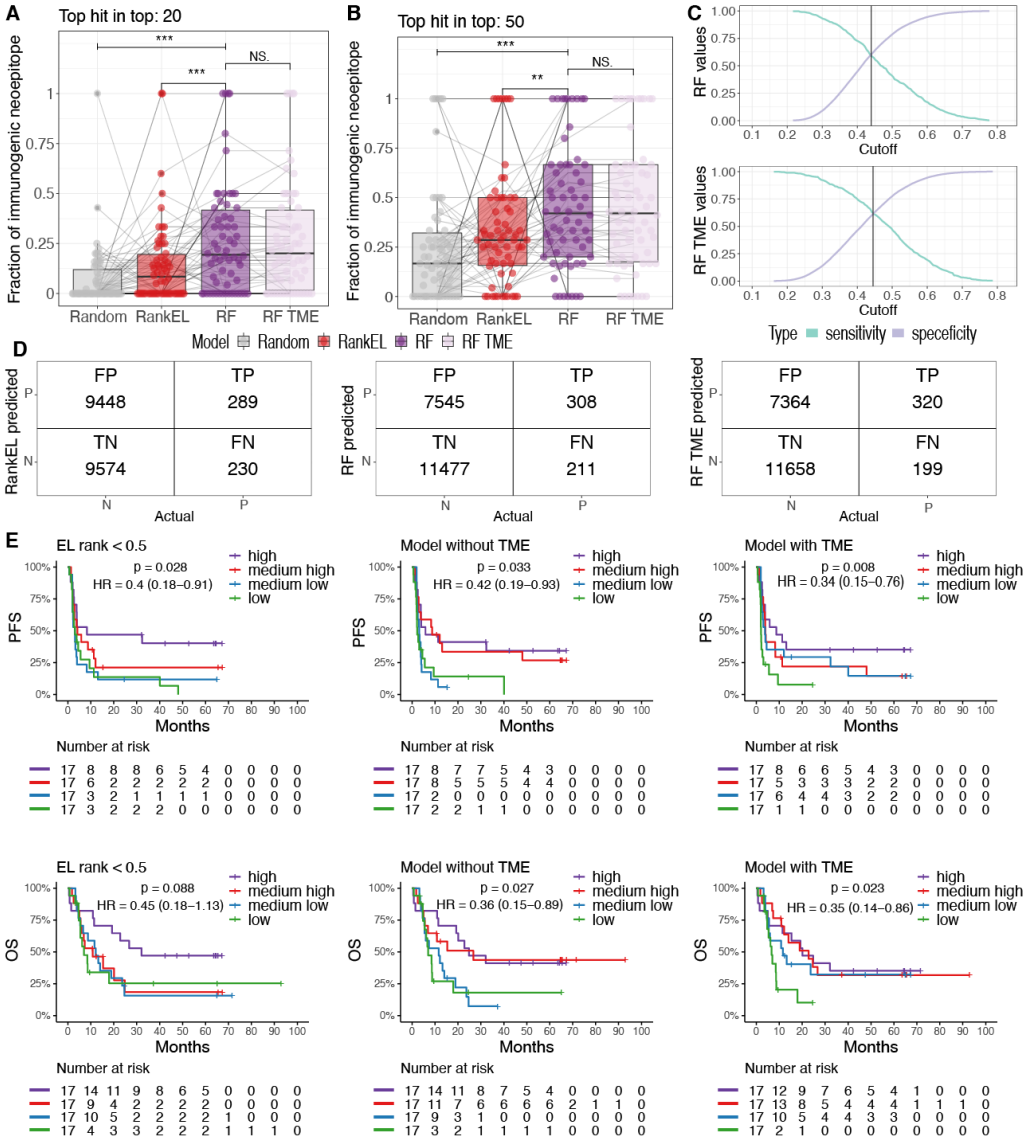


Figure 5, Patient performance and survival. A+B) The fraction of immunogenic neoepitopes in the top 20 and top 50 neoepitope candidates of the random forest (RF) model (dark purple), the RF model with CYT and HLA expression (light purple), red according to Eluted Ligand %Rank, and grey is randomly sampled peptides. A) top 20 neopeptides RF vs. RankEL (p=0.00047), RF vs. random ($p=9.7^{-7}$), RF vs. RF TME (p=0.85). B) top 50 neopeptides RF vs. RankEL (p=0.005), RF vs. random ($p=4.7^{-5}$), RF vs. RF TME (p=0.76). C+D) Sensitivity and specificity calculated for a range of different cutoffs where the point the curve crosses define the set cutoff of what is predicted immunogenic and non-immunogenic. E) Confusion matrix (left) for RankEL with threshold < 0.5 and according to the pre-selected neoepitopes with expression above 0.01. The (middle) shows the RF model without TME features with the defined threshold found in (C). The confusion matrix to the (right) is the RF model with TME features included, with the defined threshold found in (D). F) Kaplan-Meyer curves for observing all predicted neopeptides with a threshold of RankEL< 2 and Expression > 0.01, which included predicted neoepitopes that have not been screened, for example, HLA alleles that were not available and neopeptides for patients selected with a more restricted threshold. The survival analysis is made for the three categories described in the confusion matrix. The patients are separated into four groups according to the number of predicted

neoepitopes above the defined threshold. The four groups are determined according to the quantile where "high" is above the 3rd quantile, "medium high" in between the 2nd and 3rd quantile. "medium low" is between the 2nd and 1st quantile, and low is below the 1st quantile. The threshold for predicted neoepitopes for the RF model with and without TME (found in Figure 5B and C), whereas the threshold for RankEL is set to 0.5 and is also the threshold used in the confusion matrix.

Discussion

Previous studies investigating neoepitope candidates are hampered by limited data, making it challenging to learn the rules of neoantigen immunogenicity. This study evaluates above 19,000 neoepitope candidates with more than 500 immunogenic candidates verified through large-scale T cell interrogation using barcode-labelled peptide-MHC multimers. As such, this dataset is, by our knowledge, the largest experimentally verified pool of neoepitope candidates. This study uses random forest (RF) modeling to predict neoantigen immunogenicity, including 22 features covering mutation-specific, peptide sequence-specific, and patient-specific parameters (1,25)

To overcome the scarcity of validated neoepitopes, other studies have predicted neoepitope immunogenicity based on pathogen-derived epitopes (26,36). One study that investigates immunogenicity characteristics only included cancer neoepitopes and found significant differences in pMHC stability, affinity, and gene expression level (TPM) (25). The data set for this study has encountered the pMHC binding and expression level during the selection of peptides, as they were previously shown to be important parameters (25,26) and therefore focuses on features beyond these already pre-selected features. In this small dataset (25), including 30 immunogenic neoepitopes, low hydrophobicity was suggested as an important parameter for immunogenicity (25). However, this is in contrast to the findings presented here and by others (35,36), demonstrating that high hydrophobicity positively influences immunogenicity. This demonstrates the need for assessment of larger datasets, allowing data partitioning to avoid conclusions driven by few epitope sequences.

Our feature-based RF model (IMPROVE) was able to improve neoepitope prediction over single feature-based models and NNalign. Our results provide insight into the most important characteristics of an immunogenic neoepitope. The feature with the highest importance to model performance is the hydrophobicity of non-anchor residues. PropHydroAro and PRIME were also found important, which all address the hydrophobicity and/or aromatic residues. Models with backward and forward feature selections were applied to ensure that the model with all included features gained the highest performance (data not shown). Nevertheless, considering a simple model, one of the following features might be enough for a proper prediction.

Binary features, for example, one-hot-coding of the mutation type (missense, in-frame insertion, in-frame deletion, and frameshift mutations), did not show any importance within all the features explored and were not included in the final model. Likewise, the binary characteristics of the mutation, e.g. clonal or sub-clonal, which previously was shown to be important to raise a T cell response (44) did not show any importance to the IMPROVE model, likely due to the binary nature of such characteristic. Instead, the cellular prevalence given as a continuous variable was used to define clonal mutations and was included in the model instead. The cellular prevalence appeared to be the top four essential features of the model (**Figure 3B**), supporting that, indeed, mutational clonality does influence immunogenicity.

A critical issue of neoepitope analysis and prediction is the high uncertainty of the annotated immunogenicity labels. As illustrated in **Figure 3F**, the same pMHC complex can be immunogenic in one patient but negative when tested in others (sharing the same mutation and HLA). As most neopeptides are patient specific and hence are only experimentally assessed in one single experiment, this likely underestimates the true number of positive peptides. Drawing a parallel to the investigation of epitopes from infectious diseases, several studies have demonstrated such variable behavior in, for instance, influenza(45) and SARS-CoV-2 (46) infections, that can be explained by the phenomenon of immunodominance. And in this context, in the field of pathogen research, a peptide would not be labeled as certain negative, having only been tested in one individual. In the field of neopeptides, these are most often unique for each patient, and therefore, it is impossible to perform validation in multiple

patients. This place an upper bound to the reliability of the annotated immunogenicity target values of the neopeptide data and hence also a bound on the predicted accuracy of models developed for neoepitope prediction. The results and performance evaluations in the current (and earlier epitope immunogenicity assessment/prediction) work must be interpreted in the context of these limitations. In cancer patients, the TME is additionally critically influencing the ability to raise tumor-antigen specific T cell responses within a given patient (41). Our study addresses the importance of TME features in a combination of prioritizing neoantigen candidates and IMPROVE showed an increased performance by including these TME features in the prediction. Notably, the patient specific TME features add additional information about patients to the model and cannot be used alone, which also was shown by non-improved performance on the patient level from the original model to the model with the TME feature included.

We observed that the predicted neoepitope load from IMPROVE showed a better correlation to patients' clinical outcome (overall survival) than RankEL. This was further improved by the TME model and supports the improved capacity to identify immunologically relevant neoeptipes through this selected strategy. Also, in previous studies (Kristensen *et al*. and Holm *et al*.), the level of T cell recognition of neoantigens was associated with improved clinical outcome and survival(5,6) The mUC cohort showed a difference in survival, comparing the increase in number of neoepitopes specific T cells responses after three weeks of treatment, and the melanoma cohort found improved survival of patients with frequency and breadth of recognition of neoepitopes among the T cells in the infusion product. The NARTs frequency could gain an interesting aspect in regard to neoepitope prediction by ranking the neoepitopes by the frequency of the NART. However, due to the variability in samples used for screening (PBMC, TILs, and infusion product), this estimate was biased only due to the sample type. For example, the NARTS detected with either the infusion product or TILs was found with a higher frequency than the one from PBMCs. Therefore, we did not consider the frequency in the prediction strategy.

This study presents an improved model for the prediction of neoepitope immunogenicity. Such improvements are highly needed to advance the field of therapeutic neoepitope targeting. The IMPROVE model is built on a large dataset of neoepitopes evaluated for T cell recognition in cancer patients. The model improves the prediction of immunogenic neoepitopes and is capable of catching true neoepitopes in more than 2/3 of the patients only by taking the top 20 neoepitope candidates. Furthermore, we improved the model performance by considering patients' specific TME features. Which can help overcome some of the enormous patient variability which challenges the accuracy of neoepitope prediction. Even though the findings were based on broad-scale validated neopeptides, we will need verification from more data, especially more immunogenic neoepitopes, in order to improve the performance of our models. Nevertheless, our proposed models improved the ranking of neoepitope candidates, which is a critical task for immunotherapeutic targeting of neoepitopes, in e.g. personalized cancer vaccines.

Materials and methods

Patient selection

The cancer patients included in this study were enrolled in 3 different studies: a melanoma cohort that received TIL-ACT (5), a cohort of metastatic urothelial carcinoma (mUC) patients who received PD-L1 checkpoint inhibition (6), and a basket trial cohort with different cancer types and CPI treatments. Only patients with assessable patient material were included, which resulted in 70 patients in total.

Neoepitope prediction

The neoepitope prediction was performed as described in Kristensen N *et al*. and Holm *J et al*. (5,6). Shortly, paired tumor and normal WES and tumor RNAseq raw fastq files were pre-processed with trim-galore (47) version 0.4.0 to remove low-quality fragments and adapters. Burrows-Wheeler Aligner (48) version 0.7.15 was used to align the trimmed WES reads to the human reference genome (GRCh38) and MarkDuplicates from Picard-tools version 2.9.1 (49) were applied to tag duplicated reads due to technical artifacts. The quality scores were recalibrated with GATK, and the paired bam files were used as input to MuTect2 (50) from GATK (18) to detect somatic mutations. The bladder and melanoma cohort were processed following guidelines from GATK 3.8.0. GATK version 4.0.1 was used for the Basket cohort. Tumoral transcript abundance was quantified using Kallisto version 0.42.1 (51) from RNAseq. Patient-specific HLA genotype was determined using OptiType (52) version 1.2 on WES derived from normal samples. HLA-types, RNA expression, and somatic mutation VCF files were used as input to MuPeXI (9) version 1.1.2 to extract and rank neoepitope candidates. Posterior filtering of candidates was performed considering RNA expression values > 0.01 and a predicted binding Rank < 2 using NetMHCpan version 4.0 (53), except when patients had few peptide candidates, where the threshold was extended to 0.1 and RankEL < 0.5. If the expression level was insufficient to be obtained, they were reported as "-" in the MuPeXI output and were included in the peptide selection for the basket cohort. This cover in total 78 peptides from four patients.

Neopeptide experimental immunogenicity assessment

In total, 19.541 neoepitope candidates were screened for their potential to activate a T cell immune response with MHC multimer barcoding as described in (5,6,13). Predicted neopeptides and viral control peptides were synthesized by Pepscan (Pepscan Presto, The Netherlands). Peptides were dissolved to 10 mM in DMSO. Neopeptides were folded with the corresponding HLA and labeled with DNA barcodes, identifying all unique peptide-MHC (pMHC) multimers. Each patient-specific multimer library was screened with patient-derived samples, including peripheral blood mononuclear cells (PBMC) and tumor-infiltrating lymphocytes (TILs), and the melanoma cohort was also screened with the infusion product used for the TIL-ACT.

Feature calculation

Neopeptide features

The mutation consequence, as well as the sequences of mutant and wild-type peptides, were obtained with MuPeXI version 1.2 (9). The variant allele fraction (VarAlFrac) was obtained from mutect2, also given in the MuPeXI output as (Allele Frequency). The expression levels from the mutated transcript were calculated using Kallisto version 0.42.1 (51) was obtained from the MuPeXI output. No transcript was assessable for 78 neoepitopes belonging to 4

patients from the basket trial, and in these cases, the expression was obtained from Human Protein Atlas using NetMHCpanExp (54). The priority score (PrioScore) was also attained from the MuPeXI output. The likelihood of mutant and wild-type peptides binding to the patient's MHC was predicted with NetMHCpan 4.1 (53)

Both the eluted ligand % rank (RankEL) and binding affinity % rank (RankBA) predictions were retrieved. The differential agretopicity index (DAI) was calculated as the differences between the mutant and normal RankEL (Normal rank – Mutant Rank (22,23). The stability of the pMHC complex was predicted with NetMHCstabpan 1.0 (55). Additionally, NetMHCpanExp-1.0 (NetMHCExp) was employed to jointly evaluate the expression of a peptide and its likelihood of binding to its cognate MHC and generating in this way a combined prediction that takes into account both features (54). Since anchor residues are in contact with MHC, non-anchor residues are more exposed to interaction with TCRs. Therefore, the non-anchor subsequence was defined as the fourth to the penultimate residue among the predicted binding core with NetMHCpan-4.1.

The foreignness score (Foreignness) was measured as previously described in (27) and the function from antigen.garnish (26) was used to calculate the score. The similarity between the mutant peptide and the corresponding wild-type peptide or self-similarity (SelfSim) was calculated using the Kernel distance (56). Cellular prevalence (CelPrev) was calculated as previously described in (5) using Sequenza (57) version 3.0, Shixiang/copy-number (58) version 1.26.0, and PyClone (59) version 0.13.1. Transcript abundance was derived from RNAseq data using Kallisto version 0.42.1 (51). To validate the expression of mutated alleles in RNA, we mapped the RNAseq to the reference genome using STAR v2.5.3 (60), and then we retrieved all the bases with samtools mpileup (61) at the variant positions reported by MuTect2. The proportion of mutated transcripts among the covered transcripts (ValMutRNACoef) was assessed using the formula $Nreadssupporting\ variant/(Nreadsofcoverage + 100)$, We were not able to validate if the mutation was present in the RNA in 15% of the neoepiotpes due to inconsistency in mutation position in MuPeXI and mutect2.

Physiochemical descriptors of the neopeptide sequences were calculated. Molecular weight (mw), molar extinction coefficient, the relative frequency of F, W, and Y amino acids or aromaticity (Aro), instability index (Inst), and the relative frequency of V, I, Y, F, W, and L amino acids or helix (PropHydroAro) are calculated with ProteinAnalysis module from BioPython (62). The isoelectric point with EMBOSS was calculated with the Peptides package (63) in R. The mean hydrophobicity scale (64) and the proportion of different physicochemical classes of amino acids (small, aromatic, acidic, basic) were calculated for the non-anchor subsequence. The propensity for TCR recognition was calculated with PRIME (36).

Patient Features

The expression of MHC molecules (HLAexp) in the tumor cells was derived from RNAseq data. The Cytolytic Activity (CYT) is calculated as the geometric mean of GZMA and PRF1, as previously described in(40), and the expression of these genes was also derived from RNAseq data. The abundance of tumor-infiltrating immune and stromal cells was estimated with MCP-counter (11), using the RNAseq expression values obtained with Kallisto (51).

Models

Dataset

A dataset was assembled that contains the neopeptide sequences, the calculated features, and the immune response as the target value. For some neopeptides, it was not possible to obtain all the proposed features (there was no available RNAseq data for some patients, it was not

possible to calculate the cellular prevalence when tumor samples were derived as a cell lines, and MuPeXI does not report a normal peptide when a frameshift mutation has more than 4 mismatches compared to the most similar normal peptide, impeding the calculation of self-similarity). Therefore, we impute the missing values using the mean of the corresponding feature, except for the expression values in which we used NetMHCpanExp to retrieve expression values from the Human Protein Atlas (v. 20.0) database.

To avoid data leakage and overfitting of the models, a partitioning scheme was defined. We implemented a clustering algorithm to group neoepitopes based on i) shared subsequences or motifs between immunogenic neoepitopes and ii) the patient. In this way, all neoepitopes from the same patient and all similar neoepitopes are grouped together in the same partition. First, the immunogenic neoepitopes are grouped by the mentioned criteria, and then the non-immunogenic neopeptides are included in the defined partitions. If a peptide in the test-set existed in the training data, the peptide was deselected from the training. This only affects the negative peptides as the positive ones were separated by partition.

Random Forest

Random forest (RF) models were developed using the RandomForestClassifier module of scikit-learn (65)ackage in Python version 3.7.6. The hyperparameter max_depth was set to 6, n_estimator was set to 2000, min_sample_leaf was set to 6, and a nested cross-validation scheme was used. To avoid data imbalance, 500 negative data points (non-immunogenic peptide) were subsampled during training 50 times, and an ensemble score for the prediction was calculated.

NNAlign

The NNAlign version 2.1 (34)(34) method was used to train a neopeptide sequence-based model with a 5-fold with nested cross-validation, using the same partition as the RF model. The motif length for the alignment was set to 6, and the flanking region was not considered. Like the RF model, if a peptide from the test set, existed in training, the peptide was removed.

Survival analysis

To predict survival probability with the effect of immunotherapy, based upon the developed model from the study. The developed RF models were saved with "pickle dump" from pandas, and an ensemble score for each model was used as the immunogenicity score. The immunogenicity score was predicted for the remaining neopeptides with RankEL< 2 and expression > 0.01, which were not included in the study. This included, for example, peptides that have been deselected for patients with many predicted neopeptides and therefore have stricter criteria during the neopeptide selection. Additionally, some HLA alleles were not assessable, and neopeptides with these alleles were also selected for immunogenicity screening. Patient RH-08 was excluded as the sample was taken after treatment, and MM-24 was excluded as the sample was from cell-line. The prediction score found for the remaining neopeptides was added on top of the peptides that were included in the study. This covers the dataset used for the survival analysis.

The separation of the four groups is based on the quantile of the variable. Patients with predicted neoepitopes above or equal to the 3rd (75%) quantile were defined as "high". Patients with peptides between the 2nd (50%) and 3rd quantile were defined as "medium high", and patients with peptides between the 1st (25%) and 2nd were defined as "medium low". Lastly, patients below or equal to the 1st quantile were defined as "low". The hazard score and p-value were calculated based on the comparison of the high vs. low group.

Statistics

All the statistics were calculated in R (66) version 4.1.1. The Wilcoxon rank-sum test was applied to analyze the features. Spearman's correlation coefficients were used to measure the correlation between variables. To assess the performance of the models, the AUC and AUC 0.1 were calculated with the function in ROCR version 1.0.11 (67), and the difference between ROC curves was computed with roc.test with default options from pROC version 1.18.0 (68). Kaplan-Meyer curves were created with the survminer package version 0.4.9 (69), and the hazard ratios were calculated with Cox proportional hazards regression model using the survival package version 3.3.1 (70).

Author contribution

AB: Calculated features, made RF model, discussed data, created figures, and wrote the manuscript. I**B:** Calculated features, discussed data, and wrote the manuscript. **BR:** Calculated features, made partitioning for data, discussed data, and wrote the manuscript. **HMGA:** Calculated features and discussed data. **KKM:** Assisted in data analysis. **AM:** assisted in ML models. **NPK, SAT, and CH:** screened neopeptides for the melanoma cohort. **JSH:** screened neopeptides for the mUC cohort. **KHM and UKH:** screened neopeptides for the Basket cohort. **ALS and FOB:** investigated mutation call quality and revised the manuscript. **CB**: Provided technical guidance. **MN and SRH:** conceived the concept, discussed data, supervised during the study, and wrote the manuscript.

Acknowledgment

We would like to thank all the medical doctors who have included the patients and taken the samples from the patients. This included Vinicius Araújo Barbosa de lima and Ulrik Lassen, who have been responsible for taking the samples from the Basket trail, Samuel Funt, who has taken the samples from the mUC cohort, and people from CCIT for being responsible for taking the samples from the melanoma cohort. Additionally, we would like to thank Tripti Tamhane, who has produced the monomers used for screening. Finally, we would like to thank all the patients for participating in the trials.

References

1. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. Vol. 17, The Lancet Oncology. 2016.
2. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, et al. Companion Diagnostic, Pharmacogenomic, and Cancer Biomarkers Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. Mol Cancer Ther [Internet]. 2017 [cited 2018 Dec 21];16(11). Available from: www.aacrjournals.org
3. van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, et al. Erratum: Genomic correlates of response to CTLA-4 blockade in metastatic melanoma (Science (2015) 350:6257 (207-211)). Vol. 350, Science. 2015.
4. Linette GP, Carreno BM. Neoantigen Vaccines Pass the Immunogenicity Test. Vol. 23, Trends in Molecular Medicine. 2017.
5. Kristensen NP, Heeke C, Tvingsholm SA, Borch A, Draghi A, Crowther MD, et al. Neoantigen-reactive CD8+ T cells affect clinical outcome of adoptive cell therapy with tumor-infiltrating lymphocytes in melanoma. Journal of Clinical Investigation. 2022;132(2).
6. Holm JS, Funt SA, Borch A, Munk KK, Bjerregaard AM, Reading JL, et al. Neoantigen-specific CD8 T cell responses in the peripheral blood following PD-L1 blockade might predict therapy outcome in metastatic urothelial carcinoma. Nature Communications 2022 13:1 [Internet]. 2022 Apr 11 [cited 2022 Sep 23];13(1):1–17. Available from: https://www.nature.com/articles/s41467-022-29342-0
7. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science (1979). 2015;348(6230):69–74.
8. Hundal J, Carreno BM, Petti AA, Linette GP, Griffith OL, Mardis ER, et al. pVAC-Seq: A genome-guided in silico approach to identifying tumor neoantigens. Genome Med. 2016;8(1).
9. Bjerregaard AM, Nielsen M, Hadrup SR, Szallasi Z, Eklund AC. MuPeXI: prediction of neo-epitopes from tumor sequencing data. Cancer Immunology, Immunotherapy [Internet]. 2017 Sep 20 [cited 2018 Dec 28];66(9):1123–30. Available from: http://link.springer.com/10.1007/s00262-017-2001-3
10. Tappeiner E, Finotello F, Charoentong P, Mayer C, Rieder D, Trajanoski Z. TIminer: NGS data mining pipeline for cancer immunology and immunotherapy. Bioinformatics. 2017;33(19).
11. Kim S, Kim HS, Kim E, Lee MG, Shin EC, Paik S, et al. Neopepsee: Accurate genome-level prediction of neoantigens by harnessing sequence and amino acid immunogenicity information. Annals of Oncology. 2018;29(4).
12. Zhou C, Wei Z, Zhang Z, Zhang B, Zhu C, Chen K, et al. PTuneos: Prioritizing tumor neoantigens from next-generation sequencing data. Genome Med. 2019;11(1).
13. Bentzen AK, Hadrup SR. Evolution of MHC-based technologies used for detection of antigen-responsive T cells. Vol. 66, Cancer Immunology, Immunotherapy. Springer Science and Business Media Deutschland GmbH; 2017. p. 657–66.
14. Vesely MD, Schreiber RD. Cancer Immunoediting: antigens, mechanisms and implications to cancer immunotherapy. Ann N Y Acad Sci [Internet]. 2013 [cited 2022 Aug 25];1284(1):1. Available from: /pmc/articles/PMC3648872/
15. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol [Internet]. 2016 Oct 20 [cited 2022 Aug

19];17(1):1–20. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1070-5
16. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. In: Methods in Molecular Biology. 2018.
17. Keith Henry Moss UKHVAB de LESMABAMBOOAKBAMMIMSULSRH. Characteristics of neoantigen-reactive CD8+ T cells as an immune predictor for clinical outcome to checkpoint blockade therapy in a pan-cancer setting. Manuscript in preperation. 2022;
18. van der Auwera GA, O'Connor BD. Genomics in the cloud : using Docker, GATK, and WDL in Terra. O'Reilly Media. 2020.
19. Xia H, Richters M, Ramirez C, Saus CP, Cotto K, Dunn G, et al. 65. Accurate neoantigen prediction depends on mutation position relative to patient-specific MHC anchor locations. Cancer Genet. 2020;244.
20. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: Application to the MHC class i system. Bioinformatics. 2016;32(4).
21. Bjerregaard AM, Nielsen M, Jurtz V, Barra CM, Hadrup SR, Szallasi Z, et al. An analysis of natural T cell responses to predicted tumor neoepitopes. Front Immunol. 2017;8(NOV):1–9.
22. Coelho ACMF, Fonseca AL, Martins DL, Lins PBR, da Cunha LM, de Souza SJ. NeoANT-HILL: An integrated tool for identification of potential neoantigens. BMC Med Genomics. 2020;13(1).
23. Case KO, Kennedy KF, Kujtan L, Subramanian J. The role of tumour neoantigens in the differential response to immunotherapy (IO) in EGFR and BRAF mutated lung cancers: Quantity or quality? Annals of Oncology. 2019;30.
24. Ghorani E, Rosenthal R, McGranahan N, Reading JL, Lynch M, Peggs KS, et al. Differential binding affinity of mutated peptides for MHC class I is a predictor of survival in advanced lung cancer and melanoma. Annals of Oncology. 2018;29(1).
25. Wells DK, van Buuren MM, Dang KK, Hubbard-Lucey VM, Sheehan KCF, Campbell KM, et al. Key Parameters of Tumor Epitope Immunogenicity Revealed Through a Consortium Approach Improve Neoantigen Prediction. Cell. 2020 Oct 29;183(3):818-834.e13.
26. Richman LP, Vonderheide RH, Rech AJ. Neoantigen Dissimilarity to the Self-Proteome Predicts Immunogenicity and Response to Immune Checkpoint Blockade. Cell Syst. 2019 Oct 23;9(4):375-382.e4.
27. Luksza M, Riaz N, Makarov V, Balachandran VP, Hellmann MD, Solovyov A, et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. Nature. 2017;551(7681).
28. Karimnezhad A, Palidwor GA, Thavorn K, Stewart DJ, Campbell PA, Lo B, et al. Accuracy and reproducibility of somatic point mutation calling in clinical-type targeted sequencing data. BMC Med Genomics. 2020;13(1).
29. Pei S, Liu T, Ren X, Li W, Chen C, Xie Z. Benchmarking variant callers in next-generation and third-generation sequencing analysis. Vol. 22, Briefings in Bioinformatics. 2021.
30. Chen Z, Yuan Y, Chen X, Chen J, Lin S, Li X, et al. Systematic comparison of somatic variant calling performance among different sequencing depth and mutation frequency. Sci Rep. 2020;10(1).
31. Chowell D, Krishna S, Becker PD, Cocita C, Shu J, Tan X, et al. TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. Proc Natl Acad Sci U S A. 2015;112(14).

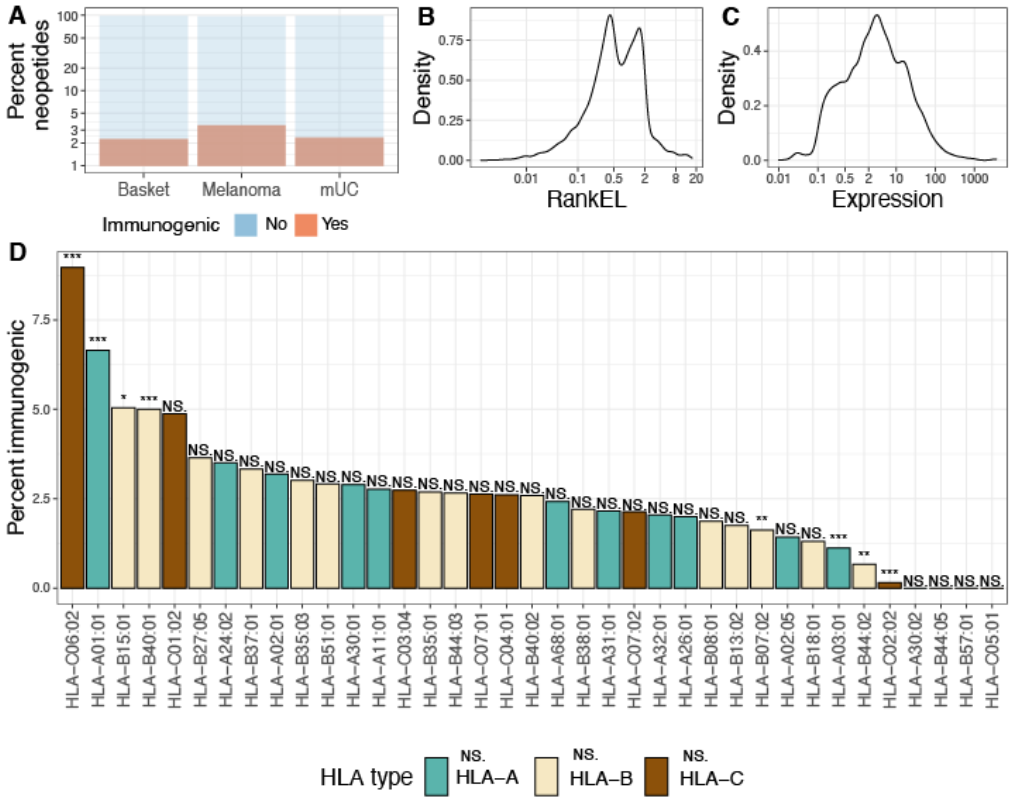
32. Gregorutti B, Michel B, Saint-Pierre P. Correlation and variable importance in random forests. Stat Comput [Internet]. 2017 May 1 [cited 2022 Aug 19];27(3):659–78. Available from: https://link.springer.com/article/10.1007/s11222-016-9646-1
33. Darst BF, Malecki KC, Engelman CD. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. BMC Genet [Internet]. 2018 Sep 17 [cited 2022 Aug 19];19(1):1–6. Available from: https://bmcgenomdata.biomedcentral.com/articles/10.1186/s12863-018-0633-8
34. Nielsen M, Andreatta M. NNAlign: A platform to construct and evaluate artificial neural network models of receptor-ligand interactions. Nucleic Acids Res. 2017;45(W1).
35. Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, de Silva AD, Sette A, et al. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. PLoS Comput Biol. 2013;9(10).
36. Schmidt J, Smith AR, Magnin M, Racle J, Devlin JR, Bobisse S, et al. Prediction of neo-epitope immunogenicity reveals TCR recognition determinants and provides insight into immunoediting. Cell Rep Med. 2021;2(2).
37. Lauss M, Donia M, Harbst K, Andersen R, Mitra S, Rosengren F, et al. Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. Nat Commun. 2017;8(1):1–11.
38. Litchfield K, Reading JL, Puttick C, Thakkar K, Abbosh C, Bentham R, et al. Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. Cell. 2021;184(3).
39. Chowell D, Yoo SK, Valero C, Pastore A, Krishna C, Lee M, et al. Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. Nat Biotechnol. 2022;40(4).
40. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. Cell [Internet]. 2015 Jan 15 [cited 2018 Dec 26];160(1–2):48–61. Available from: https://www.sciencedirect.com/science/article/pii/S0092867414016390
41. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell. 2017;171(6):1259-1271.e11.
42. Poggi A, Zocchi MR. Mechanisms of tumor escape: Role of tumor microenvironment in inducing apoptosis of cytolytic effector cells. Vol. 54, Archivum Immunologiae et Therapiae Experimentalis. 2006.
43. Schaafsma E, Fugle CM, Wang X, Cheng C. Pan-cancer association of HLA gene expression with cancer prognosis and immunotherapy efficacy. Br J Cancer. 2021;125(3).
44. McGranahan N, Furness AJS, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science (1979). 2016;351(6280):1463–9.
45. Wu C, Zanker D, Valkenburg S, Tan B, Kedzierska K, Zou QM, et al. Systematic identification of immunodominant CD8+ T-cell responses to influenza A virus in HLA-A2 individuals. Proc Natl Acad Sci U S A. 2011;108(22).
46. Saini SK, Hersby DS, Tamhane T, Povlsen HR, Amaya Hernandez SP, Nielsen M, et al. SARS-CoV-2 genome-wide T cell epitope mapping reveals immunodominance and substantial CD8+ T cell activation in COVID-19 patients. Sci Immunol. 2021;6(58).
47. Felix Krueger. TrimGalore https://github.com/FelixKrueger/TrimGalore. 2021.
48. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009 Jul;25(14):1754–60.

49. BroadInstitue. Picard https://github.com/broadinstitute/picard. BroadInstitue; 2021.
50. Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C, Lichtenstein L. Calling Somatic SNVs and Indels with Mutect2. bioRxiv. 2019;
51. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol [Internet]. 2016 May 4 [cited 2019 Jan 2];34(5):525–7. Available from: http://www.nature.com/articles/nbt.3519
52. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: Precision HLA typing from next-generation sequencing data. Bioinformatics. 2014;30(23):3310–6.
53. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. The Journal of Immunology. 2017;199(9).
54. Garcia Alvarez HM, Koşaloğlu-Yalçın Z, Peters B, Nielsen M. The role of antigen expression in shaping the repertoire of HLA presented ligands. iScience [Internet]. 2022 Sep 16 [cited 2022 Nov 19];25(9). Available from: http://www.cell.com/article/S2589004222012470/fulltext
55. Rasmussen M, Fenoy E, Harndahl M, Kristensen AB, Nielsen IK, Nielsen M, et al. Pan-Specific Prediction of Peptide–MHC Class I Complex Stability, a Correlate of T Cell Immunogenicity. The Journal of Immunology. 2016;197(4).
56. Shen WJ, Wong HS, Xiao QW, Guo X, Smale S. Towards a Mathematical Foundation of Immunology and Amino Acid Chains. 2012 May 28 [cited 2022 Nov 19]; Available from: https://arxiv.org/abs/1205.6031v2
57. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, et al. Sequenza: Allele-specific copy number and mutation profiles from tumor sequencing data. Annals of Oncology. 2015;26(1).
58. Wang Shixiang. Wang S. Copynumber. https://github.com/ShixiangWang/copynumber; 2020.
59. Roth A, Khattra J, Yap D, Wan A, Laks E, Biele J, et al. PyClone: Statistical inference of clonal population structure in cancer. Nat Methods [Internet]. 2014 [cited 2022 Aug 25];11(4):396. Available from: /pmc/articles/PMC4864026/
60. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1).
61. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2).
62. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25(11).
63. Osorio D, Rondón-Villarreal P, Torres R. Peptides: A package for data mining of antimicrobial peptides. R Journal. 2015;7(1).
64. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1).
65. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research. 2011;12.
66. Team RC. R: A language and environment for statistical computing v. 3.6. 1 (R Foundation for Statistical Computing, Vienna, Austria, 2019). Vol. 11, Scientific Reports. 2021.
67. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: Visualizing classifier performance in R. Bioinformatics. 2005;21(20).

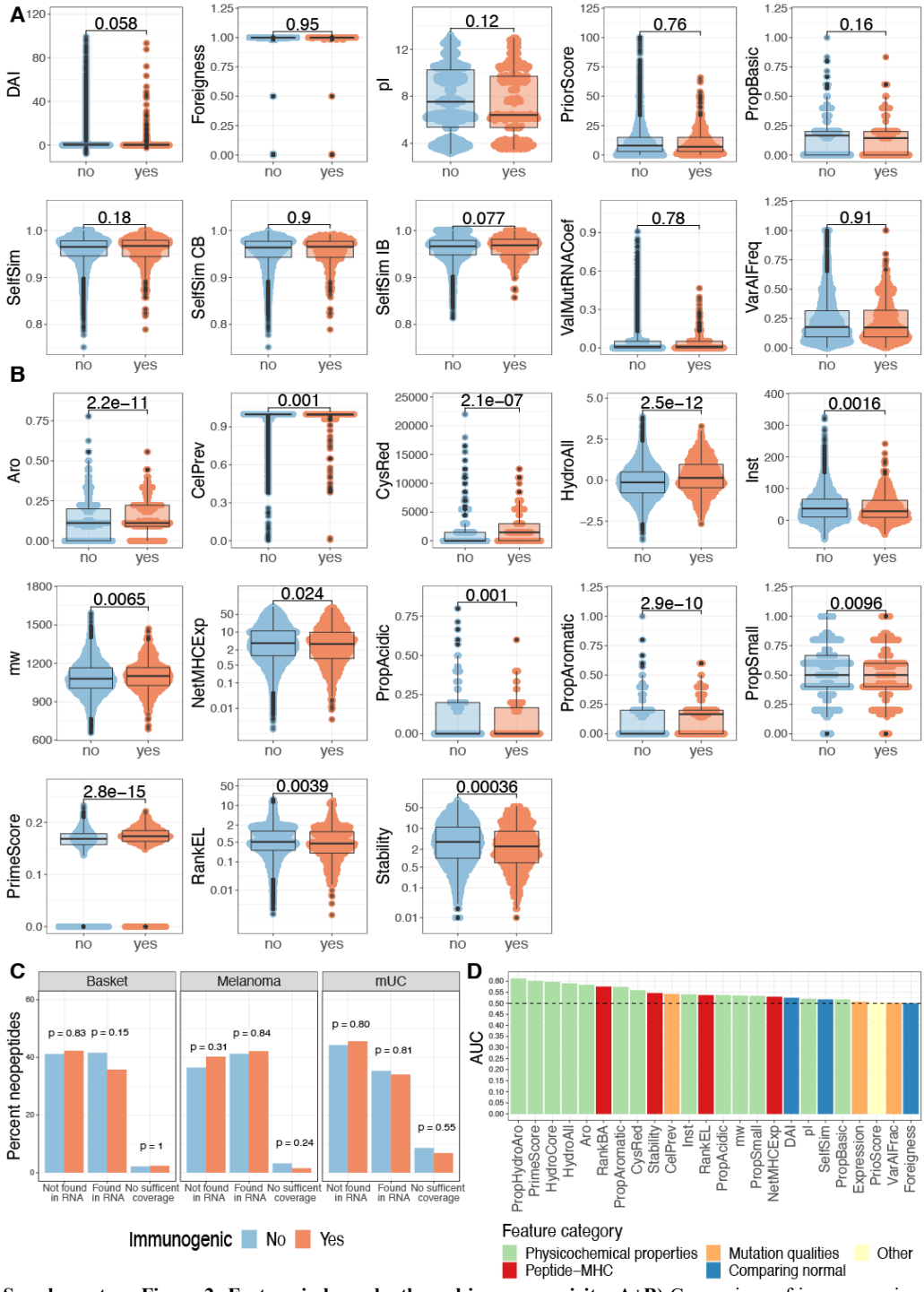
68. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12.
69. Alboukadel Kassambara MK and PB. survminer: Drawing Survival Curves using "ggplot2". R package version 0.4.9. https://CRAN.R-project.org/package=survminer. 2021.
70. Borgan Ø. Modeling Survival Data: Extending the Cox Model. Terry M. Therneau and Patricia M. Grambsch, Springer-Verlag, New York, 2000. No. of pages: xiii + 350. Price: $69.95. ISBN 0-387-98784-3. Stat Med. 2001;20(13).

Supplementary Table 1: Neopeptides features. The abbreviation used in the paper and the figures, the full name of the feature, and a description of what the feature is and how it is obtained.

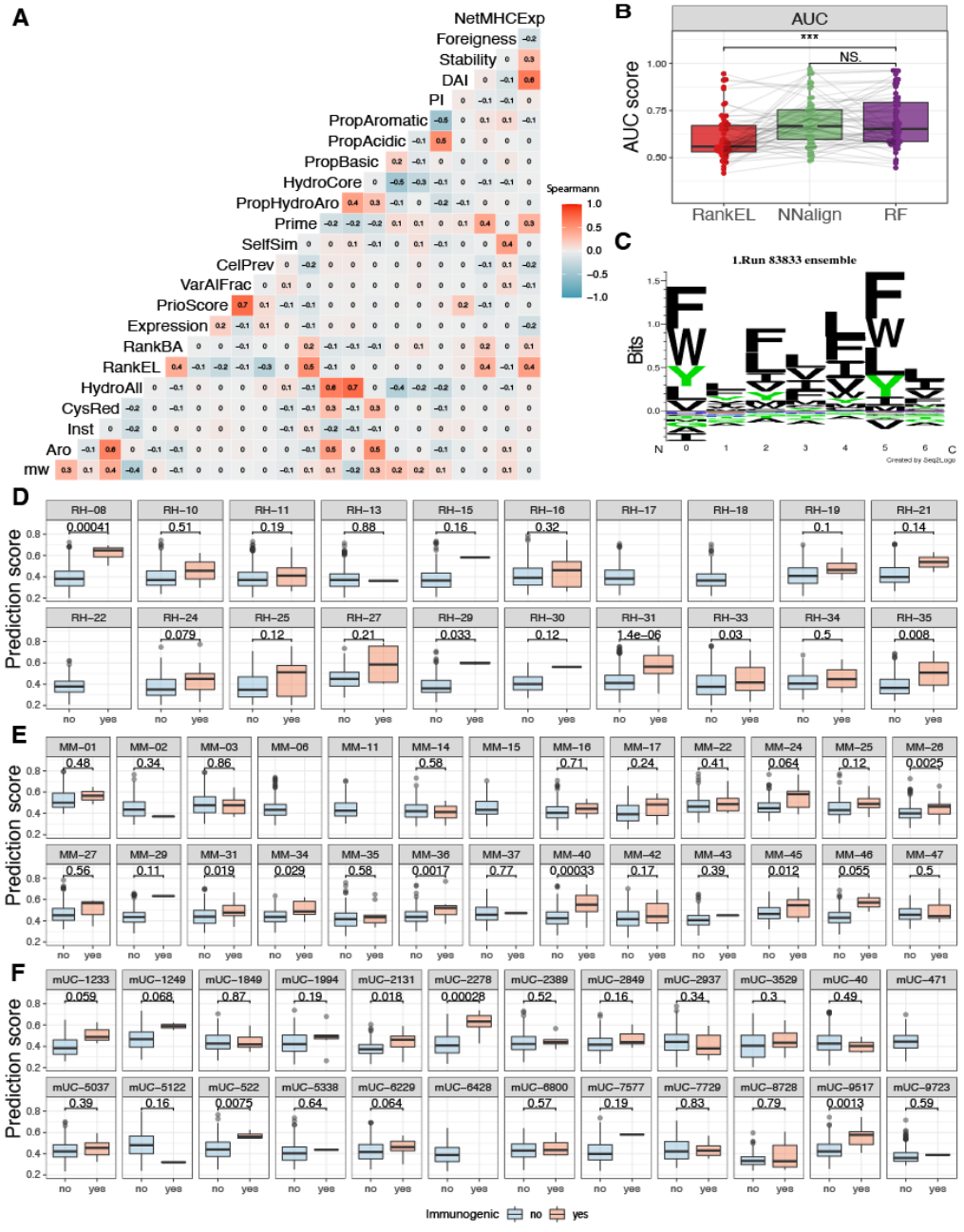
Abbreviation	Name	Description
mw	Molecular Weight	Molecular Weight of the neopeptide, from biopython (1)
aro	Aromaticity	Aromaticity according to Lobry, 1994. from biopython (1)
inst	Instabillity	Instability according to Guruprasad et al 1990 from biopython (1)
CysRed	Cysteine residues	Molar extinction coefficient of Cysteine residues(1)
RankEL	Eluted Ligand %Rank	Eluted Ligand %Rank of mutated peptide as predicted by NetMHCpan 4 .1 (2)
RankBA	Binding Affinity %Rank	Binding Affinity of mutated peptide as predicted by NetMHCpan 4.1 (2)
Expression	Expression Level	Expression level of transcript containing the mutation obtained from Kallisto (3)
NetMHCExp	NetMHCExppan	Calculated using NetMHCexppan(4)
PrioScore	Priority Score from MuPeXI	Neoepitope prioritization score as computed by MuPeXI (5)
VarAlFreq	Allele Frequency	Frequency of the allele from mutect2 (6)
CelPrev	Cellular Prevalence	The prevalence of the tumor cell from PyClone (7)
SelfSim	Self Similarity	Self-similarity to the corresponding Normal peptide calculated using kernel distance (8)
Prime	Score PRIME	Score from the tool PRIME (9)
PropHydroAro	Proportion of hydrophobic and aromatic residues	The proportion of: V, I, Y, F, W, L. given as Helix in biopython packages. (1)
HydroCore	Mean Hydrophobicity in Core without Anchor residues	Mean Hydrophobicity scale value of non anchor residues of binding core. Scale from Kyte and Doolittle, 1982.(10)
HydroAll	Mean hydrophobicity	The mean hydrophobicity of the entire peptides. Scale from Kyte and Doolittle, 1982. (10)
PropSmall	Proportion of Small amino acids in non anchor positions of binding core	
PropAro	Proportion of Small amino acids in non anchor positions of binding core	The proportion the following residues in the non anchor positions of the binding core: F, H, W, Y
PropBasic	Proportion of Small amino acids in non anchor positions of binding core	The proportion the following residues in the non anchor positions of the binding core: H, K, R
PropAcidic	Proportion of Small amino acids in non anchor positions of binding core	The proportion the following residues in the non anchor positions of the binding core: D, E
pI	Isoelectric point	Isoelectric Point computed with EMBOSS(11)
DAI	Differential Agretopicity Index	Diffrence between Normal EL rank and Mutant EL rank. Calculated by, DAI = Normal EL rank - Mutant EL rank. (12,13)
Stability	Stability score	MHC-peptide stability score, as predicted by NetMHCStabPan-1.0 (14)
Foreignness	Foreignness Score	Function from antigen garnish original from Luksza M, et al.(12,13,15,16)
ValMutRNACoef	RNA coefficient	Aligned using STAR aligner (17)
HLAexp	HLA expression	Expression from the corresponding HLA in RNA sequencing obtained from kallisto (3)
CYT	Cytolytic activity	Geometric mean of PRF1 and GZMA as presented in Rooney 2015 (18)
T cells	T cells	MCP-counter score for T cells (19)
TcellsCD8	CD8 T cells	MCP-counter score for CD8 T cells (19)
CytoxLympho	Cytotoxic lymphocytes	MCP-counter score for Cytotoxic lymphocytes (19)
Blinage	B lineage	MCP-counter score for B lineage (19)
NKcells	NK cells	MCP-counter score for NK cells (19)
Monocytes	Monocytic lineage	MCP-counter score for Monocytic lineage (19)
MyeloidDC	Myeloid dendritic cells	MCP-counter score for Myeloid dendritic cells (19)
Neutrophils	Neutrophils	MCP-counter score for Neutrophils (19)
Endothelial	Endothelial cells	MCP-counter score for Endothelial cells (19)
Fibroblast	Fibroblasts	MCP-counter score for Fibroblasts (19)



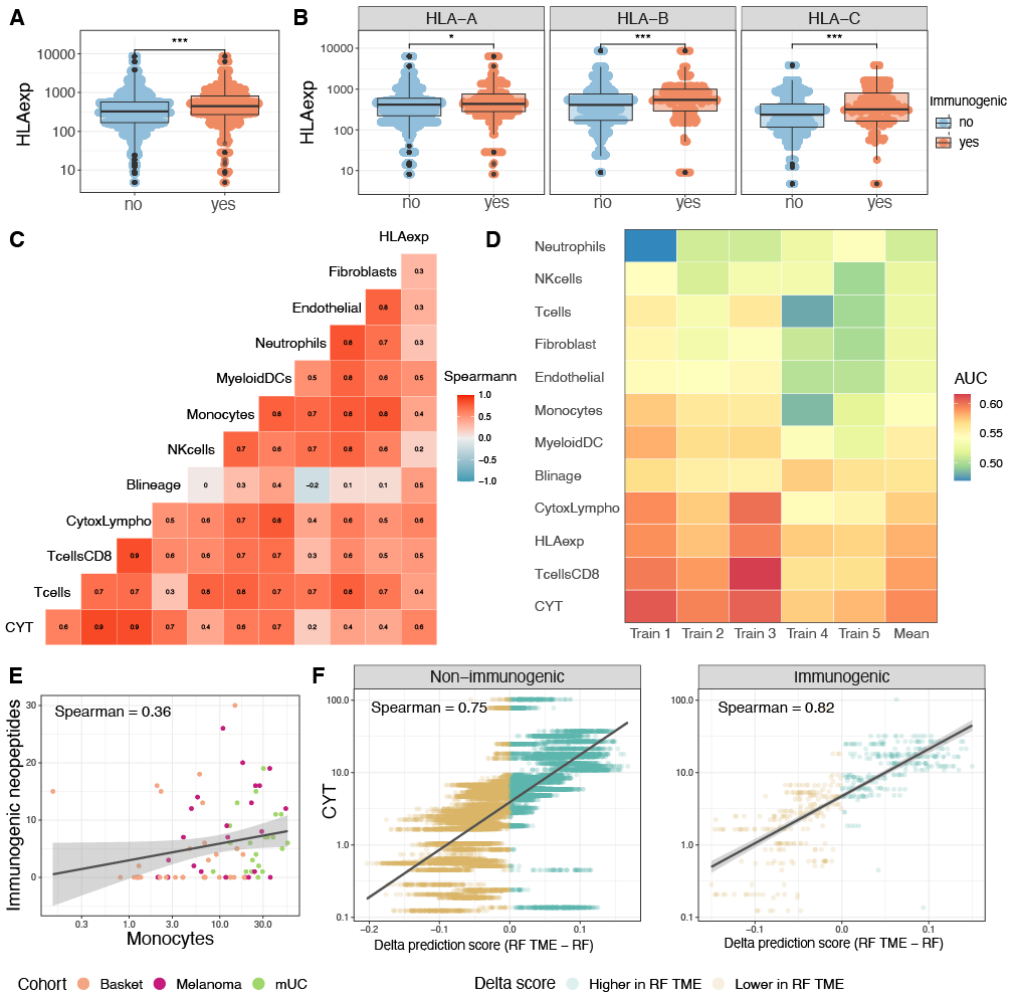
Supplementary Figure 1: Data overview. A) The percent of immunogenic and non-immunogenic neoepitopes within the three cohorts. **B)** Density of Eluted Ligand % Rank (RankEL) for all neoepitope candidates. **C)** Density of Expression level (Expression) from the corresponding genes for neoepitope candidates. **D)** The percent of immunogenic neoepitopes for each HLA allele colored by the HLA class test made by proportion z-test. p values < 0.001 = ***, p values < 0.01 = ** p values < 0.05 = * p-values > 0.5 = NS.



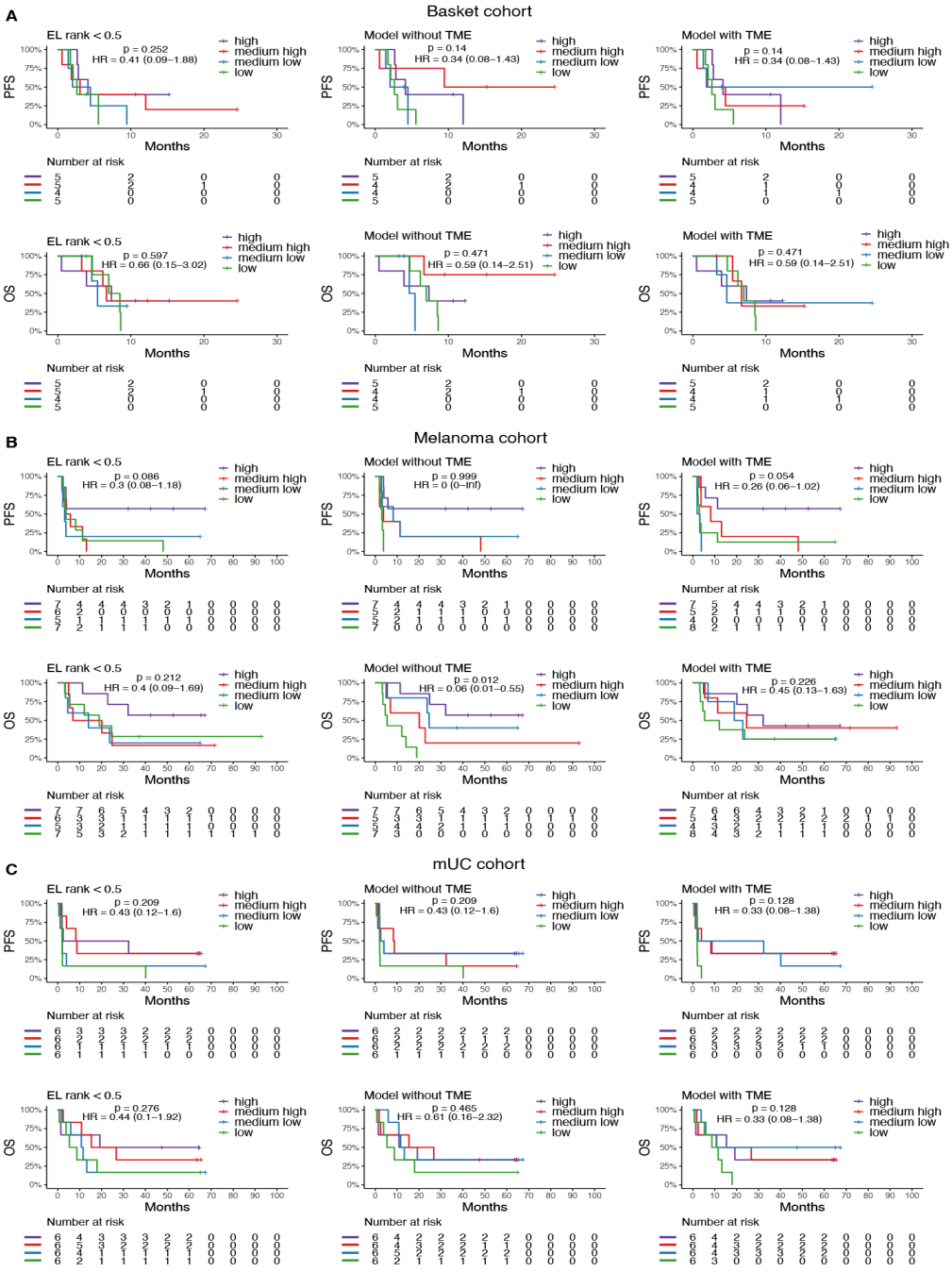
Supplementary Figure 2: Feature independently and immunogenicity. A+B) Comparison of immunogenic and non-immunogenic neoepitopes according to the values of each feature. Statistic made with Wilcox test with p-value adjustment by Bonferroni. **A)** Non-significant features, **B)** Significant features. **C)** Validation of mutations in RNAseq for each cohort no significant difference according to a proportion test. **D)** AUC for all features individually colored by feature type.



Supplementary Figure 3: RF modeling. A) Spearman correlation comparing all features. **B)** AUC per patient where RF obtained a significant improvement compared to RankEL (p-val = 0.00028), and no significant difference in Rf compared to NNalign (p = 0.89). **C)** NNalign logo plot from 5-fold cross-validation. b-**C-E)** Comparison of the prediction score from the random forest model with the immunogenic and non-immunogenic neoepitopes for each patient in each cohort. p values calculated with the non-paired Wilcoxon test. **D)** Basket cohort, **E)** Melanoma cohort, **F)** mUC cohort.



Supplementary Figure 4, random Forrest with Tumor Microenvironment (TME) features. A) HLA expression comparing immunogenic and non-immunogenic neoepitopes ($p = 2.4^{-10}$, Wilcox test). **B)** HLA expression per HLA class comparing immunogenic and non-immunogenic neoepitopes. HLA-A p= 0.02, HLA-B $p= 3.9^{-6}$, HLA-C $p= 4.6^{-6}$ (all with Wilcox test) **C)** Spearman correlation of TME feature. **D)** Performance for each feature for each partitioning and the mean across all partitions. **E-G)** Spearman correlation figure of the number of responses compared to three TME features. **E)** Monocyte linage (Monocytes). **F)** Cytolytic Activity (CYT) on the y-axis and the delta prediction score calculated by the RF TME minus the RF score separated into immunogenic and non-immunogenic neopeptides. Correlation performed with spearman correlation.



Supplementary figure 5, Survival curves per cohort. **A-C)** Kaplan-Meyer curves for the three categories (left) RankEL, middle RF model without TME features, and (right) the RF model including the TME features. Where patients are separated into four groups according to the number of predicted neoepitopes above a defined threshold, the four groups are defined according to the quantile where "high" are above the 3rd quantile, "medium high" is between the 2nd and 3rd quantile. "medium low" is between the 2nd and 1st quantile, and low is below the 1st quantile. **A)** Basket cohort, **B)** melanoma cohort, **C)** mUC cohort.

References

1. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics. 2009;25(11).
2. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. The Journal of Immunology. 2017;199(9).
3. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol [Internet]. 2016 May 4 [cited 2019 Jan 2];34(5):525–7. Available from: http://www.nature.com/articles/nbt.3519
4. Garcia Alvarez HM, Koşaloğlu-Yalçın Z, Peters B, Nielsen M. The role of antigen expression in shaping the repertoire of HLA presented ligands. iScience [Internet]. 2022 Sep 16 [cited 2022 Nov 19];25(9). Available from: http://www.cell.com/article/S2589004222012470/fulltext
5. Bjerregaard AM, Nielsen M, Hadrup SR, Szallasi Z, Eklund AC. MuPeXI: prediction of neo-epitopes from tumor sequencing data. Cancer Immunology, Immunotherapy [Internet]. 2017 Sep 20 [cited 2018 Dec 28];66(9):1123–30. Available from: http://link.springer.com/10.1007/s00262-017-2001-3
6. Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C, Lichtenstein L. Calling Somatic SNVs and Indels with Mutect2. bioRxiv. 2019;
7. Roth A, Khattra J, Yap D, Wan A, Laks E, Biele J, et al. PyClone: Statistical inference of clonal population structure in cancer. Nat Methods [Internet]. 2014 [cited 2022 Aug 25];11(4):396. Available from: /pmc/articles/PMC4864026/
8. Shen WJ, Wong HS, Xiao QW, Guo X, Smale S. Towards a Mathematical Foundation of Immunology and Amino Acid Chains. 2012 May 28 [cited 2022 Nov 19]; Available from: https://arxiv.org/abs/1205.6031v2
9. Schmidt J, Smith AR, Magnin M, Racle J, Devlin JR, Bobisse S, et al. Prediction of neo-epitope immunogenicity reveals TCR recognition determinants and provides insight into immunoediting. Cell Rep Med. 2021;2(2).
10. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1).
11. Osorio D, Rondón-Villarreal P, Torres R. Peptides: A package for data mining of antimicrobial peptides. R Journal. 2015;7(1).
12. Coelho ACMF, Fonseca AL, Martins DL, Lins PBR, da Cunha LM, de Souza SJ. NeoANT-HILL: An integrated tool for identification of potential neoantigens. BMC Med Genomics. 2020;13(1).
13. Case KO, Kennedy KF, Kujtan L, Subramanian J. The role of tumour neoantigens in the differential response to immunotherapy (IO) in EGFR and BRAF mutated lung cancers: Quantity or quality? Annals of Oncology. 2019;30.
14. Rasmussen M, Fenoy E, Harndahl M, Kristensen AB, Nielsen IK, Nielsen M, et al. Pan-Specific Prediction of Peptide–MHC Class I Complex Stability, a Correlate of T Cell Immunogenicity. The Journal of Immunology. 2016;197(4).
15. Luksza M, Riaz N, Makarov V, Balachandran VP, Hellmann MD, Solovyov A, et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. Nature. 2017;551(7681).

16. Richman LP, Vonderheide RH, Rech AJ. Neoantigen Dissimilarity to the Self-Proteome Predicts Immunogenicity and Response to Immune Checkpoint Blockade. Cell Syst. 2019 Oct 23;9(4):375-382.e4.
17. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1).
18. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. Cell [Internet]. 2015 Jan 15 [cited 2018 Dec 26];160(1–2):48–61. Available from: https://www.sciencedirect.com/science/article/pii/S0092867414016390
19. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol [Internet]. 2016 Oct 20 [cited 2022 Aug 19];17(1):1–20. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1070-5

7 EPILOGUE

The Manuscripts in this PhD contribute to novel findings in the field of immune signatures and target discovery in human tumors. All Manuscripts address topics within the field of immunotherapy and potential targets for $CD8^+$ T cells. Manuscript I contributes to findings regarding predictive biomarkers to CPI treatment and show the importance of using a combination of biomarkers rather than a single biomarker. Manuscript II and III validate a large pool of predicted neoepitopes to find the ones that could raise reactivity from NARTs and hence be classified as immunogenic neoepitopes. We also studied the correlation of TME signatures in patients with a high frequency of detected NARTs and addressed the ability of using NARTs as a biomarker to predict survival outcomes. Additionally, we looked at neoepitope features to distinguish the immunogenic neoepitopes from the non-immunogenic ones. The final Manuscript (Manuscript VI) introduces an excellent bridge between the three first Manuscripts (Manuscript I, II, and III) and uses the knowledge gained from these Manuscripts to characterize neoepitope immunogenicity. In this Manuscript, we investigated, as far as we know, the largest pool of validated neoepitope candidates and examined features that can be used to identify the immunogenic neoepitopes using machine learning. The Manuscript also introduces how the TME, which also has been considered in Manuscript I,II, and III can affect neoepitope recognition. The following section will discuss concerns and perspectives drawn from the four Manuscripts and their interconnection.

7.1 PREDICTIVE BIOMARKER

When evaluating cancer patients and their treatment options, it is critical to know whether a treatment can actually benefit the patient. Through the recent years of success with CPI therapy, it has become a first-line treatment for some patient groups (Jamy and Sonpavde, 2017; Peters et al., 2022). However, the response rate of CPI treatment reaches 50 % and leaves the remaining half of the patients non-responding (Chowell et al., 2022; Yarchoan et al., 2017). Manuscript I evaluates different biomarkers individually and in various combinations to achieve the best possible biomarker describing benefit from CPI treatment for a diverse cohort. These biomarkers are also validated in the mUC

cohort from Manuscript II. This study suggests biomarkers which have already shown great predictive effects in previous studies, including; neoepitope load, CYT, T cell diversity. Additionally, we observed that PD-L2 expression had a positive correlation with better survival probability. PD-L2 can, in the same way as PD-L1, bind to PD-1 and inhibit T cell profiliration and cytokine production (Solinas et al., 2020) A study also suggested PD-L2 as a biomarker for patients treated with anti-PD-1 CPI (Takamori et al., 2018). Interestingly, we observed the correlation between PD-L2 expression and better survival independent of treatment, also within the anti-PD-L1 treated patients. Besides that, PD-L2 can also be an indicator of immune activity and an inflamed TME (Wang et al., 2022). Obeid *et al.* showed an association between PD-L2 expression and higher frequency of tumor-infiltrating T cells (Obeid et al., 2016). Similar to our findings, they and Yearley *et al.* showed a positive correlation between PD-L2 expression and better survival probability. They also observed that some tumor types with the absence of PD-L1 expression expressed PD-L2 and by observing both PD-L1 and PD-L2 added more predictive power in predicting survival probability than observing PD-L1 expression alone (Yearley et al., 2017). All these studies and our results suggest that we should draw more attention to PD-L2 and its importance in the interaction of immune cells with cancerous cells. In addition, PD-L2 can be considered as a target for immunotherapy and the possible addition to other CPI therapies.

The number of suggested biomarkers is enormous (Bai et al., 2020; Filipovic et al., 2020; Rooney et al., 2015). Recently, the clonal TMB and LOHHLA have been suggested to be more predictive than TMB alone across various cohorts (Litchfield et al., 2021). These parameters have not been considered in this Manuscript and would be interesting to be implemented in the future.

An important keynote from Manuscript I is that a combination of biomarkers is a better predictor than examining a single biomarker. In summary, first, we observed biomarker signatures on a continuous scale and observed patient outcomes on a binary scale by comparing progressive patients with non-progressive patients. Secondly, we implemented cox regression analysis to study the suggested biomarkers from the first analysis and investigated different combinations to find the most predictive biomarkers for the survival of the patients. For this, the survival of the patients is kept on a continuous scale, but the biomarkers are transformed to a binary scale of high vs. low. The Manuscript illustrates that some borderline patients could be wrongly predicted due to the strict cut-off in the binary scale. Machine learning survival model approaches have been developed to overcome this problem and keep the survival probability and biomarkers on a continuous scale and can combine multiple of biomarkers (Acharjee et al., 2020; Pickett et al., 2021; Toth et al., 2019). We have only studied 29 patients

in this Manuscript, which might not be enough to produce a machine-learning model without a high probability of overfitting the data.

7.1.1 Predict best possible treatment option

The prediction of whether a patient can obtain benefits from different immunotherapy treatments is critical. Many studies embrace the effect of specific therapy in a cohort in aid of predicting if the treatment works or not. However, it is optimal to predict the efficacy of various treatment options and the best and most efficient treatment for patients. Studies have found that patients might fail to effectively respond in one kind of immunotherapy treatment, while succeeding another one (Rosenberg et al., 2011; Weber et al., 2013). Finding the correct first-line treatment would help to regress tumors early, resulting in increased survival probability. However, suppose treatment is received without any beneficial effect, the cancer might have progressed and develop metastasis to other regions of the body, and at that point challenging to treat (Shih et al., 2022). Predicting the best treatment candidate requires a model that has learned the associations of TME and outcome from multiple trials and searches for patterns between an environment type and survival probability.

7.2 CANCER TYPE AND TREATMENT VARIABILITY

Manuscript II and III investigate NARTs as a possible biomarker. Manuscript II showed great potential for using the number of detected NARTs three weeks after treatment as a biomarker for predicting clinical benefit. The TME before treatment also showed diversity, comparing the number of responses after treatment. This led to the possibility of using RNAseq before treatment to predict the ability to detect NART after treatment. Manuscript III focused on the frequency of the NARTs and how NART frequency can be used as a predictive biomarker for patients benefitting of immunotherapy with TIL-ACT. Both studies also showed TME diversity by comparing patients with high and low frequency of NARTs. These two Manuscripts show different views of how NARTs can predict patient outcomes concordantly; these two comparisons are not the same. Manuscript II observed the number of detected NARTs three weeks after treatment and Manuscript III showed best predictive effect by comparing the NARTs frequency given as the sum of estimated frequency pr. patients, detected in the TIL infusion product. That bias between cohorts can be due to the disease, as in Manuscript II, we investigated mUC patients, and in Manuscript

III, melanoma patients. However, another parameter that might explain the differences in the treatment. The mUC patients were treated with anti-PD-L1, which might remove the immuno-suppression on many different T cells resulting in many NARTs but with lower frequency. Whereas the melanoma cohort was treated with TIL-ACT, and the expanded infusing product might be dominated by few T cell clones, resulting in a high frequency of a few clones.

Differences were also obtained in the evaluation of neoepitope immunogenicity characteristics. In Manuscript III, we saw a difference in expression level but no difference in MHC binding capacity with eluted ligand % rank (RankEL) . Opposite in Manuscript II, no differences in expression level were found, but differences in RankEL and binding affinity % Rank (RankBA). Furthermore, mutation clonality has previously been shown to be important (McGranahan et al., 2016), but neither of the two Manuscripts could confirm that proposition. This leads us to the search of defining what characteristics are aquired by an immunogenic neoepitope.

7.3 NEOEPITOPE IMMUNOGENECITY

The prediction of immunogenic neoepitopes is difficult because of the limited amount of validated data available. Even though Manuscript IV studied above 19000 validated neoepitope candidates predicted with the same prediction technique and following the same validation approach, no remarkable performance was obtained in aid of prediction immunogenicity with the use of RF modelling. However, the created model, IMPROVE, predict immunogenicty beyond the already known important feature (RankEL and gene expression) as the neoepitopes were already pre-selected for these parameters. IMPROVE, can be used on top of these pre-selection to rank neoepitope candidate and the ranking was better than ranking the pre-selected neoeptiope candidates additionally with RankEL. Furthermore, we observed that physiochemical properties of the neopeptides such as hydrophobicity and the abundance or aromatic residues contributed the most to the model. This model was comparable to the NNAlign model (Nielsen and Andreatta, 2017), which only takes the peptide sequence into account because the NNAlign model learned some of the same important characteristics as the RF model, concordantly the hydrophobic and aromatic residues were important for an immunogenic neoepitope. This finding was also supported by PRIME (Schmidt et al., 2021), which observed similar properties and was also one of the features included in the model with high importance. Nevertheless, the feature-based RF model still out-competes the RankEL and NNAlign model, which can draw comparable conclusions as in Manuscript I. Namely, not only one parameter is enough to predict the immune system's com-

plexity, therefore the feature-based model performs better. Nevertheless, even if it was possible to predict the immunogenicity of a neoepitope from peptide-specific features with high accuracy, it needs a TIL to recognize it and hence be classified as "immunogenic". We observed from the data that identical pMHC complexes from different patients can be immunogenic in one and not in the others. This could be explained by the patient-to-patient variability and the absence of TILs, hence missing recognition of the neoepitope. Both Manuscript II and III showed the patients concentration of NARTs were associated with the TME. Furthermore, the cancer cells can escape the T cell recognition in various manners, as mentioned in Chapter 1.6, such as a down regulation of HLA allele expression (Gajewski et al., 2017; Nagarsheth et al., 2017). These observations challenge an ML-based model in performing accurate prediction as it learns from past experiences, and in these cases, the past experiences are not valid for all patients because the duplicates can give different outcomes in different patients. This reflects the next point of the Manuscript, namely, implementing some of the complexity of the immune system, the TME. We observed some features describing the TME, including CYT, which also have been reported in Manuscript I, and cell populations from MCP-counter (also reported in Manuscript II). By adding these TME features on top of the existing peptide-specific features in the RF model, it resulted in increased performance in predicting the neoepitope immunogenicity. Noteworthy, these TME features cannot be used as stand-alone features to predict neoepitope immunogenicity but adds information about the capability of discovering NARTs in the patient. The Manuscript only considers a few TME features to include, which showed promising results in improving the neoepitope prediction. However, various other TME features could have been considered for the analysis and are of interest for further work.

Conclusively, these four Manuscripts add novel insight into biomarker discovery and neoepitope prediction. However, there is still room for a lot of clarification and further analysis with more data to predict patient survival and neoepitope immunogenicity with high specificity and sensitivity.

BIBLIOGRAPHY

Acharjee, A., Larkman, J., Xu, Y., Cardoso, V. R., and Gkoutos, G. V. (2020). A random forest based biomarker discovery and power analysis framework for diagnostics research. BMC medical genomics, 13(1):1–14.

Ali, J., Khan, R., Ahmad, N., and Maqsood, I. (2012). Random forests and decision trees. International Journal of Computer Science Issues (IJCSI), 9(5):272.

Bai, R., Lv, Z., Xu, D., and Cui, J. (2020). Predictive biomarkers for cancer immunotherapy with immune checkpoint inhibitors. Biomarker Research, 8(1):1–17.

Beatty, G. L. and Gladney, W. L. (2015). Immune escape mechanisms as a guide for cancer immunotherapytailoring cancer immunotherapy. Clinical cancer research, 21(4):687–692.

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W. H., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome biology, 17(1):1–20.

Benjamin, D., Sato, T., Cibulskis, K., Getz, G., Stewart, C., and Lichtenstein, L. (2019). Calling somatic snvs and indels with mutect2. BioRxiv, page 861054.

Bentzen, A. K., Marquard, A. M., Lyngaa, R., Saini, S. K., Ramskov, S., Donia, M., Such, L., Furness, A. J., McGranahan, N., Rosenthal, R., et al. (2016). Large-scale detection of antigen-specific t cells using peptide-mhc-i multimers labeled with dna barcodes. Nature biotechnology, 34(10):1037–1045.

Besser, M. J., Shapira-Frommer, R., Itzhaki, O., Treves, A. J., Zippel, D. B., Levy, D., Kubi, A., Shoshani, N., Zikich, D., Ohayon, Y., et al. (2013). Adoptive transfer of tumor-infiltrating lymphocytes in patients with metastatic melanoma: Intent-to-treat analysis and efficacy after failure to prior immunotherapiesintent-to-treat analysis of til act and impact of ipilimumab. Clinical Cancer Research, 19(17):4792–4800.

Bjerregaard, A.-M., Nielsen, M., Hadrup, S. R., Szallasi, Z., and Eklund, A. C. (2017). Mupexi: prediction of neo-epitopes from tumor sequencing data. Cancer Immunology, Immunotherapy, 66(9):1123–1130.

Blank, C. U., Rozeman, E. A., Fanchi, L. F., Sikorska, K., van de Wiel, B., Kvistborg, P., Krijgsman, O., van den Braber, M., Philips, D., Broeks, A., et al. (2018). Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage iii melanoma. Nature medicine, 24(11):1655–1661.

Blass, E. and Ott, P. A. (2021). Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. Nature Reviews Clinical Oncology, 18(4):215–229.

Bolotin, D. A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I. Z., Putintseva, E. V., and Chudakov, D. M. (2015). MiXCR: Software for comprehensive adaptive immunity profiling. Nature Methods, 12(5):380–381.

Borghaei, H., Smith, M. R., and Campbell, K. S. (2009). Immunotherapy of cancer. European journal of pharmacology, 625(1-3):41–54.

Boulesteix, A.-L., Janitza, S., Kruppa, J., and König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2(6):493–507.

Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic rna-seq quantification. Nature biotechnology, 34(5):525–527.

Buchbinder, E. I. and Desai, A. (2016). CTLA-4 and PD-1 pathways similarities, differences, and implications of their inhibition. American Journal of Clinical Oncology: Cancer Clinical Trials, 39(1):98–106.

Burdett, N. and Desai, J. (2020). New biomarkers for checkpoint inhibitor therapy. ESMO open, 5:e000597.

Case, K., Kennedy, K., Kujtan, L., and Subramanian, J. (2019). The role of tumour neoantigens in the differential response to immunotherapy (io) in egfr and braf mutated lung cancers: Quantity or quality? Annals of Oncology, 30:v513.

Chakraborty, S., Hosen, M. I., Ahmed, M., and Shekhar, H. U. (2018). Onco-multi-omics approach: a new frontier in cancer research. BioMed research international, 2018.

Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., and Peters, S. (2019). Development of tumor mutation burden as an im-

munotherapy biomarker: utility for the oncology clinic. Annals of Oncology, 30(1):44–56.

Chen, X. and Ishwaran, H. (2012). Random forests for genomic data analysis.

Chowell, D., Yoo, S.-K., Valero, C., Pastore, A., Krishna, C., Lee, M., Hoen, D., Shi, H., Kelly, D. W., Patel, N., et al. (2022). Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. Nature Biotechnology, 40(4):499–506.

Chu, D. and Wei, L. (2019). Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation. BMC cancer, 19(1):1–12.

Coelho, A. C. M., Fonseca, A. L., Martins, D. L., Lins, P. B., da Cunha, L. M., and de Souza, S. J. (2020). neoant-hill: an integrated tool for identification of potential neoantigens. BMC Medical Genomics, 13(1):1–8.

Cooper, G. M. and Hausman, R. (2000). A molecular approach.

Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., Sher, X., Liu, X. Q., Lu, H., Nebozhyn, M., et al. (2018). Pan-tumor genomic biomarkers for pd-1 checkpoint blockade–based immunotherapy. Science, 362(6411):eaar3593.

Cruz, J. A. and Wishart, D. S. (2006). Applications of machine learning in cancer prediction and prognosis. Cancer informatics, 2:117693510600200030.

Dayhoff, J. E. and DeLeo, J. M. (2001). Artificial neural networks: opening the black box. Cancer: Interdisciplinary International Journal of the American Cancer Society, 91(S8):1615–1635.

Doroshow, D. B., Bhalla, S., Beasley, M. B., Sholl, L. M., Kerr, K. M., Gnjatic, S., Wistuba, I. I., Rimm, D. L., Tsao, M. S., and Hirsch, F. R. (2021). Pd-l1 as a biomarker of response to immune-checkpoint inhibitors. Nature reviews Clinical oncology, 18(6):345–362.

Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. Neural Computing and Applications, 29(10):685–693.

Favero, F., Joshi, T., Marquard, A. M., Birkbak, N. J., Krzystanek, M., Li, Q., Szallasi, Z., and Eklund, A. C. (2015). Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Annals of Oncology, 26(1):64–70.

Felix Krueger (2021). TrimGalore https://github.com/FelixKrueger/TrimGalore.

Ferris, S. T., Durai, V., Wu, R., Theisen, D. J., Ward, J. P., Bern, M. D., Davidson, J. T., Bagadia, P., Liu, T., Briseño, C. G., et al. (2020). cdc1 prime and are licensed by cd4+ t cells to induce anti-tumour immunity. Nature, 584(7822):624–629.

Filipovic, A., Miller, G., and Bolen, J. (2020). Progress toward identifying exact proxies for predicting response to immunotherapies. Frontiers in Cell and Developmental Biology, 8:155.

Fritsch, E. F., Rajasagi, M., Ott, P. A., Brusic, V., Hacohen, N., and Wu, C. J. (2014). Hla-binding properties of tumor neoepitopes in humanstumor neoepitopes in humans. Cancer immunology research, 2(6):522–529.

Fritz, J. M. and Lenardo, M. J. (2019). Development of immune checkpoint therapyfor cancer. Journal of Experimental Medicine, 216(6).

Gajewski, T. F., Corrales, L., Williams, J., Horton, B., Sivan, A., and Spranger, S. (2017). Cancer immunotherapy targets based on understanding the t cell-inflamed versus non-t cell-inflamed tumor microenvironment. In Advances in Experimental Medicine and Biology, volume 1036.

Ghorani, E., Rosenthal, R., McGranahan, N., Reading, J., Lynch, M., Peggs, K., Swanton, C., and Quezada, S. (2018). Differential binding affinity of mutated peptides for mhc class i is a predictor of survival in advanced lung cancer and melanoma. Annals of oncology, 29(1):271–279.

Gonzalez, H., Hagerling, C., and Werb, Z. (2018). Roles of the immune system in cancer: from tumor initiation to metastatic progression. Genes & development, 32(19-20):1267–1284.

Goodman, A. M., Kato, S., Bazhenova, L., Patel, S. P., Frampton, G. M., Miller, V., Stephens, P. J., Daniels, G. A., and Kurzrock, R. (2017). Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancerstmb predicts response to immunotherapy in diverse cancers. Molecular cancer therapeutics, 16(11):2598–2608.

Gregorutti, B., Michel, B., and Saint-Pierre, P. (2017). Correlation and variable importance in random forests. Statistics and Computing, 27(3):659–678.

Hackl, H., Charoentong, P., Finotello, F., and Trajanoski, Z. (2016). Computational genomics tools for dissecting tumour–immune cell interactions. Nature Reviews Genetics, 17(8):441–458.

Hobohm, U., Scharf, M., Schneider, R., and Sander, C. (1992). Selection of representative protein data sets. Protein Science, 1(3):409–417.

Huang, P.-W. and Chang, J. W.-C. (2019). Immune checkpoint inhibitors win the 2018 nobel prize. biomedical journal, 42(5):299–306.

Hundal, J., Carreno, B. M., Petti, A. A., Linette, G. P., Griffith, O. L., Mardis, E. R., and Griffith, M. (2016). pvac-seq: a genome-guided in silico approach to identifying tumor neoantigens. Genome medicine, 8(1):1–11.

Jamy, O. and Sonpavde, G. (2017). Emerging first line treatment options for bladder cancer: a review of phase ii and iii therapies in the pipeline. Expert Opinion on Emerging Drugs, 22(4):347–355.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. The Journal of Immunology, 199(9).

Khalilia, M., Chakraborty, S., and Popescu, M. (2011). Predicting disease risks from highly imbalanced data using random forest. BMC medical informatics and decision making, 11(1):1–13.

Khomiak, A., Brunner, M., Kordes, M., Lindblad, S., Miksch, R. C., Öhlund, D., and Regel, I. (2020). Recent discoveries of diagnostic, prognostic and predictive biomarkers for pancreatic cancer. Cancers, 12(11):3234.

Kiyotani, K., Mai, T. H., and Nakamura, Y. (2017). Comparison of exome-based hla class i genotyping tools: identification of platform-specific genotyping errors. Journal of human genetics, 62(3):397–405.

Kohavi, R. and Edu, S. (1993). A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial intelligence, 2.

Kristensen, N. P., Heeke, C., Tvingsholm, S. A., Borch, A., Draghi, A., Crowther, M. D., Carri, I., Munk, K. K., Holm, J. S., Bjerregaard, A.-M., et al. (2022). Neoantigen-reactive cd8+ t cells affect clinical outcome of adoptive cell therapy with tumor-infiltrating lymphocytes in melanoma. The Journal of clinical investigation, 132(2).

Kuksin, M., Morel, D., Aglave, M., Danlos, F.-X., Marabelle, A., Zinovyev, A., Gautheret, D., and Verlingue, L. (2021). Applications of single-cell and bulk rna sequencing in onco-immunology. European Journal of Cancer, 149:193–210.

Latchman, Y., Wood, C. R., Chernova, T., Chaudhary, D., Borde, M., Chernova, I., Iwai, Y., Long, A. J., Brown, J. A., Nunes, R., et al. (2001). Pd-l2 is a second ligand for pd-1 and inhibits t cell activation. Nature immunology, 2(3):261–268.

Lauss, M., Donia, M., Harbst, K., Andersen, R., Mitra, S., Rosengren, F., Salim, M., Vallon-Christersson, J., Törngren, T., Kvist, A., et al. (2017). Mutational and putative neoantigen load predict clinical benefit of adoptive t cell therapy in melanoma. Nature communications, 8(1):1–11.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with burrows–wheeler transform. bioinformatics, 25(14):1754–1760.

Li, X. and Wang, C.-Y. (2021). From bulk, single-cell to spatial rna sequencing. International Journal of Oral Science, 13(1):1–6.

Liang, F., Zhang, S., Wang, Q., and Li, W. (2020). Clinical benefit of immune checkpoint inhibitors approved by us food and drug administration. BMC cancer, 20(1):1–11.

Lisiecka, U. and Kostro, K. (2016). Mechanisms of tumour escape from immune surveillance. Journal of Veterinary Research, 60(4):453–460.

Litchfield, K., Reading, J. L., Puttick, C., Thakkar, K., Abbosh, C., Bentham, R., Watkins, T. B., Rosenthal, R., Biswas, D., Rowan, A., et al. (2021). Meta-analysis of tumor-and t cell-intrinsic mechanisms of sensitization to checkpoint inhibition. Cell, 184(3):596–614.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. Genome biology, 15(12):1–21.

Łuksza, M., Riaz, N., Makarov, V., Balachandran, V. P., Hellmann, M. D., Solovyov, A., Rizvi, N. A., Merghoub, T., Levine, A. J., Chan, T. A., et al. (2017). A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. Nature, 551(7681):517–520.

McGranahan, N., Furness, A. J., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S. K., Jamal-Hanjani, M., Wilson, G. A., Birkbak, N. J., Hiley, C. T., et al. (2016). Clonal neoantigens elicit t cell immunoreactivity and sensitivity to immune checkpoint blockade. Science, 351(6280):1463–1469.

McGranahan, N., Rosenthal, R., Hiley, C. T., Rowan, A. J., Watkins, T. B., Wilson, G. A., Birkbak, N. J., Veeriah, S., Van Loo, P., Herrero, J., et al. (2017). Allele-specific hla loss and immune escape in lung cancer evolution. Cell, 171(6):1259–1271.

McGranahan, N. and Swanton, C. (2017). Clonal heterogeneity and tumor evolution: past, present, and the future. Cell, 168(4):613–628.

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., Flicek, P., and Cunningham, F. (2016). The ensembl variant effect predictor. Genome biology, 17(1):1–14.

Murphy, K. and Weaver, C. (2017). Janeway 's Immunology.

Nagarsheth, N., Wicha, M. S., and Zou, W. (2017). Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy.

Nahm, F. S. (2022). Receiver operating characteristic curve: overview and practical use for clinicians. Korean journal of anesthesiology, 75(1):25–36.

Neefjes, J., Jongsma, M. L., Paul, P., and Bakke, O. (2011). Towards a systems understanding of mhc class i and mhc class ii antigen presentation. Nature reviews immunology, 11(12):823–836.

Nielsen, M. and Andreatta, M. (2017). Nnalign: a platform to construct and evaluate artificial neural network models of receptor–ligand interactions. Nucleic acids research, 45(W1):W344–W349.

Obeid, J. M., Erdag, G., Smolkin, M. E., Deacon, D. H., Patterson, J. W., Chen, L., Bullock, T. N., and Slingluff, C. L. (2016). Pd-l1, pd-l2 and pd-1 expression in metastatic melanoma: Correlation with tumor-infiltrating immune cells and clinical outcome. Oncoimmunology, 5(11):e1235107.

Oiseth, S. J. and Aziz, M. S. (2017). Cancer immunotherapy: a brief review of the history, possibilities, and challenges ahead. Journal of cancer metastasis and treatment, 3:250–261.

Perry, T. and Bader-El-Den, M. (2015). Imbalanced classification using genetically optimized random forests. In Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolutionary Computation, pages 1453–1454.

Peters, S., Paz-Ares, L., Herbst, R. S., and Reck, M. (2022). Addressing cpi resistance in nsclc: targeting tam receptors to modulate the tumor microenvironment and future prospects. Journal for Immunotherapy of Cancer, 10(7).

Pickett, K. L., Suresh, K., Campbell, K. R., Davis, S., and Juarez-Colunga, E. (2021). Random survival forests for dynamic predictions of a time-to-event outcome using a longitudinal biomarker. BMC medical research methodology, 21(1):1–14.

Poggi, A. and Zocchi, M. R. (2006). Mechanisms of tumor escape: role of tumor microenvironment in inducing apoptosis of cytolytic effector cells. Archivum immunologiae et therapiae experimentalis, 54(5):323–333.

Raine, K. M., Van Loo, P., Wedge, D. C., Jones, D., Menzies, A., Butler, A. P., Teague, J. W., Tarpey, P., Nik-Zainal, S., and Campbell, P. J. (2016). ascatngs: Identifying somatically acquired copy-number alterations from whole-genome sequencing data. Current protocols in bioinformatics, 56(1):15–9.

Raskov, H., Orhan, A., Christensen, J. P., and Gögenur, I. (2021). Cytotoxic cd8+ t cells in cancer and cancer immunotherapy. British journal of cancer, 124(2):359–367.

Reck, M., Rodríguez-Abreu, D., Robinson, A. G., Hui, R., Csőszi, T., Fülöp, A., Gottfried, M., Peled, N., Tafreshi, A., Cuffe, S., et al. (2016). Pembrolizumab versus chemotherapy for pd-l1–positive non–small-cell lung cancer. N engl J med, 375:1823–1833.

Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-Validation, pages 532–538. Springer US, Boston, MA.

Ren, X., Guo, S., Guan, X., Kang, Y., Liu, J., and Yang, X. (2022). Immunological classification of tumor types and advances in precision combination immunotherapy. Frontiers in Immunology, 13.

Richman, L. P., Vonderheide, R. H., and Rech, A. J. (2019). Neoantigen dissimilarity to the self-proteome predicts immunogenicity and response to immune checkpoint blockade. Cell systems, 9(4):375–382.

Robinson, J., Barker, D. J., Georgiou, X., Cooper, M. A., Flicek, P., and Marsh, S. G. (2020). Ipd-imgt/hla database. Nucleic acids research, 48(D1):D948–D955.

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell, 160(1-2):48–61.

Rosenberg, S. A., Packard, B. S., Aebersold, P. M., Solomon, D., Topalian, S. L., Toy, S. T., Simon, P., Lotze, M. T., Yang, J. C., Seipp, C. A., et al. (1988). Use of tumor-infiltrating lymphocytes and interleukin-2 in the immunotherapy of patients with metastatic melanoma. New England Journal of Medicine, 319(25):1676–1680.

Rosenberg, S. A. and Restifo, N. P. (2015). Adoptive cell transfer as personalized immunotherapy for human cancer. Science, 348(6230):62–68.

Rosenberg, S. A., Yang, J. C., Sherry, R. M., Kammula, U. S., Hughes, M. S., Phan, G. Q., Citrin, D. E., Restifo, N. P., Robbins, P. F., Wunderlich, J. R., et al. (2011). Durable complete responses in heavily pretreated patients with metastatic melanoma using t-cell transfer immunotherapycomplete regressions in melanoma. Clinical cancer research, 17(13):4550–4557.

Roth, A., Khattra, J., Yap, D., Wan, A., Laks, E., Biele, J., Ha, G., Aparicio, S., Bouchard-Côté, A., and Shah, S. P. (2014). Pyclone: statistical inference of clonal population structure in cancer. Nature methods, 11(4):396–398.

Saxena, M., van der Burg, S. H., Melief, C. J., and Bhardwaj, N. (2021). Therapeutic cancer vaccines. Nature Reviews Cancer 2021 21:6, 21(6):360–378.

Schaap-Johansen, A.-L., Vujović, M., Borch, A., Hadrup, S. R., and Marcatili, P. (2021). T cell epitope prediction and its application to immunotherapy. Frontiers in Immunology, 12.

Schmidt, J., Smith, A. R., Magnin, M., Racle, J., Devlin, J. R., Bobisse, S., Cesbron, J., Bonnet, V., Carmona, S. J., Huber, F., et al. (2021). Prediction of neo-epitope immunogenicity reveals tcr recognition determinants and provides insight into immunoediting. Cell Reports Medicine, 2(2):100194.

Schumacher, T. N. and Schreiber, R. D. (2015). Neoantigens in cancer immunotherapy. Science, 348(6230):69–74.

Shemesh, C. S., Hsu, J. C., Hosseini, I., Shen, B.-Q., Rotte, A., Twomey, P., Girish, S., and Wu, B. (2021). Personalized cancer vaccines: clinical landscape, challenges, and opportunities. Molecular Therapy, 29(2):555–570.

Shih, N.-C., Kung, P.-T., Kuo, W.-Y., and Tsai, W.-C. (2022). Association of treatment delay and stage with mortality in breast cancer: a nationwide cohort study in taiwan. Scientific Reports, 12(1):1–10.

Solinas, C., Aiello, M., Rozali, E., Lambertini, M., Willard-Gallo, K., and Migliori, E. (2020). Programmed cell death-ligand 2: A neglected but important target in the immune response to cancer? Translational oncology, 13(10):100811.

Somvanshi, M., Chavan, P., Tambade, S., and Shinde, S. V. (2017). A review of machine learning techniques using decision tree and support vector machine.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 71(3):209–249.

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). Optitype: precision hla typing from next-generation sequencing data. Bioinformatics, 30(23):3310–3316.

Takamori, S., Takada, K., Toyokawa, G., Azuma, K., Shimokawa, M., Jogo, T., Yamada, Y., Hirai, F., Tagawa, T., Kawahara, A., et al. (2018). Pd-l2

expression as a potential predictive biomarker for the response to anti-pd-1 drugs in patients with non-small cell lung cancer. Anticancer research, 38(10):5897–5901.

Tappeiner, E., Finotello, F., Charoentong, P., Mayer, C., Rieder, D., and Trajanoski, Z. (2017). Timiner: Ngs data mining pipeline for cancer immunology and immunotherapy. Bioinformatics, 33(19):3140–3141.

Tong, M., Wang, J., He, W., Wang, Y., Pan, H., Li, D., and Zhang, H. (2018). Predictive biomarkers for tumor immune checkpoint blockade. Cancer management and research, 10:4501.

Toth, R., Schiffmann, H., Hube-Magg, C., Büscheck, F., Höflmayer, D., Weidemann, S., Lebok, P., Fraune, C., Minner, S., Schlomm, T., et al. (2019). Random forest-based modelling to detect biomarkers for prostate cancer progression. Clinical epigenetics, 11(1):1–15.

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., et al. (2013). From fastq data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics, 43(1):11–10.

Vesely, M. D. and Schreiber, R. D. (2013). Cancer Immunoediting: antigens, mechanisms and implications to cancer immunotherapy. Annals of the New York Academy of Sciences, 1284(1):1.

Walter, S. D. (2005). The partial area under the summary ROC curve. Statistics in Medicine, 24(13).

Wang, Y., Du, J., Gao, Z., Sun, H., Mei, M., Wang, Y., Ren, Y., and Zhou, X. (2022). Evolving landscape of pd-l2: bring new light to checkpoint immunotherapy. British Journal of Cancer, pages 1–12.

Weber, J. S., Kudchadkar, R. R., Yu, B., Gallenstein, D., Horak, C. E., Inzunza, H. D., Zhao, X., Martinez, A. J., Wang, W., Gibney, G., et al. (2013). Safety, efficacy, and biomarkers of nivolumab with vaccine in ipilimumab-refractory or-naive melanoma. Journal of clinical oncology, 31(34):4311.

Wells, D. K., van Buuren, M. M., Dang, K. K., Hubbard-Lucey, V. M., Sheehan, K. C., Campbell, K. M., Lamb, A., Ward, J. P., Sidney, J., Blazquez, A. B., et al. (2020). Key parameters of tumor epitope immunogenicity revealed through a consortium approach improve neoantigen prediction. Cell, 183(3):818–834.

Winter, J. M., Yeo, C. J., and Brody, J. R. (2013). Diagnostic, prognostic, and predictive biomarkers in pancreatic cancer. Journal of surgical oncology, 107(1):15–22.

Wood, M. A., Weeder, B. R., David, J. K., Nellore, A., and Thompson, R. F. (2020). Burden of tumor mutations, neoepitopes, and other variants are weak predictors of cancer immunotherapy response and overall survival. Genome medicine, 12(1):1–16.

Yarchoan, M., Hopkins, A., and Jaffee, E. M. (2017). Tumor mutational burden and response rate to pd-1 inhibition. New England Journal of Medicine, 377(25):2500–2501.

Yearley, J. H., Gibson, C., Yu, N., Moon, C., Murphy, E., Juco, J., Lunceford, J., Cheng, J., Chow, L. Q., Seiwert, T. Y., et al. (2017). Pd-l2 expression in human tumors: relevance to anti-pd-1 therapy in cancer. Clinical cancer research, 23(12):3158–3167.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterprofiler: an r package for comparing biological themes among gene clusters. Omics: a journal of integrative biology, 16(5):284–287.

Zhao, J., Chen, A. X., Gartrell, R. D., Silverman, A. M., Aparicio, L., Chu, T., Bordbar, D., Shan, D., Samanamud, J., Mahajan, A., et al. (2019). Immune and genomic correlates of response to anti-pd-1 immunotherapy in glioblastoma. Nature medicine, 25(3):462–469.

Zhu, M., Xia, J., Jin, X., Yan, M., Cai, G., Yan, J., and Ning, G. (2018). Class weights random forest algorithm for processing class imbalanced medical data. IEEE Access, 6:4641–4652.

Zito, A., Lualdi, M., Granata, P., Cocciadiferro, D., Novelli, A., Alberio, T., Casalone, R., and Fasano, M. (2021). Gene set enrichment analysis of interaction networks weighted by node centrality. Frontiers in genetics, 12:577623.