Impact of proteasomal processing on immunopeptidome repertoires and therapeutic immune recognition.

Miranda C. Lahman

A Dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2022

Reading Committee:
Aude G. Chapuis, Chair
Jean Campbell
Derek Stirewalt

Program Authorized to Offer Degree:
Molecular Medicine and Mechanisms of Disease

University of Washington

Impact of proteasomal processing on immunopeptidome repertoires and therapeutic immune recognition

Miranda C. Lahman

Chair of the Supervisory Committee: Aude G. Chapuis
Department of Laboratory Medicine and Pathology, School of Medicine

Abstract

Acute myeloid leukemia (AML) is the most common acute leukemia in adults. AML often responds to initial chemotherapy, but a majority of patients will relapse with resistant disease leaving a critical need for more effective therapies. Hematopoietic stem cell transplantation can prevent relapse long term through donor T cell-mediated elimination of leukemic cells. However, this anti-leukemic effect is unpredictable, and transplant can have detrimental side effects, including graft versus host disease (GVHD). New therapies focus the anti-leukemic effect and bypass potential GVHD by using engineered T cells with a defined T cell receptor (TCRs) that recognize proteasome-generated peptide fragments from tumor-associated proteins. Such 'TCR-T' cells rely on the targeted peptide to be processed and then presented by the restricting human leukocyte antigen (HLA). Cells can express standard- and immuno-proteasome isoforms, which can generate 'distinct' (isoform-dependent) or 'mutual' (isoform-independent) peptides. We investigated a clinical scenario in which the targeted AML modulated proteasome composition to eliminate processing of the targeted 'distinct' peptide, leading to loss of recognition by the adoptively transferred TCR-T cells and AML progression. Instead, an alternative TCR recognizing a 'mutual' peptide could respond to the progressive AML.
Our data point to a mechanism whereby AML can evade immune recognition through modulation of proteasome isoform expression. This further implies that TCR-T cell therapy could be enhanced if tailored to target 'mutual' peptides.

To determine the relative proportions of HLA-restricted 'distinct' and 'mutual' peptides and to identify promising TCR-T cell targetable peptides less likely to evade recognition, we utilized an emerging peptide discovery platform, Artemis, that employs soluble HLAs to bypasses immunoprecipitation of endogenous HLAs. Artemis-identified peptides revealed that the proportions of 'mutual' and 'distinct' peptides were specific to each HLA allele, with HLA-A*02:01 presenting the fewest 'mutual' peptides and HLA-A*11:01 presenting the most. We recovered 37 peptides from three AML-associated proteins of interest, 19 (51%) were 'mutual' peptides and some peptides (11%) were previously identified. Our findings demonstrate that defining isoform-specific proteasomal processing is critical to optimal peptide selection for TCR-T cell therapy and Artemis is an effective platform to identify 'mutual' peptides.

Acknowledgements

It takes a village.

I received support from numerous people and truly would not be graduating without their encouragement or helping hands. I want to thank *all* those (too many to list) who had a role in this PhD. Words could never express the true magnitude of my appreciation, though I will attempt to do so here.

Thank you to my mentors: Dr. Aude Chapuis and Dr. Roland Strong. Aude, thank you for accompanying me on the marathon that is a PhD. You've been encouraging when the going gets tough and celebrated the large and small victories with me along the way. A thesaurus gives many synonyms for "mentor" including advisor, coach, guide, and teacher. Here I would also add leader, friend, talented scientist, and skilled skier. You have not only enriched my technical expertise and field knowledge, but more importantly, given me opportunities to lead and explore. Roland, thank you for your continuous support and guidance. Your scientific insights are excellent, but your workplace wisdom is what I value most. From witty replies and humorous remarks to obscure facts and management advice; I always look forward to chatting in your office.

Thank you to all my lab mates, past and present, for their help. This work would not be possible without you, especially Chapter 2. Whether it was culturing cells, discussing problematic data, lending a helping hand, or even reading this very thesis; you were there for me through every step of this journey. A special thanks to Lauren Martin for being incredibly reliable and enduring the blunders of an inexperienced PhD student.

Having brilliant individuals to call when I'm stuck on a problem is invaluable. I would like to thank the bioinformaticists and computer scientists who graciously gave their time to help write (and debug!) rudimentary code as I learned to wade through large data sets. Thank you to Dr. Valentin Voillet, Dr. Francesco Mazziotta, Garrett Sczechowski and Bruce Lahman. A special thank you to Matt Turney for the crash course in python.

I would like to acknowledge the patients and their families. The most important people in translational research are the patients who enroll in clinical trials. Their generosity and sacrifice are beyond measure and is often unrecognized. Participating in a clinical trial may benefit a patient but will almost certainly benefit those who come later. The risks these patients take to advance medicine is courageous and noble.

Outside the lab I want to thank to all my friends who kept me sane by running, biking, creating, and exploring with me. I could not have persevered without you.

Finally, I would like to recognize all the fearless women in my life who taught me anything is possible and to dream big. I admire you and respect you. Without the role models and pioneers before me, I would not be here today. Continue leading the way and inspiring the next generation of women leaders, scientists, mentors, and entrepreneurs – I promise to do the same.

Dedication

I dedicate this dissertation to my family, be it by blood or by choice.

To my family in the traditional sense of the word – I love you to the moon and back. To my parents, Beth and Bruce, for giving me every tool and opportunity to thrive. You show unwavering support, answer panicked phone calls, and provide endless encouragement in life, including when earning this degree. Thank you. To my brother, Joel, you personify grit and perseverance; you have shown me how to endure bumps in the road today for a brighter tomorrow. I am proud of you. To my Grandparents. To Lyle and Mary Lahman, for encouraging the scientist in me at every stage and always being interested in what I am studying. To Dr. Thomas Seery, for inspiring me to study medicine. To Elizabeth Seery, for raising strong independent women who raised strong independent women. I vow to continue this legacy with my own children.

To my best friend, Sydney, for always inspiring me to be better. Your brilliance, love, joy, humor, and sense of adventure are unmatched. Thank you for nerding out about science with me since 7th grade. My heart and soul adore you.

To my partner, Austin, for bringing me back to reality when I need it most. For being the words of wisdom and guidance when I feel lost. For hugging me when words are not right. For being there to celebrate my moments of brilliance and lifting me up in my moments of failure. For guiding me up mountains to add perspective. And, above all, for supporting me though this entire journey. I look forward to all our adventures, triumphs, and comebacks to come. I love you.

>>>>

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Immunotherapy for cancer

Worldwide, cancer is a leading cause of death (*1, 2*) (www.cancer.gov). Physicians and scientists have long been interested in the molecular underpinnings of cancer and in developing therapeutics to effectively target and eliminate cancerous cells. Over the decades, medicine has made monumental strides in both understanding the molecular mechanisms of cancer and treating cancer patients.

Recent therapeutic advances largely focus on harnessing the power of the immune system to control cancer cells. Such immunotherapies modulate immune cells, often T cells, to increase their potency. Advances fall into two broad categories. The first involves therapies that enhance endogenous immune cells. These therapies aim to improve the function of the patient's existing immune cells to promote elimination of cancer. One example is checkpoint blockade therapy that blocks inhibitory signals on existing T cells to increase cytotoxicity and eliminate cancerous cells (*3*). A second category is engineered cell-based immunotherapies. These therapies introduce 'new' immune cells that are genetically modified to express an engineered receptor designed to engage a specific protein/protein complex found on malignant cells (*4, 5*). Engineered cell-based therapies include chimeric antigen receptor T cells (CAR-T) and transgenic T cell receptor T cells (TCR-T). CAR-T cells have shown stunning clinical results against lymphoid leukemias, but they mostly target obligate surface proteins (*6*). TCR-T cells, on the other hand, can also recognize internal proteins allowing them to target a much wider array of tumor associated antigens. However, internal proteins must first be degraded into peptides and presented by class I human leukocyte antigen (HLA) before the TCR can engage with these HLA:peptide complexes (TCR-HLA:peptide interaction). Therapeutic tumor-specific activation occurs when the engineered high-affinity TCR recognizes the targeted antigen-derived peptide fragment that is presented by the select HLA on the cancer cell surface (**Figure 1.1**). The fact that TCFR-T cells are not limited to targeting surface proteins is particularly important in diseases like acute myeloid leukemia (AML) that express limited unique surface proteins. Surface proteins characteristic of AML, like CD34, are also shared with critical stem cell populations and

therefore cannot be safely ablated with a CAR construct. Clinical trials using TCR-T cells have shown efficacy for AML patients (*7*) and in patients with certain other cancers (*8-11*).

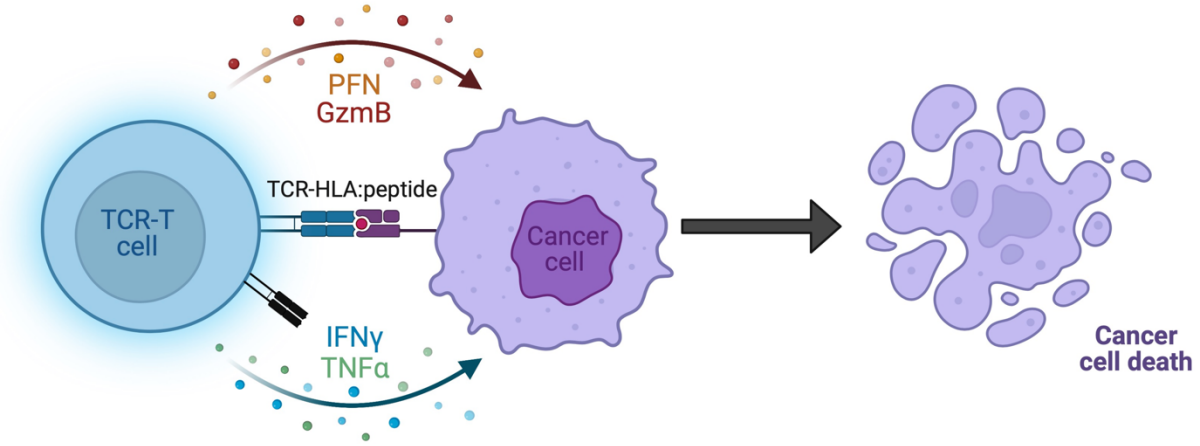


Figure 1.1 TCR-T cells recognize cancer. Engineered T cell receptor (TCR) T cells (TCR-T) express a high affinity TCR recognizing a specific antigen-derived peptide presented by a defined human leukocyte antigen (HLA) on the cell surface. This TCR-HLA:peptide interaction triggers the release of inflammatory cytokines (IFN-**γ** and TNF-**α**) as well as Perfin (PFN) an Granzyme B (GNMB) to lyse and kill the cancer cell.

Engineered cell therapies offer promising results for the treatment of AML

Majority of the work in this dissertation is applicable to multiple cancer types. However, we have chosen to use AML as our lens to view cancer immune interactions. AML is the most common acute leukemia in adults accounting for ~80 percent of acute leukemias (*12*). It is broadly characterized as having excess (≥20%) immature hematopoietic progenitor blasts in the bone marrow (*13*); blasts are abnormal cells with aberrant proliferation and arrested differentiation (**Figure 1.2**) (*14*). AML can sometimes partially recapitulate healthy differentiation acquiring phenotypes similar to healthy cells – i.e. monocyte-like AML (*14*).

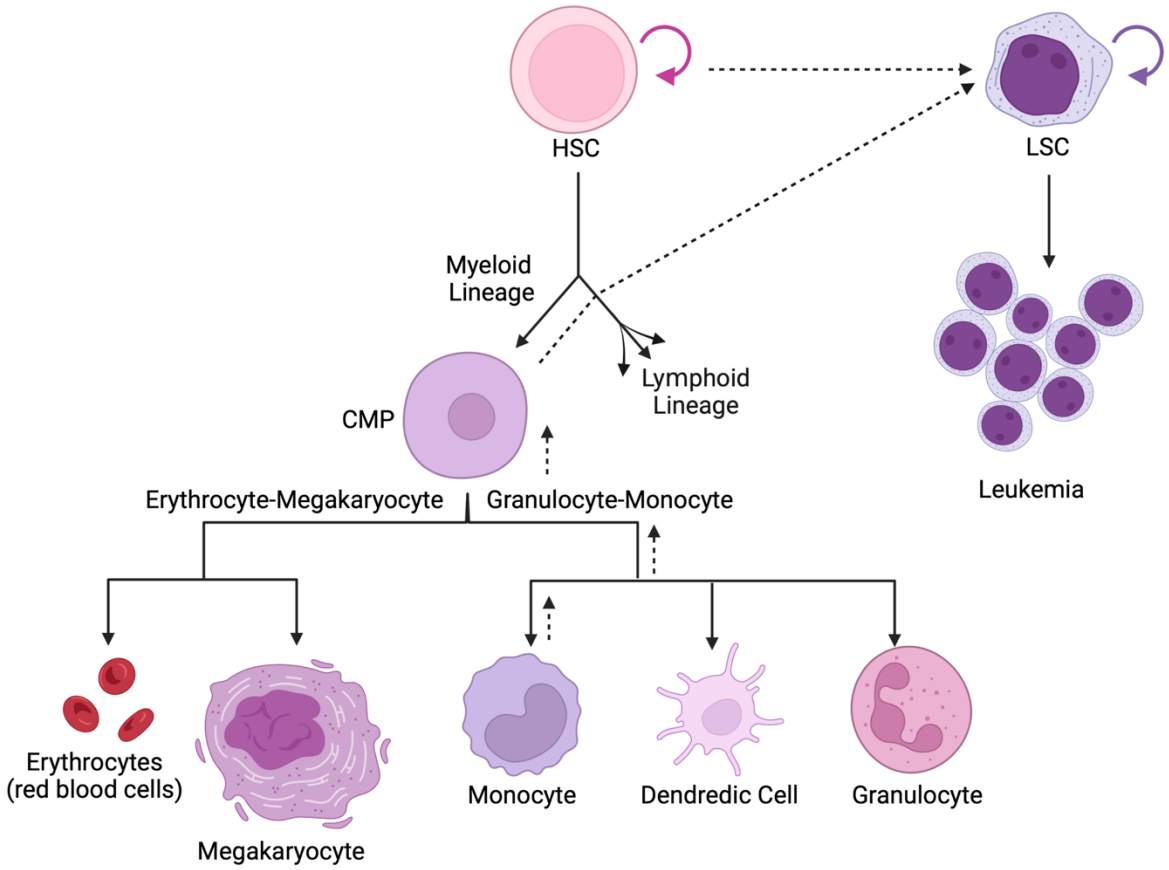


Figure 1.2 Simplified schematic of hematopoiesis. Hemopoietic stem cells (HSC) are multipotent cells capable of self renewal. HSCs can differentiate into one of two committed progenitor cells: common lymphoid progenitor or common myeloid progenitor (CMP) cells which determines if it will populate the lymphoid or myeloid lineage. CMPs will branch again to either the erythrocyte-megakaryocyte lineage to generate erythrocytes or megakaryocytes and their derivatives while Granulocyte-monocyte destined cells will form monocytes, dendritic cells, granulocytes and their derivatives. Like HSCs Leukemic stem cells (LCS) can self renew. LSCs can be derived from transformed HSCs or differentiated myeloid cells that have gained stemness (dotted lines).

However, this abnormal differentiation does not add the functional capabilities acquired through physiologic differentiation, resulting in an accumulation of dysfunctional leukemic blasts that rapidly expand to outnumber healthy hematopoietic cells and limit production of critical blood cell populations like erythrocytes (red blood cells), platelets, neutrophils, and monocytes.

AML has marked inter- and intra-tumoral heterogeneity, making it exceptionally challenging to treat. It is one of the deadliest blood cancers, with a 5-year survival rate less than 30%, substantially lower than for patients with other blood cancers (*15*). Treatment of AML still relies on chemotherapy. Current paradigms first use intense induction chemotherapy followed by several rounds of higher dose consolidation chemotherapy (*16*). Majority of patients who achieve a complete remission (CR) with chemotherapy will ultimately relapse and die of disease-related complications (*17*). Allogeneic hematopoietic cell transplantation (HCT) offers a potentially curative therapy for some patients with high-risk AML. As part of this procedure, the patient's (i.e., host) immune system is replaced by the donors (i.e., graft). Donor T cells can eliminate leukemia cells through recognition of minor histocompatibility antigens or less commonly leukemia-associated antigens presented on the cell surface (*18*) in a graft-versus-leukemia (GVL) effect. The GVL effect is entirely reliant on donor immune cells recognizing leukemic host cells resulting in variable efficacy and unpredictable before HCT. As such, many patients eventually relapse due to inability of the GVL effect to completely eradicate or control disease. Additionally, there are considerable risks associated with HCT; the most prominent is graft-versus- host disease (GVHD), where donor T cells recognize the patient's healthy cells (*19*). GVHD occurs in approximately 50% of patients post HCT and is associated with significant morbidity. Engineered T cells are one alternative to harness beneficial GVL effects without GVHD, by adoptively transferring T cells that recognize defined AML-associated antigens. Our group has successfully utilized TCR-T cell therapy in a clinical trial aimed at preventing AM relapse after HCT (*7*). However, this strategy was not sufficient to eliminate persisting AML, in part due to AML immune evasion (*20*)see also **"Lessons learned from relapsed tumors in the clinic"** below).

TCR-T cells as a model system for immune evasion

Tumor immune evasion can come from either side of the TCR-HLA:peptide interaction and untangling escape mechanisms generally requires knowing which T cells can see the tumor (T cell-intrinsic factors: e.g., TCR clonotype) and what is recognized by those T cells (tumor-intrinsic factors: e.g., antigen). In the absence of

adoptive transfer, it is often challenging to know both given discovery is co-dependent. That is, identifying a tumor-specific T cell clone requires knowing which HLA:peptide is recognized and vice versa. Under most circumstances, heterogeneity makes it challenging to know both the TCR sequence and target HLA:peptide. TCR-T cell therapies provide a controlled system where both the TCR clone and the target HLA:peptide are known. Thus, engineered cell therapies provide an opportunity to study interactions between cancer and the immune system, including during response and evasion. The Chapuis lab uses single cell RNA sequencing (scRNAseq) to probe these cancer-immune interactions. ScRNAseq is a powerful tool that first isolates single cells and then uses next-generation sequencing to profile transcriptomes of individual cells. ScRNAseq is well suited to study immune escape mechanisms as both T cell-intrinsic and tumor-intrinsic factors can be probed concurrently (*20, 21*).

ScRNAseq can track T cell clones by either the introduced TCR sequence (*7, 22*) for engineered TCR-T cells or the natural TCR CDR3 barcode for endogenous T cells. The transcriptional profile of cells expressing the engineered TCR, or a given TCR CDR3 barcode, provide insight into functional states. However, scRNAseq consumes (destroys) cells, preventing use of the same cells in downstream functional experiments, as is feasible using flow cytometry (*23*). This limits our ability to link transcriptional signatures with functional capabilities; we can only infer meaning from transcriptional signatures as function must be empirically assessed. To date, literature connecting functional states (e.g., activated, exhausted) to transcriptional signatures remains sparse, in part because traditional protein markers of functionality are not always optimal indicators at the transcriptional level (*24-26*).

When investigating tumor-intrinsic factors, it is beneficial to track the tumor throughout therapy. Analyzing multiple timepoints over the course of treatment enables researchers to compare the cancer cell populations that were therapeutically eliminated with those that survived. This can pinpoint resistance mechanisms and identify strategies to overcome/prevent future therapy resistance. Of course, there is an exceptionally broad range of possibilities that can result in acquired resistance. An asset of scRNAseq is the ability to probe multiple hypothesis using a single data set. Still, findings need to be validated. For example, scRNAseq can monitor target antigen and

HLA transcripts as a proxy for antigen expression and presentation but determining whether the HLA:peptide complex is actually presented requires protein level analysis.

While not sufficient to definitively establish mechanisms of escape, tools like scRNAseq allow us to efficiently explore a wide range of possible evasion mechanisms, including T cell-intrinsic and tumor-intrinsic possibilities, without observer bias and with relatively few patient samples making scRNAseq an asset to clinical research where patient samples are limiting.

Indeed, the Chapuis lab has used scRNAseq to reveal mechanisms of immune evasion in clinical trials (*20, 21*). In one example, the tumor selectively downregulated the TCR-T targeted HLA (*21*). ScRNAseq revealed the relapsed tumor lacked the targeted HLA-A transcripts but not HLA-B transcripts. HLA-A loss was confirmed with immunohistochemistry (*21*). In another example, detailed in Chapter Two, scRNAseq was used to probe both T cell-intrinsic and tumor-intrinsic mechanisms of resistance (*20*). On the T cell side, this patient had persistent T cells that appeared activated during remission and quiescent at relapse, function at remission and relapse was confirmed with ex vivo experiments. On the tumor side, this patient's AML retained the targeted HLA-A and antigen expression but modified peptide processing to prevent the target peptide from being produced (*20*). This example highlights the use of scRNAseq to query multiple avenues of escape and identify a less common mechanisms of escape. It likewise underscores the need to carefully select target peptides.

A needle in a haystack: finding target peptides

Effective TCR-T cell immunotherapy depends not only on targeting good tumor-associated antigens but also on selecting quality antigen-derived peptides to promote successful presentation. Ideal target antigens are uniquely and broadly expressed across tumor cells with minimal expression in healthy tissue. Mutated proteins are excellent targets, as the mutant antigen is not found in healthy cells, provided a peptide containing the mutation is processed and presented by HLA. Examples of mutated antigens in AML include frameshifts in nucleophosmin 1 ($NPM1^{mut}$) (*27-29*) and calreticulin ($CALR^{mut}$) (*30*), which encode an additional 11-47 amino acids not found in the wild type protein. Viral antigens (such as those associated with Merkle cell

carcinoma: see Appendix II) are likewise excellent TCR-T cell targets, although they are only associated with a subset of cancers. Overexpressed self-antigens can also be suitable TCR-T cell targets, although self-tolerance can make finding a high-affinity TCR challenging (*31*). Examples of targetable overexpressed self-antigens in AML include Wilms tumor antigen 1 (WT1) (*32, 33*) and cyclin A-1 (CCNA1) (*34, 35*). WT1 has little expression in healthy tissue, with 10-100 fold overexpression in AML (*32*), and has been safely targeted by our group in a TCR-T cell clinical trial (Fred Hutch 2498; NCT01640301) (*7*). While target antigen selection is critical, the antigen-derived peptide presented on the chosen HLA is the functional moiety engaging the TCR. Antigen-derived peptides are tested for immunogenicity, and those that elicit an inflammatory response, often measured by interferon-γ production, are considered for TCR-T cell therapy. Safety studies, especially for self-antigens, are imperative to ensure antigen-derived peptides are not presented on healthy tissue.

Current methods for selecting TCR-T targeted antigen-derived peptides pose challenges. Typical peptide discovery pipelines center on *in silico* predictions that determine if *any* of the theoretically possible antigen-derived peptides are expected to bind the selected HLAs. Understanding how these algorithms are trained can guide their use in peptide selection. These algorithms are trained on datasets containing (1) peptide binding affinity measurements and/or (2) peptides eluted directly off cells and detected with mass spectrometry. Cells expressing either a single HLA allele or multiple HLAs can be used as substrates for elution. Using netMHCpan as an example, earlier versions were trained on only binding affinity data (*36*). This data is biophysical and may accurately predict whether a peptide will bind the HLA but cannot estimate the probability a peptide is processed and presented. Newer versions of netMHCpan, including the most recent version 4.1 updated in 2020, are trained on a combination of peptide binding affinity data and peptides eluted off cells to better predict if a peptide will be presented (*37*). Still, biases exist in these training sets, including the HLA alleles tested (some have many more input data sets than others), peptide ionization patterns on the mass spectrometer, and length distributions (many of these training sets are heavily weighted with 9mer and 10mer peptides) (*36, 37*). These *in silico* algorithms will continue to improve as more data sets are published, which will increase the variety of

HLAs and peptide lengths in the training set. While prediction software has improved, it remains only a tool for peptide target selection. Empirical evidence is required to support pursuing a target in the lab and further transfer to the clinic.

To ensure T cell recognition and therefore peptide presentation, binding predictions must be followed by indirect *in vitro* approaches using antigen-specific T cell recognition to validate peptide immunogenicity. Such approaches center on exogenously loading synthetic peptides onto presenting cells and measuring expansion of reactive T cells. However, promising immunogenic peptides sometimes fail to show endogenous presentation. In such cases, T cells raised against the artificially presented peptide cannot meaningfully reduce tumor in pre-clinical models.

Ensuring endogenous presentation of peptides before validating antigenicity can be addressed by directly eluting peptides off presenting HLAs bound to the cell surface. This relies on immunoprecipitation of membrane-bound HLAs followed by mass spectrometry and computational deconvolution to match eluted peptides to the presenting HLA allele (*38*).

Other recent approaches, use soluble HLAs to circumnavigate immunoprecipitation (*39*) or engineered cell lines expressing a single HLA allele to avoid computational deconvolution (*40*). We employed an emerging peptide discovery platform, Artemis (*41*), to bypass standard immunoprecipitation. Artemis is depicted in **Figure 3.1** and further detailed in Chapter 3. Briefly, Artemis utilizes an innovative soluble HLA called a single chain dimer (SCD)**.** The SCD has the transmembrane portion of the HLA replaced with a 6x-histadine tag to facilitate easy capture and the beta-2-microglobulin subunit is physically linked to the alpha chain with a glycine-serine linker to fuse the HLA dimer into a single amino acid sequence, hence a single chain dimer (SCD) **(Figure 3.1A**). The SCD is secreted into the media, where it is easily collected and captured with traditional immobilized metal-ion affinity chromatography (IMAC). Captured SCD is denatured to release the presented peptide while the SCD remains bound to the IMAC column. Eluted peptides are sequenced using mass spectroscopy to reflect naturally-presented peptides (**Figure 3.1B**).

Artemis, like all peptide discovery approaches, requires subsequent validation of immunogenicity, but saves time and resources by selecting peptides that have been

processed and presented. Use of the SCD avoids technically difficult standard immunoprecipitation methods that require detergent to solubilize the membrane bound HLAs; detergents easily ionize on the mass spectrometer and can obscure detection of peptides. Moreover, classical methods have relatively low peptide recovery compared to Artemis (*41*). Peptides recovered from Artemis have nearly identical characteristics to those identified with classical methodologies as described in depth by Finton et al. (*41*), although imaging studies have not been performed to confirm the SCD localizes to the same intracellular compartments. Peptides identified with both methods share similar peptide length distributions, HLA binding motifs, HLA binding predictions and are derived from all cell compartments (*41*).

While Artemis can expedite peptide discovery and ensure natural presentation, it also provides a model system to explore antigen processing and its impact on the immunopeptidome, defined as the sum of peptides available for presentation by HLAs (*42*).

The peptide processing and presentation pathway

Antigen processing and presentation is complex. For the purposes of this dissertation, I will focus on the classical pathway for class I HLA presentation. Conventional peptide processing begins in the cytosol where misfolded and ubiquitinated proteins are degraded by the proteasome into peptides 2-25 amino acids in length (*43*), which are then transported to the endoplasmic reticulum (ER) via the transporter associated with antigen presentation (TAP) (*44*). Once in the ER, peptides may be further trimmed from the N-terminus by ER amino peptidases (ERAPs) (*45*). Simultaneously, HLAs are synthesized and folded in the ER. Unoccupied HLAs are unstable and require additional proteins to stabilize the HLA before it binds a peptide. Peptides in the ER (either direct form the proteasome digestion or trimmed by ERAP) interact with empty HLAs. If the interaction is sufficiently strong, the peptide binds the HLA, forming a stable HLA-peptide complex. The complex traffics to the surface, where it presents the internal peptide to surveying cytotoxic T- cells (*46*)

Proteasomal degradation is the first step in antigen processing and presentation and is considered the primary source of class I HLA-displayed peptides. Importantly, we

observed modulation of the proteasome to evade TCR-T cell therapy in a patient case (*20*). Thus, it is a central component to the work described in this dissertation. It is important to note that several versions of the proteasome exist. The two prominent isoforms are the standard proteasome and the immunoproteasome. The standard proteasome is expressed in most somatic cells under basal conditions. The immunoproteasome is expressed in many immune cells, including antigen presenting cells and hematopoietic cells, but can be induced in most somatic cells under inflammatory conditions with IFN-γ exposure (*47*). These isoforms only differ by three catalytic beta-subunits: β1, β2, and β5 in the standard proteasome and β1i, β2i, and β5i in the immunoproteasome. The two isoforms generate distinct but overlapping peptide repertoires (**Figure 1.3**). Two intermediate forms of the proteasomes also exist, one with a single immunoproteasome subunit (β5i, β1, β2, (sIP)), and one with two immunoproteasome subunits (β5i, β1i, β2, (dIP)) (*48*).

Proteasomes not only provide peptides for recognition by the cytotoxic T cells, but they also maintain protein homeostasis. They are evolutionarily conserved, with homologues in yeast and prokaryotes (*49-51*). The proteasome catalytic core consists of four heptameric rings, two alpha rings and two beta rings stacked together – α- β-β-α– to form a hollow, barrel shaped structure referred to as the 20S proteasome. In prokaryotes, these rings consist of homogeneous alpha and beta subunits; in eukaryotes, the rings contain 7 unique α-subunits and 7 unique β-subunits (*49*). Humans have genes encoding additional β-subunits, including the immunoproteasome β-subunits, when expressed, these subunits can be substituted into the 20S core (*49*).

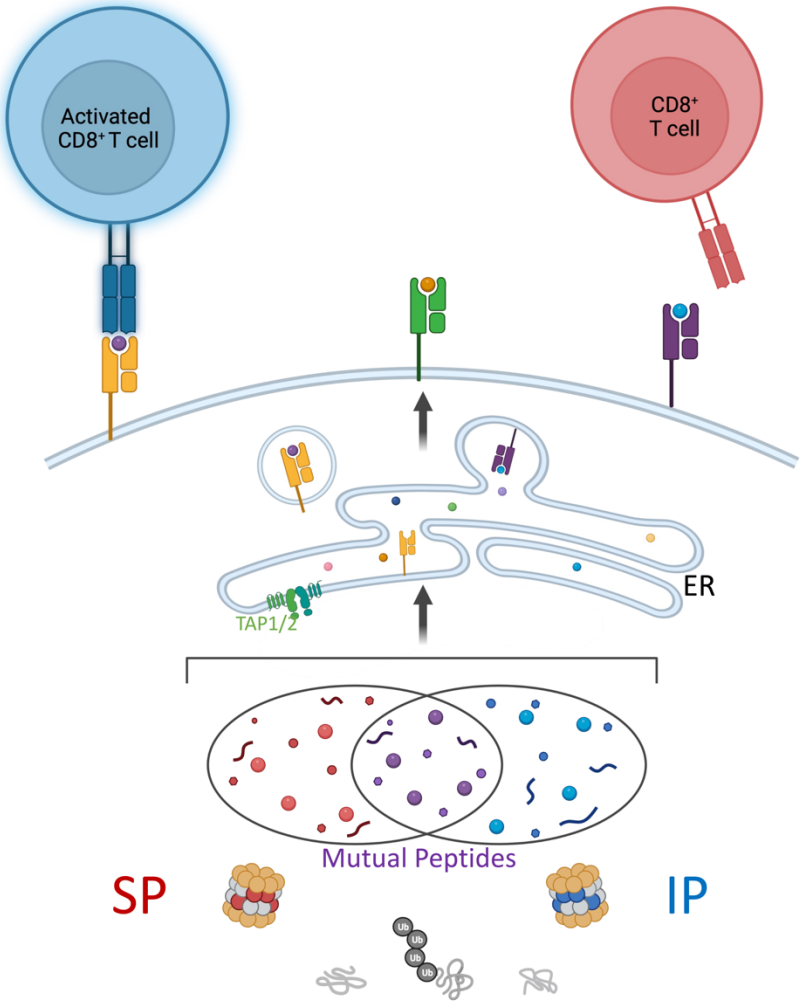


Figure 1.3 Schematic of antigen processing. Protein degradation by standard proteasome (SP) and immunoproteasome (IP) isoforms generate immunopeptidomes with 'distinct' and 'mutual' peptides. Peptides can be differentially recognized by TCRs.

Proteasome formation is a highly ordered process with several chaperones guiding each step. The alpha ring is the first component to be assembled (*50, 51*). Once the alpha ring is assembled, it serves as a scaffold for incorporation of the beta subunits. Each β-subunit incorporates in a specific order. For standard proteasome formation, the first β-subunit to incorporate is β2, followed by β3, β4, β5, β6, then β1 and lastly β7 (**Figure 1.4**). At this point half of a full proteasome complex is assembled, this half-

proteasome will associate with another half-proteasome. Pro-peptides on β1, β2, β5, β6, and β7 are then cleaved to form a mature, active proteasome (*51*). Additional regulatory caps may then bind to the 20S core.

The order of incorporation for immunoproteasome β-subunits is slightly different in comparison to the standard proteasome. When expressed, immuno-subunits β2i and β1i are the first subunits to incorporate and are mutually integrated; presence of β2i and β1i promotes subsequent incorporation of β5i to form a functional immunoproteasome (*52*) (**Figure 1.4**). Changes in proteasome isoform expression resulting in immune evasion have been described (*20, 53, 54*).

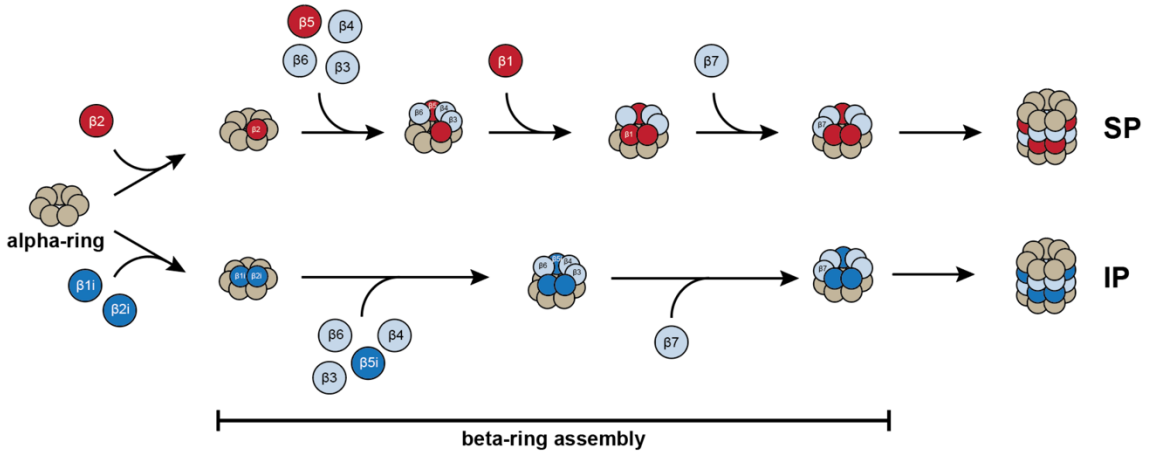


Figure 1.4 Simplified schematic of proteasome beta-ring assembly. Both SP and IP isoforms start with identical alpha-ring formation which serves as a scaffold for beta-ring formation. Top: SP beta-ring assembly. The first β-subunit to incorporate is β2 followed by β3, β4, β5, β6, then β1 and lastly β7 Bottom: IP beta-ring assembly. β2i and β1i are the first subunits to incorporate and are mutually integrated followed by β3, β4, β5, β6, and lastly β7. For both isoforms, two half-proteasomes come together to form a functional proteasome.

Lessons learned from relapsed tumors in the clinic

While contemporary TCR-T cell immunotherapies show promise, response rates remain sporadic (*20, 55*). Further research to untangle the molecular underpinnings driving response and resistance is required to improve TCR-T cell therapies. The unfortunate irony is that more can be learned from patients who do not respond to therapy or develop acquired resistance than those who have complete long-term

remissions. Comparing tumor before therapy and at relapse can reveal mechanisms of escape and offer valuable scientific insights.

Learning from therapy resistance is critical for advancing therapeutics. We observed several cases of acquired resistance in a clinical trial led by Drs. Aude Chapuis and Phil Greenberg at the Fred Hutchinson Cancer Center (FH-2498; NCT01640301). We took a deep dive into several case reports to identify mechanisms of resistance and pinpoint areas for improvement.

27 patients with high-risk AML (N=27) were enrolled and treated with TCR-T cells ($T_{TCR\text{-}C4}$) expressing a high-affinity TCR recognizing the $WT1_{126\text{-}134}$ [RMFPNAPYL] peptide presented by HLA-A2 (FH-2498; NCT01640301). In this trial, patients were enrolled in one of two arms: Arm 1 – 'prophylactic arm' with 12 patients or Arm 2 – 'treatment arm' with 15 patients. In both arms, patients received $T_{TCR\text{-}C4}$ cells after HCT and additional doses of $T_{TCR\text{-}C4}$ cells were administered if persistence of $T_{TCR\text{-}C4}$ cells fell below 3% of the total $CD8^+$ T cells. In the prophylactic arm, cells were administered after HCT but before any recurrent disease was detectable to prevent relapse. These patients performed remarkably well, with 100% relapse free survival at follow up (median follow up of 44 months) compared to 54% relapse-free survival in matched controls (N=88). Full results from Arm 1 have been published (*7*). Conversely, Arm 2 patients received $T_{TCR\text{-}C4}$ cells after relapsing post-HCT, thus $T_{TCR\text{-}C4}$ cells were administered in a therapeutic setting. While some patients from Arm 2 experienced periods of prolonged response followed by disease relapse, survival was not significantly ($p = 0.34$) improved over matched controls (**Figure1.5A**). Progressive disease in the presence of persistent $T_{TCR\text{-}C4}$ suggest immune evasion.

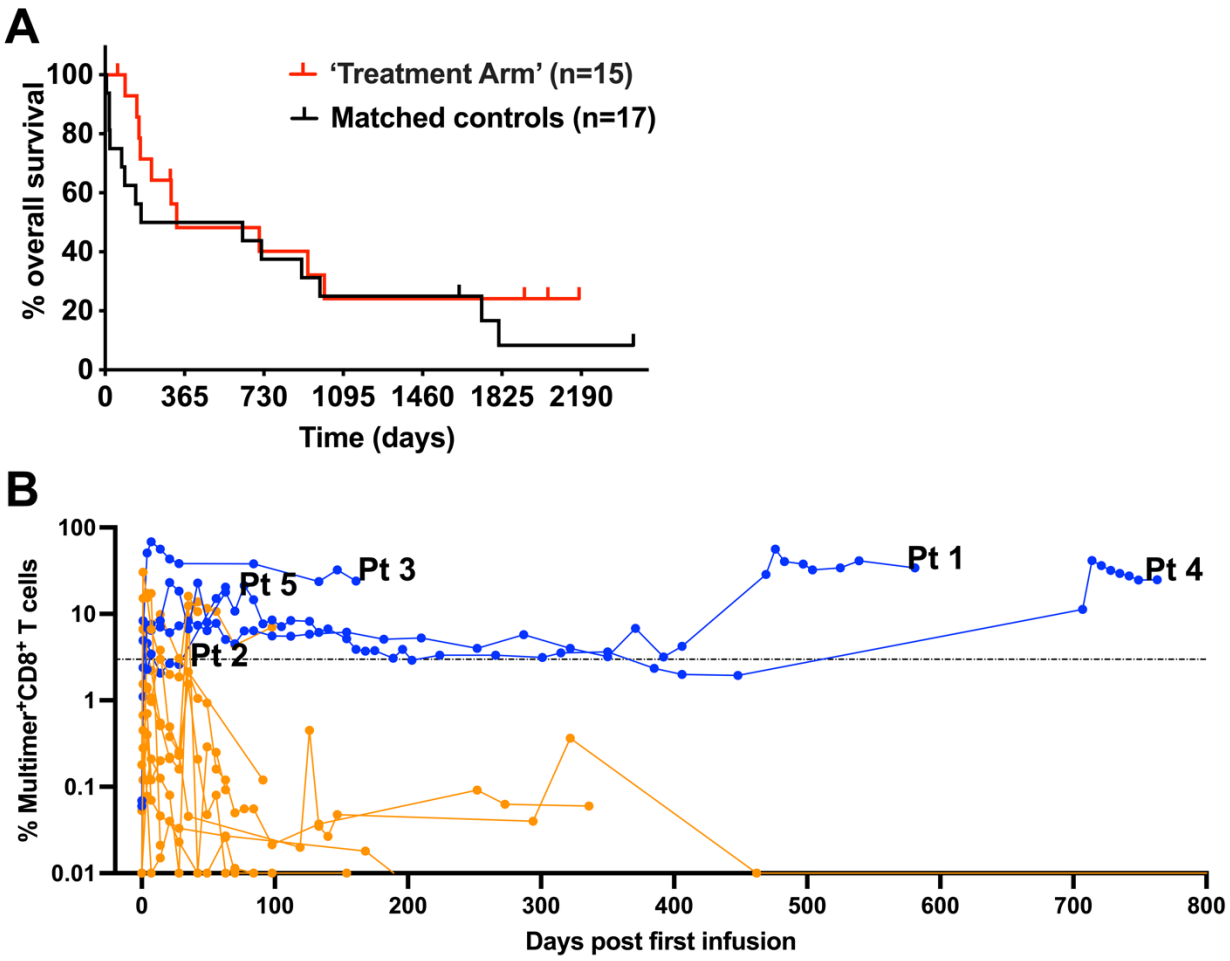


Figure 1.5 Arm 2 patient statistics. **(A)** Overall survival from relapse for Arm 2 patients and matched controls. p = 0.34 **(B)** $T_{TCR\text{-}C4}$ persistence as percent multimer$^+$ cells of CD8$^+$ cells in PBMC for patients with persisting (blue) or not (orange) therapeutic T cells. Patients in Table 1.1 are indicated.

Of the 15 patients in Arm 2, five had persistent T cells (**Figure 1.5B, blue**) with detectable, concurrent disease, creating a setting where cancer-immune interactions and mechanisms of immune escape could be studied. Because acquired resistance can stem from either T cell-intrinsic or tumor-intrinsic factors, we utilized scRNAseq to follow both $T_{TCR\text{-}C4}$ and AML cells through therapy. Upon preliminary analysis, all five patients showed different mechanisms of immune evasion (**Table 1**) highlighting the complex nature of immune surveillance and acquired resistance.

One patient (Patient 3) demonstrated immune evasion with relapsed AML, having acquired a partially differentiated phenotype. Another patient (Patient 4) had an

intriguing clinical course, tolerating an unusually high number of treatments with azacytidine. We predict azacytidine renders the AML visible to $T_{TCR\text{-}C4}$ through upregulation of WT1 (*56*). One patient (Patient 5) displayed signs of dysfunctional $T_{TCR\text{-}C4}$ cells. In depth analysis of these patients is beyond the scope of this thesis and will be discussed in an upcoming manuscript with Dr. Francesco Mazziotta. Patients 1 and 2 are a subject of this dissertation and discussed in the following chapter.

Patient	$T_{TCR\text{-}C4}$ Persistance	WT1 Expression	β1i Expression	Predicted Mechanism of Escape	Mechanism Confirmed
Pt 1	Yes	Yes	No	Anitgen presentation, AML β1i loss	Yes
Pt 2	Yes	Yes	Yes	Heterogeneity and high disease burden	Yes
Pt 3	Yes	Yes	Yes	AML differentiation, Relpase as AMML	Ongoing
Pt 4	Yes	Yes	Yes	Cyclic WT1 expression, azacitidine	Ongoing
Pt 5	Yes	Yes	Yes	$T_{TCR\text{-}C4}$ cell dysfunction	Ongoing

Table 1 Predicted mechanisms of escape. Predicted mechanisms of immune evasion are listed for five Arm 2 patients who showed persistent $T_{TCR\text{-}C4}$ cell in the presence of relapsing leukemia.

I will first describe the intriguing clinical case of AML Patient 1, specifically discovering how their AML, which initially responded to TCR-T cell therapy with $T_{TCR\text{-}C4}$, became resistant. Knowledge from this clinical case prompted a deep investigation into antigen processing. I will then use the emerging peptide discovery platform, Artemis, to investigate the contribution of the standard proteasome and immunoproteasome to the immunopeptidome, as well as reveal ‘distinct’ and ‘mutual’ peptides from AML-associated antigens. Findings from these studies may inform the next iteration of TCR-T cell target peptide discovery for improved cell-based immunotherapies. Finally, at the core of the work described in this dissertation are people; we strive to use this knowledge to improve patient outcomes.

Acknowledgements.
Figures 1.1, 1.2, 1.3, and 1.4 were created with BioRender.com

References

1. H. Sung *et al.*, Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209-249 (2021).
2. R. L. Siegel, K. D. Miller, H. E. Fuchs, A. Jemal, Cancer statistics, 2022. *CA Cancer J Clin* **72**, 7-33 (2022).
3. S. C. Wei, C. R. Duffy, J. P. Allison, Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. *Cancer Discov* **8**, 1069-1086 (2018).
4. N. P. Restifo, M. E. Dudley, S. A. Rosenberg, Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat Rev Immunol* **12**, 269-281 (2012).
5. C. H. June *et al.*, Engineered T cells for cancer therapy. *Cancer Immunol Immunother* **63**, 969-975 (2014).
6. C. J. Turtle *et al.*, CD19 CAR-T cells of defined CD4+:CD8+ composition in adult B cell ALL patients. *J Clin Invest* **126**, 2123-2138 (2016).
7. A. G. Chapuis *et al.*, T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. *Nat Med* **25**, 1064-1072 (2019).
8. A. P. Rapoport *et al.*, NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. *Nat Med* **21**, 914-921 (2015).
9. P. F. Robbins *et al.*, Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1. *J Clin Oncol* **29**, 917-924 (2011).
10. L. A. Johnson *et al.*, Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. *Blood* **114**, 535-546 (2009).
11. N. B. Nagarsheth *et al.*, TCR-engineered T cells targeting E7 for patients with metastatic HPV-associated epithelial cancers. *Nat Med* **27**, 419-425 (2021).
12. U. Creutzig, M. A. Kutny, R. Barr, R. F. Schlenk, R. C. Ribeiro, Acute myelogenous leukemia in adolescents and young adults. *Pediatr Blood Cancer* **65**, e27089 (2018).
13. J. W. Vardiman, N. L. Harris, R. D. Brunning, The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood* **100**, 2292-2302 (2002).
14. P. van Galen *et al.*, Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity. *Cell* **176**, 1265-1281 e1224 (2019).
15. R. M. Shallis, R. Wang, A. Davidoff, X. Ma, A. M. Zeidan, Epidemiology of acute myeloid leukemia: Recent progress and enduring challenges. *Blood Rev* **36**, 70-87 (2019).
16. M. S. Tallman *et al.*, Acute Myeloid Leukemia, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* **17**, 721-749 (2019).
17. A. K. Burnett *et al.*, Curability of patients with acute myeloid leukemia who did not undergo transplantation in first remission. *J Clin Oncol* **31**, 1293-1301 (2013).
18. A. M. Dickinson *et al.*, Graft-versus-Leukemia Effect Following Hematopoietic Stem Cell Transplantation for Leukemia. *Front Immunol* **8**, 496 (2017).
19. V. Ramachandran, S. S. Kolli, L. C. Strowd, Review of Graft-Versus-Host Disease. *Dermatol Clin* **37**, 569-582 (2019).

20. M. C. Lahman *et al.*, Targeting an alternate Wilms' tumor antigen 1 peptide bypasses immunoproteasome dependency. *Sci Transl Med* **14**, eabg8070 (2022).
21. K. G. Paulson *et al.*, Acquired cancer resistance to combination immunotherapy from transcriptional loss of class I HLA. *Nat Commun* **9**, 3868 (2018).
22. C. H. June, R. S. O'Connor, O. U. Kawalekar, S. Ghassemi, M. C. Milone, CAR T cell immunotherapy for human cancer. *Science* **359**, 1361-1365 (2018).
23. A. Cossarizza *et al.*, Guidelines for the use of flow cytometry and cell sorting in immunological studies (third edition). *Eur J Immunol* **51**, 2708-3145 (2021).
24. M. Stoeckius *et al.*, Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* **14**, 865-868 (2017).
25. J. D. Buenrostro, B. Wu, H. Y. Chang, W. J. Greenleaf, ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 21-21 29 29 (2015).
26. F. Mair *et al.*, A Targeted Multi-omic Analysis Approach Measures Protein Expression and Low-Abundance Transcripts on the Single-Cell Level. *Cell Rep* **31**, 107499 (2020).
27. A. Liso *et al.*, Nucleophosmin leukaemic mutants contain C-terminus peptides that bind HLA class I molecules. *Leukemia* **22**, 424-426 (2008).
28. J. Greiner *et al.*, Mutated regions of nucleophosmin 1 elicit both CD4(+) and CD8(+) T-cell responses in patients with acute myeloid leukemia. *Blood* **120**, 1282-1289 (2012).
29. D. I. van der Lee *et al.*, Mutated nucleophosmin 1 as immunotherapy target in acute myeloid leukemia. *J Clin Invest* **129**, 774-785 (2019).
30. D. Prins, C. Gonzalez Arias, T. Klampfl, J. Grinfeld, A. R. Green, Mutant Calreticulin in the Myeloproliferative Neoplasms. *Hemasphere* **4**, e333 (2020).
31. T. M. Schmitt *et al.*, Generation of higher affinity T cell receptors by antigen-driven differentiation of progenitor T cells in vitro. *Nat Biotechnol* **35**, 1188-1195 (2017).
32. L. Bergmann, U. Maurer, E. Weidmann, Wilms tumor gene expression in acute myeloid leukemias. *Leuk Lymphoma* **25**, 435-443 (1997).
33. M. A. Cheever *et al.*, The prioritization of cancer antigens: a national cancer institute pilot project for the acceleration of translational research. *Clin Cancer Res* **15**, 5323-5337 (2009).
34. D. L. Stirewalt *et al.*, Identification of genes with abnormal expression changes in acute myeloid leukemia. *Genes Chromosomes Cancer* **47**, 8-20 (2008).
35. S. Ochsenreither *et al.*, Cyclin-A1 represents a new immunogenic targetable antigen expressed in acute myeloid leukemia stem cells with characteristics of a cancer-testis antigen. *Blood* **119**, 5492-5501 (2012).
36. M. Nielsen, M. Andreatta, NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med* **8**, 33 (2016).
37. B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* **48**, W449-W454 (2020).

38. A. Nelde, D. J. Kowalewski, S. Stevanovic, Purification and Identification of Naturally Presented MHC Class I and II Ligands. *Methods Mol Biol* **1988**, 123-136 (2019).
39. E. Barnea *et al.*, Analysis of endogenous peptides bound by soluble MHC class I molecules: a novel approach for identifying tumor-specific antigens. *Eur J Immunol* **32**, 213-222 (2002).
40. S. Sarkizova *et al.*, A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat Biotechnol* **38**, 199-209 (2020).
41. K. A. K. Finton *et al.*, ARTEMIS: A Novel Mass-Spec Platform for HLA-Restricted Self and Disease-Associated Peptide Discovery. *Front Immunol* **12**, 658372 (2021).
42. J. A. Vizcaino *et al.*, The Human Immunopeptidome Project: A Roadmap to Predict and Treat Immune Diseases. *Mol Cell Proteomics* **19**, 31-49 (2020).
43. A. F. Kisselev, T. N. Akopian, K. M. Woo, A. L. Goldberg, The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J Biol Chem* **274**, 3363-3371 (1999).
44. J. C. Roder, F. Meyer, H. Pritzkow, An unusual hexanickel cage complex with mu- and mu 3-chloro bridges and an interstitial mu 6-chloride. *Chem Commun (Camb)*, 2176-2177 (2001).
45. I. Evnouchidou, M. Weimershaus, L. Saveanu, P. van Endert, ERAP1-ERAP2 dimerization increases peptide-trimming efficiency. *J Immunol* **193**, 901-908 (2014).
46. K. L. Rock, E. Reits, J. Neefjes, Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends Immunol* **37**, 724-737 (2016).
47. A. Rouette *et al.*, Expression of immunoproteasome genes is regulated by cell-intrinsic and -extrinsic factors in human cancers. *Sci Rep* **6**, 34019 (2016).
48. J. Abi Habib *et al.*, Efficiency of the four proteasome subtypes to degrade ubiquitinated or oxidized proteins. *Sci Rep* **10**, 15765 (2020).
49. K. Tanaka, The proteasome: overview of structure and functions. *Proc Jpn Acad Ser B Phys Biol Sci* **85**, 12-36 (2009).
50. S. Murata, H. Yashiroda, K. Tanaka, Molecular mechanisms of proteasome assembly. *Nat Rev Mol Cell Biol* **10**, 104-115 (2009).
51. M. J. Kunjappu, M. Hochstrasser, Assembly of the 20S proteasome. *Biochim Biophys Acta* **1843**, 2-12 (2014).
52. M. Groettrup, S. Standera, R. Stohwasser, P. M. Kloetzel, The subunits MECL-1 and LMP2 are mutually required for incorporation into the 20S proteasome. *Proc Natl Acad Sci U S A* **94**, 8970-8975 (1997).
53. E. S. Schultz *et al.*, The production of a new MAGE-3 peptide presented to cytolytic T lymphocytes by HLA-B40 requires the immunoproteasome. *J Exp Med* **195**, 391-399 (2002).
54. M. Keller *et al.*, The proteasome immunosubunits, PA28 and ER-aminopeptidase 1 protect melanoma cells from efficient MART-126-35 -specific T-cell recognition. *Eur J Immunol* **45**, 3257-3268 (2015).

55. P. Shafer, L. M. Kelly, V. Hoyos, Cancer Therapy With TCR-Engineered T Cells: Current Strategies, Challenges, and Prospects. *Front Immunol* **13**, 835762 (2022).
56. S. Harada *et al.*, Dynamic change in peripheral blood WT1 mRNA levels within three cycles of azacitidine predict treatment response in patients with high-risk myelodysplastic syndromes. *Ann Hematol* **101**, 1239-1250 (2022).

Chapter 2- Immune evasion through immunoproteasome deficiency

This chapter has been adapted from Lahman, M.C., et al. *Targeting an alternate Wilms' Tumor Antigen-1 peptide bypasses immunoproteasome dependency.* Sci. Trans. Med. 2022. PMID: 35138909

Abstract

Designing effective anti-leukemic immunotherapy requires understanding mechanisms governing tumor control or resistance. Here we report a mechanism of escape from immunologic targeting in a patient with acute myeloid leukemia (AML), who relapsed one year following immunotherapy with engineered T cells expressing a human leukocyte antigen A*02 (HLA-A2) restricted T cell receptor (TCR) specific for a Wilms' Tumor Antigen 1 peptide, $WT1_{126\text{-}134}$, ($T_{TCR\text{-}C4}$). Resistance occurred despite persistence of functional therapeutic T cells and continuous expression of WT-1 and HLA-A2 by the patient's AML cells. Analysis of the recurrent AML revealed expression of the standard proteasome, but limited expression of the immunoproteasome, specifically the beta-subunit 1i (β1i), which is required for presentation of $WT1_{126\text{-}134}$. Analysis of a second patient treated with $T_{TCR\text{-}C4}$ demonstrated specific loss of AML cells co-expressing β1i and WT1. To determine if the WT1 protein continued to be processed and presented in the absence of immunoproteasome processing, we identified an HLA-A2 restricted peptide ($WT1_{37\text{-}45}$) generated by immunoproteasome deficient cells, including WT1-expressing solid tumor lines. We isolated and tested a TCR targeting this alternative peptide. T cells expressing this TCR ($T_{TCR37\text{-}45}$) killed the first patients' relapsed AML resistant to $WT1_{126\text{-}134}$ targeting, as well as other primary AML, in vitro. $T_{TCR37\text{-}45}$ controlled solid tumor lines lacking immunoproteasome subunits both *in vitro* and in an NSG mouse model. As proteasome composition can vary in AML, defining and preferentially targeting such proteasome-independent peptides may maximize

therapeutic efficacy and potentially circumvent AML immune evasion by proteasome-related immunoediting.

Introduction

Allogeneic hematopoietic cell transplant (HCT) can cure some patients with high-risk AML, which is largely a result of donor T cells eliminating leukemia by recognizing minor histocompatibility antigens or less commonly leukemia-associated antigens (*1*). This graft-versus-leukemia (GVL) effect is sporadic, entirely reliant on donor immune cells recognizing differences between host and donor cells and is not predictable pre-HCT. Leukemic relapse post-HCT remains the leading cause of death, especially in patients who enter HCT with features of high-risk AML, in part due to the inability to ensure a GVL effect.

Adoptive therapy provides T cells re-directed towards defined tumor-associated antigens. Most notably, T cells engineered with chimeric antigen receptors (CARs) targeting surface proteins, such as CD19 or B cell maturation antigen (BCMA), have shown stunning clinical results in B cell malignancies (*2*), but identifying AML surface proteins that are unique or safely targetable remains a challenge (*3*). By contrast, T cell receptors (TCRs) can recognize a broader set of peptides, which can be derived from intracellular and surface proteins (*4*). To achieve consistent antileukemic activity, we sought to target Wilms' Tumor Antigen 1 (WT1), an intracellular transcription factor overexpressed in most leukemic cells but lowly expressed in normal adult tissues (*5*). The $WT1_{126\text{-}134}$ peptide was successfully eluted from a human leukocyte antigen (HLA)-A*02:01 positive leukemia, confirming presentation, although it is not routinely present in peptide elution sets (*6*) suggesting there may not be abundant presentation of this peptide (*7*). Still, T cells targeting $WT1_{126\text{-}134}$ have contributed to maintenance of sustained complete responses in vaccine trials and exhibited a direct anti-leukemic effect in our group's prior clinical study where donor-derived WT1-specific $CD8^+$ T cell clones were transferred into patients post-HCT (*5, 8, 9*).

Based on these results, we isolated a high affinity TCR (TCR-C4) restricted to HLA-A*02:01 and specific for the $WT1_{126\text{-}134}$ epitope and cloned it into a clinically validated lentiviral vector. Donor Epstein Barr Virus (EBV)-specific $CD8^+$ T cells were

selected as substrates to both minimize risk of graft-versus-host disease (GVHD) and enhance survival of therapeutic cells – EBV-specific $CD8^+$ T cells were transduced with TCR-C4 ($T_{TCR\text{-}C4}$) (*10*) (NCT01640301). Following prophylactic administration to patients with HLA-matched high-risk AML post-HCT, $T_{TCR\text{-}C4}$ demonstrated in vivo persistence and sustained remissions were achieved in all patients despite risk of relapse that exceeded 50% (*10*). However, a formal assessment of direct $T_{TCR\text{-}C4}$ anti-tumor activity was challenging as leukemia was undetectable at the time of $T_{TCR\text{-}C4}$ transfer.

By re-directing T cell recognition towards a known tumor epitope, adoptive therapy trials offer the added benefit of limiting the tumor-T cell interaction to a defined target. Paradoxically, the requirements for a therapeutic response can often be better elucidated not in patients with complete sustained responses, but rather in patients whose tumors acquire resistance to the immune intervention. This can reveal critical response and evasion mechanisms (*11*).

Tumor escape after CAR-T cell infusions has commonly reflected loss of the targeted antigen epitope from the cell surface (*12*). For TCR-mediated interventions, this is best paralleled by loss of expression of the targeted antigen or the HLA-restricting-allele (*13*). HLA loss has rarely been associated with post-HCT relapse in AML (*13*) and WT1 mutation was not a major contributor to immune escape in patients with AML who received WT1 peptide vaccination (*14*) or infusions of clonal WT1-specific T cells in our previous study (*9*). Instead, and specific to targeted TCR-T cell therapy, presentation of the targeted epitope can be impacted by proteasome processing (*15*). Indeed, the TCR_{C4}-targeted epitope ($WT1_{126\text{-}134}$) is reliant on the immunoproteasome (IP) (*16*), the dominant isoform in hemopoietic cells (*17*).

To investigate potential mechanisms of AML resistance to $T_{TCR\text{-}C4}$, one patient was identified who relapsed after a second transplant. After receiving $T_{TCR\text{-}C4}$, this patient experienced an apparent sustained remission but then relapsed one year later despite large numbers of persistent $T_{TCR\text{-}C4}$ cells. We performed high-dimensional analysis of the persistent T cells and recurrent AML, which confirmed $T_{TCR\text{-}C4}$ were functional, and that AML continued to express both the restricting HLA allele (HLA-A*02) and unmutated WT1. Further in-depth analysis and observations in a second patient who received $T_{TCR\text{-}C4}$ with circulating $WT1^+$ AML point to a previously

undescribed mechanism of *in vivo* AML escape from a targeted T cell response, involving changes in antigen-processing of the targeted epitope. We then pursued an alternative HLA-A*02:01-restricted epitope, $WT1_{37\text{-}45}$, which is less susceptible to the observed mechanism of resistance and thus represents a promising TCR-T cell immunotherapy target. Our data underscore the necessity of elucidating acquired resistance mechanisms to inform rational design of more effective TCR-T cell therapies not only for AML, but for immunotherapy more broadly.

Results

An unusual clinical course calls for deeper investigation

A 27-year-old patient developed recurrent AML with both extramedullary chloromas and minimal residual disease (MRD) in the bone marrow 7 months after receiving a second matched, unrelated donor HCT for relapsed refractory disease (**Figure 2.1A**). The patient was treated with local radiation and systemic salvage chemotherapy, which achieved remission as evidenced by no evaluable disease (NED), and then received an infusion of donor-derived $T_{TCR\text{-}C4}$ in Arm 2 of this study (*10*). Based on the history of relapsing after a second allo-HCT, he was at extremely high risk of early relapse and death, as AML relapse post-HCT is associated with a median survival of about 3 months (*18*). Surprisingly, the patient remained in remission without further therapy for 368 days before exhibiting recurrent AML with both a bladder chloroma and flow-cytometric evidence of MRD (0.04% AML cells) in the bone marrow. He then received reinduction chemotherapy followed by a second $T_{TCR\text{-}C4}$ infusion on day 471 after the 1st infusion with T cells derived from the same pool of transduced cells originally infused, but the leukemia continued to relentlessly progress. This unusual clinical course with a long 1-year remission after multiple relapses led us to investigate potential immunologically related mechanisms of remission and relapse.

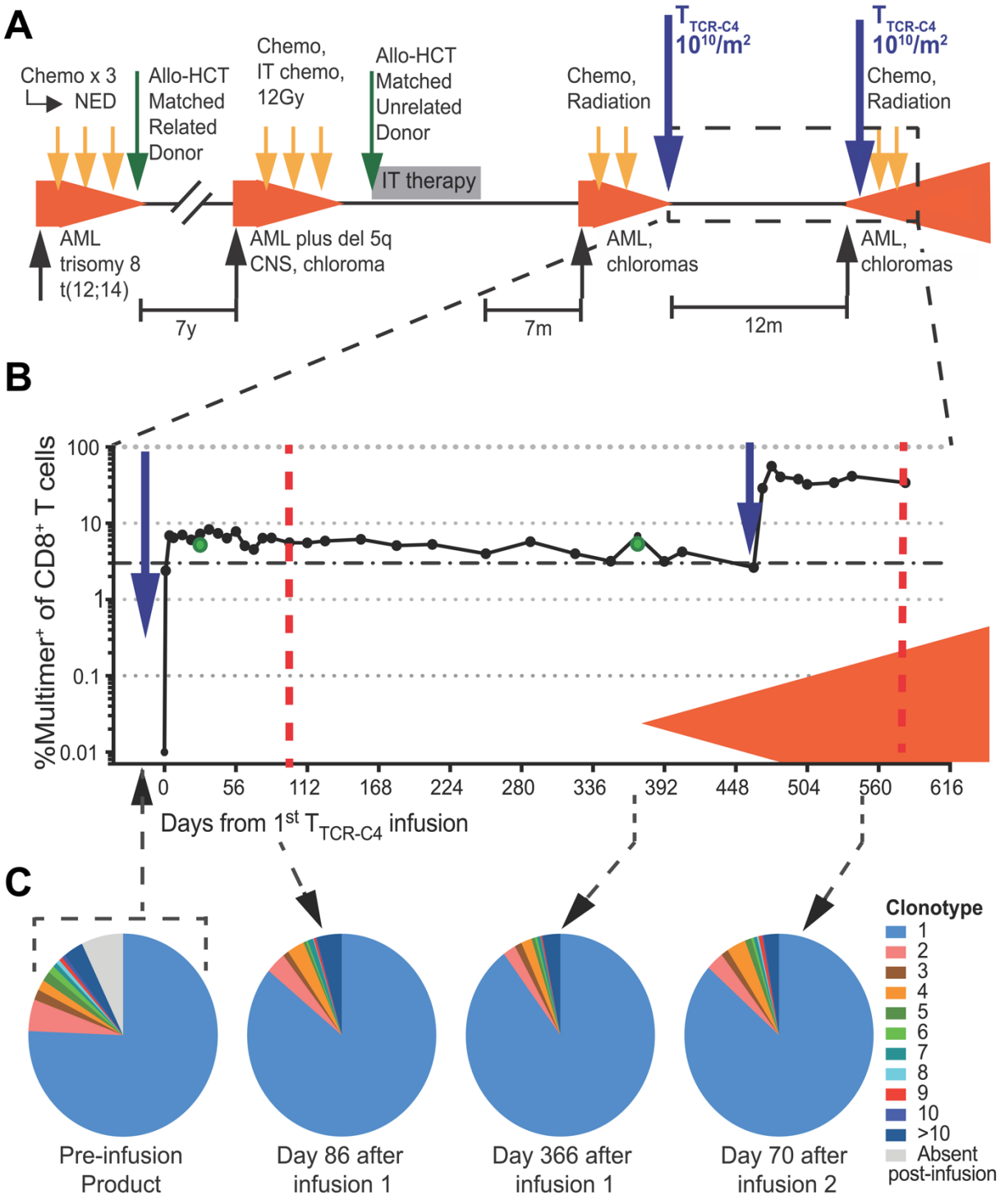

A
Chemo x 3
NED
Allo-HCT
Matched
Related
Donor
Chemo,
IT chemo,
12Gy
Allo-HCT
Matched
Unrelated
Donor
IT therapy
Chemo,
Radiation
T TCR-C4 10^10/m^2
Chemo,
Radiation
T TCR-C4 10^10/m^2
AML
trisomy 8
t(12;14)
AML plus del 5q
CNS, chloroma
AML,
chloromas
AML,
chloromas
7y
7m
12m
B
%Multimer+ of CD8+ T cells
100
10
1
0.1
0.01
0
56
112
168
224
280
336
392
448
504
560
616
Days from 1st T TCR-C4 infusion
C
Clonotype
1
2
3
4
5
6
7
8
9
10
>10
Absent
post-infusion
Pre-infusion
Product
Day 86 after
infusion 1
Day 366 after
infusion 1
Day 70 after
infusion 2

Figure 2.1 $T_{TCR\text{-}C4}$ persist in the presence of relapsed AML. (A) Timeline of patient 1's treatment regimens. Chemo, chemotherapy; radiation, radiation therapy; IT, intrathecal; NED, no evidence of disease; CNS, central nervous system. **(B)** Percent multimer$^+$ of CD8$^+$ T cells in PBMCs (solid circles) and bone marrow (green circles) collected before and at defined timepoints after $T_{TCR\text{-}C4}$ infusions are shown. The orange shaded area indicates presence of AML. Dotted red lines represent days 100 and 110 (or day 581 after the 1st) after 1st and 2nd infusion respectively. **(C)** Pie charts representing individual clonotypes composing the pre-infusion $T_{TCR\text{-}C4}$ product (left) and at indicated timepoints (right) after both infusions are shown.

Therapeutic T cells persist through remission and relapse

Both infusions of 10^{10} $T_{TCR\text{-}C4}/m^2$ (on days 0 and 471) were administered without prior lymphodepleting chemotherapy. $T_{TCR\text{-}C4}$ in peripheral blood reached peak frequencies of 8.2% and 56% of CD8$^+$ T cells after the 1st and 2nd infusion, respectively (**Figure 2.1B**). Overall, $T_{TCR\text{-}C4}$ frequency remained above 3% with sustained, robust absolute $T_{TCR\text{-}C4}$ counts (**Figure 2.2A**). Comparable high $T_{TCR\text{-}C4}$ frequencies were observed in the bone marrow at days 28 and 368 (**Figure 2.1B, green circles**). These high proportions of persistent transferred $T_{TCR\text{-}C4}$ are analogous to results obtained in patients who received $T_{TCR\text{-}C4}$ post-HCT prophylactically to prevent relapse (*10*). Both infusion products were expanded immediately prior to infusion from identical pre-infusion aliquots. Infusions primarily (>75%) consisted of a single expanded clonotype that also comprised majority of persistent $T_{TCR\text{-}C4}$ after both infusions (**Figure 2.12C and Figure 2.2B**). $T_{TCR\text{-}C4}$ surface expression of memory-associated markers (CD45RO, CD27, CD28, and CD62L (*19*)) remained similar after both $T_{TCR\text{-}C4}$ infusions (**Figure 2.3A**). Ex vivo analysis of recovered $T_{TCR\text{-}C4}$ revealed the cells were functional, maintaining cytokine production in response to targets presenting the cognate $WT1_{126\text{-}134}$ peptide during the period of apparent remission (day 114) and in the context of progressive AML (day 70 or day 541 after the 2nd or 1st infusion, respectively) (**Figure 2.3B**).

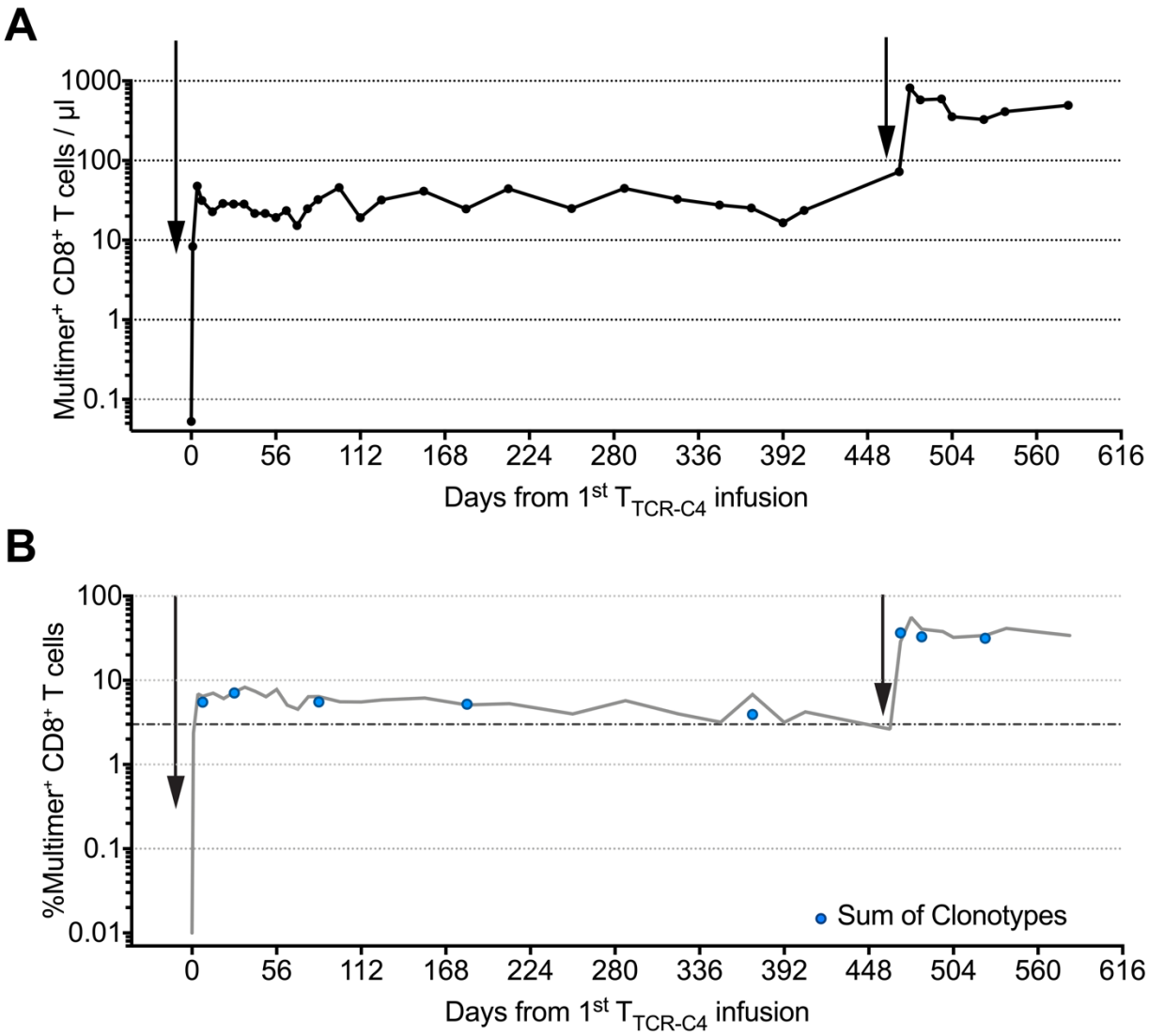


Figure 2.2 Patient 1 *in vivo* persistence kinetics of transferred $T_{TCR\text{-}C4}$. (A) Multimer^{+}CD8^{+} T cells/µl (y-axis) were enumerated in PBMCs collected after infusion (x-axis, day relative to first infusion). Time of infusions are indicated with black arrows. **(B)** The relationship over time after infusion (x-axis) between the percent multimer^{+}CD8^{+}T cells (gray line) in PBMCs and the sum of detected frequencies of individual infusion product clonotypes (blue circles) is shown. The percentage of multimer^{+}CD8^{+} T cells (y-axis) were measured in PBMCs collected after infusions (x-axis, day relative to first infusion). Time of infusions indicated with black arrows. The thick dashed line indicates the persistence cutoff defined in this trial as 3% multimer^{+}CD8^{+} T cells.

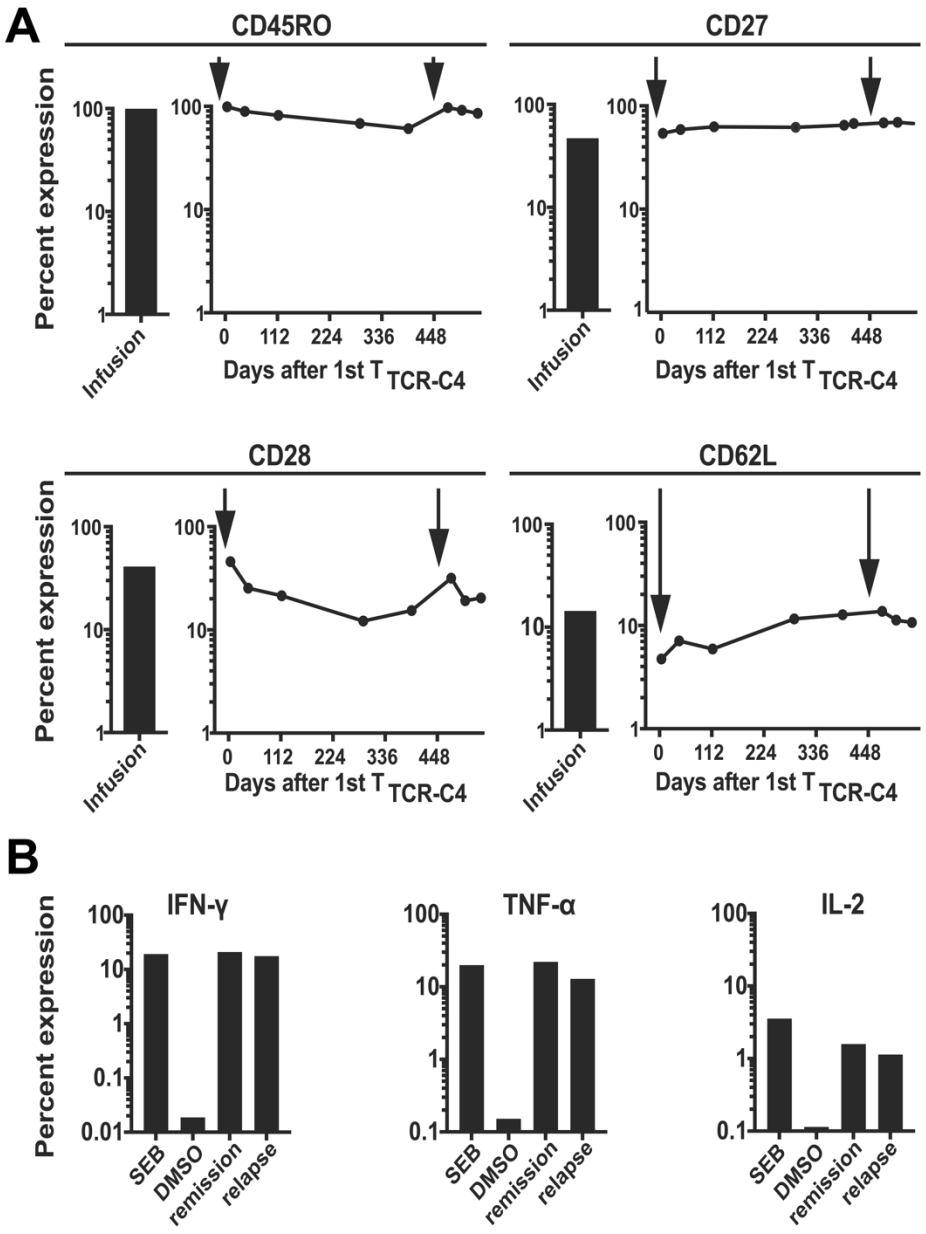


Figure 2.3 $T_{TCR\text{-}C4}$ retain phenotype and function in the presence of relapsed AML. (A) Percent expression of CD45RO, CD27, CD28 and CD62L on infusion products (bar graphs to the left) and at days 4, 42, 114, 289, 393, 478, 508 and 542 after the 1st infusion (graphs to the right) are shown. The three last timepoints are also days 7, 35 and 71 after the second infusion. Timing of infusions are indicated by black arrows. **(B)** Percent expression of IFN-γ, TNF-α and IL-2 (functional markers) in response to ex vivo stimulation $WT1_{126\text{-}134}$ peptide (1nM) during remission (day 114 after 1st infusion) and relapse (day 581 after 1st infusion, day 110 after 2nd). Maximum and minimal cytokine expression following exposure to Staphylococcal enterotoxin B (SEB) and dimethylsufoxide (DMSO) are shown for each cytokine.

Single cell RNA sequencing of $T_{TCR\text{-}C4}$ transcriptome in line with recent activation at remission

To investigate if transcriptomes of circulating $T_{TCR\text{-}C4}$ differed between clinical remission and relapse, single-cell RNA sequencing (scRNAseq) (*20*) was performed on available peripheral blood mononuclear cells (PBMCs) 100 days after the 1st $T_{TCR\text{-}C4}$ infusion (remission) and 110 days after the 2nd infusion (relapse), which represents 581 days after the 1st infusion (**Figure 2.1B, dotted red lines).** Data were generated from a single sequencing run containing both samples, reads were aggregated together, and clusters were delineated (**Figure 2.4A**) and then visualized separately, based on barcodes unique to the timepoint (**Figure 2.4B**). Surprisingly, during clinical remission, $CD8^+$ cells grouped into 2 separate and distinct clusters in an unsupervised transcriptional profile analysis (**Figure 2.4C, light blue and purple clusters**), but the light blue cluster was nearly absent at relapse (**Figure 2.4D**). During remission, 96% of $T_{TCR\text{-}C4}$ as identified by TCR-C4 transgene expression, were grouped in the light blue cluster whereas endogenous $CD8^+$ T cells grouped in the purple cluster (**Figure 2.4E)**. In contrast, during relapse in the presence of circulating blasts, 91% of $T_{TCR\text{-}C4}$ grouped with the purple endogenous $CD8^+$ cells (**Figure 2.4F**). To probe for differences between $T_{TCR\text{-}C4}$ at remission and relapse, an unsupervised differential gene expression analysis was performed. The top 10 significantly differentially expressed genes ($p<0.05$ and logFC $>|\log_2(1.5)|$) showed $T_{TCR\text{-}C4}$ cells at remission had higher expression of genes associated with T cell activation, including lymphotoxin-β [*LTB*], granzyme K [*GZMK*], human leukocyte antigen (HLA)-DR [*HLA-DR*] and CD74 [*CD74*], which mediates assembly and trafficking of HLA class II complexes – (**Figure 2.4G**) (*21-24*). Gene set analysis of effector cytokine/chemokine genes(*25*) (interferon (IFN)-γ [*IFNG*], C-X-C motif chemokine receptor 1 [*CX3CR1*], C-C motif chemokine ligand 5 [*CCL5*] and tumor necrosis factor alpha (TNF-α) [*TNF*]), cytolytic effector genes (*25*) (granzymes A [*GZMA*], B [*GZMB*] and H [G*ZMH*] and perforin [*PRF1*]), activation/proliferation genes (*26*) (Jun proto-oncogene [*JUN*], Fos proto-oncogene [*FOS*], CD69 [*CD69*], nuclear receptor subfamily 4 group A member 1 [*NR4A1*], interleukin (IL)-2 receptor subunit gamma [*IL2RG*], and IL-2 receptor subunit alpha [*IL2RA*]) and activation/exhaustion genes (programmed cell death receptor-1 [*PDCD1*], T-Cell immunoreceptor with Ig and

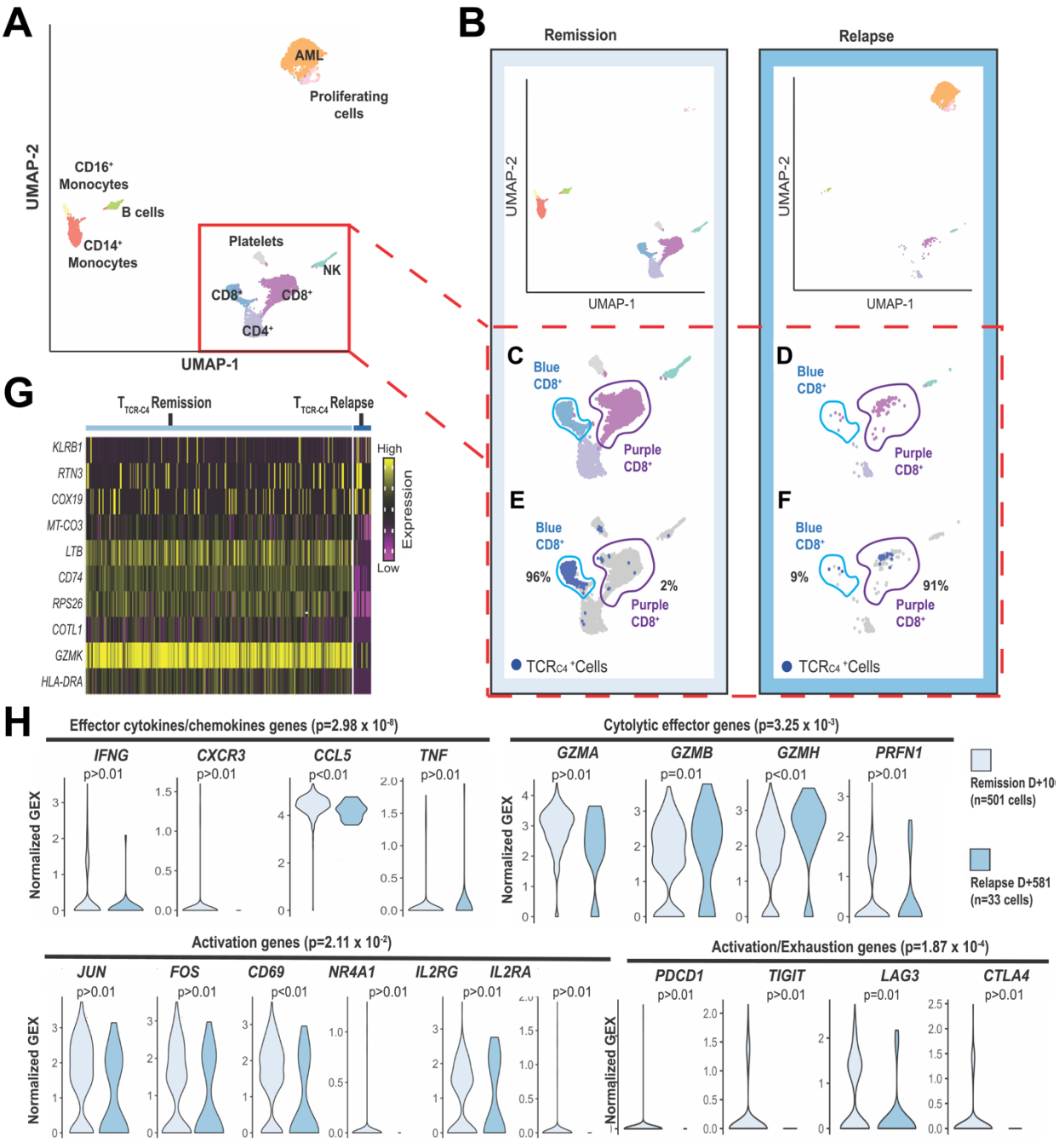

A
AML
Proliferating cells
CD16+ Monocytes
B cells
CD14+ Monocytes
Platelets
NK
CD8+
CD8+
CD4+
UMAP-2
UMAP-1
B
Remission
Relapse
UMAP-2
UMAP-1
C
Blue CD8+
Purple CD8+
D
Blue CD8+
Purple CD8+
E
Blue CD8+
96%
2%
Purple CD8+
F
Blue CD8+
9%
91%
Purple CD8+
TCRC4+Cells
TCRC4+Cells
G
TTCR-C4 Remission
TTCR-C4 Relapse
KLRB1
RTN3
COX19
MT-CO3
LTB
CD74
RPS26
COTL1
GZMK
HLA-DRA
High
Expression
Low
H
Effector cytokines/chemokines genes (p=2.98 x 10-8)
IFNG p>0.01
CXCR3 p>0.01
CCL5 p<0.01
TNF p>0.01
Normalized GEX
Cytolytic effector genes (p=3.25 x 10-3)
GZMA p>0.01
GZMB p=0.01
GZMH p<0.01
PRFN1 p>0.01
Normalized GEX
Remission D+10
(n=501 cells)
Relapse D+581
(n=33 cells)
Activation genes (p=2.11 x 10-2)
JUN p>0.01
FOS p>0.01
CD69 p<0.01
NR4A1 p>0.01
IL2RG p>0.01
IL2RA p>0.01
Normalized GEX
Activation/Exhaustion genes (p=1.87 x 10-4)
PDCD1 p>0.01
TIGIT p>0.01
LAG3 p=0.01
CTLA4 p>0.01
Normalized GEX

Figure 2.4 $T_{TCR\text{-}C4}$ present an activated transcriptional profile during remission but not during relapse. (A) UMAP visualization of PBMCs from both the remission (100 days after 1st infusion) and relapse (110 days after the second infusion or 581 days after the 1st) (red dotted lines **Figure 2.1B**) samples are visualized together. PBMCs (n=7704) clustered into populations as indicated by labels. Clustering biostatistical analysis is described in methods **(B)** UMAP visualization of the separated PBMCs obtained at remission (left) and relapse (right). **(C)** Close-ups are shown for clusters containing T cells during remission with identification of two distinct $CD8^+$ T cell clusters (light blue and purple) based on their transcriptional signatures. **(D)** Close-ups are shown for clusters containing T cells during relapse revealing the $CD8^+$ T cells are predominantly in the purple cluster. **(E)** Localization of $T_{TCR\text{-}C4}$ (dark blue dots) cells during remission are shown. **(F)** Localization of $T_{TCR\text{-}C4}$ (dark blue dots) cells during relapse are shown. Percentage of the total $T_{TCR\text{-}C4}$ cells in each cluster are indicated. **(G)** The heat map shows the 10 most differentially expressed genes (DEG) comparing $T_{TCR\text{-}C4}$ during remission (n=501) to those during relapse (n=33). **(H)** Analysis of gene-sets representing effector cytokines, cytolytic effector genes activation, and exhaustion genes were compared in $T_{TCR\text{-}C4}$ during remission (light blue) and relapse (darker blue), with each transcript in the gene-set shown as a violin plot. The shape of the violin displays frequencies of values. Model-based Analysis of Single Cell Transcriptomics (MAST) was used to determine the significance shown above each plot. Significance thresholds were set a priori at a threshold of false discovery rate of 5% and positive or negative fold change > log2(1.5).

ITIM domains, [*TIGIT*], lymphocyte activating 3 [*LAG3*], and cytotoxic T-lymphocyte associated protein 4, [*CTLA4*]) were significantly enriched in $T_{TCR\text{-}C4}$ at remission versus relapse ($p=2.98 \times 10^{-8}$, $p=3.25 \times 10^{-3}$, $p=2.11 \times 10^{-2}$ and $p=1.87 \times 10^{-4}$ respectively) (**Figure 2.4H**), suggesting repetitive stimulation or activation during remission. Although not every gene in each composite gene-set was individually statistically significant, each of the total gene-sets were significantly different. For example, expression of IFN-γ and TNF-α assessed individually did not achieve statistical significance ($p > 0.01$, $p > 0.01$ respectively) but were significant as part of the effector cytokine/chemokine gene-set ($p = 2.98 \times 10^{-8}$), which may reflect both the known decline in expression of these genes after dissociation from the activating signal (*27*) and potential decline of polyfunctionality from repetitive activation events.

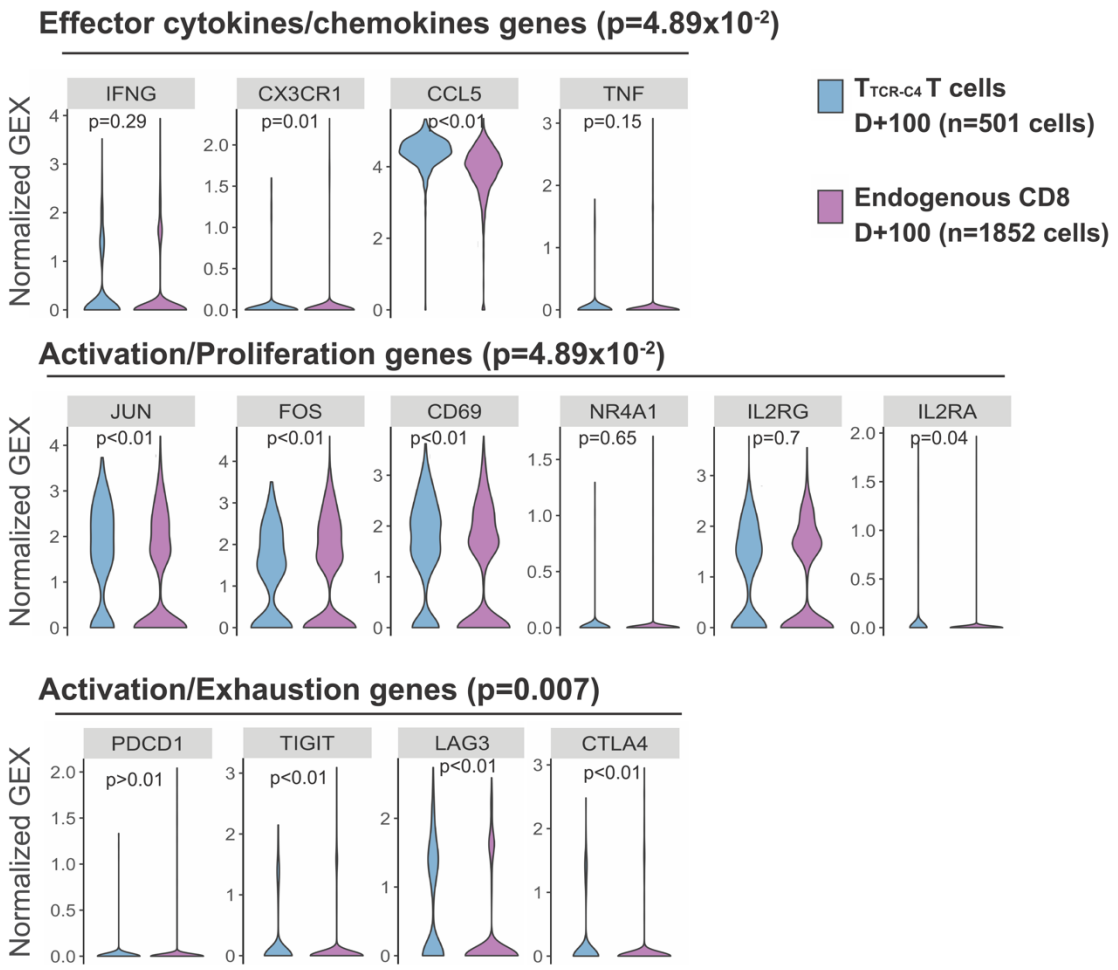


Figure 2.5 Gene set differences between $T_{TCR\text{-}C4}$ and endogenous $CD8^+$ T cells at remission. Expression of transcripts for gene-sets representing effector cytokine/chemokine production and T cell activation at remission for $T_{TCR\text{-}C4}$ (blue, n=501) and endogenous $CD8^+$ T cells (purple, n=1852). All graphs are shown as violin plots. The shape of the violin displays frequencies of values. Model-based Analysis of Single Cell Transcriptomics (MAST) was used to determine the significance shown on each plot. Significance thresholds were set a priori at a threshold of false discovery rate of 5% and positive or negative fold change > log2(1.5). Effector cytokine/chemokine genes: IFN-γ [*IFNG*], C-X-C motif chemokine receptor 1 [*CX3CR1*], C-C motif chemokine ligand 5 [*CCL5*] and TNF-α[*TNF*]. Activation/Proliferation genes: Jun proto-oncogene [*JUN*], Fos proto-oncogene [*FOS*], [*CD69*], nuclear receptor subfamily 4 group A member 1 [*NR4A1*], IL-2 receptor subunit gamma [*IL2RG*], and IL-2 receptor subunit alpha [*IL2RA*]. Activation/Exhaustion genes: programmed cell death receptor-1 [*PDCD1*], T-cell immunoreceptor with Ig and ITIM domains [*TIGIT*], lymphocyte activating-3 [*LAG-3*] and cytotoxic T-lymphocyte associated protein 4 [*CTLA4*]. shown.

Furthermore, comparison of endogenous $CD8^+$ T cells from the same sample to remission $T_{TCR\text{-}C4}$ transcriptomes again revealed significant enrichment for the effector cytokine/chemokine, activation/proliferation, and activation/exhaustion gene-sets in $T_{TCR\text{-}C4}$ ($p=4.89 \times 10^{-2}$, $p=4.89 \times 10^{-2}$ and $p=0.007$ respectively) (**Figure 2.5**), suggesting the profile represents antigen-specific events rather than global T cell activation in the host. As previously evidenced, $T_{TCR\text{-}C4}$ remained functional at both timepoints (**Figure 2.3B**), consistent with the T cells being in an activated but not exhausted state. The differences in gene expression (log2-fold change) observed at remission between $T_{TCR\text{-}C4}$ and endogenous cells were lost at relapse (**Figure 2.6**).

As $T_{TCR\text{-}C4}$ also contains an endogenous EBV-specific TCR, the patient was screened for evidence of EBV reactivation as a potential explanation for T cell activation. Very few copies of the EBV viral genome were detected immediately after the first $T_{TCR\text{-}C4}$ infusion (*10*), but none were detected by day 7 or at serial timepoints thereafter (**Table 2**). The results suggest $T_{TCR\text{-}C4}$ may have lysed and eliminated an EBV reservoir shortly after infusion. The possibility of a low degree of EBV antigen presentation leading to signaling through the endogenous EBV-specific TCR cannot be eliminated and could explain T cell activation during the prolonged period of clinical remission. However, the latter appears unlikely, as $T_{TCR\text{-}C4}$ no longer exhibited evidence of activation at relapse, a time when systemic illness might be more likely to lead to EBV reactivation. This further infers that $T_{TCR\text{-}C4}$ at relapse were not responding to leukemia or a low degree of EBV antigen. Consequently, we next evaluated the antigenicity of the patient's leukemia.

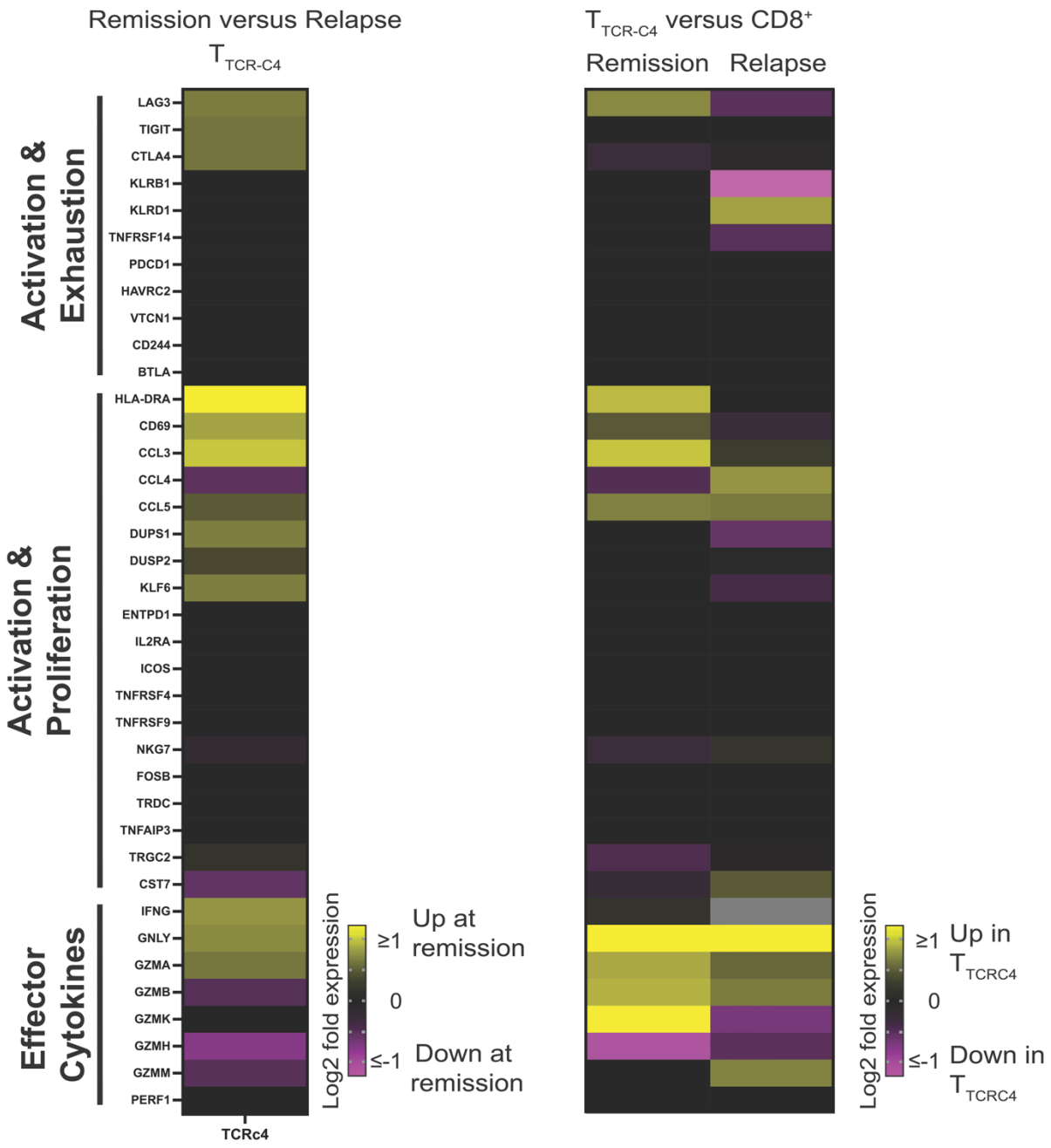


Figure 2.6 Gene expression differences between T_{TCR-C4} at remission and relapse and between T_{TCR-C4} and endogenous CD8+ T cells at both remission and relapse. Heat maps show gene expression differences (shown as log2 fold change) for a large set of activation & exhaustion genes, activation & proliferation genes, and effector cytokines. Comparisons: between T_{TCR-C4} at remission and at relapse; T_{TCR-C4} and endogenous CD8+ T cells at remission; T_{TCR-C4} and endogenous CD8+ T cells at relapse. Gray indicates Log2 fold change could not be calculated.

Days after 1st T_{TCR-C4} infusion	CMV (IU/mL)	EBV (IU/mL)
0	NDET	NDET
1	NDET	130
7	NDET	NDET
14	NDET	NDET
24	NDET	NDET
180	NDET	NDET
210	NDET	NDET
254	NDET	NDET

Table 2 Detection of CMV and EBV viremias after infusions. Patient serum was analyzed for presence of CMV and EBV virus over the course of treatment. NDET: none detected.

Relapsed AML expressed both target antigen and HLA

AML samples for immunohistochemistry (IHC) were obtained from a bone marrow aspirate during relapse after the first allo-HCT and prior to the first T_{TCR-C4} infusion (**Figure 2.7A and B)**, from a bladder chloroma at the time of relapse after the 1st T_{TCR-C4} infusion (**Figure 2.7A and C)**, and from a bone marrow aspirate 71 days after the 2nd T_{TCR-C4} infusion (**Figure 2.7A and D)**. All specimens showed abundant WT1 protein expression in the majority (>90%) of leukemic blasts (**Figure 2.7 3B to D)**. Clinical-grade next-generation sequencing performed at relapse after T_{TCR-C4} (day 58 after the 2nd T_{TCR-C4} infusion) found no WT1 sequence. However, $CD8^+$ T cells were not enriched in the chloroma (**Figure 2.7C**), consistent with the lack of an effective T cell response at relapse. To assess potential differences in epitope presentation, HLA class I protein expression was evaluated. On AML prior to the second HCT, IHC could be performed only for total HLA class I expression, which was positive (**Figure 2.7B**). HLA-A expression was evaluable in the cervical chloroma at relapse before any T_{TCR-C4}

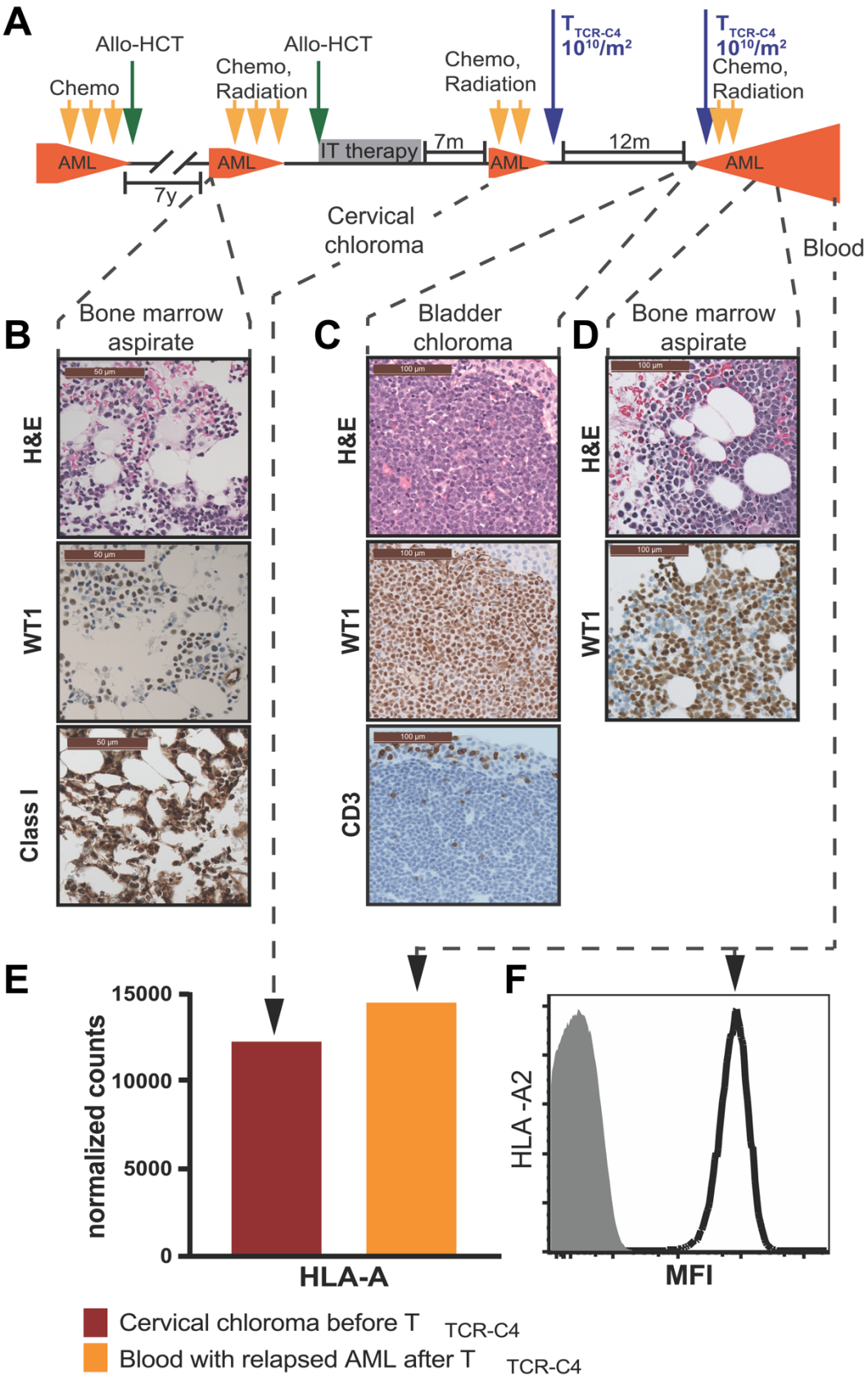

A
Allo-HCT
Chemo
Allo-HCT
Chemo, Radiation
Chemo, Radiation
T TCR-C4 10^10/m^2
T TCR-C4 10^10/m^2
Chemo, Radiation
AML
AML
IT therapy
7m
AML
12m
AML
7y
Cervical chloroma
Blood
B
Bone marrow aspirate
C
Bladder chloroma
D
Bone marrow aspirate
H&E
WT1
Class I
H&E
WT1
CD3
H&E
WT1
E
normalized counts
15000
10000
5000
0
HLA-A
F
HLA -A2
MFI
Cervical chloroma before T TCR-C4
Blood with relapsed AML after T TCR-C4

Figure 2.7 WT1 and HLA-A2 expression maintained in relapsed AML. (A) The timeline shows occurrence of bone marrow aspirate, myeloid chloroma, or blood obtained relative to the timeline of successive therapies. Presence of AML is indicated in orange. **(B)** Pre-$T_{TCR\text{-}C4}$ AML Hematoxylin and Eosin (H&E) (top panel), immunohistochemistry (IHC) of WT1 expression (middle panel) and HLA class I expression (bottom panel) is shown. Scale bars are 50 mm. **(C)** H&E (top panel), WT1 IHC expression (middle panel) and CD3 infiltration (bottom panel) is shown for bladder chloroma obtained early post-$T_{TCR\text{-}C4}$ relapse. Scale bars are 100 mm. **(D)** Post-$T_{TCR\text{-}C4}$ AML H&E (top panel) and WT1 IHC (bottom panel) are shown. Scale bars are 100 mm. **(E)** Normalized counts (y-axis) of bulk RNA expression of HLA-A obtained from the pre-$T_{TCR\text{-}C4}$ cervical chloroma and a post-$T_{TCR\text{-}C4}$ PBMC sample are shown, both containing greater than 89% AML. **(F)** HLA-A2 protein expression was detected by flow cytometry of the post-$T_{TCR\text{-}C4}$ blood. MFI, median fluorescence intensity.

infusions, using bulk RNA sequencing on residual clinical material from an archival formalin fixed, paraffin embedded (FFPE) specimen. As a control, HLA-A was also analyzed from PBMC obtained 110 days after the 2nd $T_{TCR\text{-}C4}$ infusion (the same sample on which scRNAseq was performed). HLA-A expression was detected in these samples (**Figure 2.7E**), both of which contained greater than 89% AML. As a note, the NanoString platform utilized permitted only probing for HLA-A and not for a specific allele. However, HLA-A2 expression was directly evaluable on relapsed AML cells present in PBMCs obtained 110 days after the 2nd $T_{TCR\text{-}C4}$ infusion (**Figure 2.7F**). The presence of HLA class I at all time points, and specifically of HLA-A2 at the last relapse, suggests that HLA-A*02:01 expression was retained throughout the disease. The expression of both unmutated target protein and restricting HLA allele at late relapse excludes antigen loss or mutation as well as restricting HLA loss or downregulation as the cause of AML immune escape in the presence of abundant, potentially reactive, functional $T_{TCR\text{-}C4}$.

Single cell RNA sequencing reveals low immunoproteasome in relapsed AML

We next investigated antigen processing components in relapsed AML after $T_{TCR\text{-}C4}$ infusions to account for the presumed failure of recognition. Evaluation by scRNAseq of the top 100 significant differentially expressed genes ($p<0.05$ and $logFC>|log_2(1.5)|$) in AML versus non-AML hematopoietic cells (in the same sample) revealed that the

AML cells had significantly ($p<0.01$ and logFC=-0.99) lower expression of the immunoproteasome subunit gene β1i [*PSMB9*] (**Figure 2.8, red arrow)**, which was confirmed by very low to undetectable β1i protein incorporation in relapsed AML (**Figure 2.9**).

Due to the absence of a viably frozen pre-treatment AML sample, it was difficult to directly assess if this patient's leukemia expressed the immunoproteasome before $T_{TCR\text{-}C4}$ targeting $WT1_{126\text{-}134}$ were initially infused. Therefore, we used residual clinical material from the cervical chloroma FFPE specimen (before the 1st $T_{TCR\text{-}C4}$ administration) and PBMCs obtained 110 days after the 2nd $T_{TCR\text{-}C4}$ infusion (**Figure 2.7E**) to assess changes in proteasome subunit expression. Compared to pre-$T_{TCR\text{-}C4}$, AML post-$T_{TCR\text{-}C4}$ had more transcripts encoding proteins ubiquitously present in the

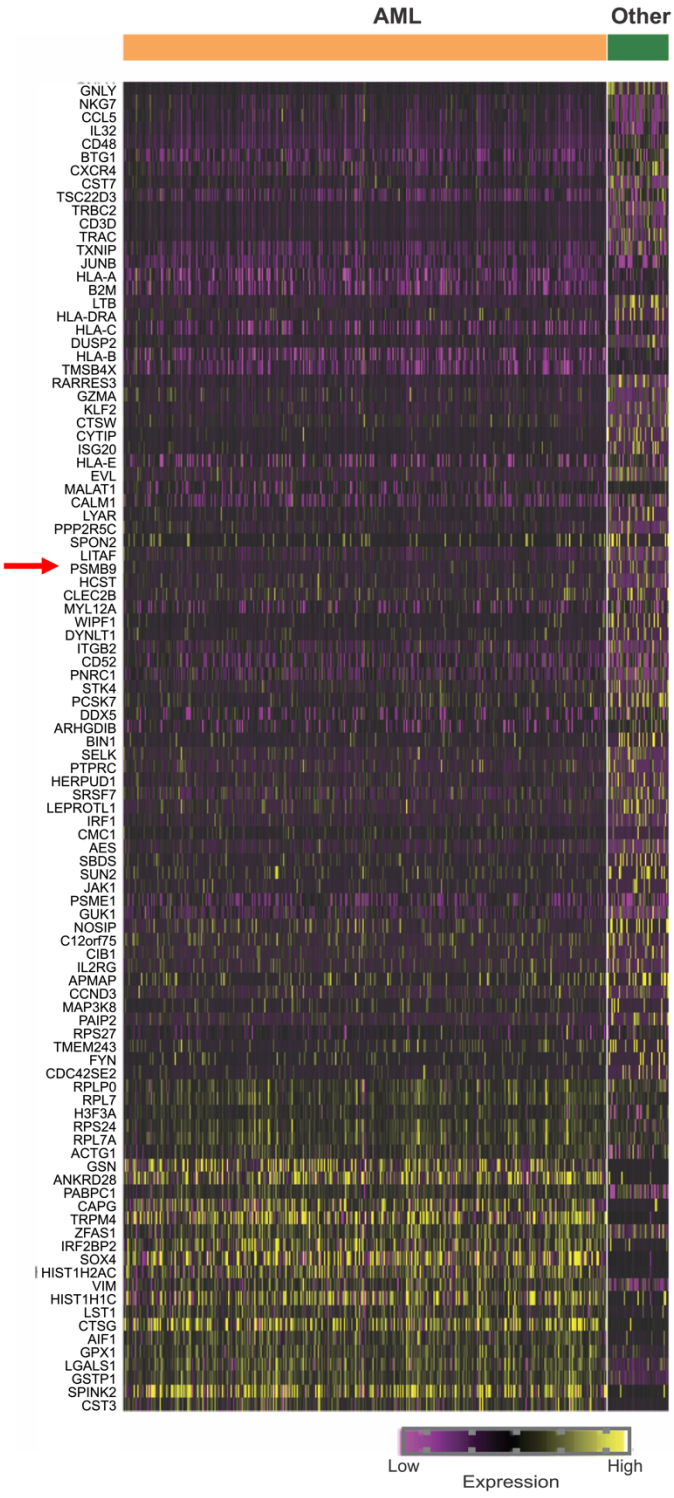


Figure 2.8 Reduced PSMB9 expression in relapsed AML. The heat map shows the top 100 significantly differentially expressed genes (DEG) ($p < 0.05$ and $\log FC > |\log_2(1.5)|$) comparing AML (n=2,447) during relapse after T cell therapy to other cells in the sample (n=306). The red arrow indicates immunoproteasome (IP) subunit β1i [*PSBM9*]

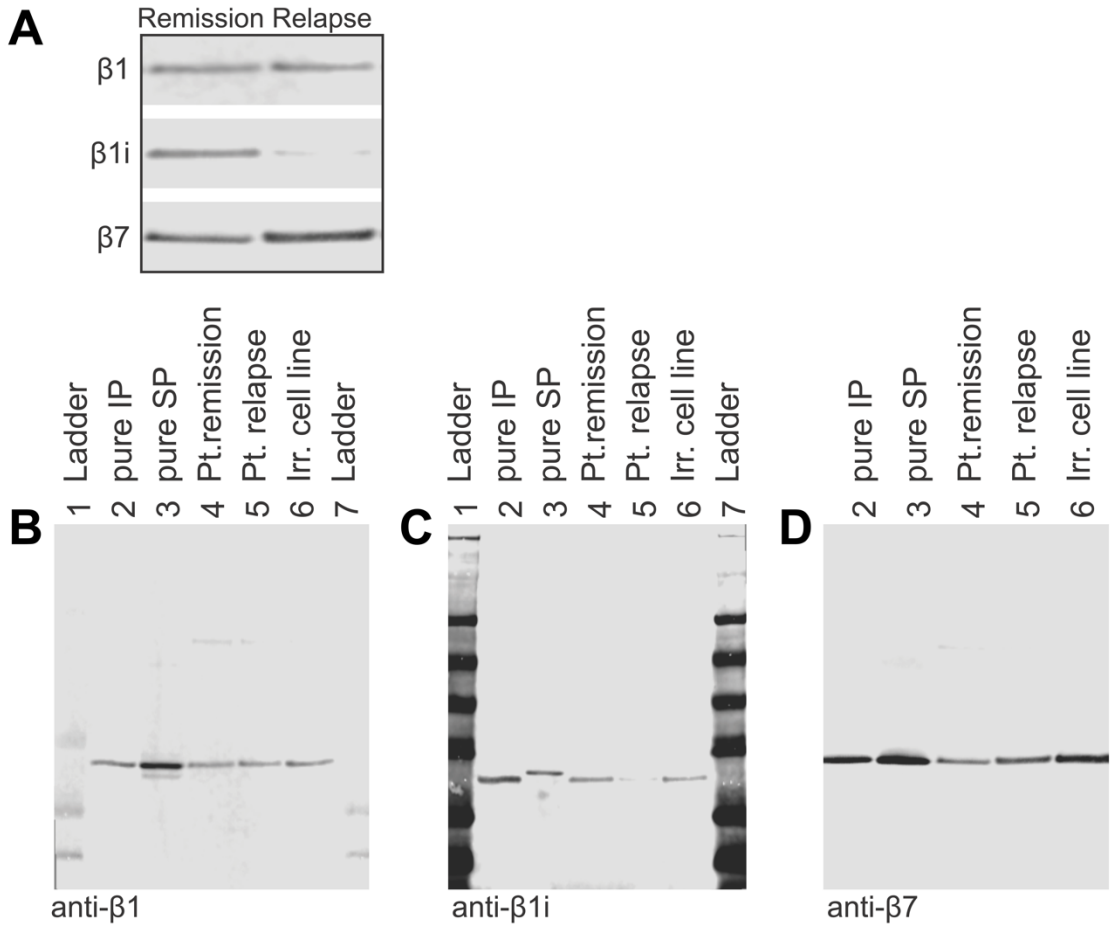


Figure 2.9 Protein expression of proteasome subunits at remission and relapse. (A) Immunoblots show standard proteasome specific subunit β1, immunoproteasome subunit β1i, and shared proteasome subunit β7. Assembled proteasome complexes were immunoprecipitated from patient 1's PBMC samples during remission (complete sample) and from relapse (CD3 depleted AML). Proteasome subunit β7 is shared between proteasome complexes and serves as a loading control of immunoprecipitated fractions. **(B to D)** Full blots are shown for **(B)** anti-β1, **(C)** anti-β1i, and **(D)** anti-β7. Lanes are 1-Ladder, 2-purified Immunoproteasome (Thermo Fisher Scientific), 3-purified Standard proteasome (Novus Biologics), 4-Patient PMBCs at remission 5- Patient AML at relapse, 6-irrelevant cell line, 7-ladder.

caps of all 26S proteasomes (ATPase 2 [*PSMC2*] and non-ATPase 7 [*PSMD7*]) (**Figure 2.10A, black bars**), suggesting comparatively increased total proteasome in AML post-$T_{TCR\text{-}C4}$ (*15*). However, AML post-$T_{TCR\text{-}C4}$ had less transcripts associated with the immunoproteasome subunits, β5i [*PSMB8*] (fold change -0.14) and β1i [*PSMB9*] (fold

change -0.67), which would restrict immunoproteasome formation. Although β2i [*PSMB10*] was not decreased (fold change +0.25), its incorporation into an immunoproteasome requires β1i, as its presence alone cannot drive immunoproteasome formation (**Figure 2.10A, blue bars**) (*28*). In contrast, although β1 was unavailable for assessment (see **Methods**), an increase in β2 [*PSMB7*] (fold change +0.63) and minimal change in β5 [*PSMB5*] (fold change -0.1) was detected, suggesting that the AML may have preferentially formed the standard proteasome (**Figure 2.10A, red bars**). The decrease in β1i resulted in minimally detectable transcripts in the post-$T_{TCR\text{-}C4}$ specimen (**Figure 2.10B)**, confirmed by absent β1i protein expression (**Figure 2.9).** Furthermore, expression of transcripts encoding proteins constituting the active proteolytic site of the immunoproteasome (β5i [*PSMB8*]*,* β1i [*PSMB9*]*,* and β2i [*PSMB10*]) were significantly ($p<0.01$) decreased, whereas those constituting the active proteolytic site of the standard proteasome (β5 [*PSMB5*], β1 [*PSMB6*] and β2 [*PSBM7*]) were significantly increased ($p<0.01$) in the AML cells compared to other hematopoietic cells (**Figure 2.11)**. Likewise, the frequency of cells expressing immunoproteasome subunits β1i [*PSMB9*]*,* and β2i [*PSMB10*] were markedly lower in the relapsed AML, whereas the frequency of cells expressing standard proteasome subunits β5 [*PSMB5*], β1 [*PSMB6*] and β2 [*PSBM7*] was increased compared to other hematopoietic cells. Immunoproteasome subunit β5i [*PSMB8*], however, opposed this trend with 50% of relapsed AML cells expressing the transcript compared to 28% of other hematopoietic cells. Reduced [*PSMB9*] expression was not a result of gene mutation, as whole exome sequencing on the relapsed AML did not reveal any pathogenic mutations in *PSMB9*. Moreover, the relapsed AML retained similar expression of IFN-γ response genes as other cells in the PBMCs, suggesting cytokine responsiveness remained intact (**Figure 2.12**).

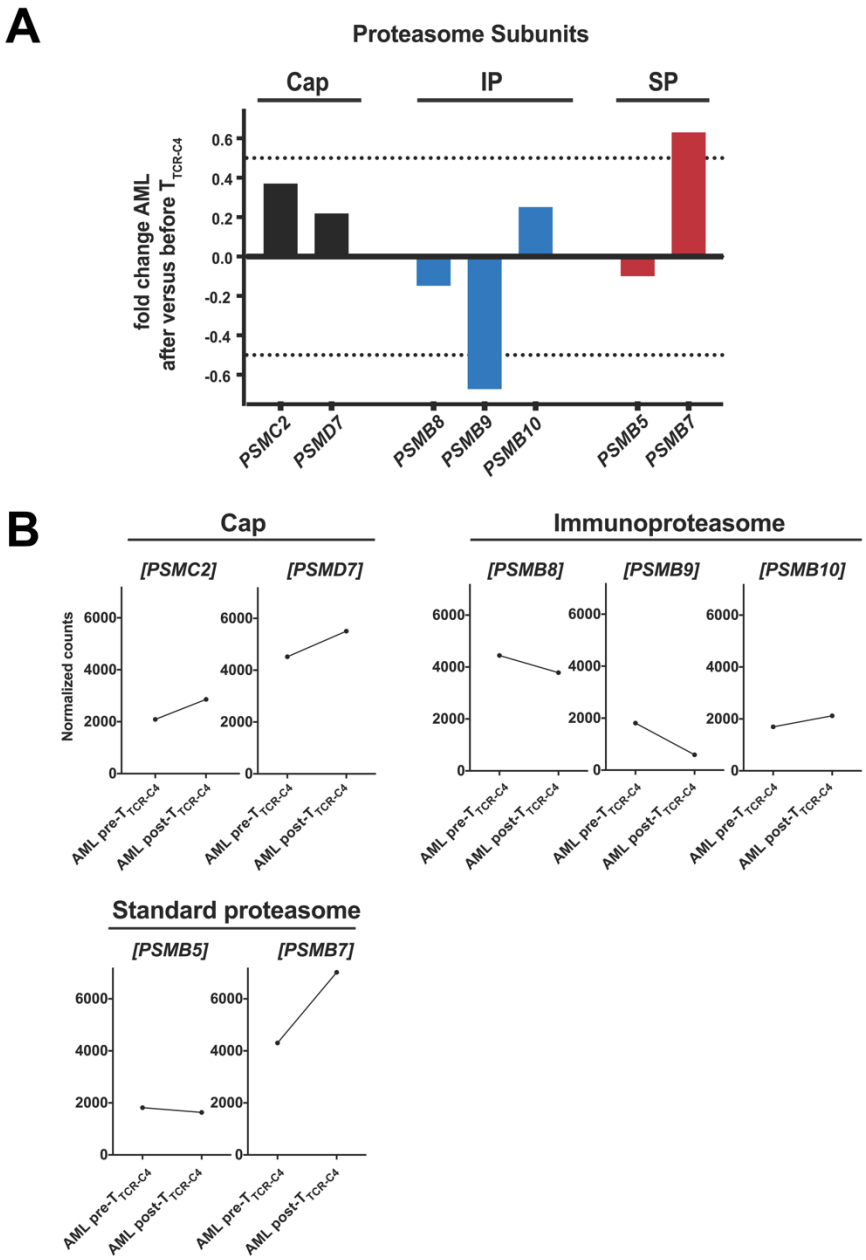


Figure 2.10 Expression of proteasome subunits. (A) The plot shows the fold change (y axis) of transcripts encoding proteasome cap ([*PSMC2*] and [*PSMD7*]; black bars), IP (β5i [*PSMB8*], β1i [*PSMB9*] and β2i [*PSMB10*]; blue bars) and standard proteasome (SP) (β5 [*PSMB5*], and β2 [*PSMB7*]; purple bars) sub-units comparing expression in the post-$T_{TCR\text{-}C4}$ obtained from PBMCs to pre-$T_{TCR\text{-}C4}$ AML obtained from a chloroma FFPE sample. β1 [*PSMB6*] was not available in the commercial RNA probe-set utilized. Bold dotted line indicates 50% change over baseline. **(B)** Transcript expression of proteasome subunit in AML before and after $T_{TCR\text{-}C4}$ infusions. Normalized counts showing expression of proteasome beta subunits and two subunits in the regulatory cap in AML samples from Patient 1. No *PSMB6* (β1) probes were available for the panel used.

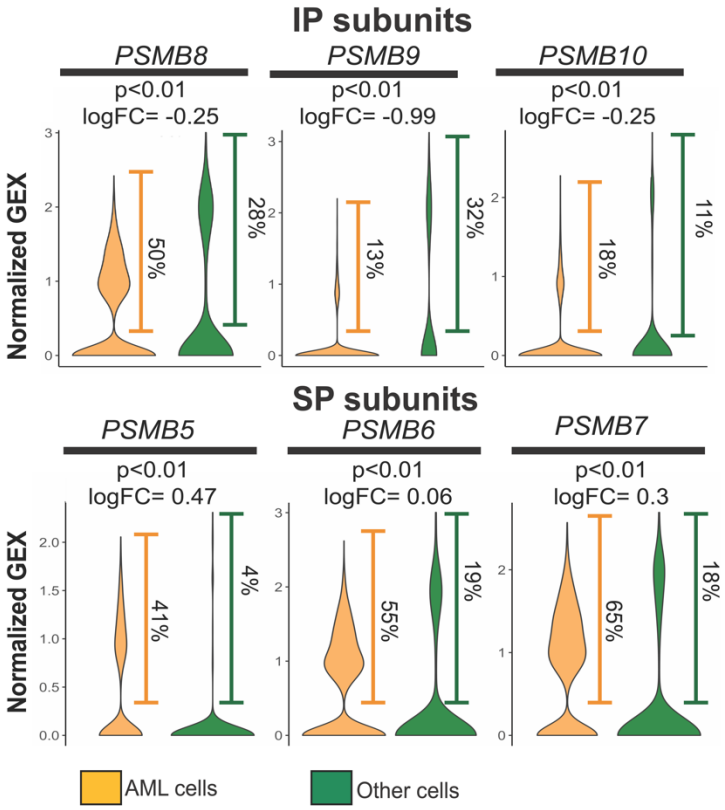


Figure 2.11 Expression of proteasome subunits in relapsed AML and other immune cells. Expression of transcripts encoding genes constituting the active proteolytic site of the IP (β5i [*PSMB8]*, β1i [*PSMB9]* and β2i [*PSMB10]*) and the SP (β5 [*PSMB5]*, β1 [*PSMB6]* and β2 [*PSMB7]*) are shown at relapse comparing AML (n=2,447, orange) and other non-malignant cells (n=306, green) in the sample. Frequencies of cells with the detectable transcript are shown to the right of each violin plot. P values are indicated as well as the log2 fold change. Positive values indicate increased and negative decreased expression in AML. MAST was used to determine significance shown above each plot. Significance thresholds were set a priori at a threshold of false discovery rate of 5% and positive or negative fold change > log2(1.5)).

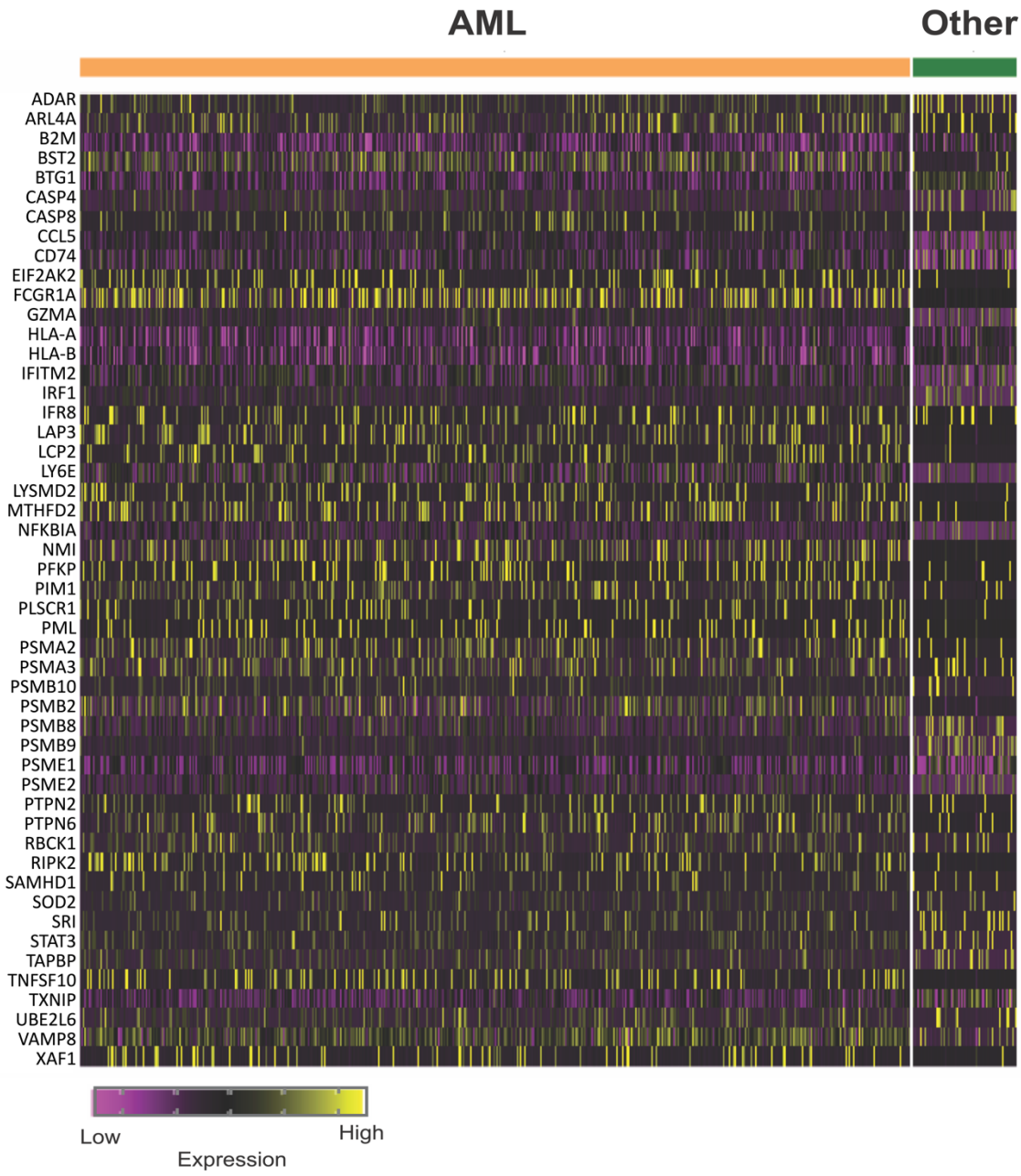


Figure 2.12 IFN-γ response genes expressed in PBMCs isolated from Patient 1 after relapse. Fifty of the most highly expressed IFN-γ response genes, in addition to IP subunits, are shown in AML and other immune cells in PBMCs isolated from patient 1.

To confirm that *PSMB9* is required for $WT1_{126\text{-}134}$ processing and $T_{TCR\text{-}C4}$ recognition, $T_{TCR\text{-}C4}$ were co-cultured with cells lines genetically modified to express the immunoproteasome, standard proteasome, an intermediate proteasome with a single

immunoproteasome subunit (β5i, β1, β2, (sIP)), or an intermediate with two immunoproteasome subunits (β5i, β1i, β2, (dIP)) (*29*). Indeed, $T_{TCR\text{-}C4}$ were found to only recognize cell lines expressing proteasome isoforms containing β1i (**Figure 2.13**).

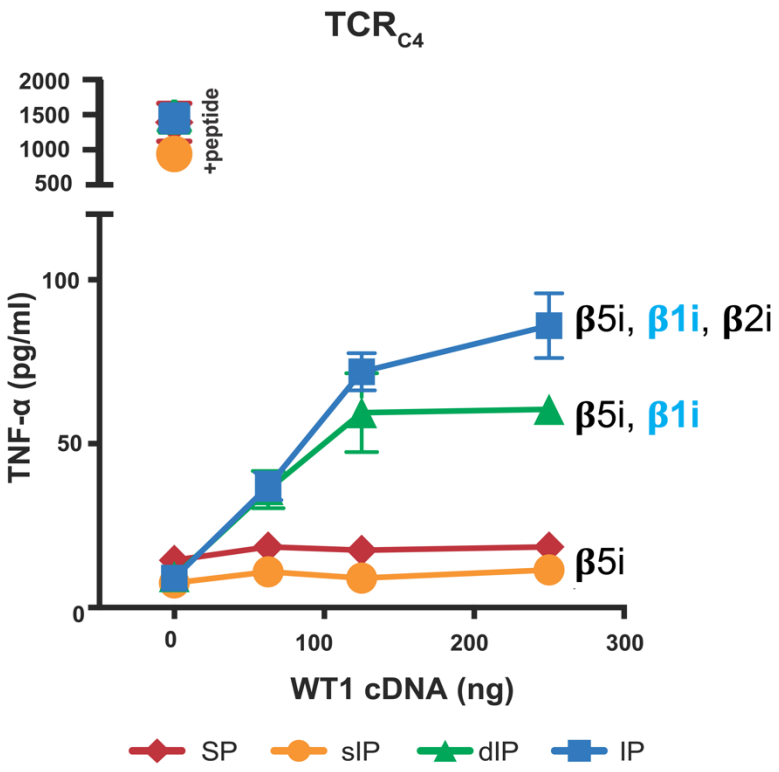


Figure 2.13 $T_{TCR\text{-}C4}$ recognizes cells expressing β1i. TNF-α production (pg/mL) was measured from T cells expressing TCR_{C4} co-cultured with cells expressing a specific proteasome isoform and transfected with varying amounts of WT1 cDNA or cells pulsed with $WT1_{126\text{-}134}$ peptide as a positive control. Data are presented as mean ± standard deviation.

To assess whether low *PSMB9* may be present in other AML specimens, which would suggest this may be a more common phenomenon, we compared the relapsed leukemia to data from six normal-donor mobilized peripheral blood stem cell samples and 38 enriched primary AML blast samples obtained from the Fred Hutch-University of Washington Hematopoietic Diseases Repository. *PSMB9* (β1i) RNA expression in this patient's relapse AML was markedly lower compared to other AML specimens, but this reduction was not rare, as eight other AMLs also had low *PSMB9* expression (**Figure 12.14**). With the caveat that adequate AML cells prior to the final relapse were

unavailable for detailed analysis, these data suggests that the absence of β1i in this patient's AML at relapse may have compromised recognition by $T_{TCR\text{-}C4}$.

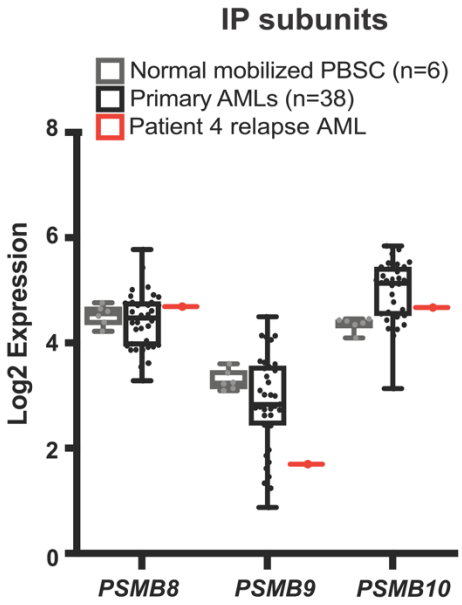


Figure 2.14 Gene expression of immunoproteasome subunits in primary AML. RNAseq on viable leukemic blasts shows expression of genes constituting the active proteolytic site of the IP (β5i [*PSMB8]*, β1i [*PSMB9]* and β2i [*PSMB10]*) in peripheral blood stem cells obtained from healthy donors (n=6, light gray), primary AML samples obtained from the Fred Hutch-University of Washington Hematopoietic Diseases Repository (black, n=38) and the patient (red, n=1). Box and whisker plot represents median, interquartile and range. Dots represent individual patient samples.

$T_{TCR\text{-}C4}$ preferentially eliminated WT1+ AML blasts co-expressing PSMB9.

A second patient (Patient 2) on this trial was identified who received $T_{TCR\text{-}C4}$ and had blasts available before and after T cell infusion (**Figure 2.15A**). This patient's AML constituted 70.2% of PBMCs at the time of $T_{TCR\text{-}C4}$ infusion. Although the absolute blast count moderately decreased immediately after infusion, the AML eventually progressed despite persistence of $T_{TCR\text{-}C4}$ at greater than 3% of peripheral $CD8^+$ T cells, and he succumbed to progressive disease 15 days after transfer (**Figure 2.15B and C**). ScRNAseq on PBMCs obtained before $T_{TCR\text{-}C4}$ infusion showed the AML blasts were heterogenous and formed seven leukemic clusters (**Figure 2.16A, B**), all expressing one or more of the AML-associated genes *CD34*, *CCND1*, *CDK6*, *FLT3*, *NPM1* and

RUNX1(**Figure 2.16C**). Five days after infusion, clusters “AML Blasts 2” and “AML Blasts 3” decreased whereas clusters “AML Blasts 1” and “monocytic AML1” expanded (**Figure 2.16D**). The expression of AML-associated genes was unchanged. HLA-A also remained expressed (**Figure 2.16E**). However, the frequency of WT1-expressing cells decreased from 27% to 2.5% after $T_{TCR\text{-}C4}$ exposure, with the most noticeable reduction in the two clusters that expressed the highest abundance of WT1 (**Figure 2.17A top**, corresponding to “AML Blasts 2” and “AML Blasts 3” in **Figure 2.16**) WT1 loss was confirmed by IHC (**Figure 2.17A, bottom)**. *PSMB9*-expressing AML cells were detected after $T_{TCR\text{-}C4}$ (**Figure 2.17B**). In all blast populations expressing WT1 and β1i, but especially in the expanding clusters (“AML Blasts 1” and “monocytic AML1”), only the proportion of cells that co-expressed WT1 and β1i (*PSMB9*) was reduced after $T_{TCR\text{-}C4}$ exposure (25.1% to 1.3% or a 19-fold reduction), whereas the proportion of cells that only expressed WT1 remained proportionately similar (24.6% to 28.6%) (**Figure 2.17C and D**). These data support the hypothesis that $T_{TCR\text{-}C4}$ exposure only efficiently eliminates $WT1^+$ blasts that also express β1i.

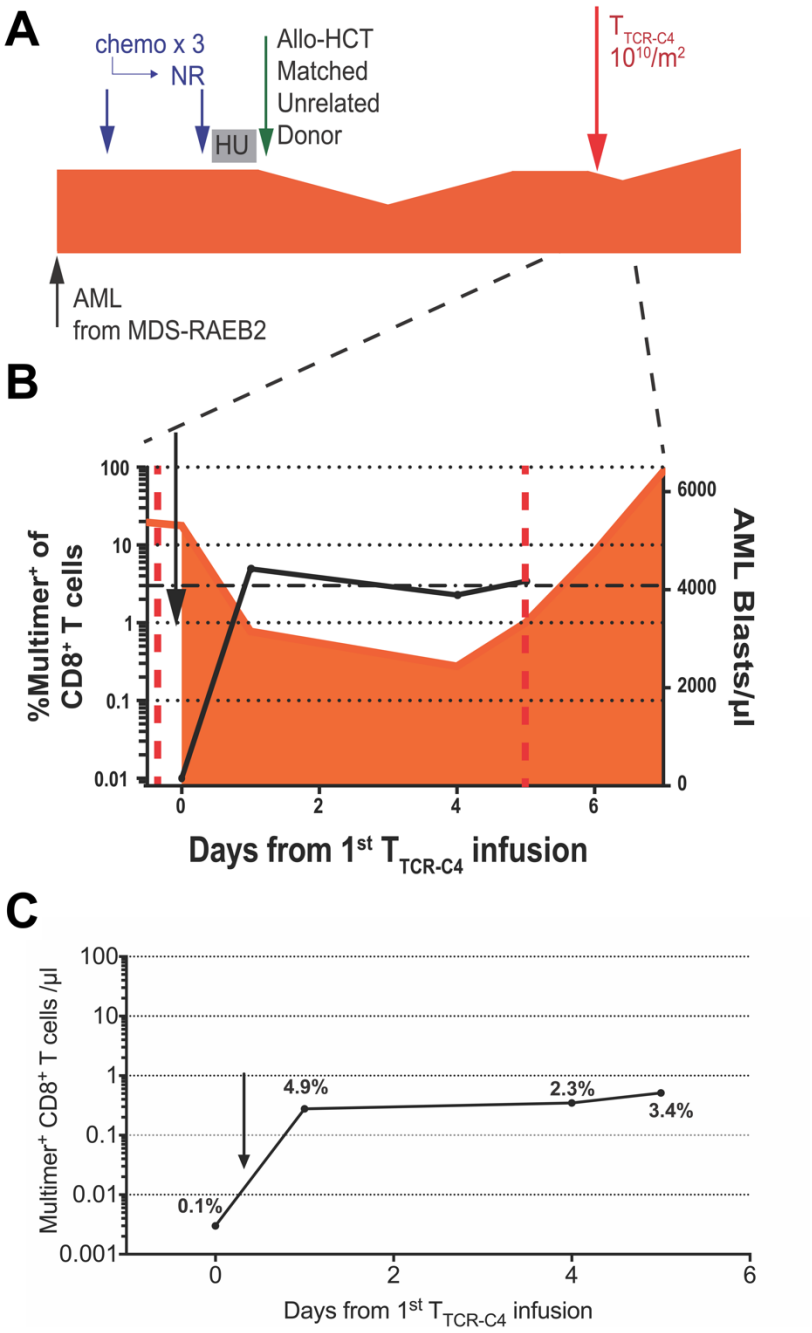


Figure 2.15 Patient 2 clinical course and $T_{TCR\text{-}C4}$ In vivo persistence kinetics. (A) Timeline of patient's treatment regimens. Chemo, chemotherapy; NR, no response; HU, hydroxyurea. **(B)** Percent multimer$^+$ of CD8$^+$ T cells is shown in peripheral blood collected before and at day +1, +4, and +5. The orange shaded area indicates presence of AML. Dotted red lines represent days +0 and +5 after $T_{TCR\text{-}C4}$. **(C)** Multimer$^+$CD8$^+$ T cells/µl (y-axis) were enumerated in PBMCs collected after infusion (x-axis, day relative to infusion). Time of infusion is indicated with black arrow. Percent multimer$^+$ of CD8$^+$ cells is shown above each timepoint.

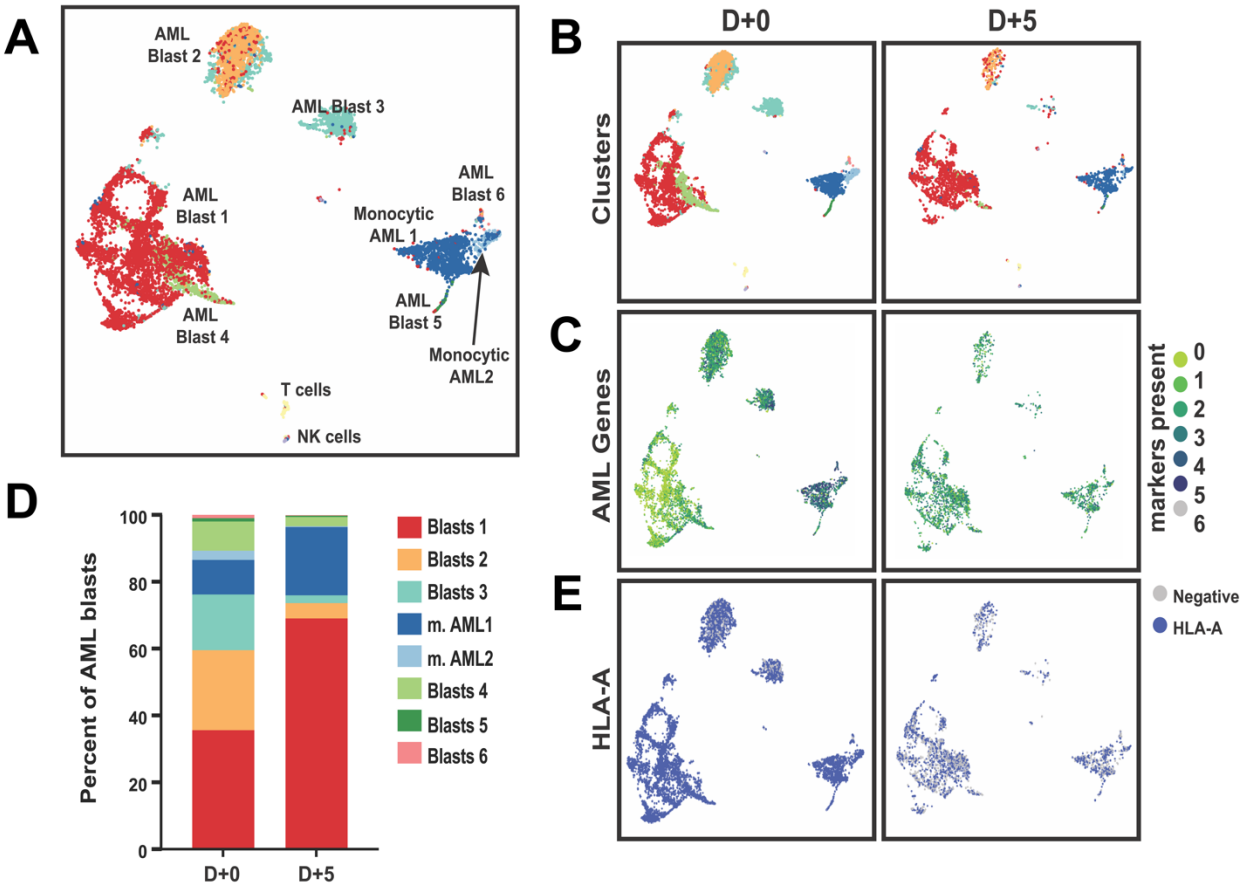
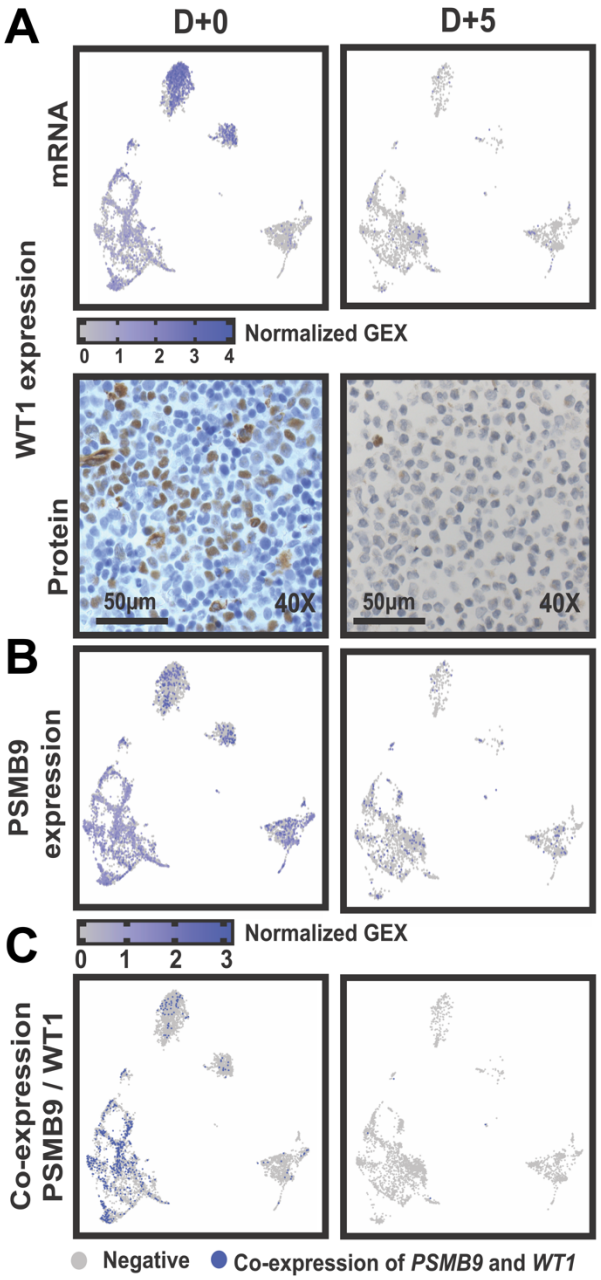


Figure 2.16 ScRNAseq of Patient 2 AML. (A) UMAP visualization is shown for scRNAseq on PBMCs from both before (D+0, n=7,381) $T_{TCR\text{-}C4}$ infusion and 5 days post 1st infusion (D+5, n=3,319) (red dotted lines in **Figure 2.15B**) aggregated into a single UMAP. (**B-E**) UMAP of AML clusters on D+0 (left) or D+5 (right) are shown. (**C**) Expression of AML specific markers *CD34*, *CCND1*, *CDK6, FLT3*, *NPM1*, *RUNX1* are shown. (**D**) Relative proportions of AML clusters at D+0 and D+5 are shown. (**E**) Expression of the gene encoding HLA-A is shown.



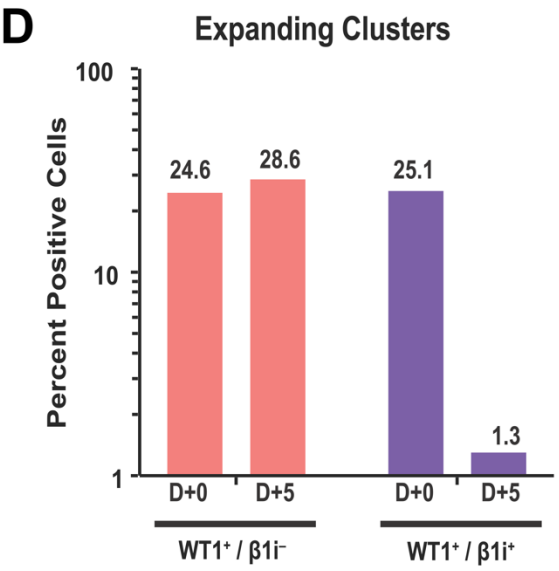


Figure 2.17 $T_{TCR\text{-}C4}$ preferentially eliminates WT1⁺ AML co-expressing β1i. (A) Expression of *WT1* transcripts from peripheral AML (top panel) and immunohistochemistry (bottom panel) of WT1 protein performed on a bone marrow biopsy (D+0) and a PBMC cell pellet (D+5) are shown. **(B)** Expression of immunoproteasome gene β1i (*PSMB9*) is shown. **(C)** Co-expression of *WT1* and *PSMB9* is shown. **(D)** The bar graph shows the percent of cells that express WT1 only (pink bars) or WT1 and β1i ([*PSMB9*]) (purple bars) before (D+0) and after (D+5) $T_{TCR\text{-}C4}$ infusion. Percent values are indicated above each column

Targeting $WT1_{37\text{-}45}$ can overcome immunoproteasome deficiency and expand therapeutic potential to solid tumors.

Identifying epitopes not requiring a specific proteasome isoform might avoid potential immunotherapy resistance and thereby broaden the clinical translation of targeting WT1. We therefore investigated whether targeting alternative HLA-A*02:01-restricted WT1 epitopes could have a therapeutic potential towards AML with low immunoproteasome expression, presumed resistant to $T_{TCR\text{-}C4}$, or solid tumors that generally lack immunoproteasome processing. We generated T cell lines from 15 healthy donors (two shown) recognizing eight previously described HLA-A*02:01-restricted WT1 epitopes (**Figure 2.18A**) Our screen identified $WT1_{37\text{-}45}$ (also identified by Ruggiero et al. (*30*)) and $WT1_{235\text{-}243}$ as promising immunogenic candidates. Cell lines targeting $WT1_{37\text{-}45}$ and $WT1_{235\text{-}243}$ were sorted on tetramer positivity (**Figure 2.18B**) and assessed for reactivity towards the HLA-A*02:01 transfected, $WT1^+$ K562 cell line which primarily expresses the standard proteasome and is not lysed by T cells targeting $WT1_{126\text{-}134}$ (**Figure 2.18C and D**) (*16*). A high affinity TCR reactive to $WT1_{37\text{-}45}$ ($TCR_{37\text{-}45}$) was isolated from the cell lines and inserted into a lentiviral backbone (see **Methods**). $CD8^+$ T cells expressing $TCR_{37\text{-}45}$ ($T_{TCR37\text{-}45}$) recognized and lysed HLA-A2- and WT1-transduced HEK-293 cell lines (endogenous standard proteasome expression) engineered or not to express the immunoproteasome (**Figure 2.19A-C**) (*29*). Of note, WT1 expression was not uniform in the cells from the WT1-transduced immunoproteasome and standard proteasome cell lines (**Figure 2.19C,** inset), which limited the maximal lytic activity achieved in this assay. $CD8^+$ T cells expressing either TCR_{C4} or $TCR_{37\text{-}45}$ lysed HLA-$A2^+WT1^+$primary AML cells (**Figure 2.20A**), but only $TCR_{37\text{-}45}$ induced caspase 3 cleavage in the first patient's relapsed, $T_{TCR\text{-}C4}$-resistant AML (**Figure 2.20B**). Furthermore, only $TCR_{37\text{-}45}$ reduced growth of the WT1 expressing solid tumor cell lines PANC-1 (*31*) and HLA-A*02:01-transduced MDA-MD-683 (*32*) in a live tumor visualization assay (**Figure 2.21A**). Exposing the PANC-1 cell line to IFN-γ for 48 hours increased β1i expression (**Figure 2.21B)** and enhanced recognition by TCR_{C4} (**Figure 2.21C**).

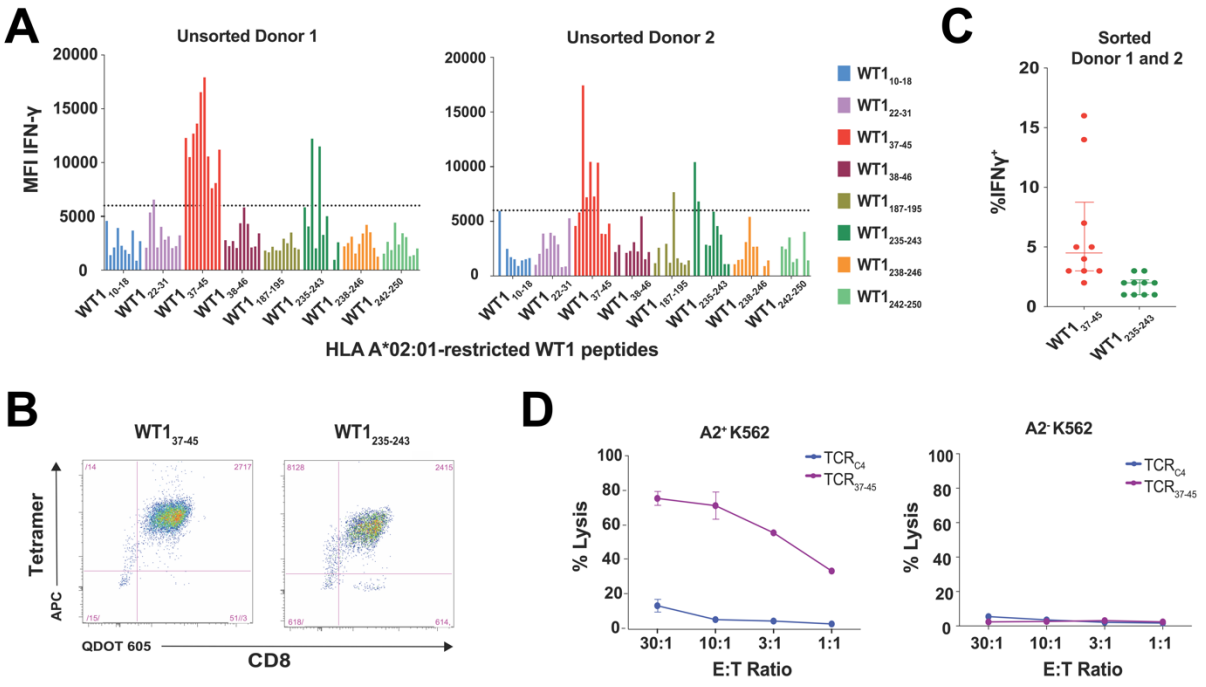


Figure 2.18 Identification of an alternate HLA-A*02:01-restricted WT1 peptide ($WT1_{37-45}$) not dependent on immunoproteasome processing. (A) Ten $CD8^+$ T cell lines from 2 HLA-A*02:01-expressing healthy donors were generated against 8 candidate HLA-A*02:01-binding WT1 peptides (x-axis). Each cell line's ability to produce IFN-γ following stimulation with the immunoproteasome-deficient but $WT1^+$ HLA-A*02:01-transduced K562 cell line at an E:T of 1:1 was assessed. Peptides were considered to have elicited high reactivity from donor T cells as measured by MFI (above the arbitrary threshold of 6000) for the IFN-g positive cells. **(B)** Representative flow plots show $CD8^+$ T cell lines targeting $WT1_{37-45}$ or $WT1_{235-243}$ that were sorted using the corresponding tetramers. **(C)** Sorted T cell lines targeting $WT1_{37-45}$ had the highest comparative reactivity, assessed by IFN-γ MFI, when cultured with HLA A*02:01-tansduced K562 cells. TCRs targeting $WT1_{37-45}$ were isolated from T cells in these lines. Data are presented as median with the interquartile range. All data points are shown. **(D)** Lysis of K562 ± expression of HLA-A*02:01 by $CD8^+$ T cells from a healthy donor transduced with TCR_{C4} or a TCR targeting $WT1_{37-45}$ (TCR_{37-45} #1) confirmed that expression of the $WT1_{37-45}$ peptide was not dependent on IP processing. Error bars show standard deviation of triplicate wells.

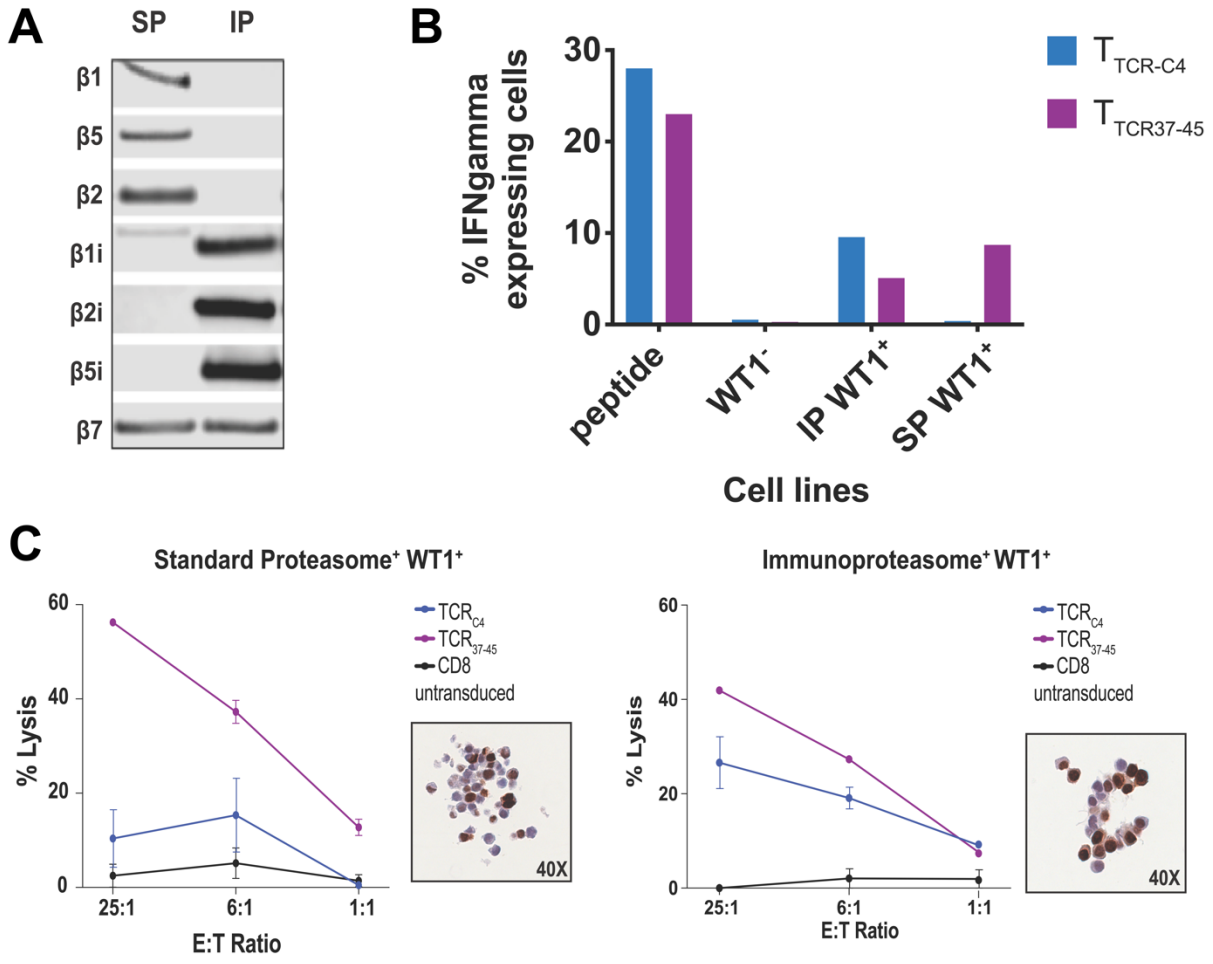


Figure 2.19 T_{TCR-C4} and $T_{TCR37-45}$ differentially recognize HLA-A2- and WT1-transduced HEK293 cells expressing the IP or SP. (A) Western blot of catalytic proteasome subunits in HEK-293 (endogenous SP expression) ± engineering to express the IP. **(B)** $CD8^+$ T cells isolated selected from a healthy, HLA-A*02^+ donor and transduced with TCR_{C4} and $TCR_{WT1-37-45}$. T_{TCRC4} and $T_{TCR37-45}$ were co-cultured with target cell lines (indicated on the x-axis). Targets are HLA-$A2^+$ T2 cells loaded with 10ng/mL respective peptide (peptide), HLA-A2-transduced HEK 293 expressing SP (WT1neg), HLA-$A2^-$ and WT1-transduced HEK 293 cells engineered to express IP (IP WT1pos), and HLA-$A2^-$ and WT1-transduced HEK 293 cells expressing SP (SP WT1 pos). Target cells were co-cultured with T_{TCR-C4} or $T_{TCR37-45}$ for 20 hours at an E:T ratio of 1.5:1. Cells assessed for IFN-γ production with flow cytometry **(C)** Percent lysis (y-axis) is shown for IP/SP-skewed HEK-293 cells lines by $CD8^+$ T cells transduced to express either TCR_{C4} (blue lines), TCR_{37-45} (red lines) and untransduced $CD8^+$ T cells (black lines) at indicated effector to target (E:T) ratios (x-axis). Inset depicts WT1 IHC of IP/SP HEK-293 cells transduced with WT1-and HLA-A*02:01. IP/SP cell lines demonstrate heterogeneity of WT1 expression. Error bars show standard deviation of triplicate wells.

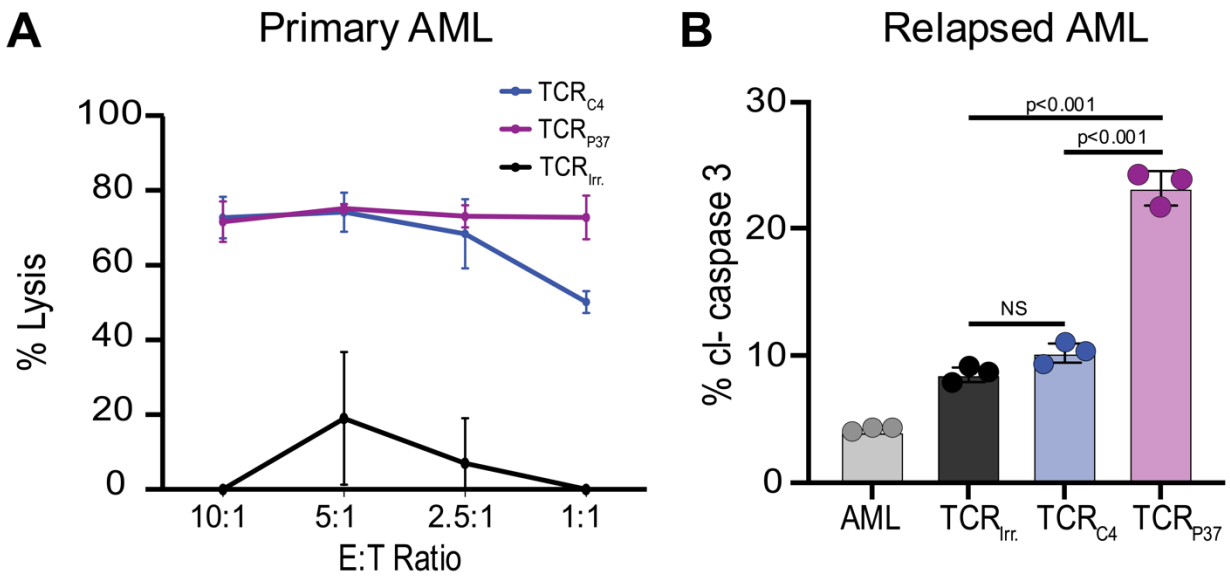


Figure 2.20 $T_{TCR37\text{-}45}$ recognize primary AML and patient relapsed AML. (A) Percent lysis (y-axis) of a primary leukemia by CD8$^+$ T cells transduced to express either an irrelevant TCR ($TCR_{irr.}$; black lines), TCR_{C4} (blue lines) or $TCR_{37\text{-}45}$ (red lines) is shown. Error bars show standard deviation of triplicate wells. **(B)** Percent AML cells expressing cleaved caspase 3 (cl-caspase 3, y-axis) is shown for Patient 1's relapsed AML either alone (gray) or cultured with CD8$^+$ T cells with the endogenous TCR removed using CRISPR-Cas9 and transduced to express $TCR_{irr.}$ (black), TCR_{C4} (blue), or $TCR_{37\text{-}45}$ (red). Error bars show standard deviation of triplicate wells.

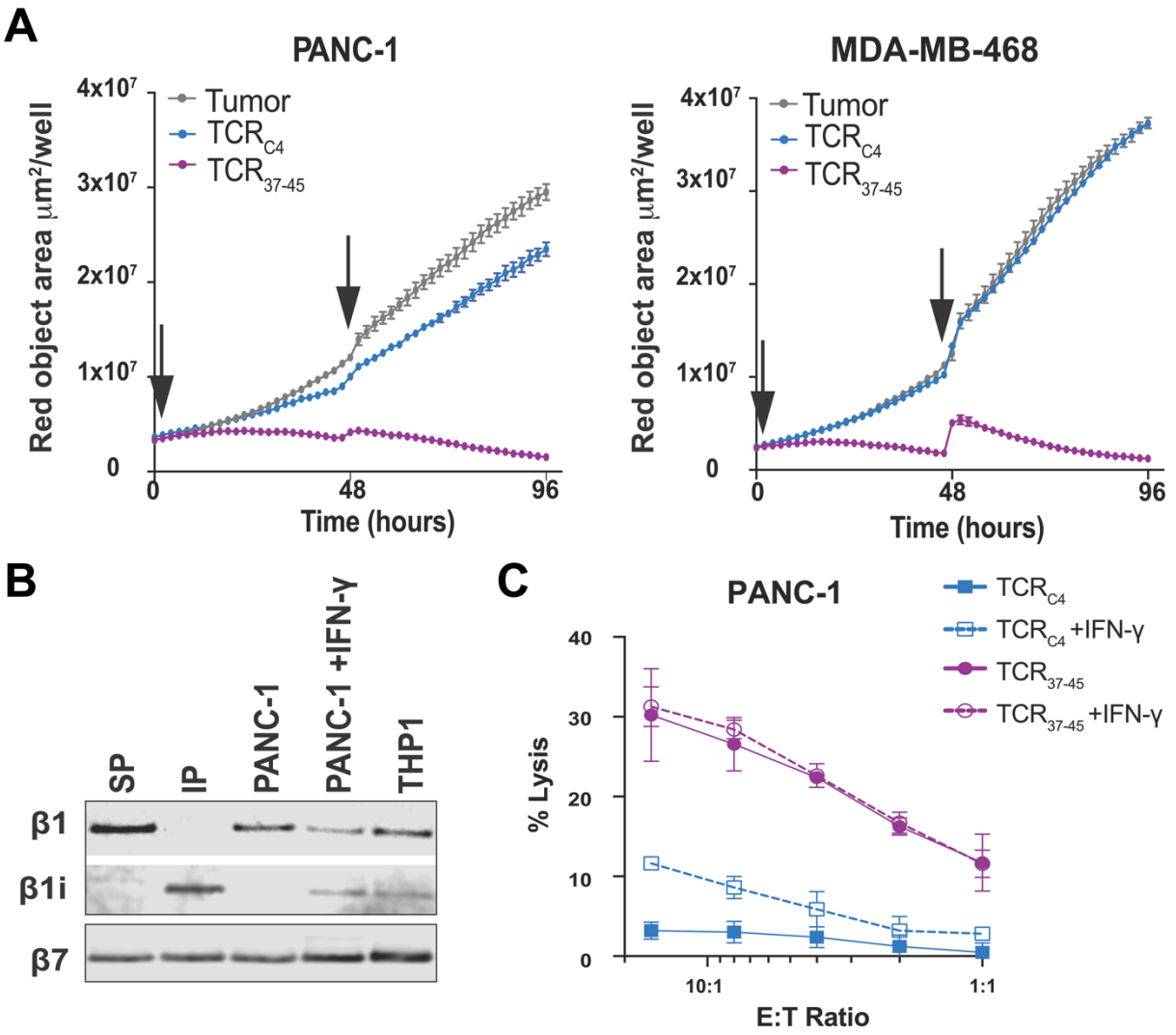


Figure 2.21 $T_{TCR37-45}$ and not T_{TCR-C4} recognize WT1 expressing solid tumors. (A) Growth kinetics for WT1^{+} cell lines PANC-1 and HLA-A*02:01-transduced MDA-MD-468 in live tumor-visualization assay in the absence (gray lines), or presence of CD8^{+} T cells transduced to express either TCR_{C4} (blue lines) or TCR_{37-45} (purple lines). Effector to target (E:T) ratio of 10:1 and arrows indicate addition of tumor cells to culture. Standard error of triplicate wells are shown. **(B-C)** 48-hour exposure to IFN-g induces β1i expression and incorporation into proteasomes in PANC-1 cells but does not completely rescue recognition by T_{TCR-C4} T cells. **(B)** Western blot of proteasome subunits β1 and β1i, respectively. PANC-1 cells were untreated or exposed to IFN-g for 48-hours. THP-1 cells are a leukemia line. All cells were lysed and assembled proteasome complexes were immunoprecipitated for immunoblot analysis. Proteasome subunit β7 is shared between both proteasome complexes and serves as a loading control of immunoprecipitated fractions. **(C)** Percent lysis is shown for PANC-1 cells that were either exposed to IFN-g for 48 hours (dotted lines) or untreated (solid lines) prior to culture with T cells expressing TCR_{C4} or TCR_{37-45} at various E:T ratios. Data are presented as mean ± standard deviation.

In anticipation of targeting $WT1_{37\text{-}45}$ for future clinical translation in both solid and liquid tumors, we sought to maximize the clinical efficacy by also modeling the engagement of both $CD4^+$ and $CD8^+$ T cells, as has proven efficient for CAR T cells approaches with a 1:1 ratio $CD8^+$:$CD4^+$ (*33*). Antigen-specific $CD4^+$ T cells, restricted by HLA class II, enhance $CD8^+$ T cell survival, proliferation and function (*34*). As HLA class II expression is not always expressed by tumors (*35*), an alternative is to impart tumor recognition to $CD4^+$ T cells through expression of a class I-restricted TCR that would co-localize engineered $CD8^+$ and $CD4^+$ T cells on the same tumor target without requiring HLA class II expression. Although high affinity TCRs can engage $CD4^+$ T cells without $CD8\alpha\beta$ binding, TCRs with such high affinity are rarely detectable, especially for targeting self-proteins such as WT1 (*36*). To engineer functional $CD4^+$ T cells targeting $WT1_{37\text{-}45}$ we incorporated $CD8\alpha\beta$ along with $TCR_{37\text{-}45}$ into a lentiviral vector (see Methods). Transduction with this vector generated cells with additional CD8 on the surface yielding 'enhanced' $CD8^+$ T cells ($eCD8^+$) and $CD4^+$ T cells co-expressing CD8 and CD4 (double positive $CD4^+$, $dpCD4^+$). We next assessed the relative contribution of the $CD8\alpha\beta$ to the function of $CD8^+$ and $CD4^+$ T cells individually and in combination by intentionally decreasing the effector to target ratio from 10:1 to 4:1 and repetitively adding tumor cells to "stress" the engineered T cells. Increased expression of CD8ab did not substantially improve tumor control by $eCD8^+$ cells compared to TCR-transduced $CD8^+$ T cells (**Figure 2.22A**), but $dpCD4^+$ did mediate a sustained antitumor response, whereas $CD4^+$ T cells transduced with the TCR alone were ineffective (**Figure 2.22B**). Furthermore, a 1:1 mixture of $eCD8^+$/$dpCD4^+$ $T_{TCR37\text{-}45}$ cells more effectively controlled the solid tumor cell line, PANC-1, than either $eCD8^+$ or $dpCD4^+$ $T_{TCR37\text{-}45}$ cells alone (**Figure 2.22C**). Of note, none of these combinations were effective in clearing PANC-1 targets when $CD8\alpha\beta$ was added to TCR_{C4} in $CD4^+$ and $CD8^+$ T cells, suggesting the addition of $CD8\alpha\beta$ does not rescue reactivity in the context of limited antigen processing (**Figure 2.22D-F**). $DpCD4^+$ T cells, when compared to $eCD8^+$ T cells alone, generally secreted abundant inflammatory cytokines (IFN-γ, TNF-α, IL-2), cytokines associated with a helper function (IL-6, IL-17, and IL-21) or a Th2 response (IL-4, IL-13) and limited IL-10 associated with regulatory function (*37*). Although co-cultures of $eCD8^+$/$dpCD4^+$ resulted in less abundant cytokines compared to $dpCD4^+$ T

cells alone, all cytokines, except for IL-6, were significantly ($p < 0.01$) greater for eCD8$^+$/dpCD4$^+$ than for eCD8$^+$ T cells alone suggesting cytokines secreted by dpCD4$^+$ in co-culture may have provided support for the effector function of eCD8$^+$ T cells (**Figure 2.23**). Together, these in vitro data suggest TCRs targeting WT1$_{37\text{-}45}$ might be efficient against AMLs and solid tumors with limited immunoproteasome expression (*15*).

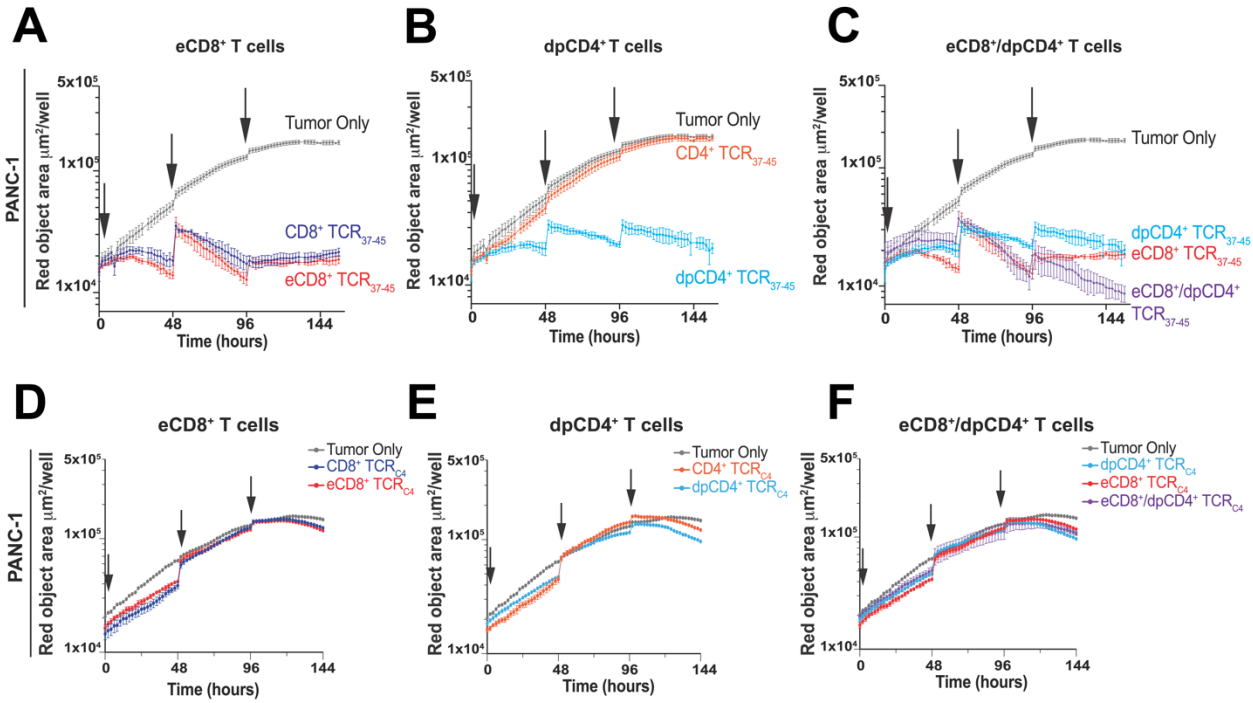


Figure 2.22 Increased abundance of CD8αβ facilitates CD4$^+$ T cell engagement for TCR$_{37\text{-}45}$ but does not rescue TCR$_{C4}$ in a PANC-1 solid tumor model. Growth kinetics are shown for WT1$^+$ cell lines PANC-1 in live tumor-visualization assay in the absence (gray lines), or presence of T cells expressing TCR$_{37\text{-}45}$ (top A-C) or TCR$_{C4}$ (bottom D-E) **(A,D)** CD8$^+$ T cells were transduced with only TCR$_{37\text{-}45}$ (light-blue line) or a poly-cistronic construct containing TCR$_{37\text{-}45}$ or TCR$_{C4}$ and CD8αβ (red line, eCD8$^+$). **(B,E)** CD4$^+$ T cells were transduced with only TCR$_{37\text{-}45}$ or TCR$_{C4}$ (blue line) or a poly-cistronic construct containing TCR$_{37\text{-}45}$ or TCR$_{C4}$ and CD8αβ (green line, dpCD4$^+$). **(C,F)** A comparison of eCD8$^+$ alone (red line), dpCD4$^+$ alone (green line) or eCD8$^+$/dpCD4$^+$ in combination (1:1 ratio, purple line) is shown. For all conditions, arrows indicate addition of tumor cells to culture. Standard error of triplicate wells are shown. 4:1 E:T ratio, total T cells was kept consistent for all conditions.

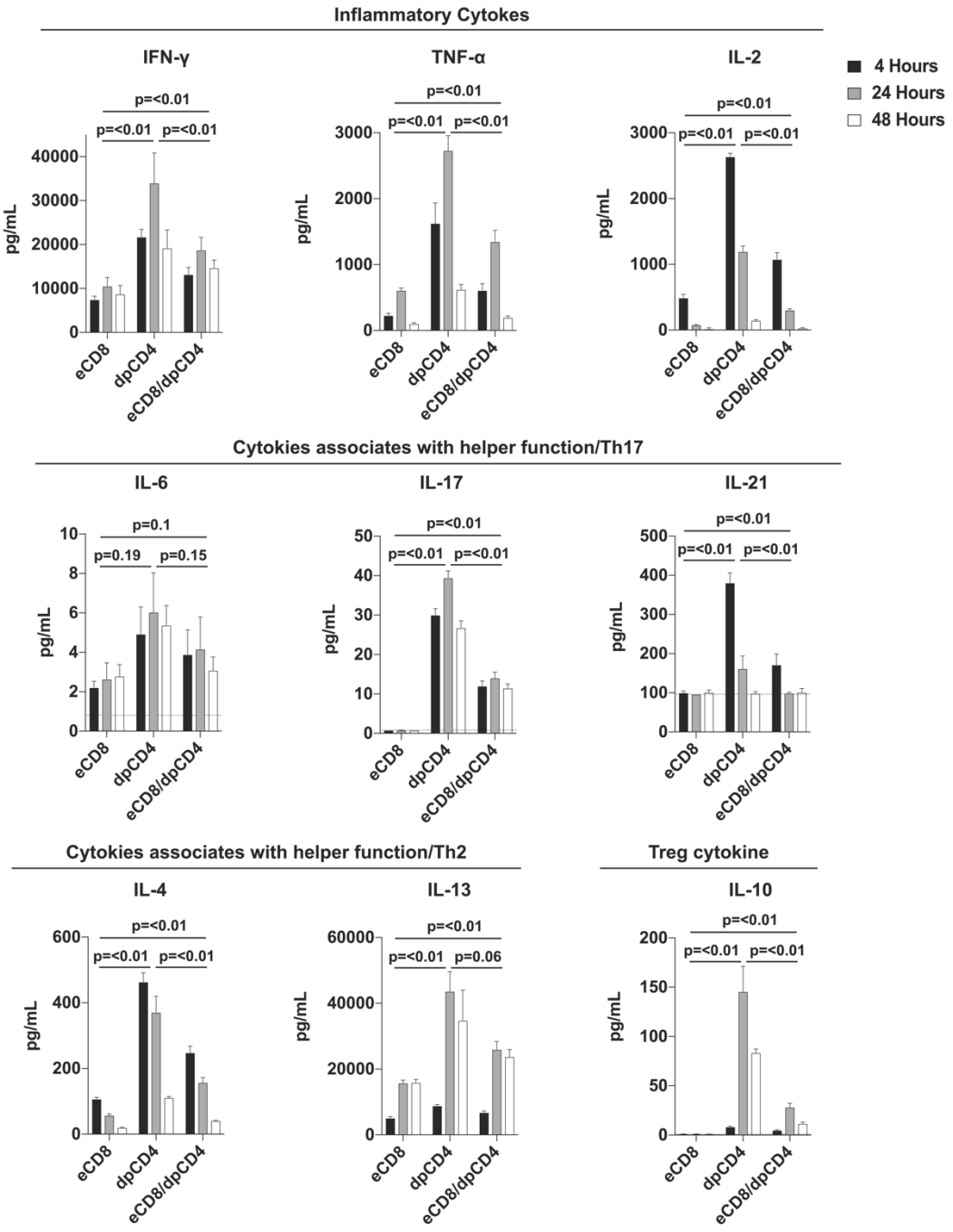


Figure 2.23 Cytokine profiles of peptide stimulated eCD8⁺, dpCD4⁺ and eCD8⁺/dpCD4⁺ $T_{TCR37-45}$ T cells. Cytokine production by $T_{TCR37-45}$ T cells was measured after 4, 24, and 48 hours of stimulation with T2 cells exogenously loaded with 10μg peptide. Except for IL-6, all cytokines are found at significantly higher concentrations in cultures of dpCD4⁺ cells compared to eCD8⁺ cells and the mixture of eCD8⁺/dpCD4⁺. The mixture of eCD8⁺/dpCD4⁺ produced significantly more cytokine than eCD8⁺ cells alone. Data are presented as mean ± standard deviation. Differences between groups were calculated using a student's t-test.

In a murine model $T_{TCR35\text{-}45}$ but not $T_{TCRC\text{-}4}$ reduce PANC-1 solid tumor.

We next investigated whether $eCD8^+/dpCD4^+$ $T_{TCR37\text{-}45}$ could reduce the burden of established PANC-1 tumors, which incorporate limited β1i into the proteasome without prolonged exposure to IFN-γ (**Figure 2.21B**). We used a NOD-*scid* IL2Rgamma^{null} (NSG) mouse model with established PANC-1 tumors to compare the activity of $eCD8^+/dpCD4^+$ with $T_{TCR37\text{-}45}$ or $T_{TCR\text{-}C4}$ or irrelevant control ($T_{TCR\text{-}irr}$) (**Figure 2.24A**). Indeed, $eCD8^+/dpCD4^+$ $T_{TCR37\text{-}45}$ effectively reduced detectable tumor in 4 of 5 mice, but $T_{TCR\text{-}C4}$ and $T_{TCR\text{-}irr}$ did not (**Figure 2.24B and C**). Thus, $T_{TCR37\text{-}45}$ was the only transduced T cell that significantly ($p = 0.03$) reduced PANC-1 tumor compared to tumor alone (**Figure 2.24D**).

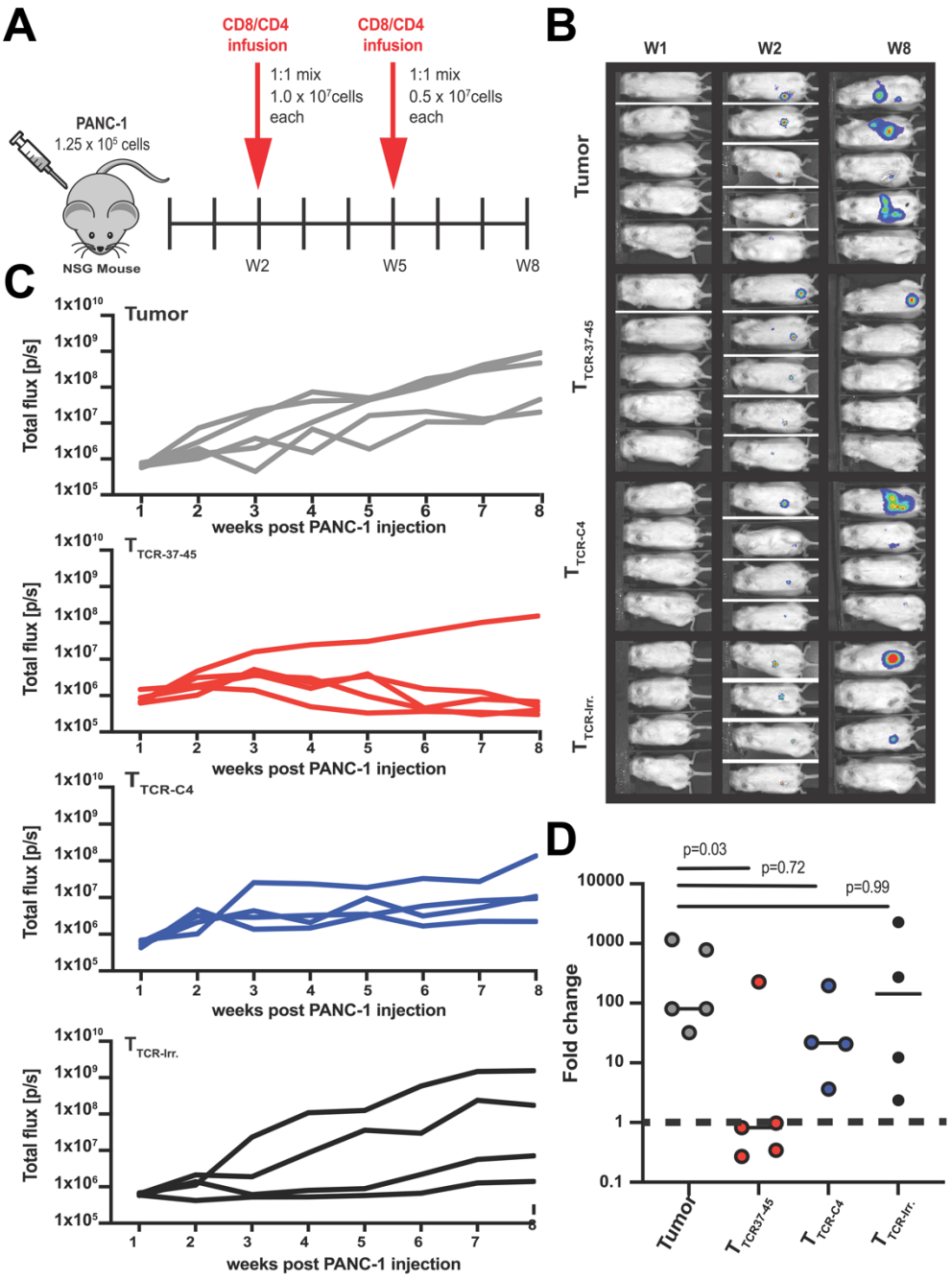


Figure 2.24 $T_{TCR37-45}$ cells reduce PANC-1 solid tumor burden in an NSG mouse model. (A) Schematic for NSG experimental treatment model. (N=2) Data from one of two biological replicates is shown. **(B)** Bioluminescence images of mice are shown at weeks 1, 2, and 8 post PANC-1_{leu} injection. All mice had detectable tumors at the time of adoptive transfer. **(C)** Total flux (p/s) for each treatment group was measured over time: Tumor (No T cells, n=5); $T_{TCR37-45}$ (n=5); T_{TCR-C4} (n=4) $T_{TCR-Irr.}$ (n=4). **(D)** Fold change relative to baseline (Week 1) at week 8 is shown for each treatment group. Data in panel D were analyzed by a Kruskal-Wallis test for non-parametric data. n.s., not significant. Horizontal bars indicate the mean.

Discussion

Here, we show evidence pointing towards a previously undescribed mechanism of AML resistance to immunologic targeting that is proteasome dependent. Proteasomes are responsible for the degradation of proteins into peptides subsequently presented by HLA class I molecules, and the enzymatic composition of the proteasome can determine precisely which peptides are presented (*38*). In our first clinical case, escape under the immunologic pressure from a focused antigen-specific T cell response may have resulted in selection of leukemic cells with altered protein processing machinery that abrogated presentation of the targeted epitope.

The standard proteasome is expressed in all somatic cells and contains a catalytic core composed of three enzymatic subunits (β1, β2, and β5). The immunoproteasome, by contrast, is expressed in hematopoietic cells and can be induced in most somatic cells by IFN-γ. It has a different catalytic core comprised of distinct enzymatic β-subunits (β1i, β2i, and β5i). The immunoproteasome and standard proteasome each produce a broad spectrum of peptides, including many identical peptides, but the immunoproteasome produces a more diverse set of peptides that are generally considered more immunogenic (*38*). Proteasomes assemble in a cooperative fashion. Subunits β2i and β1i are mutually incorporated into the pro-form of the proteasome and promote subsequent incorporation of β5i to form a functional immunoproteasome (*28*). In the absence of β2i and β1i, β2 is incorporated which inhibits β2i incorporation and prompts standard proteasome formation through incorporation of β5 then β1(**Figure 1.4**) (*39*). Formation of immunoproteasome is more efficient and takes precedence over standard proteasome formation, which is evident by the dominant presence of intact immunoproteasomes in cells that express the catalytic components of both proteasome types (*39, 40*). Unlike solid tumors that preferentially express the standard proteasome, AML and other hematopoietic cells generally express the immunoproteasome predominantly, and thus largely display peptides that have been processed by the immunoproteasome (*41*). As the $WT_{126\text{-}134}$ peptide appears to be exquisitely dependent on the presence of β1i as an integral part of the

immunoproteasome, processing of this epitope would be severely hampered following loss or downregulation of β1i.

The proteolytic machinery of tumor cells has been shown to impact the outcome of epitope-specific immunotherapies, and immunoproteasome deficiency is a feature of limited response to immunotherapy in solid tumors (*42, 43*). For example, a MAGE-A3 HLA-B40 restricted peptide was previously shown to be exclusively presented by immunoproteasome-expressing tumor cells (*44*), and the targeted $WT1_{126\text{-}134}$ epitope was shown to be poorly or not presented by solid tumor lines (*16*). By contrast, efficient presentation of the melanoma-antigen 1 (MART1) epitope, $MART1_{26\text{-}35}$, and of a renal cell carcinoma antigen, depends on standard proteasome processing and is obstructed by induction of the immunoproteasome, which no longer produces these epitopes and thus paradoxically can allow tumors to evade elimination by antigen-specific T cells that are producing IFN-γ following target recognition (*45*). Alternative processing defects can also be responsible for tumor evasion. In mice that lack expression of the endoplasmic reticulum aminopeptidase 1 (ERAP1), a component of the antigen processing machinery that trims peptides in the ER to fit in the binding groove of HLA class I molecules, HLA class I-positive, IFN-γ-responsive tumors progressed despite the presence of functional, antigen-specific cytotoxic T lymphocytes due to failure to present trimming-dependent epitopes (*46*). Consistent with our results, these findings highlight how changes in a tumor cell's antigen processing machinery can disrupt immune recognition, particularly if a single epitope is targeted.

Our study has limitations and raises questions that warrant future elucidation. We only observed one such 'resistant' case in this small initial trial and no additional HLA-$A2^+$, $WT1^+$, *PSMB9*^{low} primary AMLs could be obtained from our repository. Thus, no conclusions can be derived as to the frequency of this mechanism either under natural immunologic pressure or as a result of immunotherapy. The infused $T_{TCR\text{-}C4}$ also contained their endogenous EBV-specific TCR and signaling through this receptor cannot be formally excluded as an explanation for the activation signature observed at remission. The timing of the clinical sampling for the first patient and the unavailable 'fresh' leukemia prior to $T_{TCR\text{-}C4}$ infusion, hampered the ability to unequivocally demonstrate the recurrence was due to β1i loss. Although we observed a roughly 60%

reduction in *PSMB9* transcripts post $T_{TCR\text{-}C4}$ (**Figure 2.10A**), the unavailability of 'fresh' leukemia before $T_{TCR\text{-}C4}$ infusion did not allow for a direct comparison of incorporated proteasome subunits before and after $T_{TCR\text{-}C4}$ by immunoblot. The presence of other leukemias with low β1i and the progression of the second patient's AML with reduced WT1/β1i co-expression after $T_{TCR\text{-}C4}$ suggests there are many other potential obstacles to efficacy with engineered T cells. Establishing a murine model that can recapitulate loss of β1i as a result of immunological pressure or more generally the study of AML in the presence of T cell infusions to investigate escape mechanisms is warranted, especially as TCR gene therapy becomes more common.

The results underscore the importance of epitope selection for T cell therapy or vaccination strategies, in which the goal is to enhance the magnitude of T cell responses to selected epitopes (*47*). In particular, epitopes commonly identified in public databases reflect responses to epitopes expressed by target cells in inflammatory conditions or presented by professional antigen-presenting cells, both of which likely express the immunoproteasome, which can lead to selection of potentially sub-optimal targets. In the face of tumor heterogeneity, strategies targeting multiple proteins or epitopes using engineered T cells will likely be necessary. Targeting epitopes efficiently processed by both the standard proteasome and immunoproteasome may also increase the chances of circumventing proteasome-dependent immune-evasion (*15*). We pursued the latter strategy in searching for an epitope not exclusively dependent on the immunoproteasome. Our results suggest that selecting T cells/TCRs recognizing $WT1_{37\text{-}45}$ (See also Ruggiero et al. (*30*)) rather than $WT1_{126\text{-}134}$ should assure that changes in the relative immunoproteasome expression cannot abrogate presentation of WT1-expressing cells and allow for immune evasion, thus promoting successful clinical translation both for vaccines and TCR-engineered T cell therapies.

Materials and Methods

Clinical Protocol. The trial was approved by the Fred Hutchinson Cancer Research Center (FHCRC) Institutional Review Board, the U.S. Food and Drug Administration, and the National Institutes of Health Recombinant DNA Advisory Committee. It is registered at clinicaltrials.org as NCT1640301. Individuals with features of 'high-risk'

acute myeloid leukemia (AML) and their respective fully human leukocyte antigen (HLA)-matched (10/10) related or unrelated donors expressing HLA-A*02:01 (HLA-A2) were eligible for treatment.

Patient Selection. HLA-A2 genotype was confirmed by high-resolution typing (*48*) prior to pre-hematopoietic cell transplant (HCT) formal enrollment. High-risk features included: AML with adverse genetic abnormalities (*49*), AML beyond 1st remission, primary refractory AML, therapy-related AML, AML arising in patients with antecedent hematologic disorders, or AML with evidence of minimal residual disease (MRD) or overt disease at time of HCT (*49-51*). Exclusion criteria included: active central nervous system (CNS) disease, HIV seropositivity, grade ≥3 graft-versus-host disease (GVHD), and no available Epstein-Barr Virus (EBV)-seropositive matched donor. On this two-arm clinical trial, patients with no detectable disease post-HCT received treatment with T_{TCR-C4} on the Prophylactic Arm, and patients with evidence of disease post-HCT received T_{TCR-C4} on the Treatment Arm (*10*).

Detailed case report of Patient 1. The first individual described here was a 27-year-old man who, at the age of 15 was diagnosed with AML. His AML proved resistant induction chemotherapy. The patient finally achieved a complete remission after single agent gemtuzumab ozogamicin (Mylotarg). He underwent a first 10/10 matched related mobilized allogeneic HCT with a preparative immunodepleting regimen. HCT was successful with full engraftment and no evaluable disease (NED) detected post-HCT. His AML relapsed seven years later and was now located in the marrow (61% blasts in the marrow, in the central nervous system (67% blasts) and within a S2-S3 nerve root myelosarcoma. He received systemic re-induction chemotherapy and radiation which again achieved a complete response. He received a second reduced-intensity HCT from a 10/10 matched unrelated donor after a preparative regimen. This second HCT achieved 100% donor engraftment and NED post-HCT. He continued to receive IT MTX for 7 months post-HCT and his AML relapsed 7 months after treatment discontinuation with predominantly cervical C5 and C7 myelosarcomas and minimal marrow involvement (0.002% blasts). He received a cycle of chemotherapy and radiation. Two

months after the start of this last treatment and immediately prior to a first infusion of 10^{10} $T_{TCR\text{-}C4}/m^2$ on the Treatment Arm, he had no evaluable bone marrow AML. The patient was monitored and showed no evidence of GVHD or toxicity to cells expressing low Wilms' tumor antigen 1 (WT1) including the hematopoietic system, kidneys, lung, heart, liver, and skin were observed. He remained with NED and no additional treatments for 368 days before relapsing with a bladder myelosarcoma. His concurrent bone marrow aspirate was positive for AML blasts. He received a second infusion of 10^{10} $T_{TCR\text{-}C4}/m^2$, radiation to the bladder, and then additional chemotherapy which did not control his AML. In accordance with the patient and his family, no additional treatments were administered, and he succumbed to progressive AML a year after his 2^{nd} $T_{TCR\text{-}C4}$ infusion. Although the patient was followed closely after the last relapse, the study team respected his wishes for minimal intervention and sample collection, which limited the amount of material obtained.

Detailed Case Report of Patient 2. The second patient described was a 47-year-old man who at the age of 46 was diagnosed with a myelodysplastic syndrome with excess blasts (blasts constituted 15.6% of marrow cells). A month later he presented with circulating blasts and received chemotherapy which did not avert progressive disease, and no remission was achieved. His circulating blast count was then maintained with hydroxyurea. He was enrolled in a transplant protocol specifically designed for patients with residual leukemia and received a myeloablative regimen prior to the infusion of peripheral blood stem cells from a male matched unrelated donor. Twenty-eight days after transplant, he presented with minimal residual disease (0.28% abnormal marrow blasts). He received additional chemotherapy, but his disease continued to progress. A bone marrow biopsy performed 94 days after HCT showed 12.9% blasts. Twelve days prior to the infusion of 10^{10} $T_{TCR\text{-}C4}/m^2$ on the protocol Treatment Arm (150 days after HCT), his marrow showed 69% blasts; immediately prior to cell infusion his complete blood count included a total of 8030 white blood cells/mcl with 66% blasts (5310 cells/mcl), 9% neutrophils (720 cells/mcl) and 13% lymphocytes (1040 cells/mcl). Although his absolute blast count moderately decreased immediately after infusion, his AML subsequently progressed despite persistence of $T_{TCR\text{-}C4}$ at greater than 3% of

$CD8^+$ T cells in the peripheral lymphocyte compartment. He succumbed to progressive disease 15 days after $T_{TCR\text{-}C4}$ transfer.

Assessment of disease status. Morphology, multiparameter flow cytometry (*52*), standard cytogenetics (*53*), or genomic technologies (*54*) were routinely performed on bone marrow aspirates that were obtained from the patients. Any detectable residual disease was considered to indicate positivity for MRD.

Isolation of TCR_{C4} and lentiviral vector construction. TCR_{C4} was selected from a screen of more than 1000 T cell clones from more than 50 $HLA\text{-}A2^+$ healthy donors based on its comparatively higher functional avidity for the $WT1_{126\text{-}134}$ peptide (RMFPNAPYL) as previously described (*10*). TCR_{C4} was inserted into the pRRLSIN.cPPT.MSCV/GFP.wPRE vector obtained from the lab of Richard Morgan after excising the green fluorescent protein (GFP) gene (*55*).

Generation of $T_{TCR\text{-}C4}$. All ex vivo processing of infusion products was performed in the FHCRC cGMP Cell Processing Facility (CPF) at as previously described (*10*). Briefly, substrate $CD8^+$ T cells for generating the TCR_{C4}-transduced T cells ($T_{TCR\text{-}C4}$) were obtained from an aliquot of the granulocyte colony stimulating factor (G-CSF)-stimulated peripheral blood mononuclear cells (PBMC) donation (mobilized leukapheresis) of the patients' corresponding allogeneic stem cell donors. Cells were stimulated with the HLA-A2-restricted EBV ($EBV_{280\text{-}288}$ BMLF1 [GLCTLVAML]) peptide, transduced using the pRRLSIN.cPPT.MSCV/$TCR_{C4}\beta$-P2A-$TCR_{C4}\alpha$/wPRE lentiviral vector and sorted by flow cytometry; cells were labeled with peptide:HLA tetramers that selectively bound the WT1-specific TCR labeled with allophycocyanin (APC) and the EBV-specific TCR labeled with phycoerythrin (PE); double positive cells were selected followed by a secondary, non-specific expansion and infusion into the patient. Total production time was 4 to 6 weeks. Quality control of the infused products included assessment of inserted lentiviral vector copy number per cell (≤5 copies per cell), envelope vesicular stomatitis virus (VSV)-G DNA by qPCR as a surrogate marker for replication-competent lentivirus (RCL) (<10copies/50ng DNA) and binding to the $WT1_{126\text{-}134}$ HLA-A2-restricted

tetramer (≥30% of live cells). HLA-A2/WT1$_{126\text{-}134}$ specific tetramers (produced by the FHCRC Immune Monitoring Core facility) were used to detect $T_{TCR\text{-}C4}$ in PBMCs collected after infusions, with a staining sensitivity of 0.01% of total CD8$^+$ T cells, as previously described (*56*).

Clonotype identification in the pre-infusion product and tracking by high throughput TCRβ sequencing. To identify clonotypes composing infused $T_{TCR\text{-}C4}$, high-throughput TCRβ sequencing (HTTCS) analysis was performed on HLA-A2/WT1$_{126\text{-}134}$ tetramer$^+$ cells from the pre-infusion product to obtain >99% purity (*57*). As the TCR_{C4} sequence was codon-optimized, HTTCS could not identify the introduced WT1-specific TCR. Instead, HTTCS identified clonotypes based on the endogenous EBV-specific TCR sequence. To track the identified clonotypes, HTTCS was performed on whole PBMC obtained after transfer (*57*). Briefly, DNA was extracted from T cell products and sorted PBMC using Qiagen Maxi DNA isolation kits (QIAGEN Inc.). TCRβ complementary determining region (CDR)3 regions were amplified and sequenced from 750ng of extracted DNA by Adaptive Biotechnologies Corp using the ImmunoSEQ assay with “deep” resolution, as previously described (*58*). Raw sequence data was filtered using the Adaptive bioinformatic website based on the TCRβ V, D and J gene definitions provided by the International ImMunoGeneTics collaboration (IMGT) (*59*) using the IMGT database (www.imgt.org). Productive nucleotide sequences were used for all tracking experiments. The limit of detection of HTTCS was set at 0.001% of all TCR reads, below which frequency could not be reliably determined (*59*). Only clonotypes present in the infusion products were tracked in PBMC obtained after infusions. Infusion product ‘bystander’ clonotypes were defined as those not detected in vivo post-infusion and not expanded during the ex vivo culture process and were eliminated for further analysis (*57*). The frequency of each HTTCS-detected clonotype is based on all TCR Vβ reads which include CD4$^+$ and CD8$^+$ T cells.

Flow cytometry on T cells. $T_{TCR\text{-}C4}$ in infusion products and post-transfer PBMCs were identified by binding to the APC dye-labeled HLA-A2:WT1$_{126\text{-}134}$ tetramer (FHCRC in-house production) and analyzed by flow cytometry after staining for 30 minutes with

fluorochrome-conjugated monoclonal antibodies to CD16 (1:150 dilution, PE-Cy5, clone 3G8, Becton Dickinson. 555408), CD19 (1:80 dilution, PE-Cy5, clone HIB19, Becton Dickinson, 555414; for dump channel), CD4 (1:50 dilution, brilliant ultraviolet (BUV) 496, clone SK3, Beckton Dickinson, 612936), CD8 (1:80 dilution, Qdot 605, clone 3B5, Life Technologies, Q22157), CD28 (1:200 dilution, brilliant violet (BV) 421, clone CD28.05, BioLegend, 302930), CD27 (1:200 dilution, Fluorescein isothiocyanate (FITC), clone M-T271, BioLegend, 356404), CD62L (1:50 dilution, PE, clone SK11 Becton Dickinson, 341012) and CD45RO (1:100 dilution, Alexa Fluor 700, clone UCHL1, BioLegend, 304218). A representative example of the initial gating algorithm is shown in **Figure 2.25A**.

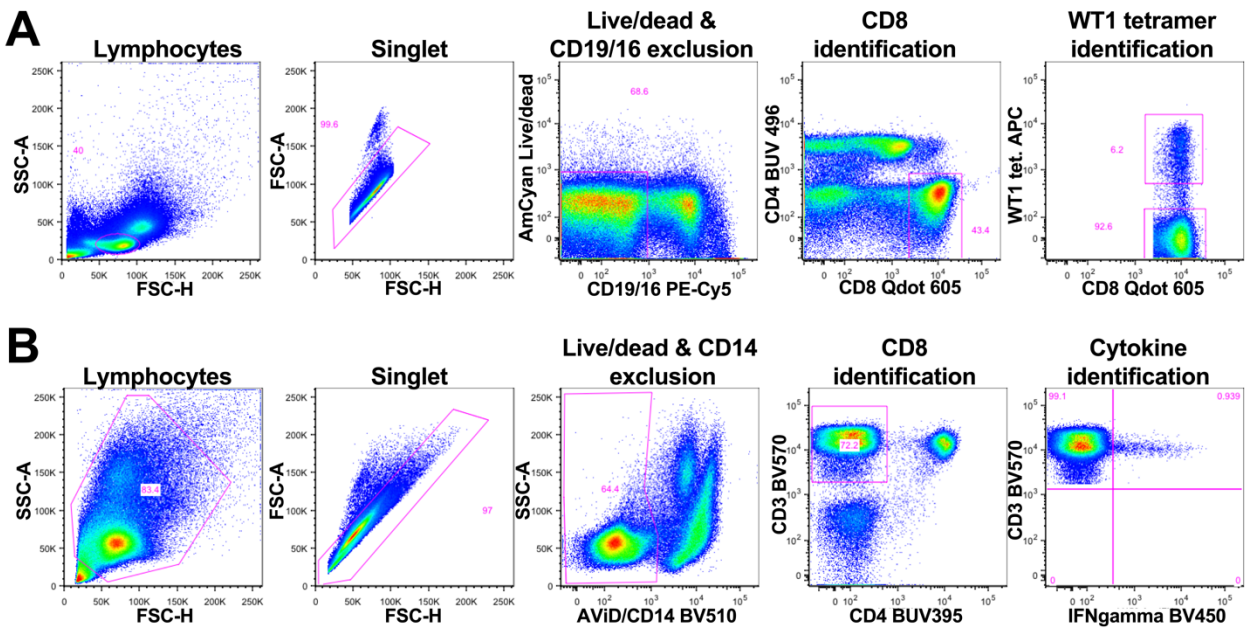


Figure 2.25 Gating strategies for surface and intracellular stains. (A) The gating strategy for surface staining is shown. Panels from left to right: Initial gating was on lymphocytes (side scatter (SSC)-area (A) versus forward scatter (FSC)-height (H)), followed by singlet exclusion (FSC-A versus FSC-H), followed by exclusion of dead cells and CD19 and CD16-expressing cells (AmCyan Live/dead versus CD19/16 PE-Cy5), followed by CD8 identification (CD4 BUV 496 versus CD8 Qdot605), followed by $WT1_{126\text{-}135}$-HLA-A2-tetramer^{+} cells identifying $T_{TCR\text{-}C4}$ (WT1-tetramer APC versus CD8 Qdot605). A representative example is shown. **(B)** The gating strategy for intracellular staining is shown. Panels from left to right: Initial gating was on lymphocyte identification (SSC-A versus FSC-H), followed by singlet exclusion (FSC-A versus FSC-H), followed by exclusion of dead cells and CD14-expressing antigen presenting cells (SSC-A versus AViD/CD14 BV510), followed by CD8 identification (CD3 BV570 versus CD4 BUV395), followed by cytokine identification (CD3 BV570 versus IFN-γ BV450). A representative example for IFN-γ is shown.

Intracellular cytokine staining. Cryopreserved PBMCs were thawed and rested overnight in RPMI-1640 medium, supplemented with 10% fetal bovine serum (FBS) (R10). Intracellular cytokine staining and stimulations were performed using a 15-color staining panel, as previously described (*60, 61*). The stimulation cocktail contained either $WT1_{126\text{-}134}$ peptide in R10 at a final peptide concentration of 1 μg/mL or R10 with an equivalent amount of dimethyl sulfoxide (DMSO) used to preserve the peptides (negative control) or staphylococcal endotoxin B (Sigma-Aldrich, 11100-45-1, positive control). Modifications to the panel included excluding interleukin (IL)-21, CD56 and

CXCR5, changing IL-2 to PE-dazzle 594 (1:100 dilution, clone MQ1-17H12, BioLegend, 500344). Cells were analyzed on an LSRII flow cytometer (Becton Dickinson), using FACS-Diva software. The resulting flow cytometry data were analyzed with FlowJo v9.9.6 (Treestar). Percent cytokine expression (% of tetramer$^+$ cells that are both cytokine$^+$ and tetramer$^+$ in the DMSO only sample [negative control]) was calculated for each sample. Percent tetramer$^+$cytokine$^+$ cells in the presence of DMSO without peptide varied among samples from 0-2.27% with a median of 0.745%. A representative example of the initial gating algorithm is shown in **Figure 2.25B**.

Serum EBV and CMV quantitation. EBV PCR was developed in-house using primers specific for the EBER gene (*62*); the limit of quantitative detection (the minimum virus concentration that gives a positive result in 95% of replicates) was 250 IU/mL (2.40 log IU/mL). Quantitative results less than 250 IU/mL are described as very low positive. CMV PCR was developed in-house using primers specific for the UL123 and gB genes (*63*); the limit of quantitative detection (the minimum virus concentration that gives a positive result in 95% of replicates) was 20 IU/mL (1.3 log IU/mL). Quantitative results less than 20 IU/mL are described as very low positive.

Single Cell RNA Sequencing (scRNAseq). ScRNAseq (*64*) was performed on samples from two patients enrolled in NCT1640301. Patient 1 ‘remission’ and ‘relapse’ samples were selected to be approximately 100 days after the last infusion to minimize changes associated with infusion of large numbers of ex vivo-manipulated cells. Patient 2 samples were selected before $T_{TCR\text{-}C4}$ infusion and five days after. PBMCs for both patients were thawed, washed, and labelled using the 10x Genomics 3’ Chromium v2.0 (Patient 1) or 5’ Chromium v2.0 (Patient 2) platform per manufacturer's instructions. Library preparation was performed as per manufacturer’s protocol with no modifications. Library quality was confirmed by Illumina TapeStation high sensitivity (evaluates library size), qubit (evaluates dsDNA quantity), and KAPA qPCR analysis (KAPA Biosystems, evaluates quantity of amplifiable transcript). Samples were mixed in equimolar fashion and sequenced on an Illumina HiSeq 2500 *rapid run* mode according to standard 10x genomics protocol.

Computational analysis for scRNAseq on Patient 1. The Cell Ranger Single-Cell Software Suite (v2.0.0) was used to perform sample demultiplexing, barcode processing and single-cell 3' gene counting (http://10xgenomics.com). First, raw base BCL files were demultiplexed using the Cell Ranger *mkfastq* pipeline into sample-specific FASTQ files. Second, these FASTQ files were individually processed using the Cell Ranger *count* pipeline. Reads, which contain cDNA inserts, were aligned to the hg38 human reference genome (Ensembl) and the known transgene codon-optimized sequence using Spliced Transcripts Alignment to a Reference (STAR) (*65*). Aligned reads were then filtered for valid cell barcodes and unique molecular identifiers (UMIs). Cell barcodes with 1-Hamming-distance from a list of known barcodes were examined. UMIs with sequencing quality score > 10% and not homopolymers were retained as valid UMIs. A UMI with 1-Hamming-distance from another UMI with more reads, for a same cell and a same gene was corrected to this UMI with more reads. The default estimated number of cells used were 10,000 cells for both 'remission' and 'relapse' timepoints. Samples from both time points were aggregated using the Cell Ranger *aggr* pipeline resulting in one gene-barcode count matrix to be used for downstream analyses. A correction for sequencing depth was performed during the aggregation.

Statistical analysis for scRNA-seq on patient 1 PBMC. Standard analyses were performed using the Seurat R package (*66, 67*). Following sequence alignment and filtering, cells with a percentage of mitochondrial genes > 20%, a unique gene counts < 200 or > 15,000 UMIs were removed. Based on these thresholds, 68 cells were discarded, resulting in a data set of 7,688 cells for downstream analyses. Only genes with at least one UMI count detected in at least three cells were kept (16,590 genes). A library-size normalization was performed for each cell. UMI counts were divided by the total number of UMI in each cell and multiplied by 10,000. Data were then log transformed and corrected for unwanted sources of variation such as the number of detected UMIs and percentage of mitochondrial content using the *ScaleData* R function as described in the Seurat manual. As high mitochondrial gene expression might be suggestive of low-quality libraries (*68*), the corrected-normalized gene-barcode matrix

was used as input to perform Principal component analysis (PCA), clustering, and Uniform Manifold Approximation and Projections for Dimension Reduction (UMAP) analyses, whereas the normalized gene-cell barcode matrix was used as input for the differential expression analysis as described below. PCA was performed using the top 2,000 most variable genes. The top 20 principal components (PCs) were then down-selected for UMAP visualization and clustering. UMAP with these parameters (n.epochs = 400, min.dist = .2, local.connectivity = 20) was performed using the *RunUMAP* R function. Cell clustering was performed using a graph-based clustering method implemented in Seurat (*FindNeighbors* and *FindClusters* R function) using a resolution of 0.25.

Differential expression analysis was performed using the R package Model-based Analysis of Single-cell Transcriptomics (MAST) (*69*) implemented within the *FindMarkers* R function from the Seurat R package. MAST provides functionality for significance testing of differential expression using a Hurdle model, gene set enrichment and facilities for visualizing patterns in residuals indicative of differential expression. It is a two-part generalized linear model that simultaneously models the rate of expression over the background of various transcripts (logistic regression), and the positive expression mean (Gaussian). These tests are two sided and adjust for the bimodal nature of scRNAseq data. Differential expression was performed between clusters, and between conditions (such as transgenic versus endogenous $CD8^+$ T cells). In each case, the normalized gene-cell barcode matrix was used as input. The model included the cellular detection rate (CDR), defined as the total of UMIs in a given cell, as a covariate to correct for biological and technical nuisance factors that can affect the number of genes detected in a cell (such as cell size and amplification bias). Genes were declared significantly differentially expressed at a false discovery rate (FDR) of 5% and a fold-change > 1.5.

Computational and statistical analysis for scRNA-seq on patient 2. A relaxed quality control (QC) strategy was used, and we chose to only exclude cells with large mitochondrial proportions as proxy for cell damage. A stringent cutoff (2 Median Absolute Deviation; MAD) was used for percent of mitochondrial genes, resulting in a data set of 10,700 cells for downstream analyses (2,477 cells were filtered out). Only

genes with at least one UMI count detected in at least three cells were retained (17,475 genes).

Data integration and standard analyses were performed using the Seurat R package (*67*). Gene expression data integration was performed using the time point as factor, resulting in two different datasets. To integrate the gene expression values, we first separately normalized each of our partitions using the *SCTransform* R function (*70*) with *vars.to.regress = "percent.mito"*. Data integration was then performed using the two R functions *FindIntegrationAnchors* and *IntegrateData* (with *dims = 1:50*). PCA was performed on the integrated dataset was performed using the *RunPCA* R function. The top 50 PCs were then used for UMAP visualization and clustering. UMAP with this parameter (*min.dist = .01*) was performed using the *RunUMAP* R function. Cell clustering was performed using a graph-based clustering method implemented in Seurat (*FindNeighbors* and *FindClusters* R function) with different resolutions. Differential expression analysis was performed using the R package, MAST (*69*), implemented within the *FindMarkers* R function from the Seurat R package as described in the *Statistical analysis for scRNA-seq on patient 1 PBMC* section

Immunohistochemistry (IHC). WT1 IHC was performed by a College of American Pathologists Clinical Laboratory Improvement Amendments (CLIA) certified laboratory with a standard clinical protocol. Paraffin-embedded core bone marrow biopsies or particle preparations were sectioned (4μM) and antigen retrieval performed by heat treatment in 10mM sodium citrate buffer (pH 6) before staining with DAKO mouse anti-WT1 monoclonal antibody (mAb) clone 6F-H2, at a 1:800 dilution. IHC for HLA Class I (HLA-A, B, C) was performed in a CLIA certified laboratory using Paraffin-embedded core bone marrow biopsies or particle preparations were sectioned (4μM). Antigen retrieval was performed at heat treatment at 95°C for 20 minutes in DIVA citrate buffer (pH 6.0), before staining with mouse monoclonal major histocompatibility complex (MHC) Class I antibody clone EMR8-5 (Abcam) at a 1:1000 dilution.

AML HLA-A2 flow cytometry. Patient PBMCs were assessed at relapse for HLA-A2 expression using Flow Cytometry. Briefly, PBMCs containing ≥89% AML were stained

for 30 minutes CD34 (1:50 dilution, BV421, clone 561, BioLegend) to identify AML blasts and for HLA-A2 (1:600 dilution, APC, clone BB7.2, Beckton Dickinson). Samples were evaluated on an FACSymphony flow cytometer and analyzed as described above.

Gene expression analysis from archived formalin fixed paraffin embedded (FFPE) samples and peripheral blood. Tumor tissues was obtained post-second HCT and pre-$T_{TCR\text{-}C4}$ from a malignant chloroma (**Figure 2.7A**) and at the time of relapse from PBMCs (day 581 after the first $T_{TCR\text{-}C4}$ infusion). Both samples contained ≥89% AML. For FFPE tissues, total RNA was extracted using the AllPrep DNA/RNA FFPE Kit (Qiagen); for PBMCs, RNA was extracted using Blood RNEasy miniprep kit (Qiagen). RNA gene expression quantification was performed using the nCounter Immunology Panel (Human V2, code set NS_Immunology_v2_c2328) that contains 579 immunology related human genes plus 15 internal housekeeping control genes per manufacturer's protocol (NanoString Technologies). The raw counts were normalized using the geometric mean of the internal controls using nSolver software (Version 4.0) and all samples passed QC parameters (**Figure 2.10B**) (*71*). For the blood sample collected after $T_{TCR\text{-}C4}$ infusion, the normalized counts for the panel-represented genes encoding the cap proteins ([*PSMC2*] and [*PSMD7*]) and the catalytic sites of the immunoproteasome (IP) and (β5i [*PSMB8*], β1i [*PSMB9*], β2i [*PSMB10*], β5 [*PSBM5*] and β2 [*PSMB7*]) were expressed as a fold-change compared to expression in the cervical chloroma sample obtained before $T_{TCR\text{-}C4}$ infusions. Of note, although the purity of both samples was ≥89% AML, the pre-$T_{TCR\text{-}C4}$ sample was a chloroma and considered to be nearly 100% AML, whereas the relapsed PBMC sample on which RNA gene expression was performed contained only 89% AML. Thus, the calculated reduction of β1i [*PSMB9*] is likely an underestimate due to immune and hematopoietic cells (expressing elevated immunoproteasome genes) in the PBMC sample.

WT1 sequencing. Clinical sequencing for oncogenic mutations including for WT1 was performed on a post-relapse bone marrow biopsy specimen from day 58 after 2nd infusion using the University of Washington OncoPlex platform. Performance details and characteristics of this technology have been previously described (*72*). Tumor

percentage in the specimen was 35%. The study quality was “high,” full length WT1 was sequenced with good coverage of all exons, and average WT1-specific coverage for this sample was 452X.

Whole Exome sequencing. Isolated AML cells were sent to University of Washington Center for Precision Diagnostics for whole exome sequencing. >95% of genes had at least 20X coverage. A single nucleotide polymorphism was detected in [*PSMB9*], which was a G to A transition at position 179 (c.179G>A) yielding a change from an arginine to histidine (p.Arg60His). This polymorphism is reported to occur in 26.8% of the population (NCBI ALFA allele frequency, global population) (*73*).

T cell isolation, activation, and transduction for cytotoxicity assays, live tumor-cell killing assays, and in vivo studies. $CD8^+$ and $CD4^+$ T cells were isolated from healthy HLA-A*02:01 donor PBMCs using the EasySep Human positive selection kits (STEMCELL Technologies) following the manufacturer’s protocol. T cells were seeded at 10^6 T cells/ml with 50 IU/ml IL-2 and were activated with either CTS Dynabeads CD3/CD28 (Life Technologies), T Cell TransAct, human (Miltenyi Biotec), or 30ng/ml OKT3 antibody according to manufacturer’s instructions. After 24 to 48 hours of stimulation, the stimulant (beads, TransAct, or OKT3) was removed and T cells were transduced with lentiviral particles encoding TCR_{C4}, $TCR_{37\text{-}45}$ or an irrelevant virus-specific TCR ($TCR_{Irr.}$). In some cases, transduction efficiency was enhanced by spinfecting T cells at 30°C and 1055 x g for 90 minutes with 5 µg/mL polybrene (Sigma-Aldrich) or 10mg/ml protamine sulfate. Because $TCR_{37\text{-}45}$ is dependent on CD8 expression in assays involving use of $CD4^+$ cells, the CD8α and CD8β chains were added to the existing lentiviral constructs to generate double positive $CD4^+$ ($dpCD4^+$) and ‘enhanced’ $CD8^+$ ($eCD8^+$) T cells. Cells were maintained every 2 to 3 days with IL-2 and media supplemented with 5 to 10% human serum, L-glutamine (Thermo Fisher Scientific) or glutamax (Thermo Fisher Scientific) and antibiotics (penicillin (200U/ml) and streptomycin (0.1mg/ml) (Thermo Fisher Scientific). After seven days, T cells were stained for 30 min and then selected for $CD8^+$ (1:100, BV421, clone RPA-T8, Becton Dickinson, 562428) and tetramer-binding (APC, FHCRC) using flow cytometry and

subsequently placed in non-specific rapidly expanding protocol (REP) (*74*) for 9 to 12 days before use. Cells used for ^{51}Cr release assays, live-tumor assays, recognition assays, and in vivo studies were used after sort purification. Cells used in the AML cytotoxicity assays underwent Cas9 electroporation before WT1-specific TCR transduction to knockout both their TCRα and TCRβ chains of the endogenous TCR (see CRISPR methods below).

Assessment of $WT1_{126-134}$ dependence on specific immunoproteasome subunits. Four HEK-293EBNA cell lines, each expressing one proteasome subtype (*29, 75*), were used to study processing of the peptide $WT1_{126-135}$ by the four proteasome isoforms. The day before transfection, 30,000 HEK-293 cells were seeded in collagen-treated 96-well plates (Greiner) in 100µl Iscove's Modified Dulbecco Medium (IMDM) plus 10% FBS supplemented with essential amino acids, penicillin (200U/ml) and streptomycin (0.1mg/ml). The next day, 293 cells were transfected in duplicate using LT-transit (Mirus Bio), 50 ng HLA-A*02:01 and the indicated concentration of the pCMV6-XL4-WT1 variant D cDNA (Origene). As control, 1µM peptide $WT1_{126-135}$ was pulsed on HLA-A2 transfected cells for 30 minutes. Peptide was removed before addition of T_{TCR-C4}. After 24 hours, transfected cells were incubated with 10,000 T_{TCR-C4} in 200µl IMDM medium supplemented with 10% FBS, essential amino acids, penicillin (200U/ml), streptomycin (0.1mg/ml) and 25 U/ml recombinant IL-2. After 20 hours of co-culture, the culture supernatant was collected, and tumor necrosis factor (TNF)-α released by T_{TCR-C4} was measured using WEHI -164 cl13 cells as previously described (*76*).

Collection, sorting, and RNA extraction of normal donor peripheral blood stem cells and primary AMLs. For our study, we had access to historical data from six normal donor mobilized peripheral stem cells and peripheral blood or bone marrow AML samples from 38 patients. All donors and patients signed the informed consent for Fred Hutch protocol #1690: Fred Hutch/University of Washington Hematopoietic Diseases Repository. Cryopreserved samples were thawed, and RNA was extracted as previously described (*77*). Briefly, thawed cells were stained for 30 minutes with CD45 (APC-H7, clone 581, BioLegend, 560178) to identify lymphoid and myeloid populations, as well as CD34 (APC, Clone 581, BioLegend, 343510) and CD117 (PE, clone 104D2,

BioLegend, 313204) to identify the immunophenotype of leukemic blasts. Viable AML blasts were sorted on a BD FACS Aria II instruments for viable AML blasts. Cells from the $CD34^+$ enriched normal donor mobilized peripheral stem cells and leukemic blast populations were then lysed in RLT-Plus buffer (Qiagen) supplemented with β-mercaptoethanol (Sigma-Aldrich), and RNA were extracted by using the AllPrep DNA/RNA Mini kit (Qiagen) and quantified by using the Nanodrop spectrophotometer (Thermo Fisher Scientific). RNA integrity number (RIN) was determined on an Agilent 2200 TapeStation (Agilent Technologies).

RNA sequencing on viable leukemia blasts and normal mobilized peripheral blood stem cells. RiboErase (Roche) was utilized to deplete ribosomal RNA, and transcriptome libraries were generated using a KAPA Stranded RNA-Seq Library Preparation Kit. Paired-end 50 bp or 75 bp reads were sequenced on either an Illumina HiSeq 2500 instrument or NovaSeq 6000 instruments (Illumina). A pipeline of established bioinformatics tools, including FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), STAR2 (*65*), RSeQC (*78*), Subread/featureCounts (*79*), was used by the Fred Hutch bioinformatics shared resource to assess data quality, map reads to GRCh37/hg19 and compute gene-level expression based on GENCODE gene annotation V31 (*80*). The expression matrix was processed by edgeR (*81*) and normalized using the Trimmed Mean of M-values (TMM) method (*82*).

Identification of the HLA-A*0201-restricted WT1-specific $TCR_{37\text{-}45}$.

1. **Generation of WT1 epitope-specific T cell lines.** Antigen-specific T cell lines were generated from two donors that were specific for the following WT1 peptides: $WT1_{10\text{-}18}$ (ALLPAVPSL), $WT1_{22\text{-}31}$ (GGCALPVSGA), $WT1_{37\text{-}45}$ (VLDFAPPGA), $WT1_{38\text{-}46}$ (LDFAPPGAS), $WT1_{187\text{-}195}$ (SLGEQQYSV), $WT1_{235\text{-}243}$ (CMTWNQMNL), $WT1_{238\text{-}246}$ (WNQMNLGAT), and $WT1_{242\text{-}250}$ (NLGATLKGV). Ten lines were generated from each donor targeting each epitope (160 total) as previously described (*74*). Briefly, dendritic cells (DCs) were derived from the plastic adherent fraction of PBMCs after culture for two days (days -2 to 0) in

media supplemented with 1000 U/ml IL-4 and 800 U/ml Granulocyte-Macrophage colony stimulating factor (GM-CSF). On day -1, a maturation cocktail containing 10ng/ml TNF-α, 10ng/ml IL-1β, 1000 U/ml IL-6 and 1000ng/ml prostaglandin E2 (PGE_2) was added. On day 0, DCs were harvested, washed, and pulsed with peptide. $CD8^+$ T cells were isolated from PBMCs using anti-CD8 microbeads and stimulated with peptide-pulsed DCs in the presence of 30ng/ml IL-21. Cultures were fed every 2 to 3 days by exchanging half of the medium and adding 12.5 U/ml IL-2, 5ng/ml IL-7 and 5ng/ml IL-15. T cells were re-stimulated every 10 days by culturing at a 1:2 ratio with irradiated, peptide-pulsed, autologous PBMCs.

2. **Sort-purification of WT1-specific T cell lines.** T cell lines from the most responsive donor PBMCs were stained with peptide-HLA-A*02:01 tetramers specific for $WT1_{37\text{-}45}$ and $WT1_{235\text{-}243}$. The optimal tetramer concentration for staining was determined by titrating each tetramer on a cell population containing antigen-specific T cells and evaluating specific staining above background (non-specific) T cells. The tetramer-stained cells were sorted for high tetramer expression and expanded using a modified REP with IL-2 (50 U/mL), OKT3 (1 mg/mL), and 100-fold excess of irradiated feeder cells in 8 mL of RMPI-1640 supplemented with 10% heat inactivated human serum, 25 mM HEPES, 12 mM L-glutamine, and 55 μM β-mercaptoethanol (CTL media). Cultures were maintained with CTL media and 50 U/mL IL-2 every 2 to 3 days.
3. **Evaluation of antigen-specific cytokine production in response to $HLA\text{-}A2^+$ K562 cells.** Sorted antigen-specific T cells were co-cultured at a 1:1 ratio with HLA-A2-transduced K562 cells ($A2^+$ K562). After 6-hour incubations in the presence of golgi-inhibitors (BD GolgiPlug and GolgiStop), cells were surface-stained with anti-CD8 and then fixed (BD Cytofix/Cytoperm) before intracellular labelling with anti-interferon (IFN)-γ in BD Perm/Wash buffer. The cells were then analyzed by flow cytometry to determine the percentage of IFN-γ^+ cells for each sample.
4. **Isolation and expression of a $WT1_{37\text{-}45}$-specific TCR.** $WT1_{37\text{-}45}$-specific T cells that exhibited the most robust IFN-γ response when cultured with $HLA\text{-}A2^+$ K562s, were further sorted for the highest population of tetramer binding T cells

(about 1% of the total population). The sorted tetramer^{hi} cells were lysed, RNA was isolated, and full-length TCRα- and TCRβ- encoding genes (*TRA* and *TRB*) were amplified by RACE PCR (TakaraBio). A dominant TCRA/TCRB clonotype was evident after sequencing, which was confirmed to be $WT1_{37\text{-}45}$ specific by expression in Jurkat cells followed by $WT1_{37\text{-}45}$ specific tetramer staining. This TCR ($TCR_{WT1\text{-}37\text{-}45}$ #1) was then synthesized as a codon-optimized P2A-linked lentiviral expression construct in the pRRLSIN.cPPT.MSCV/GFP.wPRE vector after excising the GFP gene (*55*).

5. **Lysis of A2$^+$ K562 cells by Primary T cells expressing WT1-specific TCRs**. Target HLA-A2$^+$ K562 cells or control (HLA-A2$^-$) K562 cells were labeled with ^{51}Cr (Perkin Elmer) overnight, washed and incubated in triplicate with effector T cells expressing either TCR_{C4} or $TCR_{WT1\text{-}37\text{-}45}$ #1 at various effector to target (E:T) ratios. Supernatants were harvested for γ-counting after a 4-hour incubation and percent specific lysis was calculated.

Immunoblot of proteasome catalytic subunits from patient sample or HEK-293 cell lines expressing specific proteasome isoforms. For the relapsed patient sample, PBMCs were depleted of CD3$^+$ cells using EasySep CD3$^+$ selection kit II (StemCell) following manufacturer's instructions. CD3$^+$ cells were discarded leaving an AML- enriched blast population. Blasts or cell lines were then counted and lysed at 40M cells/mL for 15 minutes on ice using Pierce IP lysis buffer (Thermo Fisher Scientific) with 1X proteasome inhibitor (Thermo Fisher Scientific). Lysates were centrifuged at 13,000G for 10 minutes to clear.

Assembled proteasome complexes were isolated using co-immunoprecipitation with and anti-α2 antibody, MCP21 (Enzo Lifesciences, BML-PW8105) and immunoblotted with different immunoprecipitation beads, as previously described (*83*) (*29*). Briefly, Dynabeads M-270 epoxy beads (Invitrogen) were functionalized with MCP21 at 35 μg antibody per mg Dynabeads following manufacturer's instructions. Cleared lysate was quantified using a BCA assay Kit (Thermo Fisher Scientific) and 700 ng protein was incubated with 1mg functionalized beads to capture assembled

proteasome complexes. Complexes were eluted in 0.1M Glycine buffer pH 2 to 2.5 and neutralized with equal volume Tris pH 7.5.

Proteasome composition was assessed by immunoblot. Elution fractions were separated on a denaturing 4-12% polyacrylamide gel and then transferred to a PVDF (Invitrogen) membrane, using the mini-blot module (Invitrogen). Membranes were blocked for 1 hour at room temperature in TBS containing 5% of dry milk and 0.1% Tween 20, and washed three times in TBS, 0.1% Tween and incubated with the relevant primary antibody. Standard proteasome (SP) and immunoproteasome beta subunits were probed overnight in a solution of TBS, 5% milk powder, and 0.1% Tween-20 for β1[patient AML: (1:500 dilution, Santa Cruz Bio, sc-374405), cell lines: (1:500 dilution, Enzo Life sciences, BLM-PW8140)], β2 (1:250 dilution, Enzo Lifesciences, BML- PW8145), β5 (1:1000 dilution, Enzo Lifesciences, BML- PW8895), β1i (1:500 dilution, Enzo Lifesciences, BML- PW8840), β2i(1:250 dilution Novus, NBP1-8660-), β5i (1:1000 dilution, Enzo Lifesciences, BML- PW8355) and the shared beta-subunit β7 (1:1000 dilution, Enzo Lifesciences, BML- PW8135) was used as a loading control.

Transduction of HEK-293EBNA cells expressing the IP or SP with WT1 and HLA-A*02:01. HEK-239EBNA cell lines expressing either the IP or SP (*83*) were serially transduced with lentiviral constructs containing HLA-A*02:01 then WT1 (pCVL-A) (*84*). Cells were cultured on CELLCOAT Type I Collagen plates (Greiner bio-one) in IMDM medium supplemented with L-Glutamine, 1%, non-essential ammino acids (NEAA), 1X, and 10% FBS. Hygromycin and puromycin were added to select for IP expressing cells. Cells were first transduced with HLA-A*02:01; positive cells (APC) were isolated using the Sony MA900. HLA-A2$^+$ cells were then transduced a second time with full-length WT1 (variant D) in the pCVL-A (*84*) co expressing eGFP. Cells were again stained for HLA-A2 and double positive HLA-A2$^+$GFP$^+$ cells were selected using the Sony MA900 WT-1 expression was further confirmed by IHC

Cytokine analysis of eCD8$^+$ and dpCD4$^+$ $T_{TCR37\text{-}45}$ cells. PBMCs were thawed from a healthy donor. CD8$^+$ and CD4$^+$ T cells were selected using CD8 and CD4 positive selection kits (Stem Cell Technologies). CD8$^+$ and CD4$^+$ T cells were transduced with a

polycistronic construct containing $TCR_{37\text{-}45}$ and CD8αβ to yield $eCD8^+$ and $dpCD4^+$ T cells respectively. 5.0 $x10^5$ cells of either $eCD8^+$, $dpCD4^+$, or a combination $eCD8^+/dpCD4^+$(1:1 ratio; 2.5 $x10^5$ cells each) were co-cultured with T2 cells exogenously loaded with 10μg $WT1_{37\text{-}45}$ peptide at an E:T ratio of 5:1 for 4, 24, or 48 hours before supernatant was collected for analysis. Cytokine concentrations were determined using the R&D systems Luminex (*85*) platform according to manufacturer's instructions. Based on the inability of $CD4^+$ T cells to control tumor without co expression of the CD8αβ co-receptor (**Figure 2.22 and Figure 2.23**), suggesting a lack of TCR activation through TCR signaling, we did not assess the cytokine profile of $CD4^+$ T cells alone. Similarly, $CD8^+$ cells expressing only $TCR_{37\text{-}45}$ showed identical tumor control as $eCD8^+$ T cells and were not assessed separately.

Cytotoxicity assays of HEK-293EBA IP and SP cell lines. Cytotoxic activity of mock-transduced, $T_{TCR\text{-}C4}$ -transduced and $TCR_{37\text{-}45}$ -transduced $CD8^+$ T were examined by assessing the capacity of the T cells to lyse WT1- and HLA-A2-transduced HEK-293-EBNA cell lines expressing either the IP or SP in a 4 hour ^{51}Cr release assay at the indicated effector to target (E:T) ratios as previously described (*86*).

AML cell death assays. The endogenous TCR can non-specifically react with the patient sample (only HLA-A2 is matched) necessitating an additional CRISPR step to view specific TCR signaling. Because TCR_{C4}, $TCR_{37\text{-}45}$, and $TCR_{Irr.}$ are codon optimized, they are not susceptible to the crisprRNA (crRNA) used. To measure T cell-induced AML lysis, CRISPR edited T cells were transduced with $TCR_{37\text{-}45}$, TCR_{C4}, and TCR_{Irr} and stained with carboxyfluorescein succinimidyl ester and cultured with primary AML cells at the indicated E:T ratios in triplicates for 18 hours as previously described (*87*). Briefly, after the culture period, surviving AML cells were identified by staining for CD45 (1:600 dilution, PE Texas Red, clone 2D1, BioLegend), CD34 (1:50 dilution, BV421, clone 561, BioLegend) and CD38 (1:160 dilution BUV395, clone HB7, Beckton Dickinson) for 30 minutes and quantified with Flow Count beads (Molecular Probes). Percent killing was calculated by the formula: (absolute number of AML targets in co-culture/absolute number of AML targets in target only control well) x 100.

To measure T cell induced AML apoptosis, T cells with endogenous TCR a and b chains knocked out with CRISPR were transduced with $TCR_{37\text{-}45}$, TCR_{C4}, or TCR_{irr} and cultured with PBMC from patient 1's relapsed AML at a 3:1 E:T ratio (3.0 x 10^6 T cells, 1.0 x 10^6 AML) for 1 hour. The cells were harvested and stained with a viability dye (TONBO Biosciences #13-0870-T100) and cell surface markers for 30 minutes, CD34 (1:50 dilution, BV421, BioLegend, 343610), CD8 (1:100 dilution, APC-Cy7, BD Biosciences, 557760) and CD45 (1:300 dilution, PE, BioLegend, 368510). The cells were fixed with 4% Formaldehyde (CellSignal, 47746) and permeabilized with 100% Methanol (CellSignal, 13604) according to manufacturer's instructions, followed by intracellular stain for 1 hour for cleaved caspase-3 (1:50 dilution, FITC, CellSignal, 96695) and flow cytometry analysis. Percent cleaved-caspase 3 expressing $CD8^-CD34^+$ cells were quantified.

CRISPR-Cas9 Gene editing by electroporation. $CD8^+$ and $CD4^+$ T cells were selected from healthy PBMCs using the EasySep Human CD8 Positive Selection kit II and Human $CD4^+$ T Cell isolation kit (STEMCELL Technologies). T Cells were seeded at 10^6 T cells/mL with 50 IU/mL IL-2 and activated with T Cell TransAct, human (Miltenyi Biotec) according to manufacturer's instructions; TransAct was removed after 48 hours. Cells were then electroporated with Cas9 complexed with CRISPR-RNA (crRNA) targeting the TCRα and TCRβ to knock out the endogenous TCR, sequences for TRAC Exon 1 and TRBC Exon 1 were 5'-AGAGTCTCTCAGCTGGTACA-3' and 5'-GGAGAATGACGAGTGGACCC -3' respectively. To generate gRNAs, crRNAs and tracrRNAs were thawed, mixed at 1:1 v/v ratio (160mM stock) and incubated at 37°C for 30 minutes. To create the ribonucleoprotein (RNP) mixture, Cas9 protein (MacroLab, Berkeley, 61µM stock) and poly-L-glutamic acid (Sigma-Aldrich, 100 mg/mL in H_2O stock) were both added to the gRNA at a 1:1 v/v ratio (Poly-L-glutamic Acid:Cas9 protein). Poly-L-glutamic acid increases stability of the Cas9 (*88*). T cells were resuspended in Lonza P3 Buffer (1x10^6 per 20µL) and 4µL of the RNP mixture was added to each sample. The Cell/RNP mixture was transferred to a 16 well nucleocuvette strip (Lonza) and electroporated using the pulse code EH115 (Lonza, Nucleofector Device). Following electroporation, 80µL of prewarmed medium was

added and the samples recovered for 15 minutes at 37°C. Cells were then transferred to 1mL of prewarmed media and allowed to rest for 1 hour at 37°C before transduction with lentivirus encoding TCR_{C4}, $TCR_{37\text{-}45}$ or $TCR_{Irr.}$. An aliquot of untransduced cells was reserved to assess TCRα/TCRβ knockout efficiency by flow cytometry 3 days later, efficiency was determined as percent $CD3^-$. Cells were allowed to grow for 7 days after transduction, then FACS sorted for $CD8^+$tetramer$^+$ (BV421, APC), then placed into REP for 7 to 9 days.

Live tumor-cell killing assays. T cell killing capacity was monitored using the IncuCyte S3 Live-Cell Analysis System (Essen BioScience). Two human solid tumor lines - pancreatic cancer cell line, Panc-1 (ATCC, CRL-1469), and breast cancer cell line, MDA-MB-468 (ATCC, HTB-26) tumors were transduced with the IncuCyte NucLight Red Lentivirus Reagent (Sartorius) and seeded (10,000 cells/well) in 48-well plates (Corning). When comparing WT1-specific TCRs, wells were cocultured in triplicate at an E:T ratio of 10:1 with either no T cells, or with TCR-transduced $CD8^+$ T cells. When investigating if $CD4^+$ T cells could help sustain an anti-tumor response, wells were cocultured in triplicate at an E:T ratio of 4:1 with no T cells, $CD8^+$ T cells, dp$CD4^+$ T cells or a $CD8^+$/dp$CD4^+$ T cell mixture. $CD8^+$ and $CD4^+$ T cells were transduced with a lentiviral construct containing CD8$\alpha\beta$ along with $TCR_{37\text{-}45}$. For all assays, red fluorescence was recorded at 2-hour intervals for 96 hours or 156 hours. Tumor rechallenge (10,000 cells/well) was conducted at 48 hours and 96 hours. Images were analyzed using the IncuCyte S3 v. 2019b software (Essen BioScience) with total red object area (μm^2/well) corresponding to the degree of tumor cell death over time.

In vivo $T_{TCR37\text{-}45}$ therapy model. Experiments were performed following the approved IACUC protocol 50898. In accordance with best practices, the ARRIVE checklist was followed for in vivo studies. The HLA-$A2^+$ $WT1^+$ pancreatic cancer cell line Panc-1 (ATCC, CRL-1469), was transduced to express a firefly luciferase (PANC-1_{Luc}); individual clones were assessed for high luciferase expression. Clone#9 was selected for a high degree of transgene expression and used in subsequent in vivo studies. Female NSG (NOD-*scid* IL2Rgamma^{null}, Jackson Laboratory 005557) mice (*89*) of

similar age (six to eight weeks) were injected intraperitoneally (I.P.) with 1.25 x 10^5 PANC-1_{Luc} cells. As a 'prevention' model of adoptive T cell therapy, 3 days later mice either received no T cells at all (tumor only), or therapeutic $T_{TCR37\text{-}45}$ T cells consisting of 2x10^6 each eCD8$^+$ and dpCD4$^+$. For the 'treatment' model of adoptive T cell therapy, tumor burden was measured 14 days post tumor injection and mice with no detectable tumor were excluded from the study. Remaining mice with detectable tumor were enrolled in the study and randomly assigned to treatment groups. a one-way ANOVA with Tukey's multiple comparison test was used to confirm there was no significant difference in tumor burden between the groups ($p \geq 0.64$). Mice either received no T cells at all (tumor only), or therapeutic T cells ($T_{TCR37\text{-}45}$, T_{TCRC4} ,or $T_{TCRirr.}$) consisting of 10x10^6 each eCD8$^+$ and dpCD4$^+$. Mice in both studies were imaged and weighed weekly. The mice were administered 1µL of Luciferin per gram of body weight at 15 mg/mL of concentration I.P. under isoflurane anesthesia. After 15 minutes the mice were imaged while still under isoflurane anesthesia. At week 4, $T_{TCR37\text{-}45}$ treated mice received a second injection of $T_{TC37\text{-}45}$ T cells (2.0x10^6 each CD4$^+$ and CD8$^+$). Mice were euthanized on week 8.

References

1. A. M. Dickinson *et al.*, Graft-versus-Leukemia Effect Following Hematopoietic Stem Cell Transplantation for Leukemia. *Front Immunol* **8**, 496 (2017).
2. F. Lim, S. O. Ang, Emerging CAR landscape for cancer immunotherapy. *Biochem Pharmacol* **178**, 114051 (2020).
3. F. Perna *et al.*, Integrating Proteomics and Transcriptomics for Systematic Combinatorial Chimeric Antigen Receptor Therapy of AML. *Cancer Cell* **32**, 506-519 e505 (2017).
4. J. D. Stone, D. M. Kranz, Role of T cell receptor affinity in the efficacy and specificity of adoptive T cell therapies. *Front Immunol* **4**, 244 (2013).
5. H. Sugiyama, WT1 (Wilms' tumor gene 1): biology and cancer immunotherapy. *Jpn J Clin Oncol* **40**, 377-387 (2010).
6. R. Narayan *et al.*, Acute myeloid leukemia immunopeptidome reveals HLA presentation of mutated nucleophosmin. *PLoS One* **14**, e0219547 (2019).
7. C. Berlin *et al.*, Mapping the HLA ligandome landscape of acute myeloid leukemia: a targeted approach toward peptide-based immunotherapy. *Leukemia* **29**, 647-659 (2015).
8. S. Ochsenreither *et al.*, "Wilms Tumor Protein 1" (WT1) peptide vaccination-induced complete remission in a patient with acute myeloid leukemia is

accompanied by the emergence of a predominant T-cell clone both in blood and bone marrow. *J Immunother* **34**, 85-91 (2011).
9. A. G. Chapuis *et al.*, Transferred WT1-Reactive CD8+ T Cells Can Mediate Antileukemic Activity and Persist in Post-Transplant Patients. *Science translational medicine* **5**, 174ra127 (2013).
10. A. G. Chapuis *et al.*, T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. *Nat Med* **25**, 1064-1072 (2019).
11. P. Sharma, S. Hu-Lieskovan, J. A. Wargo, A. Ribas, Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell* **168**, 707-723 (2017).
12. M. Ruella, M. V. Maus, Catch me if you can: Leukemia Escape after CD19-Directed T Cell Immunotherapies. *Comput Struct Biotechnol J* **14**, 357-362 (2016).
13. M. Waterhouse *et al.*, Genome-wide profiling in AML patients relapsing after allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant* **17**, 1450-1459 e1451 (2011).
14. A. Busse *et al.*, Mutation or loss of Wilms' tumor gene 1 (WT1) are not major reasons for immune escape in patients with AML receiving WT1 peptide vaccination. *J Transl Med* **8**, 5 (2010).
15. N. Vigneron, J. Abi Habib, B. J. Van den Eynde, Learning from the Proteasome How To Fine-Tune Cancer Immunotherapy. *Trends Cancer* **3**, 726-741 (2017).
16. A. Jaigirdar, S. A. Rosenberg, M. Parkhurst, A High-avidity WT1-reactive T-Cell Receptor Mediates Recognition of Peptide and Processed Antigen but not Naturally Occurring WT1-positive Tumor Cells. *J Immunother* **39**, 105-116 (2016).
17. D. J. Kuhn, R. Z. Orlowski, The immunoproteasome as a target in hematologic malignancies. *Semin Hematol* **49**, 258-262 (2012).
18. N. M. Moukalled, M. A. Kharfan-Dabaja, What is the role of a second allogeneic hematopoietic cell transplant in relapsed acute myeloid leukemia? *Bone Marrow Transplant* **55**, 325-331 (2020).
19. N. P. Restifo, L. Gattinoni, Lineage relationship of effector and memory T cells. *Curr Opin Immunol* **25**, 556-563 (2013).
20. E. Papalexi, R. Satija, Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* **18**, 35-45 (2018).
21. J. L. Gommerman, J. L. Browning, C. F. Ware, The Lymphotoxin Network: orchestrating a type I interferon response to optimize adaptive immunity. *Cytokine Growth Factor Rev* **25**, 139-145 (2014).
22. Y. Guo, J. Chen, T. Zhao, Z. Fan, Granzyme K degrades the redox/DNA repair enzyme Ape1 to trigger oxidative stress of target cells leading to cytotoxicity. *Mol Immunol* **45**, 2225-2235 (2008).
23. B. M. Peterlin, Transcriptional regulation of HLA-DRA gene. *Res Immunol* **142**, 393-399 (1991).
24. B. Schroder, The multifaceted roles of the invariant chain CD74--More than just a chaperone. *Biochim Biophys Acta* **1863**, 1269-1281 (2016).
25. T. C. Greenough *et al.*, A Gene Expression Signature That Correlates with CD8+ T Cell Expansion in Acute EBV Infection. *J Immunol* **195**, 4185-4197 (2015).

26. C. Y. Chen, L. W. Forman, D. V. Faller, Calcium-dependent immediate-early gene induction in lymphocytes is negatively regulated by p21Ha-ras. *Mol Cell Biol* **16**, 6582-6592 (1996).
27. M. K. Slifka, F. Rodriguez, J. L. Whitton, Rapid on/off cycling of cytokine production by virus-specific CD8+ T cells. *Nature* **401**, 76-79 (1999).
28. M. Groettrup, S. Standera, R. Stohwasser, P. M. Kloetzel, The subunits MECL-1 and LMP2 are mutually required for incorporation into the 20S proteasome. *Proc Natl Acad Sci U S A* **94**, 8970-8975 (1997).
29. J. Abi Habib *et al.*, Efficiency of the four proteasome subtypes to degrade ubiquitinated or oxidized proteins. *Sci Rep* **10**, 15765 (2020).
30. C. E. Ruggiero E., Prodeus A., Magnani Z.I., Camisa B., Merelli I.,, S. L. Politano C., Potenza A., Cianciotti B.C., Manfredi F., Di Bono M., Vago L.,, M. S. Tassara M., Ponzoni M., Sanvito F., Liu D., Balwani I., Galli R., Genua, O. R. M., O'Connell D., Dutta I., Yazinski S.A., McKee M., Arredouani M.,, C. F. Schultes B., Bonini C., CRISPR-based gene disruption and integration of high-avidity, WT1-specific T cell receptors improve anti-tumor T cell function. *Science of Translational Medicine*, (2021).
31. W. Glienke, L. Maute, J. Wicht, L. Bergmann, Wilms' tumour gene 1 (WT1) as a target in curcumin treatment of pancreatic cancer cells. *Eur J Cancer* **45**, 874-880 (2009).
32. D. M. Loeb *et al.*, Wilms' tumor suppressor gene (WT1) is expressed in primary breast tumors despite tumor-specific promoter methylation. *Cancer Res* **61**, 921-925 (2001).
33. C. J. Turtle *et al.*, CD19 CAR-T cells of defined CD4+:CD8+ composition in adult B cell ALL patients. *J Clin Invest* **126**, 2123-2138 (2016).
34. R. E. Tay, E. K. Richardson, H. C. Toh, Revisiting the role of CD4(+) T cells in cancer immunotherapy-new insights into old paradigms. *Cancer Gene Ther*, (2020).
35. J. Thibodeau, M. C. Bourgeois-Daigneault, R. Lapointe, Targeting the MHC Class II antigen presentation pathway in cancer immunotherapy. *Oncoimmunology* **1**, 908-916 (2012).
36. M. Ecsedi, M. S. McAfee, A. G. Chapuis, The Anticancer Potential of T Cell Receptor-Engineered T Cells. *Trends Cancer* **7**, 48-56 (2021).
37. I. Raphael, S. Nalawade, T. N. Eagar, T. G. Forsthuber, T cell subsets and their signature cytokines in autoimmune and inflammatory diseases. *Cytokine* **74**, 5-17 (2015).
38. R. M. Spaapen, J. Neefjes, Immuno-waste exposure and further management. *Nat Immunol* **13**, 109-111 (2012).
39. M. Bai *et al.*, Assembly mechanisms of specialized core particles of the proteasome. *Biomolecules* **4**, 662-677 (2014).
40. S. Khan *et al.*, Immunoproteasomes largely replace constitutive proteasomes during an antiviral and antibacterial immune response in the liver. *J Immunol* **167**, 6859-6868 (2001).
41. A. Rouette *et al.*, Expression of immunoproteasome genes is regulated by cell-intrinsic and -extrinsic factors in human cancers. *Sci Rep* **6**, 34019 (2016).

42. P. G. Coulie, B. J. Van den Eynde, P. van der Bruggen, T. Boon, Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer* **14**, 135-146 (2014).
43. S. C. Tripathi *et al.*, Immunoproteasome deficiency is a feature of non-small cell lung cancer with a mesenchymal phenotype and is associated with a poor outcome. *Proc Natl Acad Sci U S A* **113**, E1555-1564 (2016).
44. E. S. Schultz *et al.*, The production of a new MAGE-3 peptide presented to cytolytic T lymphocytes by HLA-B40 requires the immunoproteasome. *J Exp Med* **195**, 391-399 (2002).
45. M. Keller *et al.*, The proteasome immunosubunits, PA28 and ER-aminopeptidase 1 protect melanoma cells from efficient MART-126-35 -specific T-cell recognition. *Eur J Immunol* **45**, 3257-3268 (2015).
46. A. Textor *et al.*, Preventing tumor escape by targeting a post-proteasomal trimming independent epitope. *J Exp Med* **213**, 2333-2348 (2016).
47. S. Krishna, K. S. Anderson, T-Cell Epitope Discovery for Therapeutic Cancer Vaccines. *Methods Mol Biol* **1403**, 779-796 (2016).
48. M. M. Horowitz, High-resolution typing for unrelated donor transplantation: how far do we go? *Best Pract Res Clin Haematol* **22**, 537-541 (2009).
49. H. Dohner *et al.*, Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* **129**, 424-447 (2017).
50. D. Araki *et al.*, Allogeneic Hematopoietic Cell Transplantation for Acute Myeloid Leukemia: Time to Move Toward a Minimal Residual Disease-Based Definition of Complete Remission? *J Clin Oncol* **34**, 329-336 (2016).
51. P. Armand *et al.*, Validation and refinement of the Disease Risk Index for allogeneic stem cell transplantation. *Blood* **123**, 3664-3671 (2014).
52. Y. Zhou, B. L. Wood, Methods of Detection of Measurable Residual Disease in AML. *Curr Hematol Malig Rep* **12**, 557-567 (2017).
53. M. Festuccia *et al.*, Minimal Identifiable Disease and the Role of Conditioning Intensity in Hematopoietic Cell Transplantation for Myelodysplastic Syndrome and Acute Myelogenous Leukemia Evolving from Myelodysplastic Syndrome. *Biol Blood Marrow Transplant* **22**, 1227-1233 (2016).
54. M. Fang *et al.*, Prognostic impact of discordant results from cytogenetics and flow cytometry in patients with acute myeloid leukemia undergoing hematopoietic cell transplantation. *Cancer* **118**, 2411-2419 (2012).
55. S. Jones *et al.*, Lentiviral vector design for optimal T cell receptor gene expression in the transduction of peripheral blood lymphocytes and tumor-infiltrating lymphocytes. *Hum Gene Ther* **20**, 630-640 (2009).
56. A. G. Chapuis *et al.*, Transferred melanoma-specific CD8+ T cells persist, mediate tumor regression, and acquire central memory phenotype. *Proc Natl Acad Sci U S A*, (2012).
57. A. G. Chapuis *et al.*, Tracking the Fate and Origin of Clinically Relevant Adoptively Transferred CD8+ T Cells In Vivo. *Sci Immunol* **2**, (2017).
58. J. Zhou, M. E. Dudley, S. A. Rosenberg, P. F. Robbins, 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* **28**, 53-62 (2005).

59. P. C. Tumeh *et al.*, PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568-571 (2014).
60. H. Horton *et al.*, Optimization and validation of an 8-color intracellular cytokine staining (ICS) assay to quantify antigen-specific T cells induced by vaccination. *J Immunol Methods* **323**, 39-54 (2007).
61. G. Moncunill, H. Han, C. Dobano, M. J. McElrath, S. C. De Rosa, OMIP-024: pan-leukocyte immunophenotypic characterization of PBMC subsets in human samples. *Cytometry A* **85**, 995-998 (2014).
62. A. P. Limaye, M. L. Huang, E. E. Atienza, J. M. Ferrenberg, L. Corey, Detection of Epstein-Barr virus DNA in sera from transplant recipients with lymphoproliferative disorders. *J Clin Microbiol* **37**, 1113-1116 (1999).
63. M. Boeckh *et al.*, Optimization of quantitative detection of cytomegalovirus DNA in plasma by real-time PCR. *J Clin Microbiol* **42**, 1142-1148 (2004).
64. G. X. Zheng *et al.*, Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **8**, 14049 (2017).
65. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
66. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420 (2018).
67. T. Stuart *et al.*, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).
68. T. Ilicic *et al.*, Classification of low quality cells from single-cell RNA-seq data. *Genome Biol* **17**, 29 (2016).
69. G. Finak *et al.*, MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* **16**, 278 (2015).
70. C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, 296 (2019).
71. M. H. Veldman-Jones *et al.*, Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. *Cancer Res* **75**, 2587-2593 (2015).
72. C. C. Pritchard *et al.*, Validation and implementation of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens. *J Mol Diagn* **16**, 56-67 (2014).
73. J. E. Park *et al.*, PSMB9 codon 60 polymorphisms have no impact on the activity of the immunoproteasome catalytic subunit B1i expressed in multiple types of solid cancer. *PLoS One* **8**, e73732 (2013).
74. W. Y. Ho, H. N. Nguyen, M. Wolfl, J. Kuball, P. D. Greenberg, In vitro methods for generating CD8+ T-cell clones for immunotherapy from the naive repertoire. *J Immunol Methods* **310**, 40-52 (2006).
75. B. Guillaume *et al.*, Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proc. Natl. Acad. Sci. USA* **107**, 18599-18604 (2010).

76. N. Vigneron *et al.*, A peptide derived from melanocytic protein gp100 and presented by HLA-B35 is recognized by autologous cytolytic T lymphocytes on melanoma cells. *Tissue Antigens* **65**, 156-162 (2005).
77. E. L. Pogosova-Agadjanyan *et al.*, Impact of Specimen Heterogeneity on Biomarkers in Repository Samples from Patients with Acute Myeloid Leukemia: A SWOG Report. *Biopreserv Biobank* **16**, 42-52 (2018).
78. L. Wang, S. Wang, W. Li, RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184-2185 (2012).
79. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
80. A. Frankish *et al.*, GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* **47**, D766-D773 (2019).
81. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
82. M. D. Robinson, A. Oshlack, A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25 (2010).
83. B. Guillaume *et al.*, Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proc Natl Acad Sci U S A* **107**, 18599-18604 (2010).
84. A. D. Bandaranayake *et al.*, Daedalus: a robust, turnkey platform for rapid production of decigram quantities of active recombinant proteins in human cell lines using novel lentiviral vectors. *Nucleic Acids Res* **39**, e143 (2011).
85. L. Findlay *et al.*, Improved in vitro methods to predict the in vivo toxicity in man of therapeutic monoclonal antibodies including TGN1412. *J Immunol Methods* **352**, 1-12 (2010).
86. S. R. Riddell *et al.*, Restoration of viral immunity in immunodeficient humans by the adoptive transfer of T cell clones. *Science* **257**, 238-241 (1992).
87. I. Jedema, N. M. van der Werff, R. M. Barge, R. Willemze, J. H. Falkenburg, New CFSE-based assay to determine susceptibility to lysis by cytotoxic T cells of leukemic precursor cells within a heterogeneous target cell population. *Blood* **103**, 2677-2682 (2004).
88. D. N. Nguyen *et al.*, Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat Biotechnol* **38**, 44-49 (2020).
89. X. Cao *et al.*, Defective lymphoid development in mice lacking expression of the common cytokine receptor gamma chain. *Immunity* **2**, 223-238 (1995).

Chapter 3- Artemis

Abstract

Cytotoxic $CD8^+$ T cells recognize and eliminate malignant cells through recognition of antigen-derived peptides presented on class I HLA and proteasomal degradation of proteins is considered the primary source presented peptides. Cells can express standard proteaseome and/or immunoproteasome isoforms which generate 'distinct' (isoform-dependent) or 'mutual' (isoform-independent) peptides. Transgenic T cell receptor (TCR) T cell (TCR-T) immunotherapy utilizes $CD8^+$ T cells recognizing a specific HLA-presented peptide to eliminate diseased cells presenting that peptide. However, cancer cells can evade TCR-T cell immunotherapy by modulating proteasome composition to prevent production of the 'distinct' peptide and thereby avoid recognition by TCR-T cells.

TCR-T cell therapy could be enhanced if tailored to target 'mutual' peptides. We utilized an emerging peptide discovery platform, Artemis, which employs soluble HLAs to bypass immunoprecipitation of endogenous HLAs. Artemis performed in proteasome-specific cell lines (expressing either the standard proteasome or immunoproteasome) revealed the proportions of 'mutual' and 'distinct' peptides were specific to each of the three HLA alleles tested with HLA-A*02:01 having the fewest 'mutual' peptides and HLA-A*11:01 having the most. Furthermore, we recovered 37 peptides from three antigens associated with acute myeloid leukemia, 19 (51%) were 'mutual' peptides. Our findings reveal that HLA isoforms can influence the ratio of 'mutual' versus 'distinct' peptides. Defining proteasomal processing of tumor-associated antigens could enhance peptide selection for optimal targeting with TCR-T cell therapy and Artemis is an effective platform to identify 'mutual' peptides.

Introduction

Cytotoxic ($CD8^+$) lymphocytes are known to eliminate abnormal cells including those that are malignant or virally-infected. Specificity of cell-mediated cell killing is primarily achieved through recognition of antigenic peptides (*1, 2*). Peptides 8-14 amino

acids in length, bind class I HLA in the endoplasmic reticulum and are trafficked to the surface where the HLA-peptide complex can engage a T-cell receptor (TCR) on $CD8^+$ lymphocytes (*2, 3*). This TCR:HLA-peptide interaction is highly specific and integral to disease clearance making it the focus of several vaccination strategies and T cell-based immunotherapies, including transgenic TCR T-cells (TCR-T) that express a high-affinity TCR recognizing a defined HLA-peptide complex found on the malignant cell surface. While TCR-T cell immunotherapies have demonstrated clinical potential (*4-7*), identifying reliable HLA-peptide targets for therapeutic TCR-Ts remains cumbersome.

To identify whether the selected HLA presents a given peptide, many groups, including ours (*8*), first utilize *in silico* predictions to determine if a theoretical peptide binds the HLA (e.g. netMHCpan (*9*), MHCflurry (*10*)). To ultimately validate whether the peptide is naturally presented first requires generating T cell lines that recognize the target HLA-peptide and then culturing those T cells with tumor lines expressing antigen. Briefly, peptides are synthesized and exogenously loaded onto presenting cells. T cells from HLA-matched donors are then co-cultured with the peptide-loaded presenting cells and reactive T cells proliferate. After a few rounds of stimulation, T cell lines reactive to the HLA-peptide complex are generated. Reactivity of these T cells lines toward HLA-matched tumor lines expressing the parent protein are assessed to confirm natural peptide presentation. This process relies on (1) successful T cell generation and (2) indirect *in vitro* recognition (e.g. ELIspot, chromium release, Incucyte assays). Another option is to directly probe the peptides presented on the target cell surface. This relies on immunoprecipitation of membrane-bound HLAs followed by mass spectrometry and computational deconvolution to match eluted peptides to the presenting HLA (*11*). This process is labor intensive and technically challenging. Recent advances include using engineered cells expressing a single HLA to avoid computational deconvolution (*12*) and soluble HLAs to circumnavigate immunoprecipitation (*13*). While *in silico* peptide binding predictions now include algorithms trained on data sets with peptides eluted directly off cells using immunoprecipitation, they do not currently account for potential abnormalities in upstream processing of the peptide, which have been associated with several cancers (*14, 15*). Larger, more varied, and comprehensive data sets that

include processing and presentation information are necessary to improve practical application of these predictive tools for TCR-T peptide discovery.

In this work we employed an innovative peptide discovery platform, Artemis (*16*), that utilizes a soluble HLA called a single chain dimer (SCD). The SCD has the transmembrane portion of the HLA replaced with a 6x-histidine tag to facilitate easy capture. The beta-2-microglobulin subunit is physically linked to the alpha chain with a glycine-serine linker to fuse the HLA dimer into a single amino acid sequence (**Figure 3.1A**). Because the SCD lacks a transmembrane domain, the peptide-SCD complex is secreted into the media, where it is easily collected and captured using standard immobilized metal-ion affinity chromatography (IMAC). The captured SCD is denatured to release the peptide while the SCD remains bound to the IMAC column. Mass spectroscopy is then used to sequence eluted peptides that reflect naturally presented peptides (**Figure 3.1B**). Peptides recovered from Artemis have nearly identical characteristics to those identified with classical immunoprecipitation as described by Finton et al. (*16*) Peptides identified with Artemis share similar length distributions and HLA binding motifs to those recovered using traditional immunoprecipitation.

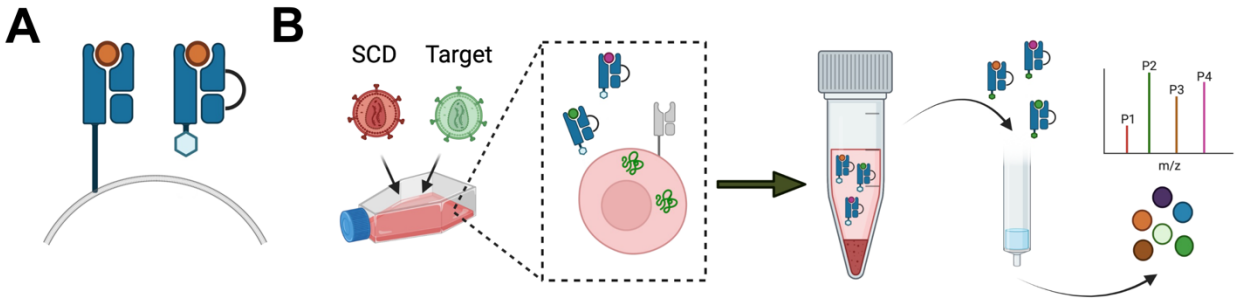


Figure 3.1 Artemis discovery pipeline. (A) Schematic of the single chain dimer (SCD). **(B)** Schematic of the ARTEMIS process. Briefly, cells are transduced with lentivirus expressing a selected SCD (mCherry reporter) and optional antigen of interest (GFP reporter). Cells are then sorted on mCHerry and GFP co-expression and double cells are expanded. Supernatant from expanded cells containing secreted SCDs is collected; SCDs are then captured using immobilized metal-ion affinity chromatography and denatured to release the presented peptides. Peptides are then sequenced on a mass spectrometer. Sequenced peptides are mapped back to the human proteome using ProteomeDiscover 2.0, which predicts the parent protein.

Furthermore, Artemis and traditional immunoprecipitation have comparable distributions of peptides predicted to be strong-binding, weak-binding, and non-binding as determined by netMHCpan4.1. Both methodologies recover peptides derived from proteins in all cell compartments at similar frequencies and recover similar peptide repertoires; 80% of peptides detected using immunoprecipitation are also observed with Artemis (*16*).

For $CD8^+$ T cells to clear malignant or infected cells, target cells must have intact upstream processing machinery to generate peptides for presentation. Proteasomal degradation is considered the primary source of class I HLA-displayed peptides (*17*) though other proteases, such as endoplasmic reticulum aminopeptidase (ERAP), can trim peptides after proteasomal degradation and before HLA presentation (*18*).

The proteasome core (20S) is composed of four heptamer rings stacked symmetrically– alpha-beta-beta-alpha – with the catalytic subunits located in the beta rings. The standard proteasome (SP) is expressed in somatic cells and many solid tumors; it contains enzymatic subunits β1, β2 and β5. The immunoproteasome (IP)is expressed in immune cells, hematopoietic cells, and many liquid tumors but can be induced in most somatic cells with IFN-γ exposure (*19*). Its catalytic core contains subunits β1i, β2i, and β5i. Thus, the SP and IP possess different catalytic subunits that are reported to generate ‘distinct’ (isoform-dependent) or ‘mutual’ (isoform independent) peptides (*20, 21*).The relative proportions of ‘mutual’ versus ‘distinct’ peptides are ill-defined and the mechanisms governing observed differences remain unclear. It has been hypothesized that IPs have unique cleavage preferences (*22*), but more recent studies suggest that IPs have an increased catalytic rate and generate more peptides, rather than divergent cleavage preferences (*23*).

Tumors can alter proteasome expression to evade immune recognition (*14, 15, 24*). ‘Mutual’ peptides offer potentially robust therapeutic T cell targets as they will be produced regardless of which proteasome isoform is expressed. Herein, we use the Artemis platform to investigate contributions of the SP and IP to the immunopeptidome and identify ‘mutual’ peptides from tumor associated antigens.

Results

HLA selection

Extensive allelic diversity in class I HLAs makes empirical investigation of the presented immunopeptidome across all HLAs impractical. Therefore, we selected a subset of representative HLAs for our investigation. Five HLA alleles offer ≥ 90% coverage for most ethnicities indicating a small subset of HLA alleles may be used to study most of the population (*25*). Three HLA-A alleles – HLA-A*02:01, HLA-A*24:02 and HLA-A*11:01 – cover most (71%) of the United States population (**Figure 3.2A**). Per records of the 2020 census, Asian (4.8%), Black (12.3%), Hispanic (16.4%) and White (63.8%) ethnicities account for 97.3% of the population (**Figure 3.2B**). Even within each ethnicity, these three HLA-A alleles cover >30% of each respective sub-population. (**Figure 3.2C**).

HLAs will show strong binding preferences for peptides containing select anchor residues, usually at the second and final residue of the binding peptide (*26*). Biochemical properties of the anchor residues are specific to each HLA (*26*). In addition to selecting alleles that cover most of the U.S. population, we chose alleles with diverse anchor residues to minimize allele-specific binding preferences obscuring differences from proteasomal cleavage. We selected HLAs from three HLA super types with final anchor residue preferences of varied chemical characteristics including small hydrophobic residues (HLA-A*02:01), large aromatic/hydrophobic residues (HLA-A*24:02) and positively charged residues (HLA-A*11:01). Diversity of anchor residue preferences at the second amino acid position are prominent in HLA-B alleles, i.e., HLA-B*07 prefers a proline and HLA-B*44 favors acidic residues (*26*). However, ERAP cleaves peptides from the n-terminus (*27*) and differences a position 2 may be more reflective of ERAP cleavage than proteasomal processing. HLA-B alleles were thus excluded and SCDs corresponding to HLA-A*2:01 (HLA-$A2_{SCD}$), HLA-A*24:02 (HLA-$A24_{SCD}$) and HLA-A*11:01 (HLA-$A11_{SCD}$) (*16*) were selected for analysis.

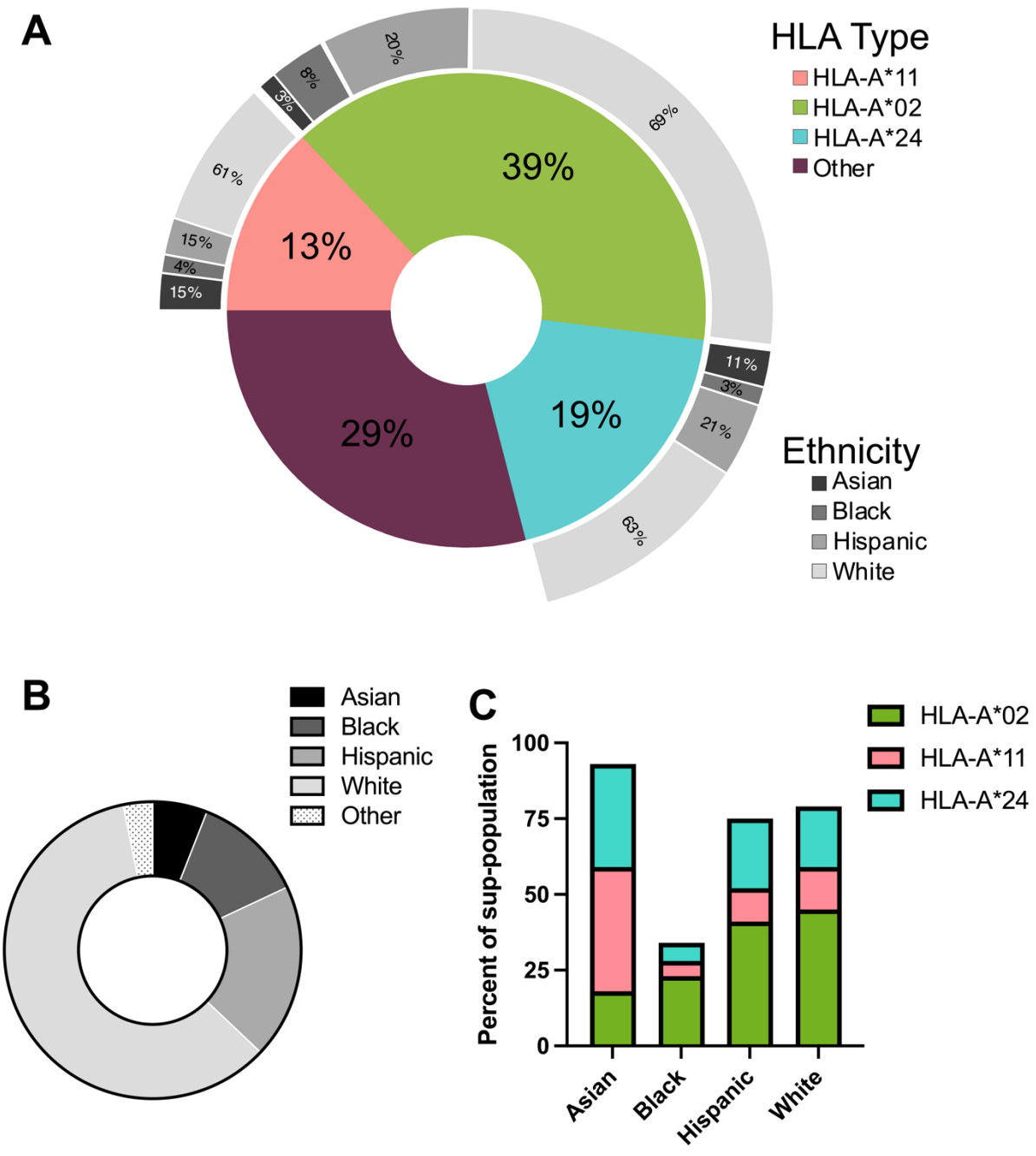


Figure 3.2 Distribution of select HLA alleles in the United States population. (A) Colored pie chart indicates distribution of the HLA alleles selected for these experiments in the US population. These three alleles achieve 71% population coverage. Grey bars outside each sector indicate the distribution of four major ethnicities within the given HLA type for the US population. **(B)** Distribution of the four major ethnicities in the US population. **(C)** The distribution of each selected HLA within each of the four major ethnicities. Greater than 30% of each ethnicity is covered with these three alleles.

Artemis immunopeptidome coverage

Antigen presentation is dynamic and inherently stochastic (*2*), making it difficult to capture the entire immunopeptidome. We sought to determine how many Artemis replicates were required to cover a substantial portion analyzing 4-6 biological replicates for each HLA_{SCD} allele and aggregating total unique peptides after each run. After approximately three replicates the rate of new peptide identification slows (*16*) (**Figure 3.3**). The total number of observed peptides varied with the HLA allele tested, with HLA-$A11_{SCD}$ displaying the most followed by HLA-$A24_{SCD}$ and lastly HLA-$A2_{SCD}$ (**Figure 3.3)**. We performed at least four biological replicates of each HLA_{SCD} proteasome composition to investigate SP/IP contributions to the immunopeptidome.

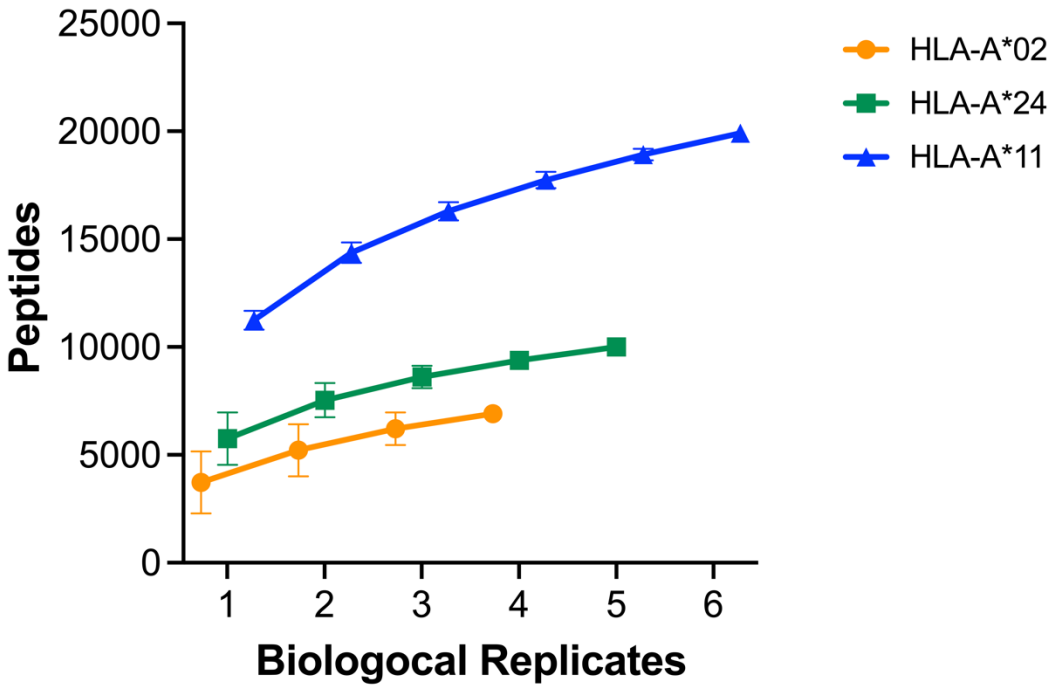


Figure 3.3 Artemis immunopeptidome coverage for three HLAs. Sum of unique peptides (y-axis) identified from 1-6 biological replicates (x-axis) for each HLA_{SCD} queried. Replicates shown were performed in unmodified HEK293E cells.

Artemis in proteasome-restricted cell lines

To investigate the immunopeptidome of each proteasome isoform, we required cell lines to exclusively express either the SP or the IP. HEK293E cells grow rapidly and densely in suspension increasing the amount of SCD that can be collected and makes them ideal for Artemis. Unmodified, HEK293E cells express the SP (**Figure 3.4A**) (*28*).

To express the IP, HEK293E cells were transduced to overexpress the IP β-subunits β1i, β2i and β5i as previously described (*28*). IP subunits have a higher affinity for assembly chaperones making IP formation faster and more efficient than SP, and the IP becomes the dominant isoform (*29, 30*). These lines were successfully used by our group in a previous publication (*14*) (**Figure 3.4A**).

When we transduced SP and IP cells with each HLA_{SCD}, we observed a slight decrease in growth kinetics (**Figure 3.4B**) and viability for the IP cells (61%) compared to SP cells (76%) (**Figure 3.4C**). Fewer peptides were detected from IP cells (**Figure 3.4D**), but no significant difference was observed in endogenous HLA expression ($p = 0.70$, **Figure 3.4E**) or transduction efficiency ($p = 0.56$, **Figure 3.4F**). IP cells have increased SCD reporter expression ($p<0.01$, **Figure 3.4G**) but there was no difference in HLA-$A24_{SCD}$, as measured by western blot (**Figure 3.4H**). The limited peptide yields from IP cells could stem from reduced viability and limited growth rather than a true biological difference in peptide recovery. IP lines were heavily modified with lentivirus to express five additional proteins including three additional IP beta-subunits, an HLA_{SCD} and an antigen of interest. High viral burdens are known to negatively impact viability (*31*).

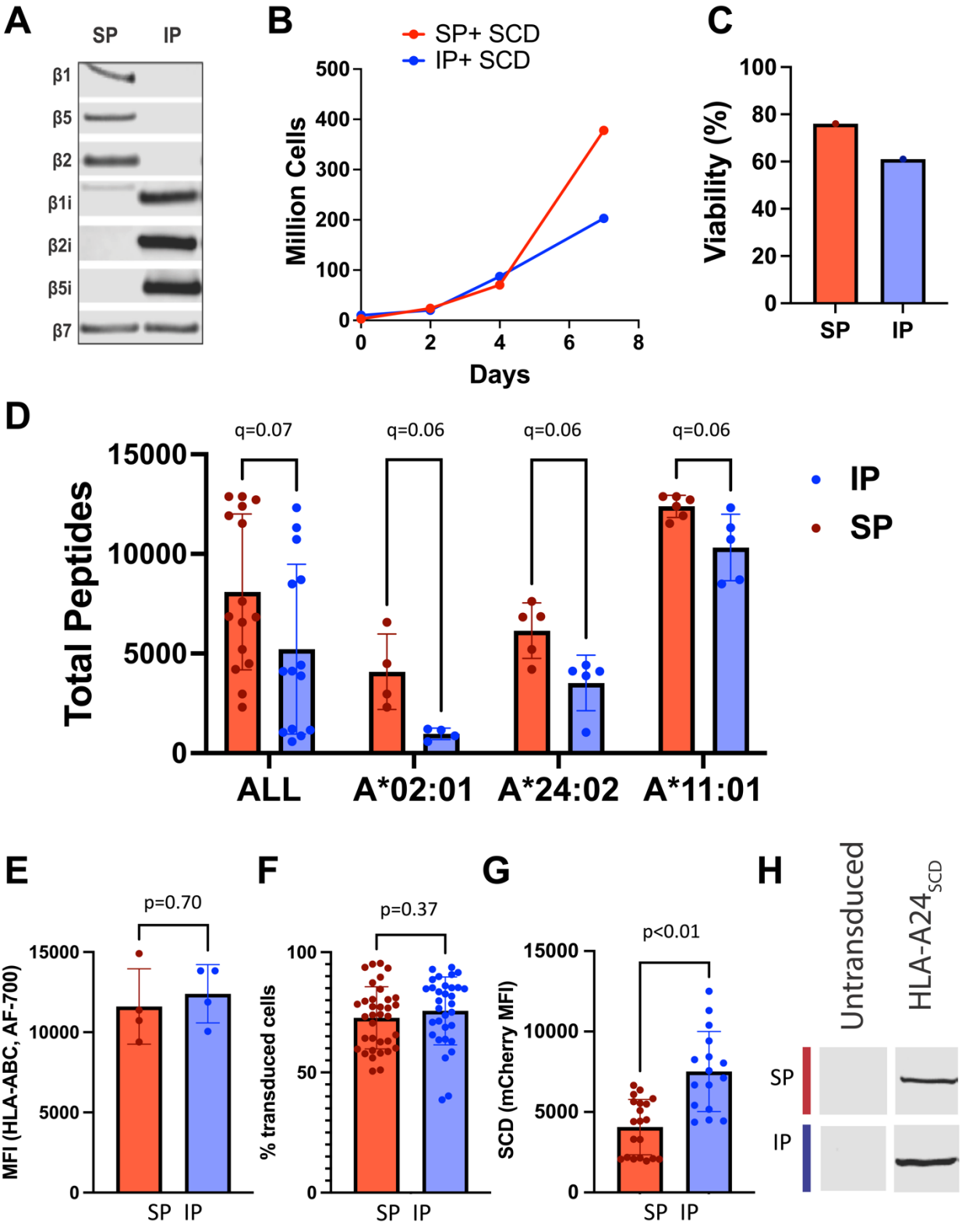

A
SP
IP
β1
β5
β2
β1i
β2i
β5i
β7
B
SP+ SCD
IP+ SCD
Million Cells
Days
C
Viability (%)
SP
IP
D
q=0.07
q=0.06
q=0.06
q=0.06
Total Peptides
IP
SP
ALL
A*02:01
A*24:02
A*11:01
E
p=0.70
MFI (HLA-ABC, AF-700)
F
p=0.37
% transduced cells
G
p<0.01
SCD (mCherry MFI)
H
Untransduced
HLA-A24SCD
SP
IP

Figure 3.4 Standard and immunoproteasome cells in Artemis. **(A)** Immunoblot of SP and IP expressing HEK293E cells. **(B)** Growth curves of SP and IP cells transduced with HLA-A24$_{SCD}$. **(C)** Viability of SP and IP cells after 4 days in culture without replenishing media. **(D)** Total peptides detected in each Artemis replicate. No significance in peptide yields was detected between SP or IP cells. Multiple unpaired student t-test with the Welch correction, q-values listed are adjusted p-values accounting for a 5% false discovery rate. **(E)** HLA expression on SP and IP cells as measured by flow cytometry. Replicates were stained independently. Unpaired student's t-test, n=4 **(F)** Transduction efficiency of SCDs as measured by flow cytometry using the fluorescent reporter (mCherry) in the SCD construct. Unpaired student's t-test, SP: n=35, IP: n=32 **(G)** SCD expression as measured by the fluorescent reporter (mCherry). Unpaired student's t-test, SP: n=20, IP: n=16 **(H)** immunoblot of SP and IP cells transduced with HLA-A24$_{SCD}$. MFI= median fluorescent intensity.

We used the Artemis pipeline to perform 4-6 biological replicates for each of the six experimental conditions: 1. SP + HLA-A2$_{SCD}$, n=4, 2. SP + HLA-A24$_{SCD}$, n=5, 3. SP + HLA-A11$_{SCD}$, n=6, 4. IP + HLA-A2$_{SCD}$, n=4, 5. IP + HLA-A24$_{SCD}$, n=5, and 6. IP + HLA-A11$_{SCD}$, n=5. Peptides from experimental replicates were aggregated and duplicates were removed. The resulting lists of unique peptides were used for downstream analyses.

We observed no substantial difference in peptide characteristics from IP and SP cells. For each HLA$_{SCD}$, Artemis binding motifs of 9mer peptides matched the widely accepted motifs from netMHCpan 4.1 for both IP and SP lines, as well as all 9mers recovered with Artemis (**Figure 3.5**). Likewise, the distribution of peptides 8-16 amino acids in length was similar for both IP and SP cells. 9mer and 10mer peptides were the most frequent (**Figure 3.5)** and is consistent with previously published work (*3, 16, 32*). Of note, some HLAs, including HLA-A11 in our subset, will accommodate longer peptides (*3*). Certain alleles repeatedly presented more peptides than others across proteasome isoforms with HLA-A11$_{SCD}$ yielding the most peptides, followed by HLA-A24$_{SCD}$ and then HLA-A2$_{SCD}$ (**Figure 3.6**).

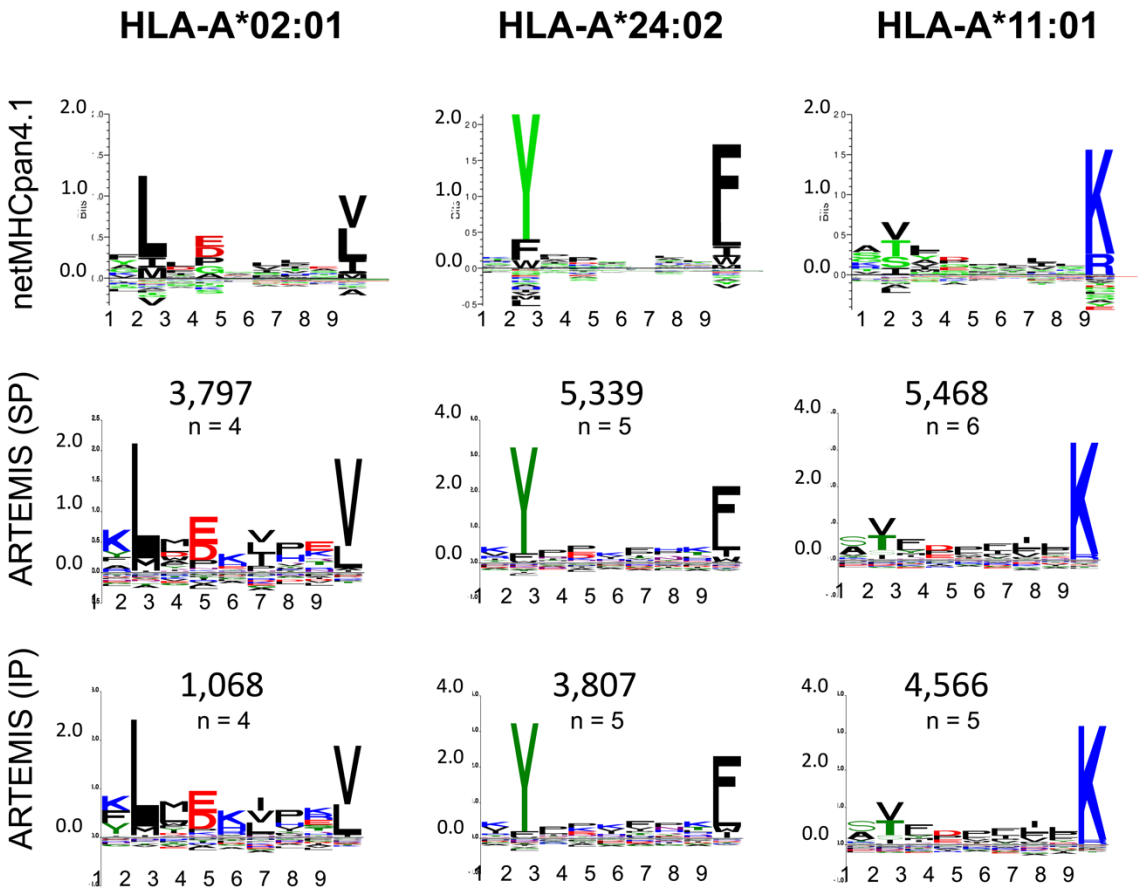


Figure 3.5 Sequence logos of HLA specific binding motifs for 9mer peptides recovered from Artemis. Row 1: Reference logos for each HLA allele quired available through the netMHCpan motif viewer portal. Row 2 and 3: Custom logos of 9mer peptides recovered from Artemis were generated using pymotif as described in the methods. Peptides from all replicates were aggregated. Resulting peptide yields are listed above each logo.

•Motif portal: http://www.cbs.dtu.dk/services/NetMHCpan/logos.ps.php

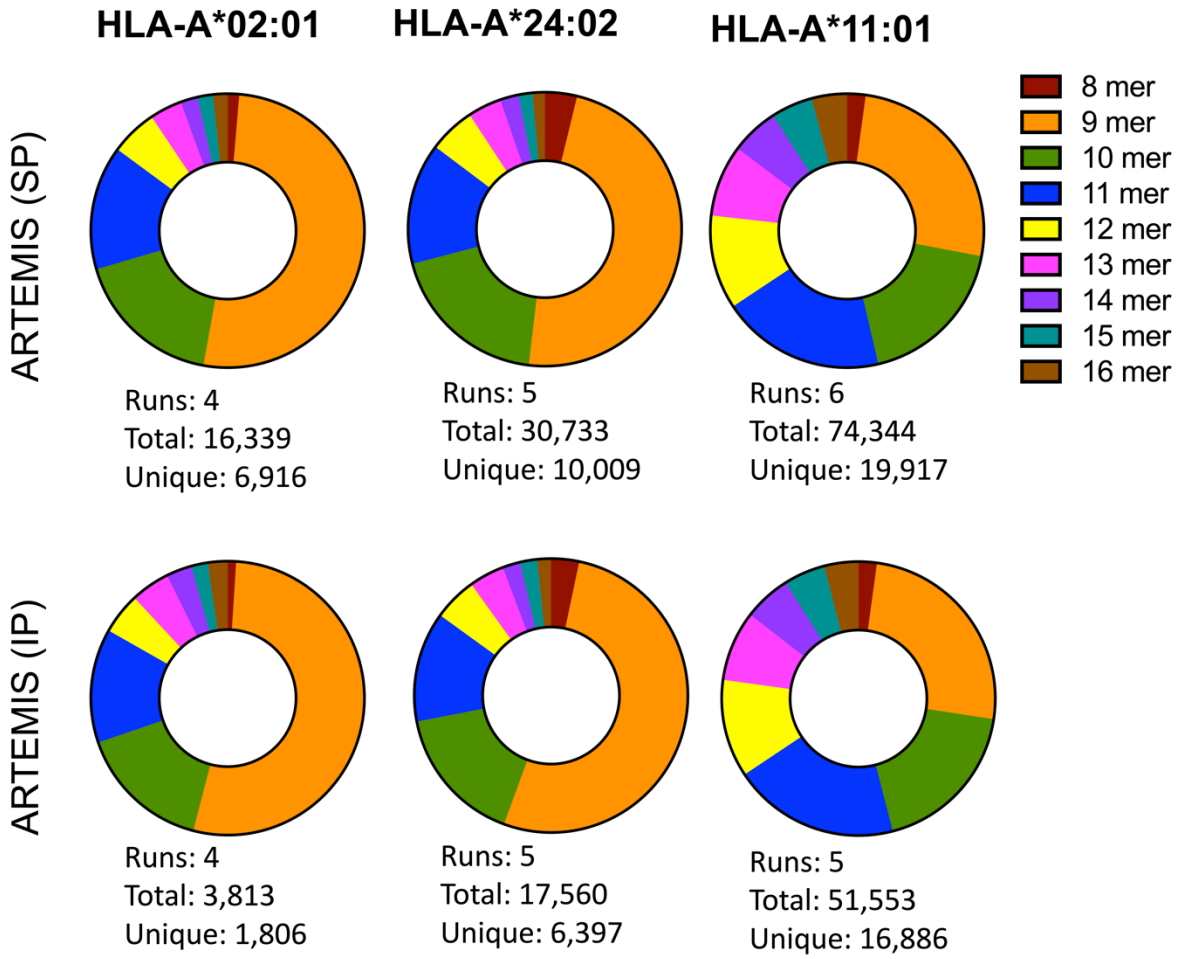


Figure 3.6 Observed length distributions for peptides recovered from each HLA_{SCD} allele for SP an IP cells. Length distributions of unique peptides from aggregated biological replicates are shown.

To determine if the observed peptides are predicted to be strongly, weakly, or not at all bound to the given HLA we utilized the web-based portal netMHCpan4.1 (*9*). As expected, most of the 9mers and 10mers were predicted to bind strongly with a rapid decrease in the frequency of strong binders as length increased above 12 amino acids. Almost no observed 8mers were predicted to strongly bind (**Figure 3.7**). SP and IP cells produced similar frequencies of strong-, weak- and non-binding peptides (**Figure 3.7**). This does not support the notion the IP can produce more strong binding peptides than the SP.

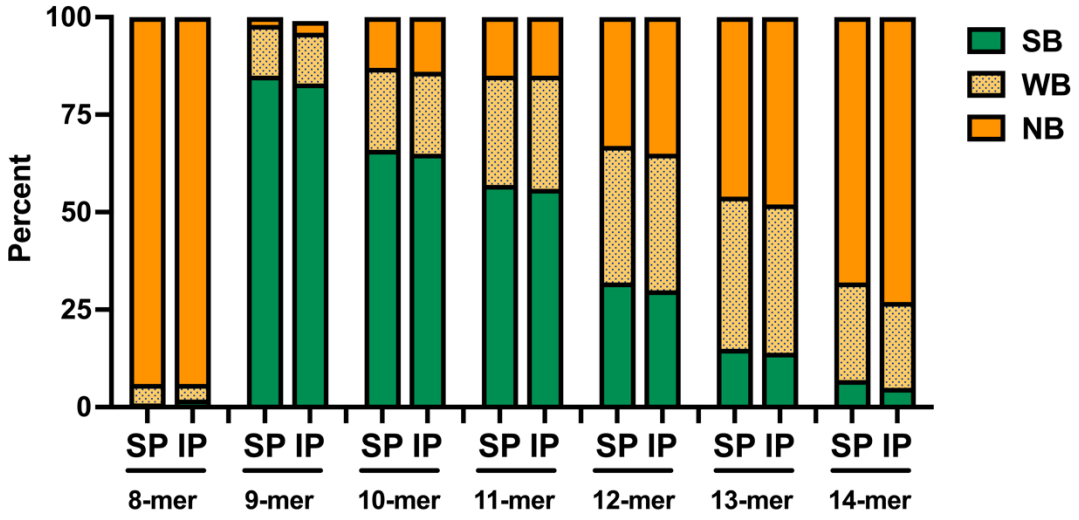


Figure 3.7 Predicted binding affinity for 8mer-14mer peptides recovered from SP and IP cells. Peptides of the specified length and derived from individual HLA_{SCD} were merged into a single peptide list for SP and IP cells. Binding affinity to the respective HLA for each peptide was determined using the netMHCpan 4.1 binding prediction web portal selecting 'peptides' as the input with HLA-A*2:01, HLA-A*24:-2 and HLA-A*11:01 alleles selected. Peptides were classified as strong binders (SB), weak binders (WB) and non-binders (NB) for each HLA. Strong binders and weak binders were tallied, and remaining peptides were assumed to be non binders.

Artemis reproducibility in SP and IP cells

Repeated sampling of biological replicates continued to reveal new peptides even after up to six Artemis replicates (**Figures 3.3, 3.8, and 3.9**) likely reflecting the dynamic nature of the immunopeptidome (*2*) and sampling stochasticity. However, the number of new peptides recovered drops steeply after the second replicate (**Figure 3.8**), implying a few replicates could capture the majority of commonly presented peptides. For example, Artemis performed on SP cells transduced with $HLA\text{-}A24_{SCD}$ yielded an average of 5760 (SD ± 1214) peptides from the first replicate and 1176 (SD ± 1008) from the second. The fourth and fifth replicates identified only 778 (SD ± 415) and 621 (SD ± 342) new peptides respectively (**Figure 3.8**). Moreover, when all SP + $HLA\text{-}A24_{SCD}$ biological replicates are aggregated, each replicate has 321-1242 unique peptides (**Figure 3.9**). Additional replicates would continue to increase the total number of observed peptides and approach complete coverage of the immunopeptidome.

However, a trend of diminishing returns after 3-4 replicates (*16, 32*). This trend is repeated across all HLA_{SCD} alleles and proteasome isoforms (**Figures 3.8 and 3.9**).

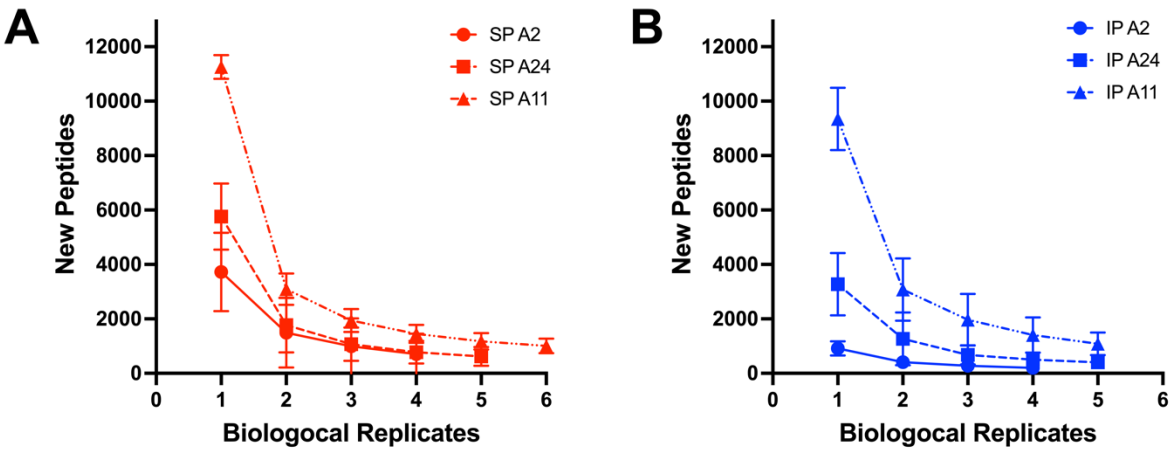


Figure 3.8 Artemis reproducibility. Average new peptides discovered in each subsequent biological replicate for **(A)** SP and **(B)** IP cells. Selecting replicates with the most unique peptides as replicate 1 would minimize the number of new peptides discovered in subsequent replicates; to reduce this bias a simulation was run with every iteration of replicate order. In other words, if there were 5 biological replicates 120 scenarios were simulated with every possible order of replicates. E.g., 1,2,3,4,5; 2,3,4,5,1; 3,4,5,1,2; etc. For HLA-$A11_{SCD}$ 6 biological replicates were performed with 720 possible orders. Because GraphpadPrism can only accommodate 512 y-values, a random sample of 512 scenarios were selected. Biological replicates for SP+HLA-$A2_{SCD}$ n=4; SP+HLA-$A24_{SCD}$ n=5; SP+HLA-$A11_{SCD}$ n=6; IP+HLA-$A2_{SCD}$, n=4; IP+HLA-$A24_{SCD}$ n=5; and IP+HLA-$A11_{SCD}$ n=5.

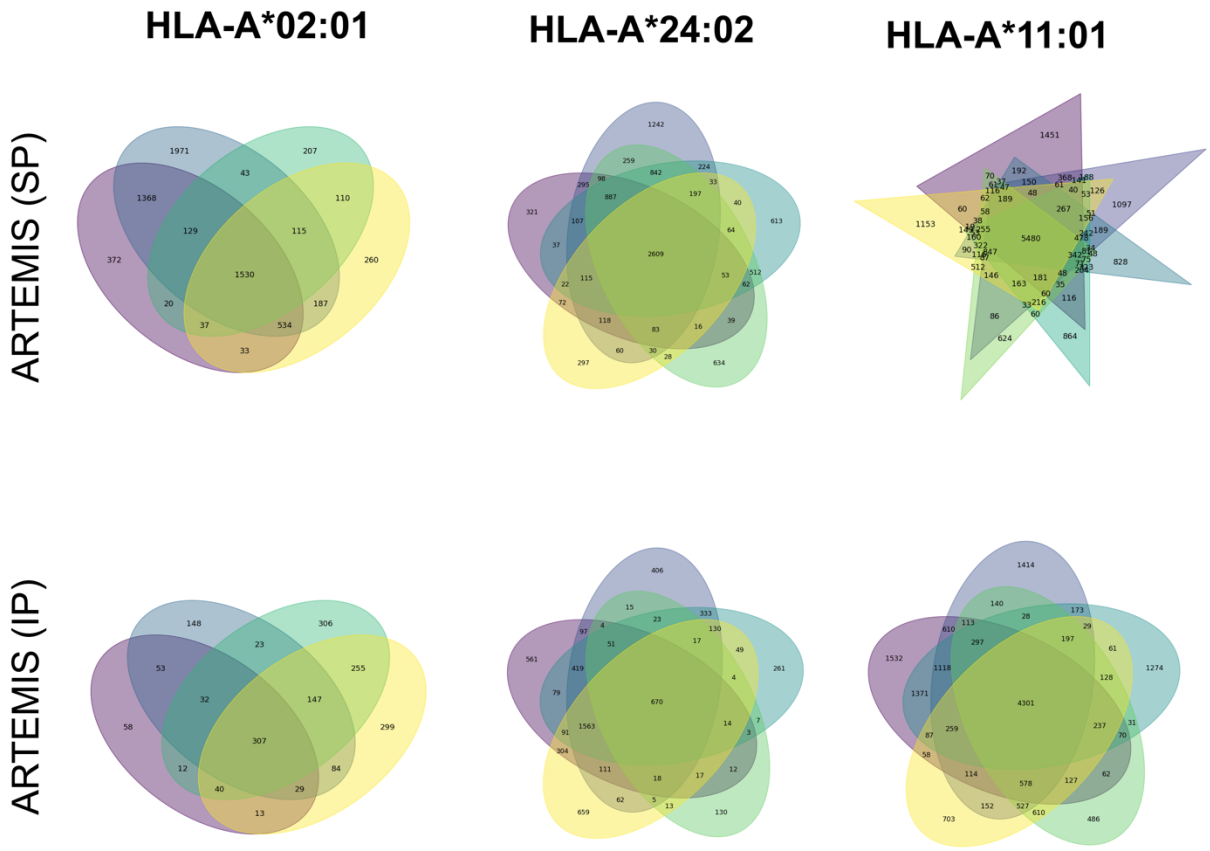


Figure 3.9 Venn diagrams of each biological replicate for SP and IP cells transduced with HLA-A2$_{SCD}$, HLA-A24$_{SCD}$ and HLA-A11$_{SCD}$. Peptides exclusive to each biological replicate were detected. Biological replicates for SP+HLA-A2$_{SCD}$ n=4; SP+HLA-A24$_{SCD}$ n=5; SP+HLA-A11$_{SCD}$ n=6; IP+HLA-A2$_{SCD}$, n=4; IP+HLA-A24$_{SCD}$ n=5; and IP+HLA-A11$_{SCD}$ n=5.

Defining the 'mutual' immunopeptidome

Although binding motifs, length distributions and predicted binding affinity did not vary between SP and IP cells, we wanted to determine if the sequences of identified peptides differed with proteasome composition. Previous literature indicates SPs and IPs generate distinct but overlapping peptide repertoires (*20, 21, 33, 34*) and clinical cases of immune evasion through modulation of proteasome subunits (*14, 24*) hint at a biologically-relevant difference. However, the extent of peptide overlap between proteasome isoforms for specific tumor antigens remains unclear.

Mass spectrometry (MS) has intrinsic variability, in part due to the complexity of the technique (*35*). Reproducibility of protein identifications from the same mixture can be up to ~70-80% across machines, when streamlined analysis protocols are in place. Even technical replicates (duplicate samples from the same tube) show variability with only ~80-85% reproducibility (*16, 36*). However, peptide reproducibility showed only 35-60% overlap between technical replicate pairs (*35*). To determine the overlap for our experiments, we compared the immunopeptidomes from three representative biological replicates for each HLA_{SCD}-proteasome combination; when possible, sets with similar numbers of peptides were used. Different HLA_{SCD} alleles were not compared as anchor residues vary between HLAs tested which sharply reduces common peptides (*16*). The overlap of biological replicates was calculated (see Methods) (*16*) for each pair in a representative set of three SP or IP runs (within the same HLA). SP replicates had an average of 79.8% overlap (SD $\pm$ 6.17%, range = 66.5-88.0%) (**Figure 3.10**, top row) while IP replicates had an average of 75.8% overlap (SD $\pm$ 5.79%, range = 66.2-82.5%) (**Figure 3.10**, bottom row). Although variation in MS combined with stochastic antigen presentation might predict limited overlap, we report overlaps similar to those previously observed between replicates (*16, 35, 36*).

We next compared the overlap between SP- and IP-expressing cells for each HLA_{SCD} allele, aggregating duplicates and aggregating all runs for each HLA and proteasome combination. We observed 32.1%, 46.9% and 60.5% SP:IP overlap for HLA-$A2_{SCD}$, HLA-$A24_{SCD,}$ and HLA-$A11_{SCD}$ respectively (**Figure 3.11A**). This was notably lower than the overlaps within the same proteasome isoform (**Figure 3.10**). Absolute peptide recovery from IP replicates was considerably lower than from SP replicates. This could contribute to a reduced overlap, as the calculation used gives each input equal weight and will skew lower when the two sets are of very different sizes (*16*). Reduced peptide yields from IP cells may be a technical artifact stemming from the reduced growth kinetics of IP cells, described above. To account for this discrepancy, a subset of peptides in each SP data set was randomly selected to match the total peptides found in the corresponding IP condition (see Methods: down sampling). Interestingly, the SP:IP overlap was further reduced to 20.2%, 35.9% and 56.0% overlap for HLA-$A2_{SCD}$, HLA-$A24_{SCD,}$ and HLA-$A11_{SCD}$ respectively (**Figure 3.11B**). This suggests that the SP:IP

overlap is HLA specific and that the fraction of 'mutual' peptides in the immunopeptidome could vary depending on which HLAs are present on the cell.

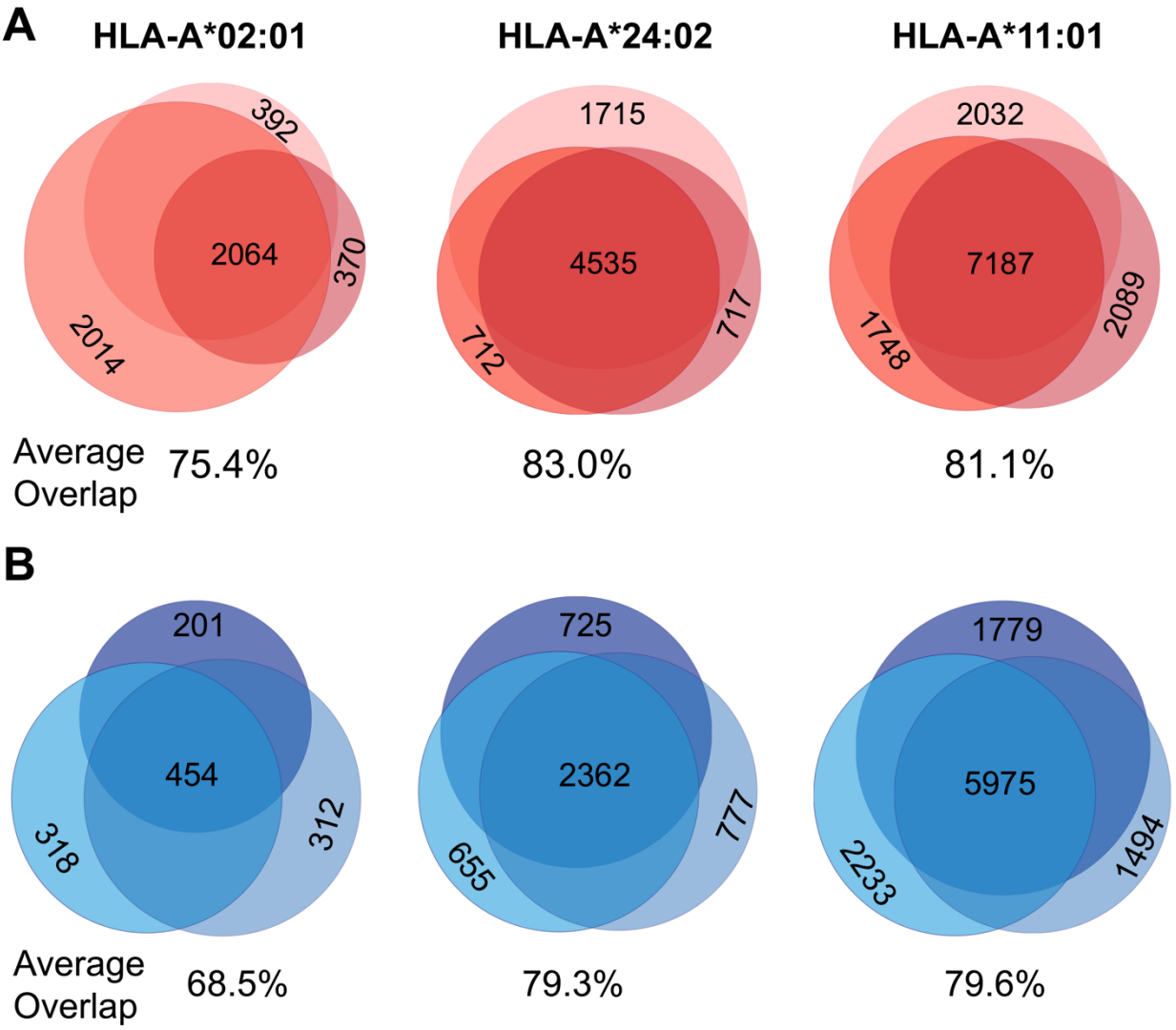


Figure 3.10 Venn diagram overlaps of three representative biological replicates for HLA-A2$_{SCD}$, HLA-A24$_{SCD}$, and HLA-A11$_{SCD}$ in SP and IP cells. Three biological replicates were selected with similar numbers of recovered peptides. Percent overlaps were calculated for each pair and average overlaps are listed below respective venndiagrams. Venn diagrams were generated using the BioVenn online portal (http://www.biovenn.nl) and overlaps were calculated as described in the methods. **(A)** Venn diagrams for SP cells and **(B)** Venn diagrams for IP expressing cells.

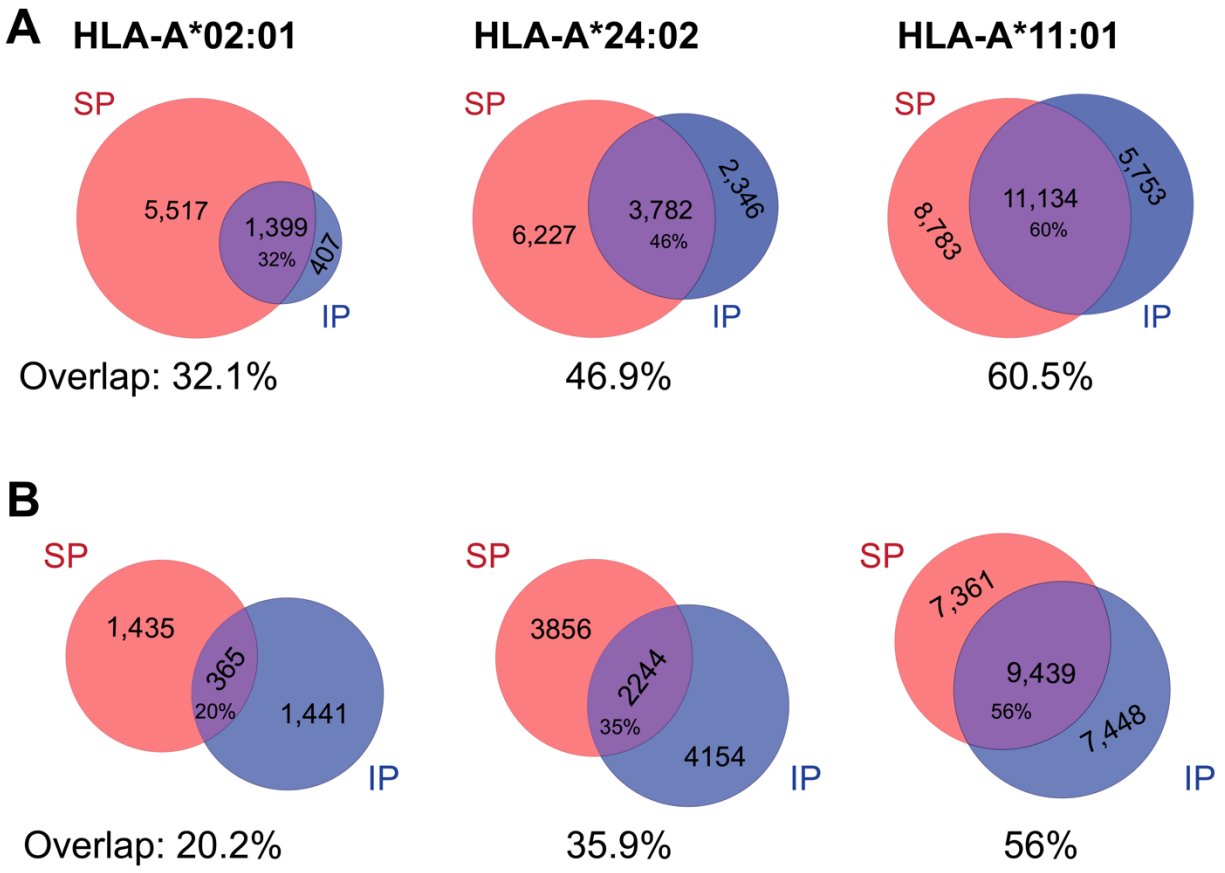


Figure 3.11 Immunopeptidome overlap for SP and IP cells. (A) Overlap of all unique peptides detected in SP and IP cells for each respective HLA_{SCD}. Biological replicates for SP+HLA-$A2_{SCD}$ n=4; SP+HLA-$A24_{SCD}$ n=5; SP+HLA-$A11_{SCD}$ n=6; IP+HLA-$A2_{SCD}$, n=4; IP+HLA-$A24_{SCD}$ n=5; and IP+HLA-$A11_{SCD}$ n=5. **(B)** Overlap of down sampled SP subsets with all recovered IP peptides. Average overlap is listed below each Venn diagram.

To see if there was any difference in 'distinct' versus 'mutual' peptide sequences, we looked for divergent patterns in the sequence logos for 9mer SP 'distinct, 'mutual', and IP 'distinct' peptides across all three HLA alleles tested (**Figure 3.12**). The peptides from SP or IP cells were aggregated for each HLA_{SCD}. The observed SP:IP overlap appeared to be a weighted average from all three HLAs with relatively high overlap (52%) likely due to the large contribution of peptides from HLA-$A11_{SCD}$, which saw the highest allele-specific overlap at 56% (**Figure 3.12, top**). There was a slight deviation between the sequence logo for the SP 'distinct' peptides and the sequence logo for both

the ‘mutual’ and IP ‘distinct’ peptides. This included more prominent lysine (K) at position 1, leucine (L) at position 2, aspartic acid (D) and glutamic acid (E) at position 4, and valine (V) at position 9. These amino acid preferences mirror the binding motif for HLA-A*02:01. Their relative higher proportion in the SP sequence logo likely reflect the low HLA-A2$_{SCD}$ yields in the IP cell line and therefore low contribution of HLA-A*02:01 peptides to the IP peptide pool. Increasing the number of peptides in the IP dataset is required to draw more definitive conclusions.

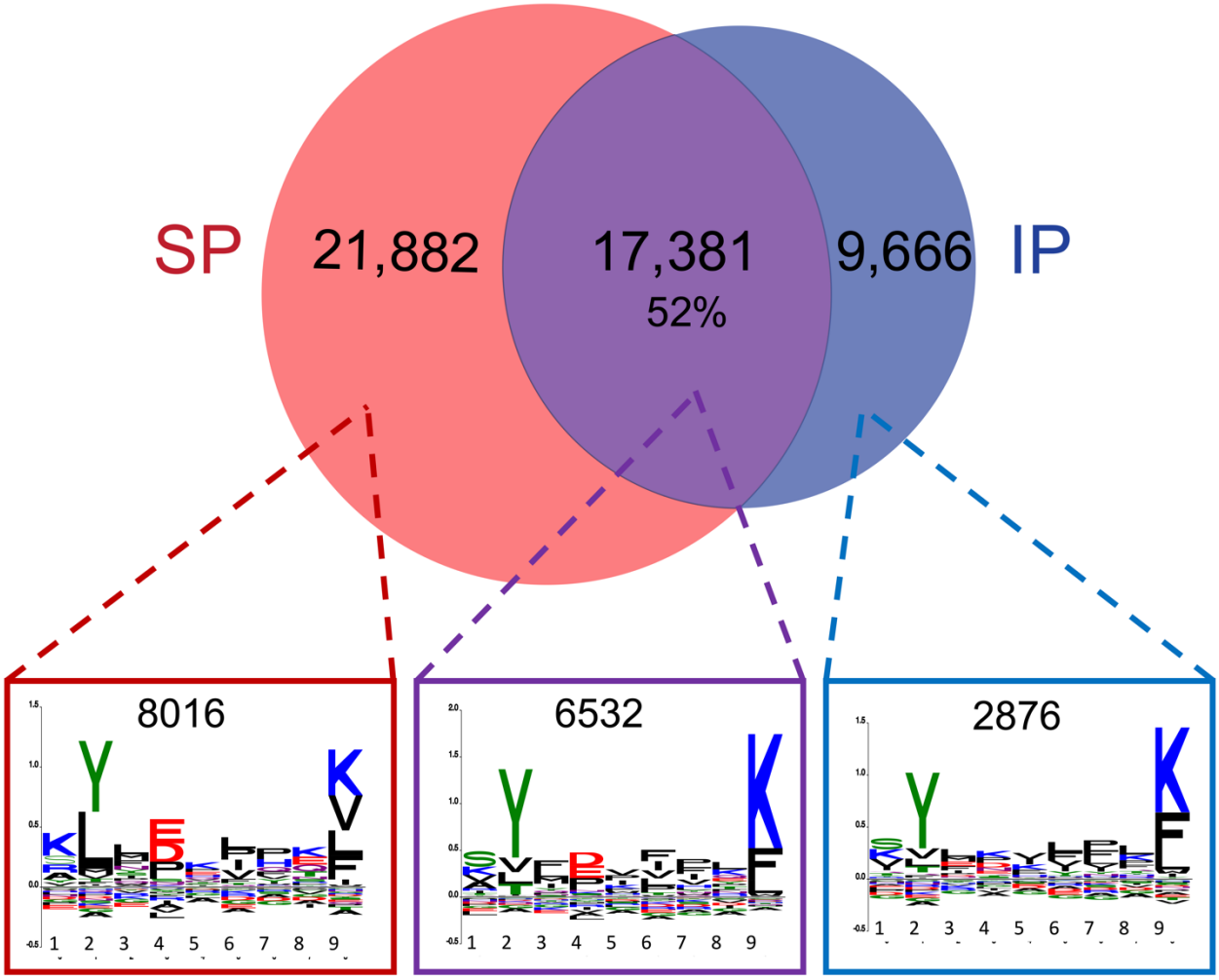


Figure 3.12 Sequence logos of ’mutual’ and ‘distinct’ peptides. Top: overlap of all unique peptides detected in SP and IP lines. Total numbers of unique peptides are listed. Bottom: Sequence logos of SP ‘distinct’, ‘mutual’ and IP ‘distinct’ 9mer peptides. Number of peptides used to generate the sequence logos are listed.

Mechanisms driving IP and SP immunopeptidome divergence are debated. Hypotheses include divergent cleavage specificities (*37-39*) and differing catalytic rates (*23*). Proteasomes cleave the peptide bond c-terminal to the final amino acid of the resulting peptide. Cleavage of a 9mer peptide is illustrated in **Figure 3.13**. For this example 9mer (**Figure 3.13A**), the flanking amino acid n-terminal to P1 is labeled P1' and the amino acid c-terminal to P9 is labeled P9', to illustrate their relationship to the resulting peptide. Because IP and SP specific β-subunits share considerable homology, 59% -71% (*23*), the flanking amino acids could influence cleavage specificities. Therefore, we looked at residues flanking the observed peptide (P1' and P9') as predicted based on alignment to the human proteome. Unique 9mers were aggregated across the three HLA_{SCD} alleles for SP and IP and predicted flanking residues were added to the sequence. Motifs for the resulting 11mers (9mer +2 flanking residues) showed no preference for specific residues flanking the cleavage site (**Figure 3.13B**). The P9 position did show a deviation between for SP and IP, a valine (V) was enriched at P9 for SP but not IP (**Figure 3.13B**), again likely driven by having very few HLA-A2 restricted peptides in the IP dataset. As HLA-A2 prefers a valine (V) or leucine (L) at P9, having few HLA-A2 restricted peptides means these residues are underrepresented in the IP dataset. Viewing presented peptides and not direct proteasomal products could obscure catalytic differences between SP and IP by overshadowing the differences with HLA binding preferences.

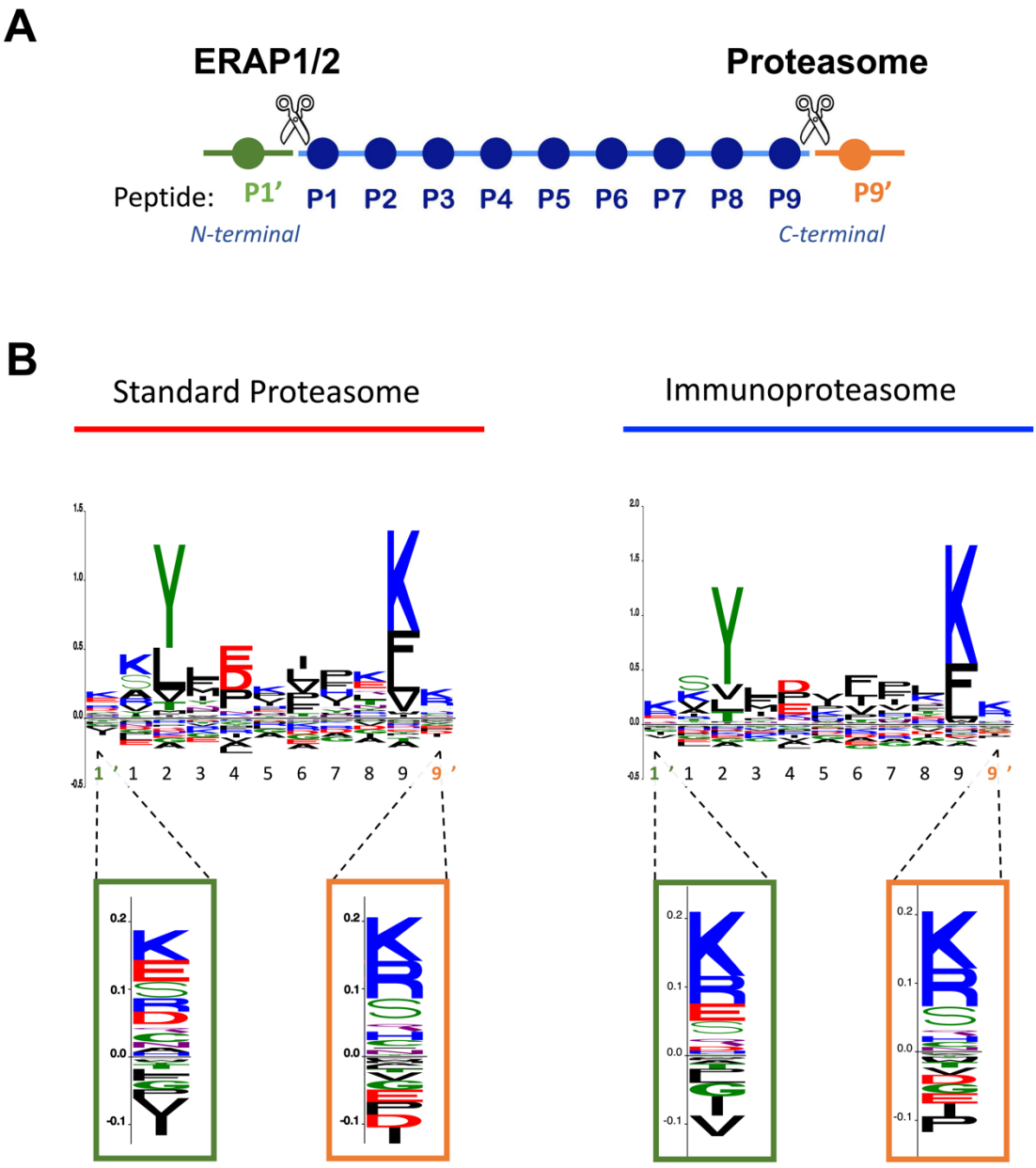


Figure 3.13 Cleavage specificities for IP and SP cells. (A) Schematic of peptide processing depicting cleavage sites for the proteasome and ERAP. **(B)** Sequence logos of aggregated SP and IP cells with one flanking residue on each terminus. Subsets enlarge P1' and P9' flanking residues. No difference in cleavage patterns were observed.

Utilizing ARTEMIS to identify 'mutual' peptides of clinical interest.

Observed reduction in SP and IP overlap, in conjunction with clinical reports of immune evasion through proteasome isoform substitution, imply that proteasomal processing should be considered when selecting clinical targets. We therefore used Artemis to identify 'mutual' peptides from cancer antigens of interest.

Three acute myeloid leukemia (AML) antigens were selected. these include an over expressed self-antigen that was previously pursued as a TCR-T cell target, Wilms Tumor-1 (WT1) (*4, 14, 40, 41*). A promising over expressed self-antigen not yet pursued as an AML TCR-T cell target is Cyclin-A1 (CCNA1) (*42*), and promising mutated antigen in AML is Calreticulin (CALR), with a frameshift mutation yielding 44 new amino acids on the c-terminus (CALRmut) (*43*).

Previous Artemis work showed exposure to IFN-γ improved peptide yields (*16*), but IFN-γ also modulates the proteasome. In an initial experiment to see if we could detect peptides from WT1, we exposed HEK293E (SP) cells transduced with HLA-$A2_{SCD}$ and WT1 to inflammatory cytokines, INFγ and TNF-α, according to the previously published Artemis protocol. Under these conditions we recovered five HLA-A2 restricted, WT1-derived peptides (**Table 3**). This included the $WT1_{37\text{-}45}$ [VLDFPAGGA] peptide identified by our group (*14*) and others (*40*) as a viable TCR-T cell target and determined to be a 'mutual' peptide through T cell recognition of the SP and IP cells. Due to exposure to cytokine exposure, we cannot conclusively determine if this peptide is produced by the IP, SP or both. Although we did observe an increase in HLA-$A2_{SCD}$-restricted peptides when exposed to IFN-γ (7349 peptides), subsequent experiments were not subjected to cytokine to prevent changes in proteasome composition and allow us to delineate 'distinct' versus 'mutual' peptides.

Proteasome	Ag.	HLA-A*02
SP/IP	WT1	FGPPPPSQA NLYQMTSQLEC NMTKLQLAL SLGEQQYSV VLDFAPPGA

Table 3 HLA-$A2_{SCD}$-restricted, WT1-derived peptides under cytokine exposure. Peptides derived from WT1 using and identified Artemis when SP cells are exposed to inflammatory cytokines.

Incomplete co-transduction of HLA_{SCD} (mCherry reporter) and antigen (GFP reporter) indicated a need to enrich for cells expressing both the SCD and target antigen (**Figure 3.14A**). Transduced cells were sorted once or twice to achieve >80% purity of double-positive cells (**Figure 3.14B**) and thereby improve the probability of detecting peptides from target antigens. Artemis was performed for each HLA_{SCD} and antigen combination at least twice. Peptides derived from antigens of interest were identified for each HLA_{SCD} (**Figures 3.15-3.17**). Unfortunately, yields of HLA-$A2_{SCD}$ restricted peptides were low in the IP line and we were unable to recover any peptides from WT1. The combination of three peptide-limiting factors, *1.* low peptide recovery for WT1 compared to other antigens, *2.* reduced peptide yields in IP versus SP cells, and *3.* fewer peptides for HLA-$A2_{SCD}$ compared to HLA-$A24_{SCD}$ or HLA-$A11_{SCD}$, likely contributed to no detectable WT1 derived, HLA-A2 restricted peptides from the IP line.

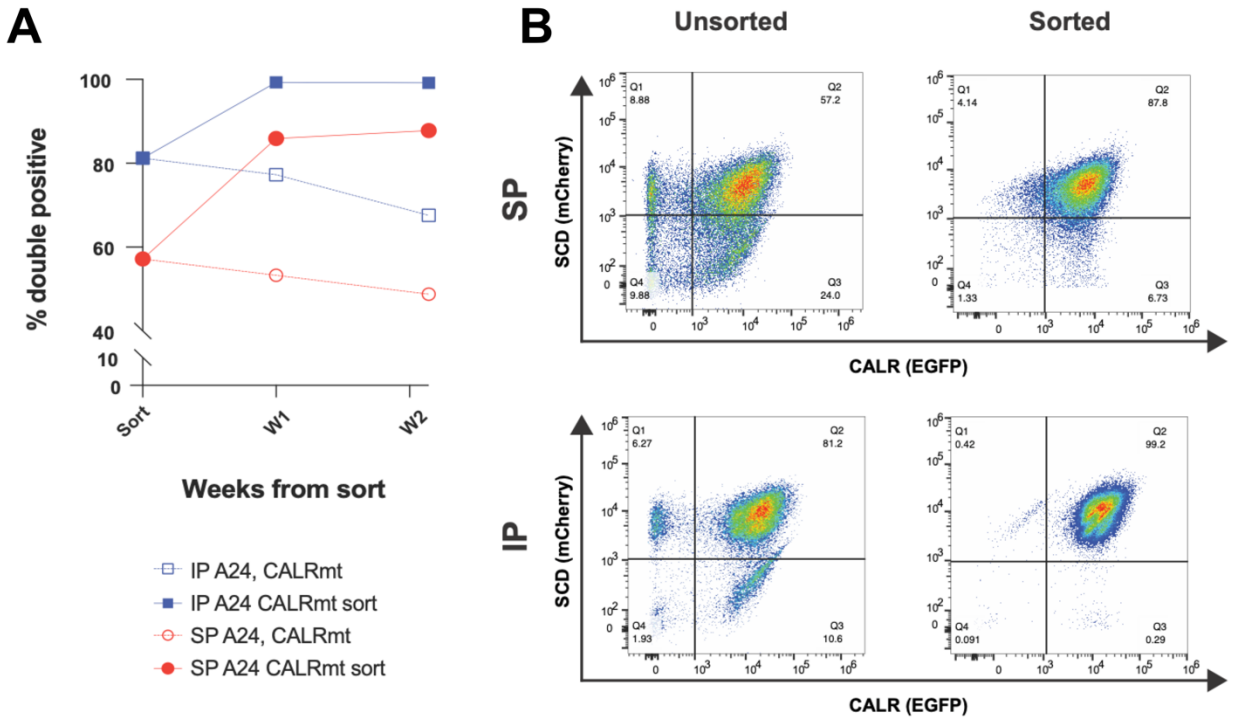


Figure 3.14 Enriching for HLA_{SCD} and antigen expressing cells. **(A)** Frequency of enriched double positive cells and unsorted cells over time. Both IP and SP cell lines show a decrease in the frequency of double positive cells over time. Cells sorted for HLA_{SCD} and antigen co-expression continue to show high frequencies of double positive cells. **(B)** Representative flow plots of cell lines at sorting and two weeks after a single sort.

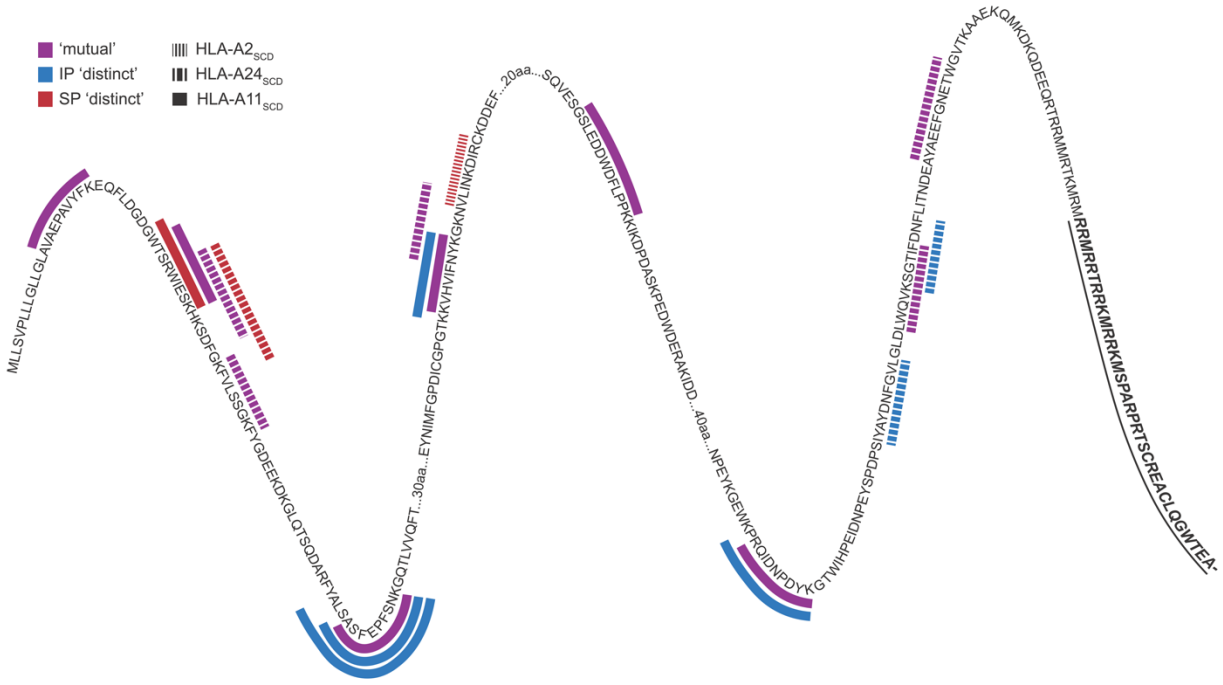


Figure 3.15 CALRmut-derived peptides across HLA alleles for IP and SP cells. Protein sequence is displayed n-terminus (left) to c-terminus (right). Where peptides occur are indicate with lines. 'Mutual' peptides are depicted in purple, IP 'distinct' peptides in blue, and SP 'distinct' peptides in red. The presenting HLA_{SCD} is indicated as a dashed line (HLA-A2), a blocked line (HLA-A24), or a solid line (HLA-A11). Note: that no peptides in the mutant portion of the protein were detected.

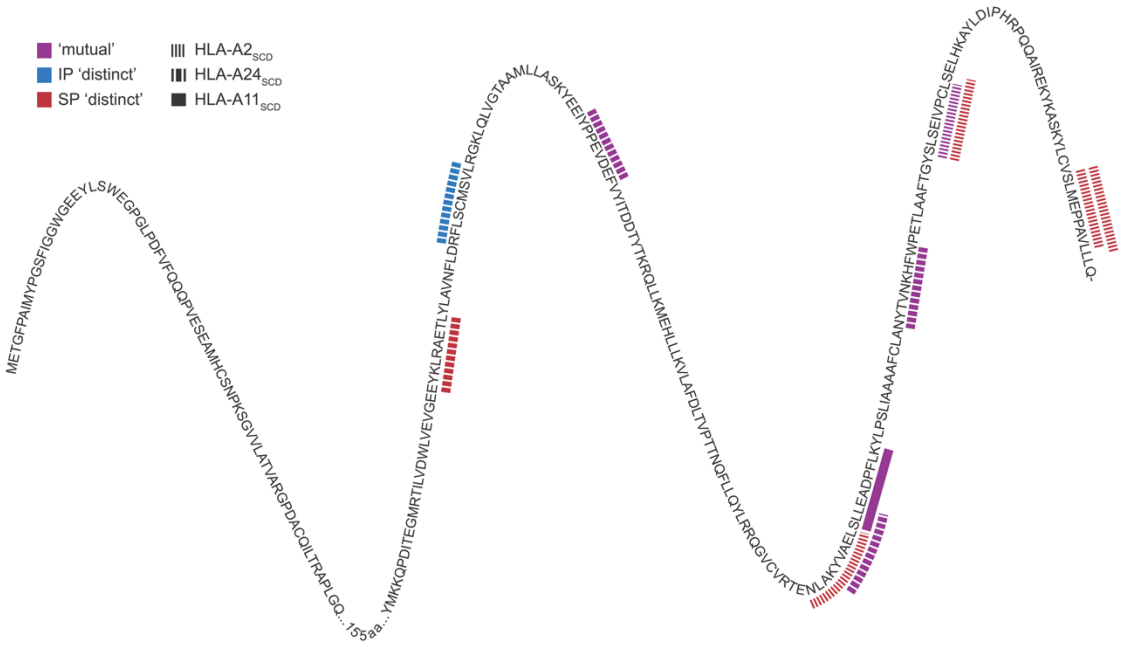


Figure 3.16 CCNA1-derived peptides across HLA alleles for IP and SP cells. Protein sequence is displayed n-terminus (left) to c-terminus (right). Where peptides occur are indicate with lines. 'Mutual' peptides are depicted in purple, IP 'distinct' peptides in blue, and SP 'distinct' peptides in red. The presenting HLA_{SCD} is indicated as a dashed line (HLA-A2), a blocked line (HLA-A24), or a solid line (HLA-A11).

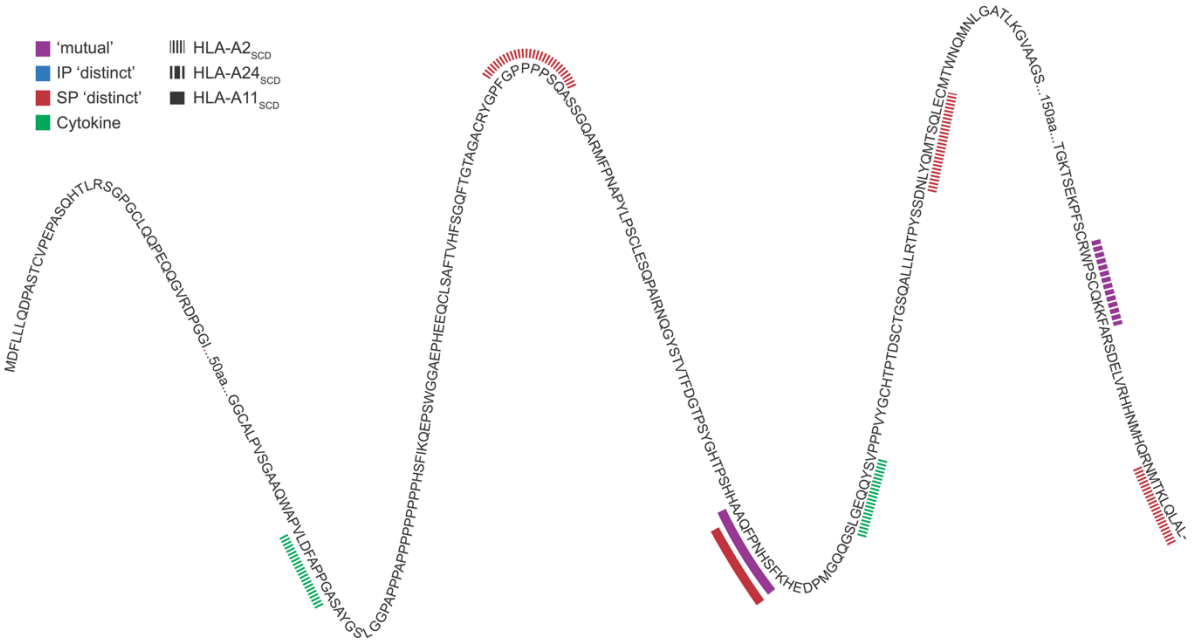


Figure 3.17 WT1-derived peptides across HLA alleles for IP and SP cells. Protein sequence is displayed n-terminus (left) to c-terminus (right). Where peptides occur are indicate with lines. 'Mutual' peptides are depicted in purple, IP 'distinct' peptides in blue, SP 'distinct' peptides in red and peptides recovered after cytokine exposure are green. The presenting HLA_{SCD} is indicated as a dashed line (HLA-A2), a blocked line (HLA-A24), or a solid line (HLA-A11).

91 unique antigen-derived peptides were identified across the three HLA alleles and proteasome-isoforms. As the selected HLAs should not present the same peptide, 54 were determined to be contaminating peptides, i.e., appearing in more than one HLA_{SCD} data set and not expected to bind any HLA_{SCD} (*16*). Contaminating peptides tended to be longer in length and with acidic residues at the C-terminus, and therefore not predicted to bind any of the selected HLAs. Although antigen presenting cells can cross-present exogeneous peptides on class I HLAs (*44*), HEK293E cells are not specialized presenting cells and are not expected to efficiently cross present peptides. Contaminant peptides were likely due to overexpressing the antigen (*16*). CALRmut-derived peptides were detected from HLA-$A2_{SCD}$-transduced IP cells, although all were considered contaminants. Contaminating peptides were culled and the remaining 37 peptides are listed in **Table 4,** 19 were 'mutual' peptides (**Table 4**, purple bold). Again, we observed differences in the frequency of 'mutual' peptides for each HLA_{SCD} allele

(**Table 4**, purple) with HLA-$A2_{SCD}$ having the fewest 'mutual' peptides and HLA-$A11_{SCD}$ having the most, suggesting that the repertoire of 'mutual' peptides is HLA-specific. Although these numbers do align with findings in **Figure 3.11**, the strength of this correlation was limited by the very few peptides recovered from each antigen. Of note, several similar peptides differed by only a single amino acid, such as [AQFPNHSFK] and [AAQFPNHSFK]. Per our filtering criteria outlined in the methods (*16*), these were considered real and unique peptides.

We next wanted to know where the 'distinct' or 'mutual' peptides were found in the parent protein. Each tested antigen is displayed as a linear sequence (**Figures 3.15-3.17)** with 'mutual' peptides in purple, IP 'distinct' peptides in blue and SP 'distinct' peptides in red. Antigens with more recovered peptides began to show 'hot spots' similar to previous work (*45*). Unfortunately, we were unable to recover any peptides from the mutant c-terminal sequence of CALRmut (**Figure 3.15**, bold and underlined).

Of the 37 recovered peptides, four were previously described to be immunogenic (**Table 4**, yellow highlight) supporting the use of Artemis as a novel platform to identify presented peptides and clinically-relevant targets. Standard immunoprecipitation was used to elute both $WT1_{185\text{-}193}$ [FGPPPPSQA] and $WT1_{242\text{-}250}$ [AQFPNHSFK] off WT1-expressing cells including primary AML cells and an ovarian cancer line (*46*) . Likewise, $CCNA1_{415\text{-}423}$ [SLSEIVPCL] has been eluted off primary AML cells (*47*). $WT1_{490\text{-}498}$ [RWPSCQKKF] has been shown to be immunogenic (*48, 49*) and used in peptide vaccines (*50*). Artemis also identified peptides not previously described, though immunogenicity still needs to be confirmed.

Proteasome	Ag.	HLA-A*02	HLA-A*24	HLA-A*11
SP	CCNA1	NLAKYVAEL	EYKLRAETL	**SLLEADPFLK**
		SLMEPPAVL	**IYPPEVDEF**	
		SLMEPPAVLL	**KYVAELSLL**	
		SLSEIVPCL	**NYTVNKHFW**	
		SLSEIVPCLS		
	CALRmt	VLINKDIRC	**KFVLSSGKF**	**ASFEPFSNK**
			NYKGKNVLI	**KVHVIFNYK**
			LWQVKSGTIF	**RQIDNPDYK**
			AYAEEFGNETW	**WTSRWIESK**
			RWIESKHKSDF	**AVAEPAVYFK**
			RWIESKHKSDFGKF	GWTSRWIESK
				GSLEDDWDFLPPK
	WT1	NMTKLQLAL	**RWPSCQKKF**	**AAQFPNHSFK**
		NLYQMTSQLEC		AQFPNHSFK
		FGPPPPSQA		
IP	CCNA1	**SLSEIVPCL**	**IYPPEVDEF**	**SLLEADPFLK**
			KYVAELSLL	
			NYTVNKHFW	
			RFLSCMSVL	
	CALRmt		**KFVLSSGKF**	**ASFEPFSNK**
			KSGTIFDNF	**KVHVIFNYK**
			NYKGKNVLI	**RQIDNPDYK**
			IYAYDNFGVL	**WTSRWIESK**
			LWQVKSGTIF	**AVAEPAVYFK**
			AYAEEFGNETW	KKVHVIFNYK
			RWIESKHKSDF	SASFEPFSNK
			RWIESKHKSDFGKF	KPRQIDNPDYK
				GSLEDDWDFLPPK
				YALSASFEPFSNK
	WT1		**RWPSCQKKF**	**AAQFPNHSFK**

Table 4 Artemis recovered 'distinct' and 'mutual' peptides from three AML associated antigens, WT1, CCNA1, and CALRmut. Peptides derived from antigens of interest that also had sequences matching the corresponding HLA binding motif were deemed well presented peptides and good candidates for future analysis. Mutual peptides are in purple. Previously described immunogenoic peptides are highlighted in yellow.

Discussion

Proteasomes are the primary source of peptides presented by class I HLA and can determine precisely what peptides are presented (*20*) for immune recognition. Indeed, there are multiple clinical case reports (*14, 24*) of proteasome modulation to escape immune surveillance highlighting the need to understand how each proteasome isoform impacts immunopeptidome repertoires. Here we use an emerging peptide discovery pipeline, Artemis, to outline the relative contributions from SP and IP to the immunopeptidome.

Intra-proteasome analysis showed a high degree of similarity and substantial overlap between biological replicates for both SP (79.8%) and IP (75.8%) but several hundred to a couple thousand peptides were replicate-exclusive reflecting inherent stochasticity. Inter-proteasome analysis showed lower overlap between the peptides from IP and SP cells, and specificity to the presenting HLA allele with HLA-A*02:01 had the least overlap (20.2%) followed by HLA-A*24:02 (35.9%) and HLA-A*11:01 (56.0%). Additional investigation is needed to determine if these overlaps are statistically significant, and we are currently working with a statistician to answer this question. Furthermore, the increased SP:IP overlap correlates with increased peptide yields. It is unclear if the variable overlaps observed for individual HLAs are due to underlying biology or under sampling for HLAs with lower peptide yields. Several of the clinical case reports describing shifts in proteasome composition to evade immune recognition involve HLA-A*02:01-restricted peptides (*14, 15*), which showed the lowest overlap in our data and may facilitate this escape phenomena. This prominence, however, may be because HLA-A*02:01 is a common restricting HLA in T cell therapy. In any case, these moderate overlaps between SP and IP peptides highlight a need to consider proteasomal processing when selecting peptides for targeted immunotherapies.

Fewer peptides were detected from IP than SP cells, which may reflect diminished growth and viability of the IP cells with high viral burdens (*31*); reducing the viral load or carefully monitoring cell growth in future replicates could increase yields from IP cells. Our work also provides evidence contradicting the existing hypothesis that the IP has different catalytic properties to generate peptides with higher binding efficiencies than the SP (*51*). We showed similar frequencies of strong binding, weak

binding and nonbinding peptides for SP and IP cells. We were also unable to find any difference in cleavage specificities between peptides from IP and SP cells. Expanding analyses to flanking residues beyond those immediately proximal to the peptide could reveal distal IP or SP cleavage specificities. Furthermore, we could expand analyses to include additional HLAs with anchor residues in the center of the peptide to potentially reveal SP/IP c-terminal cleavage specificities.

We have demonstrated that Artemis can be used to identify peptides from cancer-associated antigens of interest and to quickly distill a list of candidate peptides for further validation. Already, 4/37 antigen-derived peptides have been described as immunogenic (*46, 48*), suggesting translational potential (*14, 40*). How efficiently and accurately Artemis identifies immunogenic clinical targets requires immunogenic validation of the remaining peptides.

Artemis identified ‘distinct’ and ‘mutual’ peptides with 19 of the 37 (51.4%) antigen derived-peptides appearing as ‘mutual’. Additional experiments are needed to confirm if ‘distinct’ peptides are indeed isoform-dependent and not ‘mutual’ peptides that appear ‘distinct’ because of sampling stochasticity. This will be done with T cells recognizing ‘distinct’ and ‘mutual’ peptides; lines that respond to both SP- and IP-restricted cells will be considered ‘mutual’ peptides.

T cell-based therapies are becoming more prevalent in the clinic. As these immunotherapies gain momentum, rapidly identifying high quality TCR-T cell targets will be increasingly relevant. Correct proteasomal processing of candidate peptides is a necessary consideration when picking TCR-T target peptides. Peptide discovery using the Artemis platform demonstrated remarkable reproducibility and ability to identify ‘mutual’, immunogenic peptides.

Materials and Methods

HLA frequencies. HLA frequencies in the US population were calculated using publicly available data sets from the allele frequency net database (allelefrequencies.net) and assuming Harty-Weinberg equilibrium.

Immunoblot of proteasome catalytic subunits from HEK293E cell lines expressing specific proteasome isoforms. Assembled proteasome complexes were isolated using co-immunoprecipitation with an anti-α2 antibody, MCP21 (Enzo Lifesciences, BML-PW8105) and immunoblotted with different immunoprecipitation beads as previously described (*34*) (*28*). Briefly, Dynabeads M-270 epoxy beads (Invitrogen) were functionalized with MCP21 at 35 µg antibody per mg Dynabeads following manufacturer's instructions. Cleared lysate was quantified using a BCA assay Kit (Thermo Fisher Scientific) and 700 ng protein was incubated with 1mg functionalized beads to capture assembled proteasome complexes. Complexes were eluted in 0.1M Glycine buffer pH 2 to 2.5 and neutralized with equal volume Tris pH 7.5.

Proteasome composition was assessed by immunoblot. 10µL of the elution was separated on a denaturing 4-12% polyacrylamide gel and then transferred to a PVDF (Invitrogen) membrane, using the mini-blot module (Invitrogen). Membranes were blocked for 1 hour at room temperature in TBS containing 5% of dry milk and 0.1% Tween 20 and washed three times in TBS 0.1% Tween and incubated with the relevant primary antibody. SP and IP β-subunits were probed overnight in a solution of TBS, 5% milk powder, and 0.1% Tween-20 for β1 (1:500 dilution, Enzo Life sciences, BLM-PW8140), β2 (1:250 dilution, Enzo Lifesciences, BML- PW8145), β5 (1:1000 dilution, Enzo Lifesciences, BML- PW8895), β1i (1:500 dilution, Enzo Lifesciences, BML-PW8840), β2i(1:250 dilution Novus, NBP1-8660-), β5i (1:1000 dilution, Enzo Lifesciences, BML- PW8355) and the shared beta-subunit β7 (1:1000 dilution, Enzo Lifesciences, BML- PW8135) which was used as a loading control.

Surface HLA-expression: Expression of endogenous, membrane bound HLA was assessed using flow cytometry. SP and IP cells were stained with a pan HLA antibody (W6/32, Alexafluor-700, 1:800, BioLegend) and viability stain (Aqua Live/dead stain, BV510, 1:1000, Invitrogen). Cells were assessed on a NovoCyte Quanteon flow cytometer using standard operating procedures. Collected data was processed using FlowJo 10.8.1.

Immunoblot of SCDs. 1.0 x10^6 HEK293E cells expressing SP or IP were transduced with lentivirus to express SCD (mCherry reporter). After 24 hours lentivirus was removed and replaced with fresh media. After 5 days cells were FACS sorted to enrich for SCD^+ cells. Cells were grown and seeded at a starting concentration of 0.5 x10^6 cells /mL and allowed to grow for 5 days. Supernatant was removed and run on a denaturing 4-12% polyacrylamide gel followed by transfer to a PVDF (Invitrogen) membrane, using the mini-blot module (Invitrogen). Membranes were blocked for 5 minutes at room temperature in EveryBlot blocking buffer (BioRad). Membranes were washed three times in TBS, 0.1% Tween and incubated with and Anti-His tag primary antibody (HIS.H8, Invitrogen, 1:1000) at 4ºC overnight. Membranes were washed three times in TBS, 0.1% Tween and incubated with a IRDye® 800CW secondary antibody (1:20,000; LI-COR, 926-32212) for 30 minutes at room temperature before a second round of washes and imaging using the LI-COR Odyssey system

Artemis. All experiments described in this chapter were conducted using HEK293EBNA (HEK293E) cells expressing either the standard proteasome or immunoproteasome (*14, 28*). Our initial Artemis experiment with HLA-$A2_{SCD}$ and WT1 transduced SP cells was performed following the Artemis pipeline and exposing cells to cytokine as previously described (*16*). All subsequent experiments followed previous Artemis protocol with two additional deviations. First, the SCD construct was moved to a new vector backbone in which the UCOE site was removed and the SFFV promoter replaced with a CMV promoter. Second, cells were not exposed to inflammatory cytokines to ensure the proteasome composition was held constant.

1.0 x10^6 HEK293E cells expressing either the SP or IP were transduced with lentivirus to express SCD (mCherry reporter) and the antigen of interest (GFP reporter). After 24 hours lentivirus was removed and replaced with fresh media. After 3-7 days cells were FACS sorted to enrich for $SCD^+Antigen^+$ cells. One week later, sort efficiency was checked, and cells were re-sorted if purity was below 80%. Enriched cells were grown in Gibco Freestyle media supplemented with 10% fetal bovine serum, 1X L-glutamine, 1X pen-strep, and 1X non-essential amino acids until the culture reached a total volume 400 mL for HLA-$A24_{SCD,}$ and HLA-$A24_{SCD}$ or a total volume 800 mL HLA-

$A2_{SCD}$. Cells were then left to grow to maximum density and were harvested when viability dropped between 5-30%.

Artemis was then performed as previously described (*16*). Briefly, cells were spun down, supernatant was collected and filtered before being incubated with 200μL washed Ni-NTA beads for one hour at 37ºC. The slurry was allowed to drip through a gravity column and washed ten times with 5 mL 1X DPBS. Peptides were eluted in 600μL 5M Gu-HCL for 30 min before collection. Elution was repeated and fractions were pooled. Samples were desalted using C28 column (Brand) and run on the mass spectrometer.

Mass Spectrometry. Peptides were analyzed on a tribrid Orbitrap Fusion mass spectrometer (Thermo Fisher) As described in Finton et al. (*16*). Briefly, desalted peptides were either analyzed as-is or fractionated using a high-pH reverse phase chromatography with 8 fractions collected.

MS analyses were performed with a Thermo Scientific Easy-nLC 1000 nano-flow liquid chromatography system (Thermo Scientific) coupled to the Fusion mass spectrometer using a trap-and-column configuration. A three second cycle time was selected between master full scans in the Orbitrap mass analyzer and ions were selected for fragmentation. Selected ions were dynamically excluded for 15 seconds with an exclusion mass width of ±10 ppm to prevent repeatedly sequencing prominent peptides. Fusion data were analyzed using Proteome Discoverer 2.2 (Thermo Scientific) searching against a 2018 Uniprot human database that included common contaminants.

MS Data Analysis*. Peptide filtering*: Data from the MS was analyzed following the protocol previously described in Finton et al. (*16*). Peptides identified from the MS were filtered to remove 1) peptides derived from common contaminating source proteins, 2) fragmented peptides defined as having RPC retention times within 30 seconds of a longer, encompassing peptide, 3) non-human peptides derived from GFP and BSA, and 4) peptides shorter than 8 amino acids or longer than 16 amino acids. Peptides were then filtered to a 5% false discovery rate (FDR).

Removing duplicate peptides: For all peptide analyses other than total peptide counts, duplicate peptides were removed. When multiple biological replicates were aggregated, duplicate peptides between sets were also removed.

Binding Motifs: Sequence logos were generated from filtered, unique peptide lists as described in Finton et al. (*16*). Briefly each amino acid is represented by its single letter code. The frequency of a particular amino acid at each position was compared to the frequency of amino acids in the UNIPROT database. Amino acids deviating most from the reference frequency are depicted larger. Positive values indicate an amino acid that is enriched, negative values indicate an amino acid that is depleted. The sum of all symbols (positive and negative) is the relative entropy in bits, compared to the reference. Logos were created in SVG using the Python package palmotif (https://github.com/agartland/palmotif/).

Peptide binding predictions: Binding predictions were made with NetMHCpan, version 4.1, using the default settings on the web portal.

Venn Diagrams: Filtered, unique peptides lists were used to generate all Venn diagrams. For diagrams with <3 components the online BioVenn portal was employed. For diagrams with >3 components the Python package pyvenn was utilized (http://github.com/lankyCyril/pyvenn/).

Venn overlap: Venn overlap percentages comparing two datasets were calculated with the formula: overlap percentage=((2×A2)/(A1+(2×A2)+A3))×100 where A1 is the number of peptides in set 1 but not in set 2, A2 is the number of peptides in common between sets 1 and 2, and A3 is the number of peptides in set 2 but not in set 1. Calculating percent overlap in this standard manner skews the percentage lower when the two datasets being compared are very different sizes.

Down-sampling: Calculating the overlap between two data sets of very different sizes using the above calculation can artificially lower the resulting overlap. To minimize this when comparing SP and IP data sets, we needed to reduce the larger SP data set to match the IP counterparts. For each HLA_{SCD}, a subset of SP peptides was randomly selected from the aggregated list of unique peptides. The number of peptides selected was matched to the corresponding IP condition: HLA-$A2_{SCD}$ = 1,800 peptides; HLA_{SCD} = 6,100, HLA-$A11_{SCD}$ = 16,800.

Antigen Sequences. Nucleic acid sequences for antigen sequences were acquired from the NCBI data base. If necessary, base pairs were modified to remove the restriction enzyme sites for XhoI or BAMHI used for cloning. CCNA1 (NM_003914.4); WT1(NM_024426.6); For the Calreticulin mutant a 55 base pair deletion at exon 9 (*43*) was used to generate the frameshift mutation (NM_004343.4 55-bp del; c.1092_1143):

Acknowledgements.

We would like to thank Dr. Benoit Van den Eynde for providing the SP and IP restricted cells used in these experiments. Figure 3.1 was created with BioRender.com

References

1. T. Boon, P. G. Coulie, B. Van den Eynde, Tumor antigens recognized by T cells. *Immunol Today* **18**, 267-268 (1997).
2. K. L. Rock, E. Reits, J. Neefjes, Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends Immunol* **37**, 724-737 (2016).
3. C. M. Sylvester *et al.*, Early Childhood Behavioral Inhibition Predicts Cortical Thickness in Adulthood. *J Am Acad Child Adolesc Psychiatry* **55**, 122-129 e121 (2016).
4. A. G. Chapuis *et al.*, T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. *Nat Med* **25**, 1064-1072 (2019).
5. A. P. Rapoport *et al.*, NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. *Nat Med* **21**, 914-921 (2015).
6. K. L. Johnson, I. G. Ovsyannikova, C. J. Mason, H. R. Bergen, 3rd, G. A. Poland, Discovery of naturally processed and HLA-presented class I peptides from vaccinia virus infection using mass spectrometry for vaccine development. *Vaccine* **28**, 38-47 (2009).
7. P. Shafer, L. M. Kelly, V. Hoyos, Cancer Therapy With TCR-Engineered T Cells: Current Strategies, Challenges, and Prospects. *Front Immunol* **13**, 835762 (2022).
8. A. T. Teck *et al.*, Cancer testis antigen Cyclin A1 harbors several HLA-A*02:01-restricted T cell epitopes, which are presented and recognized in vivo. *Cancer Immunol Immunother* **69**, 1217-1227 (2020).
9. B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* **48**, W449-W454 (2020).
10. T. J. O'Donnell *et al.*, MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst* **7**, 129-132 e124 (2018).

11. A. Nelde, D. J. Kowalewski, S. Stevanovic, Purification and Identification of Naturally Presented MHC Class I and II Ligands. *Methods Mol Biol* **1988**, 123-136 (2019).
12. S. Sarkizova *et al.*, A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat Biotechnol* **38**, 199-209 (2020).
13. E. Barnea *et al.*, Analysis of endogenous peptides bound by soluble MHC class I molecules: a novel approach for identifying tumor-specific antigens. *Eur J Immunol* **32**, 213-222 (2002).
14. M. C. Lahman *et al.*, Targeting an alternate Wilms' tumor antigen 1 peptide bypasses immunoproteasome dependency. *Sci Transl Med* **14**, eabg8070 (2022).
15. M. Keller *et al.*, The proteasome immunosubunits, PA28 and ER-aminopeptidase 1 protect melanoma cells from efficient MART-126-35 -specific T-cell recognition. *Eur J Immunol* **45**, 3257-3268 (2015).
16. K. A. K. Finton *et al.*, ARTEMIS: A Novel Mass-Spec Platform for HLA-Restricted Self and Disease-Associated Peptide Discovery. *Front Immunol* **12**, 658372 (2021).
17. K. L. Rock *et al.*, Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* **78**, 761-771 (1994).
18. I. Evnouchidou, M. Weimershaus, L. Saveanu, P. van Endert, ERAP1-ERAP2 dimerization increases peptide-trimming efficiency. *J Immunol* **193**, 901-908 (2014).
19. A. Rouette *et al.*, Expression of immunoproteasome genes is regulated by cell-intrinsic and -extrinsic factors in human cancers. *Sci Rep* **6**, 34019 (2016).
20. R. M. Spaapen, J. Neefjes, Immuno-waste exposure and further management. *Nat Immunol* **13**, 109-111 (2012).
21. E. Z. Kincaid *et al.*, Mice completely lacking immunoproteasomes show major changes in antigen presentation. *Nat Immunol* **13**, 129-135 (2011).
22. R. E. Toes *et al.*, Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. *J Exp Med* **194**, 1-12 (2001).
23. M. B. Winter *et al.*, Immunoproteasome functions explained by divergence in cleavage specificity and regulation. *Elife* **6**, (2017).
24. E. S. Schultz *et al.*, The production of a new MAGE-3 peptide presented to cytolytic T lymphocytes by HLA-B40 requires the immunoproteasome. *J Exp Med* **195**, 391-399 (2002).
25. K. Gulukota, C. DeLisi, HLA allele selection for designing peptide vaccines. *Genet Anal* **13**, 81-86 (1996).
26. D. Gfeller *et al.*, The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. *J Immunol* **201**, 3705-3716 (2018).
27. A. Hearn, I. A. York, K. L. Rock, The specificity of trimming of MHC class I-presented peptides in the endoplasmic reticulum. *J Immunol* **183**, 5526-5536 (2009).

28. J. Abi Habib *et al.*, Efficiency of the four proteasome subtypes to degrade ubiquitinated or oxidized proteins. *Sci Rep* **10**, 15765 (2020).
29. M. Bai *et al.*, Assembly mechanisms of specialized core particles of the proteasome. *Biomolecules* **4**, 662-677 (2014).
30. S. Khan *et al.*, Immunoproteasomes largely replace constitutive proteasomes during an antiviral and antibacterial immune response in the liver. *J Immunol* **167**, 6859-6868 (2001).
31. E. Copreni, L. Palmieri, S. Castellani, M. Conese, A VSV-G Pseudotyped Last Generation Lentiviral Vector Mediates High Level and Persistent Gene Transfer in Models of Airway Epithelium In Vitro and In Vivo. *Viruses* **2**, 1577-1588 (2010).
32. D. Ritz *et al.*, High-sensitivity HLA class I peptidome analysis enables a precise definition of peptide motifs and the identification of peptides from cell lines and patients' sera. *Proteomics* **16**, 1570-1580 (2016).
33. J. Chapiro *et al.*, Destructive cleavage of antigenic peptides either by the immunoproteasome or by the standard proteasome results in differential antigen presentation. *J Immunol* **176**, 1053-1061 (2006).
34. B. Guillaume *et al.*, Two abundant proteasome subtypes that uniquely process some antigens presented by HLA class I molecules. *Proc Natl Acad Sci U S A* **107**, 18599-18604 (2010).
35. D. L. Tabb *et al.*, Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* **9**, 761-776 (2010).
36. U. K. Aryal *et al.*, A proteomic strategy for global analysis of plant protein complexes. *Plant Cell* **26**, 3867-3882 (2014).
37. K. Tanaka, The proteasome: overview of structure and functions. *Proc Jpn Acad Ser B Phys Biol Sci* **85**, 12-36 (2009).
38. K. Tanaka, T. Chiba, The proteasome: a protein-destroying machine. *Genes Cells* **3**, 499-510 (1998).
39. I. A. York *et al.*, The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8-9 residues. *Nat Immunol* **3**, 1177-1184 (2002).
40. E. Ruggiero *et al.*, CRISPR-based gene disruption and integration of high-avidity, WT1-specific T cell receptors improve antitumor T cell function. *Sci Transl Med* **14**, eabg8027 (2022).
41. I. Tawara *et al.*, Safety and persistence of WT1-specific T-cell receptor gene-transduced lymphocytes in patients with AML and MDS. *Blood* **130**, 1985-1994 (2017).
42. S. Ochsenreither *et al.*, Cyclin-A1 represents a new immunogenic targetable antigen expressed in acute myeloid leukemia stem cells with characteristics of a cancer-testis antigen. *Blood* **119**, 5492-5501 (2012).
43. D. Prins, A. R. Green, Mutant CALR functions: gains and losses. *Blood* **136**, 6-7 (2020).
44. J. D. Colbert, F. M. Cruz, K. L. Rock, Cross-presentation of exogenous antigens on MHC I molecules. *Curr Opin Immunol* **64**, 1-8 (2020).

45. L. Jing *et al.*, Prevalent and Diverse Intratumoral Oncoprotein-Specific CD8(+) T Cells within Polyomavirus-Driven Merkel Cell Carcinomas. *Cancer Immunol Res* **8**, 648-659 (2020).
46. R. A. van Amerongen *et al.*, WT1-specific TCRs directed against newly identified peptides install antitumor reactivity against acute myeloid leukemia and ovarian carcinoma. *J Immunother Cancer* **10**, (2022).
47. R. Narayan *et al.*, Acute myeloid leukemia immunopeptidome reveals HLA presentation of mutated nucleophosmin. *PLoS One* **14**, e0219547 (2019).
48. T. Azuma *et al.*, Identification of a novel WT1-derived peptide which induces human leucocyte antigen-A24-restricted anti-leukaemia cytotoxic T lymphocytes. *Br J Haematol* **116**, 601-603 (2002).
49. H. Ohminami, M. Yasukawa, S. Fujita, HLA class I-restricted lysis of leukemia cells by a CD8(+) cytotoxic T-lymphocyte clone specific for WT1 peptide. *Blood* **95**, 286-293 (2000).
50. S. Morita *et al.*, A phase I/II trial of a WT1 (Wilms' tumor gene) peptide vaccine in patients with solid malignancy: safety assessment based on the phase I data. *Jpn J Clin Oncol* **36**, 231-236 (2006).
51. D. A. Ferrington, D. S. Gregerson, Immunoproteasomes: structure, function, and antigen presentation. *Prog Mol Biol Transl Sci* **109**, 75-112 (2012).

Chapter 4- Conclusions and future directions

Acquired resistance is considered disease progression after a period of initial response; cancer with acquired resistance becomes more aggressive and more difficult to treat (*1*). This dissertation detailed a clinical case of acquired resistance that ultimately revealed immune evasion through changes in proteasome expression. This prompted a deeper investigation into the contribution of proteasome isoforms to the immunopeptidome and a search for 'mutual' peptides with therapeutic potential. We observed diverse mechanisms of immune evasion in this limited clinical trial, encompassing T cell dysfunction in the face of abundant antigen, loss/lack of targeted protein, and proteasome modulation. In addition to mechanisms of T cell immunotherapy resistance that have already been described, such as HLA loss (*2*), mechanisms observed in this trial underscore the necessity of additional research to fine tune therapeutic T cells and assemble an arsenal of tactics to bypass tumor escape. Contemporary, high throughput technologies like scRNAseq simultaneously profile immune-intrinsic and tumor-intrinsic evasion tactics to help us understand cancer-immune interactions and develop potent, resilient immunotherapies. For my graduate work, I have concentrated on a single, less characterized mechanism of tumor immune escape – modulating proteasome composition.

Proteasomes in TCR-T cell immunotherapy

As the primary source of class I HLA presented peptides, proteasomes are hypothesized to be central to immune recognition (*3-5*). Indeed, instances of immune evasion via changes in proteasome isoform expression have been reported in both melanoma (*6*) and AML (*7*). Likewise, multiple immunogenic peptides have been shown to rely on either IP (*7, 8*) or SP processing (*6, 9*), including the IP-specific $WT1_{126\text{-}134}$ peptide recognized by $T_{TCR\text{-}C4}$ in our clinical trial. Targeting 'mutual' peptides could limit this avenue of resistance. To date, shifts in proteasome expression have not been widely described as mechanisms of resistance to TCR-T cell therapy, possibly due to a limited number of patients treated with TCR-T cell therapies and the challenge of determining the proteasomal composition for each tumor or tumor sub-population. In the

example studied, the TCR that targeted $WT_{126\text{-}34}$, skewed the relapsed AML towards SP expression (*7*). Before employing Artemis, we isolated another HLA A2-restricted WT1-specific TCR targeting $WT1_{37\text{-}45}$ that was able to recognize both IP and SP cells as well as the patient's relapsed, $T_{TCR\text{-}C4}$ resistant AML, confirming $WT1_{37\text{-}45}$ is a 'mutual' peptide (*7*).

The method of using T cells raised against epitopes to interrogate peptide processing is labor intensive and partially relies on TCR biology. Using a streamlined methodology to identify those peptides upfront constitutes a clear advantage. I recovered the $WT_{37\text{-}45}$ peptide using Artemis on SP cells expressing HLA -$A2_{SCD}$ and transduced with the full of WT1, but the SP cells were exposed to INFγ and TNF-α preventing us from using Artemis to validate $WT_{37\text{-}45}$ as a 'mutual' peptide. We were able to isolate 'distinct' and 'mutual' peptides derived from WT1, Calreticulin and CCNA1 suggesting that Artemis is a viable strategy for 'mutual' peptide discovery – moving forward.

While shifts in proteasome composition are not currently established as a frequent mechanism of immune evasion, targeting 'mutual' peptides still offers therapeutic advantages. Targeting 'mutual' peptides could build more robust therapies armored against such immune evasion, and TCR-Ts recognizing 'mutual' peptides will likely be effective against a wider range of cancers expressing either IP, SP or a mixture. TCR-T cells targeting a 'mutual' peptide are not limited to treating tumors with specific proteasome expression, such as predominant IP expression in AML (*10*). Confirming antigen processing and presentation then becomes an important step in preclinical validation. Determining if a peptide is 'mutual' can be addressed in the initial stages of peptide selection (e.g. with Artemis) and/or later when confirming antigenicity in SP- and IP-restricted cells (e.g. with T cell recognition assays).

Furthering our understanding of the 'mutual' immunopeptidome

We observed substantially reduced immunopeptidome overlap between SP and IP cells compared to intra-proteasome overlaps for each HLA allele tested. This is consistent with previous murine studies where peptides eluted off splenocytes from wild type (presumed IP expressing) and IP-knockout (presumed SP expressing) mice

showed 50% overlap (*11*). These results likewise corroborate observed differences in proteasomal peptide processing both *in vitro* (*8, 9*) and *in vivo* (*6, 7*). We report for the first time, the extent of SP:IP overlap is specific for each HLA allele. This is not entirely surprising, given that each HLA subtype has different binding preferences (*12*).

Although reduced overlap between IP and SP cells matched previous literature (*11*), the mechanisms driving these differences remain unclear. It has been hypothesized that IPs preferentially cleave after certain residues, such as hydrophobic residues, that frequently serve as anchor residues to promote HLA binding and therefore could improve antigen presentation (*13*). Our analyses detected no difference in cleavage preferences. Isoform specific β-subunits share considerable homology – 62% between β1 and β1i, 59% between β2 and β2i and 71% between β5 and β5i (*14*). This yields proteasomes with highly similar crystal structures and may not explain unique catalytic properties. Still, empirical evidence (e.g., our observed inter-proteasome overlaps, the above murine studies, and clinical cases of immune evasion) support isoform-specific repertoires and recent studies point towards a quantitative difference rather than a qualitative difference (*14*), with the IP producing more peptides. While we did not test cleavage rates, we detected fewer peptides from IP cells which would not support this notion. However, slow cell growth kinetics could account for this discrepancy.

If catalytic differences are not driving IP:SP immunopeptidome divergence, what could account for the clearly distinct peptide repertoires? To prevent having a mixed proteasome composition in the SP cells, we did not expose cells to cytokines after the initial experiment when $WT1_{37\text{-}45}$ was detected. Instead, we used cell lines overexpressing IP subunits to induce IP formation. By not using IFN-γ to induce IP expression, which in turn modulates other IFN-γ responsive genes, we only modulated the proteasome and observed differences are likely due to specific isoforms. Perhaps the IP more strongly associates with a regulatory cap or another protein that influences protein degradation, similar to how β5i more strongly associates with UMP1 (a proteasome assembly chaperone) to drive faster IP formation (*15*). Co-immunoprecipitation of proteasomes could reveal if there are any other protein-protein interactions that could explain divergent immunopeptidomes.

Assets and challenges using Artemis for peptide discovery

Artemis captures peptides representative of the presented immunopeptidome with nearly identical metrics (e.g., HLA binding motifs, etc.) to peptides published in datasets from the widely accepted traditional immunoprecipitation methodology. Moreover, several antigen-derived peptides identified with Artemis have been previously described, implying that Artemis can be a tool for target peptide discovery. Of interest is the $WT1_{185\text{-}193}$ [FGPPPPSQA] peptide that existing algorithms predicted not to bind HLA-A2 but was found with Artemis as well as in an immunoprecipitation dataset (*16*). Conversely, several known peptides were not recovered from Artemis, including the clinically-targeted $WT1_{126\text{-}134}$. This could, in part, be due to stochastic sampling of a peptide with infrequent presentation.

Some Artemis-recovered peptides were already reported to be immunogenic, i.e., recognized by TCR-T cells that also killed cell lines expressing the targeted protein. This strongly suggests Artemis can be an effective tool for identifying clinically targetable peptides. Still, further validation of Artemis as a clinical peptide-discovery platform is required, including testing immunogenicity of the remaining peptides recovered using Artemis. To do this, T cells lines recognizing the Artemis-identified, antigen-derived peptides will be isolated and then cultured with SP or IP cells expressing both the restricting HLA and antigen. If T cells recognizing peptides classified as ‘distinct’ by Artemis are found to engage only IP or SP cells, it would confirm Artemis can accurately identify ‘mutual’ or ‘distinct’ peptides. The number of peptides eluted from the Artemis pipeline that induce recognition/killing by cognate TCRs will instruct the sensitivity of the technology.

The streamlined Artemis pipeline is not only poised to be a promising tool for target peptide discovery, but it may improve existing peptide prediction tools and deepen our biological understanding of how peptide processing impacts the immunopeptidome. Artemis could serve as an excellent method for generating large-scale training data sets for machine learning algorithms aimed at predicting if a peptide produced and also presented. Current programs are trained on peptides eluted from the cell surface (*17*), but do not account for variation in upstream processing. Some

programs can assess if a peptide is likely to be generated by the proteasome, but many lack information comparing proteasomes isoforms, are computationally intensive, have difficult user interfaces (*18*), and do not always account for downstream presentation. Using Artemis for training datasets does have limitations. Artemis cannot be used to elute peptides off primary tumor samples, requires large culture volumes, and prefers suspension lines capable of growing densely. Artemis also relies on MS for data acquisition which is costly.

Artemis offers a controlled model system to probe the underlying biology governing peptide processing and presentation. This includes the contribution of proteasome isoforms described here, but is also easily expanded to include intermediate proteasomes, and can support investigations of how other proteins like ERAP, TAP, or therapeutic inhibitors impact immunopeptidomes. Artemis serves as a streamlined tool ideally posed for therapeutic peptide discovery, production of large-scale training datasets, and biological investigations of peptide processing.

From Protein to Peptide: the TCR-T target peptide discovery pipeline

Effective TCR-T cell immunotherapy depends on strategic selection of tumor-associated proteins as well as processing and presentation of the targeted peptide. Ideal tumor-associated proteins are uniquely and broadly expressed across tumor cells with minimal expression in healthy tissue. Viral antigens and tumor-specific mutations are clear examples; over expressed self-antigens can also fall into this category if expression is limited on healthy tissue. Quality targets likewise promote tumor proliferation and oncogenicity such that loss is unfavorable in tumor progression. Concurrent with tumor antigen selection, is selection of the presenting HLA allele, which is dependent on the target patient population Once a tumor-associated protein and presenting HLA are selected, the target peptide must be identified. Artemis serves as an excellent platform for such peptide discovery. A SCD of the presenting HLA would be co-transduced with the antigen of interest into SP or IP cells and run through the Artemis pipeline ($n \geq 2$). Resulting antigen-derived peptides would then be assessed for immunogenicity. Preference would be given to ‘mutual’ peptides. Those that proved capable of stimulating T cells would be candidates for targeting with TCR-T cell therapy

and further validated for endogenous presentation. Meanwhile, a high affinity TCR from the peptide-reactive T cells could be isolated and moved through rigorous safety studies. Following the initial discovery steps, the candidate TCR would be subjected to extensive preclinical tests and external reviews before being translated to the clinic.

In Sum

Clinical observations from cases of acquired resistance provide valuable mechanistic insights into driving factors, such as proteasomal processing of target peptides, to help guide future iterations of therapy. The biochemical mechanisms dictating IP and SP immunopeptidome differences remain obscure. However, empirical evidence shows that the proteasome isoform present can result in differential recognition by T cells and should be considered when selecting TCR-T cell therapy target peptides. Artemis is an excellent tool for identifying 'mutual', immunogenic peptides but future validation experiments are required. Furthermore, Artemis is a promising high-throughput pipeline that could yield large datasets to train machine learning algorithms for peptide processing and presentation predictions.

References

1. D. S. Chen, I. Mellman, Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321-330 (2017).
2. K. G. Paulson *et al.*, Acquired cancer resistance to combination immunotherapy from transcriptional loss of class I HLA. *Nat Commun* **9**, 3868 (2018).
3. B. J. Van den Eynde, T. Boon, Tumor antigens recognized by T lymphocytes. *Int J Clin Lab Res* **27**, 81-86 (1997).
4. B. J. Van den Eynde, S. Morel, Differential processing of class-I-restricted epitopes by the standard proteasome and the immunoproteasome. *Curr Opin Immunol* **13**, 147-153 (2001).
5. N. Vigneron, J. Abi Habib, B. J. Van den Eynde, Learning from the Proteasome How To Fine-Tune Cancer Immunotherapy. *Trends Cancer* **3**, 726-741 (2017).
6. M. Keller *et al.*, The proteasome immunosubunits, PA28 and ER-aminopeptidase 1 protect melanoma cells from efficient MART-126-35 -specific T-cell recognition. *Eur J Immunol* **45**, 3257-3268 (2015).
7. M. C. Lahman *et al.*, Targeting an alternate Wilms' tumor antigen 1 peptide bypasses immunoproteasome dependency. *Sci Transl Med* **14**, eabg8070 (2022).
8. A. Jaigirdar, S. A. Rosenberg, M. Parkhurst, A High-avidity WT1-reactive T-Cell Receptor Mediates Recognition of Peptide and Processed Antigen but not

Naturally Occurring WT1-positive Tumor Cells. *J Immunother* **39**, 105-116 (2016).
9. S. Morel *et al.*, Processing of some antigens by the standard proteasome but not by the immunoproteasome results in poor presentation by dendritic cells. *Immunity* **12**, 107-117 (2000).
10. A. Rouette *et al.*, Expression of immunoproteasome genes is regulated by cell-intrinsic and -extrinsic factors in human cancers. *Sci Rep* **6**, 34019 (2016).
11. E. Z. Kincaid *et al.*, Mice completely lacking immunoproteasomes show major changes in antigen presentation. *Nat Immunol* **13**, 129-135 (2011).
12. D. Gfeller *et al.*, The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. *J Immunol* **201**, 3705-3716 (2018).
13. R. E. Toes *et al.*, Discrete cleavage motifs of constitutive and immunoproteasomes revealed by quantitative analysis of cleavage products. *J Exp Med* **194**, 1-12 (2001).
14. M. B. Winter *et al.*, Immunoproteasome functions explained by divergence in cleavage specificity and regulation. *Elife* **6**, (2017).
15. R. S. Marshall, R. D. Vierstra, Dynamic Regulation of the 26S Proteasome: From Synthesis to Degradation. *Front Mol Biosci* **6**, 40 (2019).
16. R. A. van Amerongen *et al.*, WT1-specific TCRs directed against newly identified peptides install antitumor reactivity against acute myeloid leukemia and ovarian carcinoma. *J Immunother Cancer* **10**, (2022).
17. B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* **48**, W449-W454 (2020).
18. B. R. Weeder, M. A. Wood, E. Li, A. Nellore, R. F. Thompson, pepsickle rapidly and accurately predicts proteasomal cleavage sites for improved neoantigen identification. *Bioinformatics*, (2021).

Appendix I: Quality is King: Fundamental Insights into Tumor Antigenicity from Virus-Associated Merkel Cell Carcinoma

This chapter has been adapted from Lahman, M.C., et al. *Quality is King: Fundamental Insights into Tumor Antigenicity from Virus-Associated Merkel Cell Carcinoma.* J. Invest. Derm. 2021 PMID: 33863500

Abstract

Merkel cell carcinoma (MCC) is a rare skin malignancy that is a paradigm cancer for solid tumor immunotherapy. MCCs associated with Merkel cell polyomavirus (virus-positive; VP-MCC) or chronic UV exposure (virus-negative; VN-MCC) are anti-PD(L)1 responsive, despite VP-MCC's low mutational burden. This suggests antigen quality, not merely mutation quantity, dictates immunotherapy responsiveness, and cell-based therapies targeting optimal antigens may be effective. Despite VP-MCC's antigenic homogeneity, diverse T cell infiltration patterns are observed, implying microenvironment plasticity and multifactorial contributions to immune recognition. Moreover, VP-MCC exemplifies how anti-tumor adaptive immunity can provide tumor burden biomarkers for early detection and disease monitoring.

Introduction

Merkel cell carcinoma (MCC) is a rare skin cancer with increasing incidence. It disproportionally affects elderly and immunocompromised persons, indicating immune suppression favors disease progression (*1, 2*). Risk factors include ultraviolet (UV) light exposure and fair skin, similar to other skin malignancies and consistent with UV-derived mutations driving oncogenesis (*1, 3, 4*). In 2008, clonal integration of the Merkel cell polyomavirus (MCPyV) revealed a second oncogenic driver, (*5*) sub-dividing MCC into virus positive (VP-MCC; ~80% of cases in U.S.A), and UV-associated MCCs (virus-negative; VN-MCC) (**Figure AI.1**) (*5-7*). VP-MCCs are driven by obligate expression of conserved MCPyV-derived oncoproteins, large and small-T antigen, with few additional genomic mutations (*8-10*). This makes VP-MCC genetically quite similar across patients

and considerably more homogeneous than most cancers. Uniform (VP-MCC) and diverse (VN-MCC) variants of the same cancer make MCC a natural model for studying cancer-immune interactions.

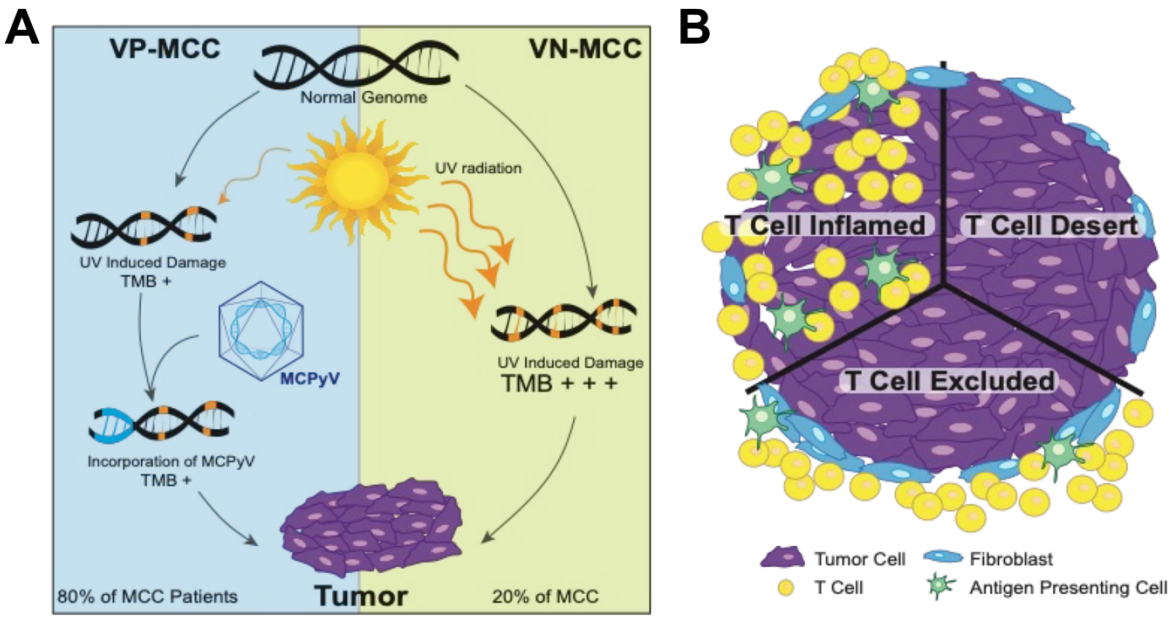


Figure AI.1 Merkel Cell Carcinoma (MCC) pathogenetic pathways and immune infiltrate patterns. (A) Schematic of MCC dual pathways to pathogenesis. MCC polyomavirus drives oncogenesis in virus-positive MCC **(**VP-MCC; left, blue). VP-MCC has a low median tumor mutation burden (TMB) of 1.2 mutations/MB. Many UV mutations are present in virus negative MCC (VN-MCC; right, green); corresponding to a very high median TMB of 63.1 mutations/MB. **(B)** Illustrations of the three major cancer immune infiltrate patterns, all observed in MCC and partially predictive of anti PD-1 response. T cell Desert-no observed T cells in the tumor. T cell Excluded-T cells restricted to the tumor periphery. T cell Inflamed-T cells infiltrate the tumor.

Immune checkpoint blockade (ICB) (*11, 12*) of the PD-1 axis has shown widespread clinical benefit across cancer types (*13-17*). ICB improves outcomes for approximately 50% of MCC patients – one of the highest solid tumor response rates (*18, 19*) – and NCCN guidelines now recommend ICB therapy as first-line treatment for metastatic MCC (*20*). MCC has a relatively even distribution of ICB responders (30-50%), non-responders (25-30%), and escapees (25-30%) (*14, 21*) across both pathogenic pathways, again rendering MCC a prototype cancer.

In this review, we use MCC as a paradigm cancer for understanding ICB response/resistance. We first explore how MCC – with dual pathogenic pathways and

even distribution of responders, non-responders and escapees – can be used to model molecular mechanisms by assessing epitope quality and immune cell-infiltrate patterns. We then discuss how cell-based immunotherapies for MCC provide a pertinent model for refining immunotherapy development. Lastly, with MCC as an example, we consider strategies utilizing immune cells for disease detection.

Highly immunogenic tumor-specific epitopes facilitate ICB response

While ICB has revolutionized cancer therapy, outcomes vary and, currently, cannot be accurately predicted. For example, higher expression of PD-1 ligand, CD274 – commonly denoted PD-L1, on tumor or tumor-associated immune cells correlates with improved response but does not reliably separate responders from non-responders (*15*). For VP- MCC, VN-MCC, and other immunogenic solid tumors (e.g. HIV-associated Kaposi Sarcoma; HIV-KS), post-ICB remissions are observed even with undetectable PD-L1 expression. A more widely used predictor of ICB response is median tumor mutational burden (TMB), which positively correlates with ICB response (*16*). Not all cancers follow this correlation emphasizing the need to precisely understand mechanisms governing response/resistance so that accurate predictors can be identified (*13, 22*). Deciphering immune interactions is challenging as tumor mutations vary between tumors, and different T cell clones can recognize identical mutations when presented by human leucocyte antigen [HLA]). Uniform (VP-MCC) and diverse (VN-MCC) variants provide a uniquely suitable platform in which to investigate this heterogeneity in the context of both high and low TMB.

Tumor-specific antigen <u>quality</u> is associated with response to PD-1 axis blockade

The positive correlation between TMB and ICB response implies heavily mutated tumors are more immunogenic and subsequently more likely to respond (*16, 23-26*). The molecular basis of this observation remains unclear, particularly with regard to a tumor's immunopeptidome, i.e., the sum of presented epitopes (intracellularly derived peptides presented on HLA) recognized by cytotoxic T cells that subsequently mount an immune response. Is this correlation between TMB and ICB response driven by TMB-high tumors having numerous, minimally immunogenic epitopes or a few, highly

immunogenic epitopes? MCC's dual etiology in conjunction with other virus-associated cancers provide an excellent opportunity to address this question and suggest that epitope quality is consequential.

VN-MCC has a high TMB (median 63.1 mutations/Megabase (MB); range 31-133), consistent with other UV-associated skin cancers.(*9*) Conversely, VP-MCC has an exceptionally low TMB (median 1.2 mutations/MB; range 0.0-2.6) and few other genomic aberrations (*9, 10, 27*). VN-MCC and VP-MCC have similar overall response rates (ORR) (44-54%) to anti-PD-1 blockade. For VN-MCC, this high ORR is predicted by median TMB (*10, 16, 28-31*). VP-MCC has a substantially better ORR than TMB alone would predict (*9, 10, 14, 32*). While VP-MCC has low TMB and subsequently few predicted neo-epitopes, it does contain MCPyV viral proteins (*33, 34*) that generate highly immunogenic epitopes. Its viral protein sequence is foreign and thus MCPyV-specific T cells recognize its HLA-presented epitopes generally with higher avidity, or ability to efficiently kill VP-MCC cells, as these T cells have not undergone thymic negative selection (*35*). Furthermore, the MCPyV T antigen locus codes for proteins under 900 amino acids in length (*36*). This small, finite, viral epitope repertoire and high ORR suggests limited but highly immunogenic epitopes may be sufficient for favorable responses.

To explore if tumor-specific antigen quality influenced ICB ORRs, we analyzed virus and non-virus-associated cancers separately. As a starting point, Dr. Mark Yarchoan kindly shared his published data set comparing PD-1 ORR and TMB across several cancers (*16*). Five virus-associated cancers were identified – VP-MCC, HIV-KS (HHV-8 driven), virus-positive head and neck (VP-H&N), cervical, and anal cancers (latter three HPV driven). The specific etiology was not available for the HIV-KS, cervical or anal cancers; these are >90% virus-driven (*37-39*). However, a skewing of the associated response rates could not be formally excluded. Since virus-associated etiology was not available for VP-MCC and VP-H&N in the Yarchoan dataset, these data were removed and replaced with data from studies that explicitly divided patients based on cancer etiology (*16, 28, 40-42*). Log transformations, linear regressions and tests of linear combinations of regression coefficients were performed using R.

When virus-associated cancers were removed from the data set leaving only non-virus-associated cancers, the correlation between TMB and ORR increased (0.488 to 0.849; n = 25). On the other hand, the virus-associated cancers (n = 5) showed a negative correlation (R^2 = -0.619), suggesting that TMB may not be an adequate predictor for all cancers (**Figure AI.2**). Thus, separating virus-positive and virus-negative cancers will be critical moving forward in deepening our understanding of TMB and ORR.

All virus-associated cancers analyzed exhibited higher ORR than TMB alone might predict. For the same TMB, virus-associated cancers have an average of 25.3% higher ORR than non-virus cancers (95% CI: 2.4-48.2; p = 0.03). Non-virus-associated cancers require 30-times greater TMB to elicit responses similar to those in virus-associated cancers (*10, 16, 28, 40-42*). On average, non-virus-associated cancers (n = 25) had a TMB of 11.6 and ORR of 17.4 while virus-associated cancers (n = 5) had a TMB of 5.7 and ORR of 39.4. While it cannot be deduced that multiple less immunogenic neoepitopes could also be effective in tumors with high TMB, this data suggests that a small number of highly immunogenic epitopes can elicit ICB responses, relative to a large number of less immunogenic epitopes.

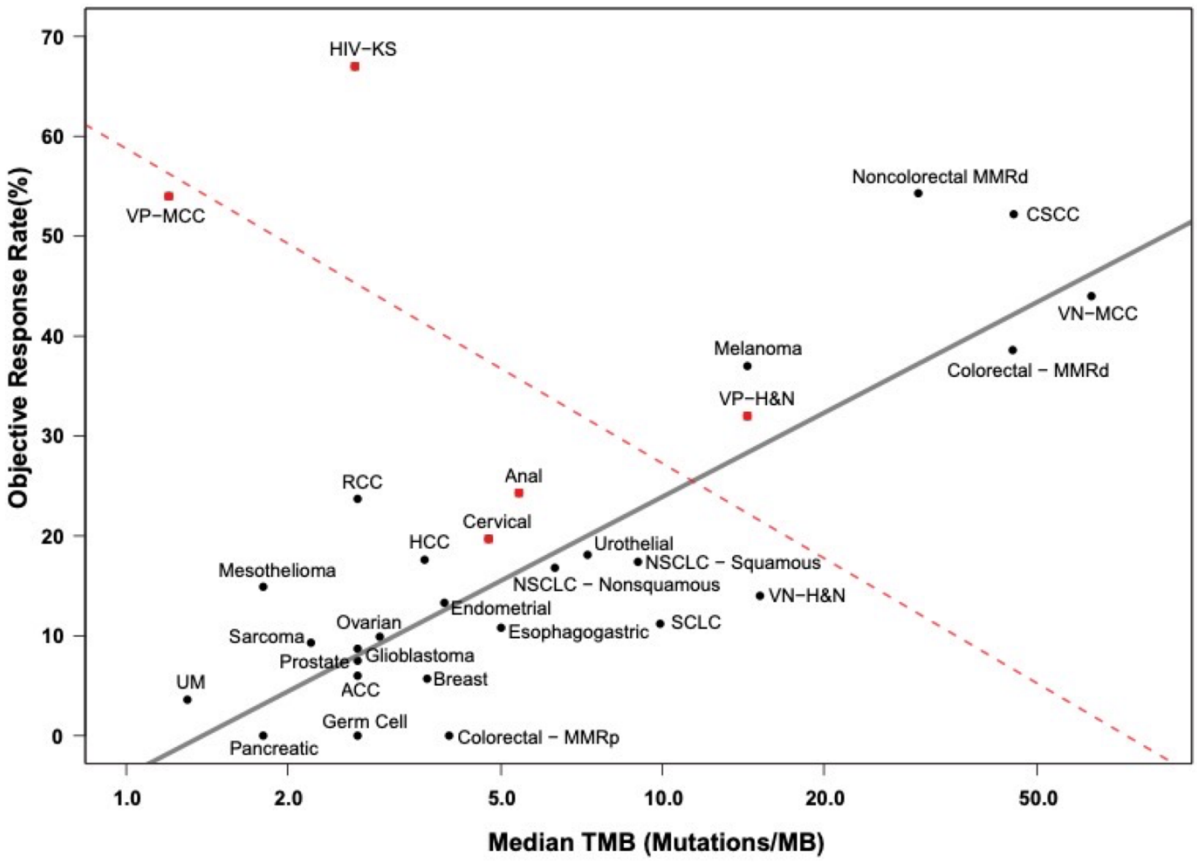


Figure AI.2 Correlation between overall response rate (ORR) to immune checkpoint blockade (ICB) therapy and median tumor mutational burden (TMB) for non-virus and virus-associated cancers. High TMB positively correlates with ICB ORR in non-virus-associated cancers (black dots, bold line) $R^2 = 0.849$ (n = 25) but negatively correlates in virus-associated cancers (Red squares, dotted line) $R^2 = -0.619$ (n = 5) A difference of 17.8 between these two slopes (95% CI: (10.2, 25.6); $p < 0.001$) indicates the association between ORR and TMB is significantly different in virus-associated cancers compared to non-virus cancers. Note that all virus-associated cancers have higher ORR than TMB would predict and generally higher ORR than the non-virus cancer counterparts with the same TMB. Virus Positive Merkel cell carcinoma (VP-MCC); Virus Negative Merkel Cell Carcinoma (VN-MCC); Human Immunodeficiency Virus associated Kaposi Sarcoma (HIV-KS); Renal Cell Carcinoma (RCC); Hepatocellular Carcinoma (HCC); Adrenocortical Carcinoma (ACC); Uveal Melanoma (UM); Non-Small Cell Lung Cancer (NSCLC); Small Cell Lung Cancer (SCLC); Virus-Positive Head and Neck (VP-H&N); Virus-Negative Head and Neck (VN-H&N); Cutaneous Squamous Cell Carcinoma (CSCC), Mismatch Repair Deficient (MMRd), Mismatch Repair Proficient (MMRp).

Oligoclonal expanded TIL suggest limited quantity of immunogenic epitopes

In melanoma, a virus-negative cancer with high TMB, very few tumor-derived epitopes are neoantigens (<0.1%) or unique tumor-associated antigens (1.7-3%). Analysis of tumor-infiltrating lymphocytes (TIL) from these same melanomas showed expansion of T cell clones specific to the few neo-epitopes/unique tumor-associated epitopes, again implying limited but potent epitopes can drive T cell responses (*43*). Furthermore, high-throughput genomic and proteomic methods have shown relatively few of the total possible neo-epitopes can mediate tumor rejection (*44, 45*). Identifying these 'needle in the haystack' epitopes driving immunogenic tumor rejection will be critical for developing broadly applicable vaccination and/or engineered cell therapies. VP-MCC offers a unique advantage for piloting such therapies due to a small antigenic space (MCPyV T antigens) to test approaches and shared, tumor-specific antigens across patients assure broad applicability.

Tumor-specific T cell response patterns are plastic and independent of tumor-associated antigen expression

Similar to TMB, TIL patterns can partially predict ICB response.(*25, 46, 47*) Three distinct patterns described by Mellman and others are observed in MCC (*48*) and other solid tumors (*46*) (**Figure AI.1B**): 1) T cell desert – no T cells detected in or near tumor. These tumors rarely respond to ICB and are common in relapsed metastases; (*9, 46, 49*) 2) T cell exclusion – T cells concentrate on edges but fail to infiltrate the tumor due to a physical and/or chemical barrier. These tumors sometimes respond to ICB therapy; (*9, 49, 50*) 3) T cell inflamed – T cells observed throughout the tumor. This phenotype is frequently seen during response and is typically associated with improved prognosis (*15, 46, 47, 49*).

Determining whether TIL patterns are a consequence of the microenvironment or variations of tumor epitope immunogenicity is exceptionally challenging. Indeed, large-scale investigations of these features have been unfruitful (*51, 52*). Similarly, allelic diversity in presentation molecules, particularly class I HLA, can influence immunogenicity, further complicating studies (*53-55*). The field does recognize that immunogenic antigens (e.g. viral proteins) play a critical role in activating lymphocytes

and recruiting T cells to tumors (*56, 57*). Here, VP-MCC provides an informative natural system in which obligate viral oncoproteins and low TMB imply relatively homogeneous immunogenicity. Therefore, in this system, TIL pattern differences are not directly related to the tumor antigen itself (MCPyV) but instead associated with microenvironment fluctuations or HLA polymorphisms rendering the same antigen (MCPyV) more or less immunogenic. Surprisingly, all three TIL patterns have been recorded in the same VP-MCC patient at different stages of response and resistance to T cell immunotherapy, (*49*) suggesting complex microenvironment factors dictate T cell infiltration. Moving forward, sophisticated systems analyses will be necessary to precisely identify which factors drive or inhibit T cell infiltration.

MCC serves as a model cancer in which to understand factors mediating immunotherapy response and resistance

Many patients' tumors progress after starting ICB treatment (primary resistance) or after a period of initial response (acquired resistance) (*58*). For MCC, ~30% of initial responders progress by 2 years post ICB initiation (*32*).

Primary resistance is generally a direct consequence of a patient's T cells being unable to respond to tumor, rendering ICB ineffective (*58, 59*). This may be driven by a number of factors: tumor lacking target antigen(s) (*60*), dysfunctional antigen presentation (*61, 62*), advanced/irreversible tumor-specific T cell dysfunction, and/or lack of tumor-specific T cells. Primary-resistant tumors presumably experience little immunogenic pressure from ICB, making the tumor unlikely to undergo major shifts; therefore, primary resistant tumors are not often utilized in studying escape mechanisms. Tumors with initial response followed by acquired resistance offer an opportunity to rigorously examine mechanisms of resistance. Identifying the root cause of acquired resistance remains challenging, as multiple contributors can produce the same clinical phenotype.

Immune escape mechanisms fit into two broad categories immune-cell intrinsic or tumor intrinsic. Immune-cell intrinsic mechanisms include changes to innate immune cells such as reduced neutrophil chemotaxis (*63*) and toll-like receptor 9 (TLR9) down regulation (*64*), and changes to adaptive immune cells such as T- cell exhaustion (PD-1, CTLA-4, TIM3) (*65*), and T-cell migration/infiltration patterns (*66-70*). Tumor intrinsic

changes include HLA loss (*49, 71, 72*) and proliferation of a pre-existing, resistant clone (*58*).

Precisely deciphering if resistance is immune-cell intrinsic or tumor intrinsic is complex (*58*). It requires simultaneously knowing which T cells can 'see' tumor (immune-cell intrinsic factor, T cell receptor (TCR) clonotypes) and what is recognized by T cells (tumor intrinsic factor, antigen/epitope). Often it is difficult to know both, given that discovery is co-dependent, i.e., identifying a tumor-specific T cell clone requires knowing which HLA-presented epitope is recognized, and vice versa. Again, VP-MCC homogeneity provides an exceptional model system.

Shared oncoprotein in VP-MCC facilitates tracking of tumor-specific T cells

In VP-MCC, continuous expression of a known, shared viral oncoprotein MCPyV Large-T antigen (LT-Ag) facilitates identification of LT-Ag specific T cells, using fluorophore-labeled epitope:HLA tetramers (*73-75*). Tetramers tiled across LT-Ag discern MCC-specific T cells from other, quiescent T cells and can reveal epitope spread; i.e., nascent immune responses against epitopes not targeted by initial immunotherapy (*76*). The phenotype or TIL pattern of tumor-specific T cells can be further examined to inform molecular drivers of tumor-immune interactions and, ultimately, immunotherapy resistance.

Clinical trials using adoptively transferred MCPyV-specific T cells in VP-MCC (*49, 77*) provide an opportunity to directly study cancer-immune interactions with easily identifiable T cells recognizing a defined epitope. Briefly, T cells are isolated from patients' peripheral blood, MCC-specific cells are expanded *ex vivo* against a known epitope (e.g. LT-$Ag_{92\text{-}101}$), purified by specific tetramer-binding and re-infused into the patient (*49, 77*). Transferred, polyclonal MCC-specific T cells all recognize the same epitope (e.g. LT- $Ag_{92\text{-}101}$). Persistence is monitored with LT- $Ag_{92\text{-}101}$-tetramers and additional markers, including PD-1, can identify functional phenotypes. Similarly, tetramers tiled across LT-Ag can elucidate antigen spread. Responding patients who received MCC-specific T cells and avelumab (anti-PD-L1) showed expansion of both MCC-Specific T cells and T cells recognizing other LT-Ag epitopes (*76*). Adoptive T cell trials provide valuable details characterizing tumor-specific T cells throughout ICB

therapy, insights into T cell-intrinsic or tumor-specific resistance mechanisms,(*49*) and novel epitopes for adoptive T cell therapy (*76, 78, 79*).

MCC cell-based immunotherapies offer insight to cancer-immune interactions

Cell-based therapies facilitate identifying, tracking, and profiling of tumor-specific immune cells. Like adoptive T cell transfer, TCR-transduced T cells (TCR-T) uniformly recognize a defined epitope:HLA complex. Multiple MCPyV epitopes have been identified and many are being considered as TCR-T targets for VP-MCC patients, including in combination therapy trials (NCT03747484) (*34, 80*). Unfortunately, therapies targeting MCPyV oncoproteins are only applicable to VP-MCC patients. To broaden the clinical scope, alternate targets are being considered, including cancer testes antigens, (*81*) and aberrantly glycosylated proteins (e.g. MUC-1)(*82*). Targeting surface proteins with chimeric antigen receptor T cells (CAR-T) recognizing glypican-3 (GPC3) is also being pursued in a Phase I clinical trial. Other immune cells can likewise be therapeutic. An ongoing phase 2 clinical trial uses activated natural killer cells (aNK) and ALT-803 (IL-15) to treat MCC patients (NCT02465957). Cell-based therapies offer naturally controlled systems with unprecedented opportunities to observe response/resistance mechanisms, cancer-immune interactions, and tumor microenvironments *in vivo*. Lessons learned will likely be broadly applicable, offering insights for next generation immunotherapies.

Acquired resistance mechanisms provide rationale for therapeutic combinations

Tumor cells can adapt to escape immune pressure. In MCC, loss or reduction of class I HLA correlates with worse prognosis, independent of immunotherapy intervention (*49, 83*). Increased prevalence of cytotoxic CD8+ T cells which recognize specific epitope:HLA complexes (*84*) is a positive prognostic indicator, (*85-87*) suggesting CD8+ T cells play a prominent role in disease clearance. Of note, many viruses evolved specific mechanisms to evade immune recognition, including down-regulation of class I HLAs (*88, 89*). Evasion mechanisms employed by virus-associated cancers could differ and some may not apply to all cancer types but understanding these evasion mechanisms will undoubtedly inform cancer-immune interactions more

broadly. Helper CD4+ T cells are also emerging as integral to coordinating effective immune responses, including by producing cytokines that upregulate class I HLA (*90-93*). Indeed, engineered CD4+ TCR-T cells are being considered as therapeutics (*94*). MCC can effectively dampen CD8+ T cell recognition by decreasing class I diversity or expression, making CD4+ T cell employment a reasonable strategy to drive immune response in tumors with limited class I (*49, 83, 95*).

PD-1/PD-L1 ICB-resistant tumors often show increased expression of other negative immune regulators, including Tim-3, Lag3, CTLA-4 and TIGIT, suggesting that inhibiting more than one immune regulator could improve responses and/or expand the patient population benefiting from ICB (*96-99*). Indeed, combination ICB therapy can improve clinical response (*76, 77, 100*). Another combination therapy trial (NCT02584829) is testing if dual or triple ICB treatment regimens combined with cell-based immunotherapies are synergistic in MCC.

ICB therapies place selective pressure on cancers with pre-existing genomic instability and heterogeneity, which can lead to resistance/relapse. Understanding these escape mechanisms is essential to broadening immunotherapy benefits (*49, 101, 102*). Targeting multiple resistance mechanisms will likely be necessary to improve outcomes and minimize late immunotherapy escape (*58*). Engineered cell therapies in VP-MCC offer remarkable cancer and immune cell homogeneity. Together they provide an unprecedented opportunity to clarify underlying mechanisms and to validate advanced therapeutic approaches.

Utilizing the immune system for more than therapy

Our immune system amplifies signals, a characteristic which can be used to monitor disease (*103, 104*). For example, antibody titers are commonly used in clinics to determine immune status, (*105*) and could be extended to oncology for tracking tumor burden or B/T cell clonal expansion.

MCC demonstrates tumor antigenicity that can be exploited for disease monitoring

Foreign proteins derived from MCPyV, and certain other cancer-associated antigens can induce antibody production that is measurable in peripheral blood (*103,*

106). Circulating MCPyV T antigen-specific antibodies are rare in healthy individuals but present in most VP-MCC patients. Oncoprotein-specific antibody titers fall during clinical response and spike prior to relapse, offering an easy, affordable, non-invasive, and predictive method for monitoring disease which is now included in MCC management guidelines (*20, 104*). Similarly, in non-small cell lung cancer, circulating tumor-specific antibodies non-invasively inform disease state by differentiating benign and malignant lung nodules (*107*). Changes in antibody titers alert physicians prior to clinical relapse, guiding early intervention for optimal patient outcomes (*56, 104*).

T cells and B cells can likewise be employed to detect tumor burden and direct treatment (*108-113*). Variable regions of B cell receptors (BCR) or TCRs serve as a natural barcode generated in B/T cell development. In B-cell and T-cell malignancies, BCR or TCR clonotype sequencing is used to monitor tumor progression via expansion/contraction of a cancerous lymphocytic clone (*108, 112*) and to detect minimal residual disease (*108, 113*). These same sequencing methods track TCRs for monitoring persistence and efficacy of engineered T cell therapies (*109, 111*).

Developing antibody arrays or immune cell monitoring for additional cancers could promote early detection, direct therapy, and improve patient outcomes. Employing these methodologies to evaluate hard-to-screen cancers could replace current, highly invasive procedures and costly imaging.

Conclusions

ICB therapy success has greatly improved outcomes for patients with many previously lethal malignancies, yet many obstacles remain. Analyses revealed an unmet need to accurately predict response, which requires a deeper understanding of mechanisms governing response, resistance and relapse. The MCC model suggests cancer epitope quality is associated with successful ICB response. Diverse immune infiltrate patterns demonstrate microenvironment plasticity in MCC and underscore a need for predictive diagnostics. Broadening the range of validated immunotherapeutic targets will be imperative to reducing late therapy escape, as demonstrated by improved response with combination regimens in MCC. Using MCC as a paradigm cancer not

only broadly informs molecular immunology and cancer immunotherapy, but also supports using endogenous immune responses to monitor disease and direct treatment.

References

1. J. Albores-Saavedra *et al.*, Merkel cell carcinoma demographics, morphology, and survival based on 3870 cases: a population based study. *J Cutan Pathol* **37**, 20-27 (2010).
2. I. Penn, M. R. First, Merkel's cell carcinoma in organ recipients: report of 41 cases. *Transplantation* **68**, 1717-1721 (1999).
3. M. Heath *et al.*, Clinical characteristics of Merkel cell carcinoma at diagnosis in 195 patients: the AEIOU features. *J Am Acad Dermatol* **58**, 375-381 (2008).
4. S. Q. Wong *et al.*, UV-Associated Mutations Underlie the Etiology of MCV-Negative Merkel Cell Carcinomas. *Cancer Res* **75**, 5228-5234 (2015).
5. H. Feng, M. Shuda, Y. Chang, P. S. Moore, Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* **319**, 1096-1100 (2008).
6. J. C. Becker *et al.*, MC polyomavirus is frequently present in Merkel cell carcinoma of European patients. *J Invest Dermatol* **129**, 248-250 (2009).
7. A. Kassem *et al.*, Frequent detection of Merkel cell polyomavirus in human Merkel cell carcinomas and identification of a unique deletion in the VP1 gene. *Cancer Res* **68**, 5009-5013 (2008).
8. R. Houben *et al.*, Merkel cell polyomavirus-infected Merkel cell carcinoma cells require expression of viral T antigens. *J Virol* **84**, 7064-7072 (2010).
9. T. C. Knepper *et al.*, The Genomic Landscape of Merkel Cell Carcinoma and Clinicogenomic Biomarkers of Response to Immune Checkpoint Inhibitor Therapy. *Clin Cancer Res* **25**, 5961-5971 (2019).
10. G. Goh *et al.*, Mutational landscape of MCPyV-positive and MCPyV-negative Merkel cell carcinomas with implications for immunotherapy. *Oncotarget* **7**, 3403-3415 (2016).
11. D. R. Leach, M. F. Krummel, J. P. Allison, Enhancement of antitumor immunity by CTLA-4 blockade. *Science* **271**, 1734-1736 (1996).
12. Y. Ishida, Y. Agata, K. Shibahara, T. Honjo, Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J* **11**, 3887-3895 (1992).
13. N. Galanina, A. M. Goodman, P. R. Cohen, G. M. Frampton, R. Kurzrock, Successful Treatment of HIV-Associated Kaposi Sarcoma with Immune Checkpoint Blockade. *Cancer Immunol Res* **6**, 1129-1135 (2018).
14. P. T. Nghiem *et al.*, PD-1 Blockade with Pembrolizumab in Advanced Merkel-Cell Carcinoma. *N Engl J Med* **374**, 2542-2552 (2016).
15. R. S. Herbst *et al.*, Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563-567 (2014).
16. M. Yarchoan, A. Hopkins, E. M. Jaffee, Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N Engl J Med* **377**, 2500-2501 (2017).
17. J. R. Brahmer *et al.*, Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* **366**, 2455-2465 (2012).

18. K. G. Paulson, M. Lahman, A. G. Chapuis, I. Brownell, Immunotherapy for skin cancer. *Int Immunol*, (2019).
19. J. C. Becker *et al.*, Merkel cell carcinoma. *Nat Rev Dis Primers* **3**, 17077 (2017).
20. C. K. Bichakjian *et al.*, Merkel Cell Carcinoma, Version 1.2018, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* **16**, 742-774 (2018).
21. H. L. Kaufman *et al.*, Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial. *Lancet Oncol* **17**, 1374-1385 (2016).
22. E. J. Lipson *et al.*, PD-L1 expression in the Merkel cell carcinoma microenvironment: association with inflammation, Merkel cell polyomavirus and overall survival. *Cancer Immunol Res* **1**, 54-63 (2013).
23. D. T. Le *et al.*, Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357**, 409-413 (2017).
24. D. T. Le *et al.*, PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* **372**, 2509-2520 (2015).
25. R. Cristescu *et al.*, Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* **362**, (2018).
26. N. A. Rizvi *et al.*, Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124-128 (2015).
27. K. G. Paulson *et al.*, Array-CGH reveals recurrent genomic changes in Merkel cell carcinoma including amplification of L-Myc. *J Invest Dermatol* **129**, 1547-1555 (2009).
28. P. W. Harms *et al.*, The Distinctive Mutational Spectra of Polyomavirus-Negative Merkel Cell Carcinoma. *Cancer Res* **75**, 3720-3727 (2015).
29. M. Yarchoan *et al.*, PD-L1 expression and tumor mutational burden are independent biomarkers in most cancers. *JCI Insight* **4**, (2019).
30. A. M. Goodman *et al.*, Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. *Mol Cancer Ther* **16**, 2598-2608 (2017).
31. A. M. Goodman, E. S. Sokol, G. M. Frampton, S. M. Lippman, R. Kurzrock, Microsatellite-Stable Tumors with High Mutational Burden Benefit from Immunotherapy. *Cancer Immunol Res* **7**, 1570-1573 (2019).
32. P. Nghiem *et al.*, Durable Tumor Regression and Overall Survival in Patients With Advanced Merkel Cell Carcinoma Receiving Pembrolizumab as First-Line Therapy. *J Clin Oncol*, JCO1801896 (2019).
33. J. G. Berry *et al.*, Association of Home Respiratory Equipment and Supply Use with Health Care Resource Utilization in Children. *J Pediatr* **207**, 169-175 e162 (2019).
34. J. G. Iyer *et al.*, Merkel cell polyomavirus-specific CD8(+) and CD4(+) T-cell responses identified in Merkel cell carcinomas and blood. *Clin Cancer Res* **17**, 6671-6680 (2011).
35. N. P. Croft *et al.*, Most viral peptides displayed by class I MHC on infected cells are immunogenic. *Proc Natl Acad Sci U S A* **116**, 3112-3117 (2019).
36. M. E. Spurgeon, P. F. Lambert, Merkel cell polyomavirus: a newly discovered human virus with oncogenic potential. *Virology* **435**, 118-130 (2013).

37. Y. Chang *et al.*, Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* **266**, 1865-1869 (1994).
38. D. A. Joseph *et al.*, Understanding the burden of human papillomavirus-associated anal cancers in the US. *Cancer* **113**, 2892-2900 (2008).
39. J. M. Walboomers *et al.*, Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* **189**, 12-19 (1999).
40. T. Y. Seiwert *et al.*, Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. *Clin Cancer Res* **21**, 632-641 (2015).
41. L. Q. M. Chow *et al.*, Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. *J Clin Oncol* **34**, 3838-3845 (2016).
42. P. Nghiem *et al.*, Durable tumor regression and overall survival (OS) in patients with advanced Merkel cell carcinoma (aMCC) receiving pembrolizumab as first-line therapy. *J Clin Oncol* **36**, (suppl; abstr 9506) (2018).
43. S. Kalaora *et al.*, Combined Analysis of Antigen Presentation and T-cell Recognition Reveals Restricted Immune Responses in Melanoma. *Cancer Discov* **8**, 1366-1375 (2018).
44. H. Ebrahimi-Nik *et al.*, Mass spectrometry driven exploration reveals nuances of neoepitope-driven tumor rejection. *JCI Insight* **5**, (2019).
45. M. Yadav *et al.*, Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* **515**, 572-576 (2014).
46. D. S. Chen, I. Mellman, Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321-330 (2017).
47. T. Gruosso *et al.*, Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. *J Clin Invest* **129**, 1785-1800 (2019).
48. K. G. Paulson *et al.*, Transcriptome-wide studies of merkel cell carcinoma and validation of intratumoral CD8+ lymphocyte invasion as an independent predictor of survival. *J Clin Oncol* **29**, 1539-1546 (2011).
49. K. G. Paulson *et al.*, Acquired cancer resistance to combination immunotherapy from transcriptional loss of class I HLA. *Nat Commun* **9**, 3868 (2018).
50. N. A. Giraldo *et al.*, Multidimensional, quantitative assessment of PD-1/PD-L1 expression in patients with Merkel cell carcinoma and association with response to pembrolizumab. *J Immunother Cancer* **6**, 99 (2018).
51. V. Thorsson *et al.*, The Immune Landscape of Cancer. *Immunity* **48**, 812-830 e814 (2018).
52. D. Tamborero *et al.*, A Pan-cancer Landscape of Interactions between Solid Tumors and Infiltrating Immune Cell Populations. *Clin Cancer Res* **24**, 3717-3728 (2018).
53. P. J. Goulder, D. I. Watkins, Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat Rev Immunol* **8**, 619-630 (2008).
54. P. Kiepiela *et al.*, Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature* **432**, 769-775 (2004).
55. N. B. Crux, S. Elahi, Human Leukocyte Antigen (HLA) and Immune Regulation: How Do Classical and Non-Classical HLA Alleles Modulate Immune Response to

Human Immunodeficiency Virus and Hepatitis C Virus Infections? *Front Immunol* **8**, 832 (2017).
56. N. J. Miller *et al.*, Merkel cell polyomavirus-specific immune responses in patients with Merkel cell carcinoma receiving anti-PD-1 therapy. *J Immunother Cancer* **6**, 131 (2018).
57. S. F. Gameiro *et al.*, Treatment-naive HPV+ head and neck cancers display a T-cell-inflamed phenotype distinct from their HPV- counterparts that has implications for immunotherapy. *Oncoimmunology* **7**, e1498439 (2018).
58. P. Sharma, S. Hu-Lieskovan, J. A. Wargo, A. Ribas, Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell* **168**, 707-723 (2017).
59. Z. Pos *et al.*, Longitudinal study of recurrent metastatic melanoma cell lines underscores the individuality of cancer biology. *J Invest Dermatol* **134**, 1389-1396 (2014).
60. M. M. Gubin *et al.*, Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* **515**, 577-581 (2014).
61. F. M. Marincola, E. M. Jaffee, D. J. Hicklin, S. Ferrone, Escape of human solid tumors from T-cell recognition: molecular mechanisms and functional significance. *Adv Immunol* **74**, 181-273 (2000).
62. A. Sucker *et al.*, Genetic evolution of T-cell resistance in the course of melanoma progression. *Clin Cancer Res* **20**, 6593-6604 (2014).
63. P. Akhbari, D. Tobin, K. Poterlowicz, W. Roberts, J. R. Boyne, MCV-miR-M1 Targets the Host-Cell Immune Response Resulting in the Attenuation of Neutrophil Chemotaxis. *J Invest Dermatol* **138**, 2343-2354 (2018).
64. N. Shahzad *et al.*, The T antigen locus of Merkel cell polyomavirus downregulates human Toll-like receptor 9 expression. *J Virol* **87**, 13009-13019 (2013).
65. O. K. Afanasiev *et al.*, Merkel polyomavirus-specific T cells fluctuate with merkel cell carcinoma burden and express therapeutically targetable PD-1 and Tim-3 exhaustion markers. *Clin Cancer Res* **19**, 5351-5360 (2013).
66. O. K. Afanasiev *et al.*, Vascular E-selectin expression correlates with CD8 lymphocyte infiltration and improved outcome in Merkel cell carcinoma. *J Invest Dermatol* **133**, 2065-2073 (2013).
67. M. Dowlatshahi *et al.*, Tumor-specific T cells in human Merkel cell carcinomas: a possible role for Tregs and T-cell exhaustion in reducing T-cell responses. *J Invest Dermatol* **133**, 1879-1889 (2013).
68. H. Sihto *et al.*, Tumor infiltrating immune cells and outcome of Merkel cell carcinoma: a population-based study. *Clin Cancer Res* **18**, 2872-2881 (2012).
69. L. Feldmeyer *et al.*, Density, Distribution, and Composition of Immune Infiltrates Correlate with Survival in Merkel Cell Carcinoma. *Clin Cancer Res* **22**, 5553-5563 (2016).
70. R. A. Clark *et al.*, Human squamous cell carcinomas evade the immune response by down-regulation of vascular E-selectin and recruitment of regulatory T cells. *J Exp Med* **205**, 2221-2234 (2008).
71. K. G. Paulson *et al.*, Downregulation of MHC-I expression is prevalent but reversible in Merkel cell carcinoma. *Cancer Immunol Res* **2**, 1071-1079 (2014).

72. C. Ritter *et al.*, Reversal of epigenetic silencing of MHC class I chain-related protein A and B improves immune recognition of Merkel cell carcinoma. *Sci Rep* **6**, 21678 (2016).
73. S. R. Burrows *et al.*, Peptide-MHC class I tetrameric complexes display exquisite ligand specificity. *J Immunol* **165**, 6229-6234 (2000).
74. A. G. Chapuis *et al.*, Transferred WT1-reactive CD8+ T cells can mediate antileukemic activity and persist in post-transplant patients. *Sci Transl Med* **5**, 174ra127 (2013).
75. J. Greiner *et al.*, High-dose RHAMM-R3 peptide vaccination for patients with acute myeloid leukemia, myelodysplastic syndrome and multiple myeloma. *Haematologica* **95**, 1191-1197 (2010).
76. K. G. Paulson *et al.*, Augmentation of adoptive T-cell therapy for Merkel cell carcinoma with avelumab. *Journal of Clinical Oncology* **35**, 3044-3044 (2017).
77. A. G. Chapuis *et al.*, Regression of metastatic Merkel cell carcinoma following transfer of polyomavirus-specific T cells and therapies capable of re-inducing HLA class-I. *Cancer Immunol Res* **2**, 27-36 (2014).
78. J. L. Gulley *et al.*, Role of Antigen Spread and Distinctive Characteristics of Immunotherapy in Cancer Treatment. *J Natl Cancer Inst* **109**, (2017).
79. M. A. Cheever *et al.*, The prioritization of cancer antigens: a national cancer institute pilot project for the acceleration of translational research. *Clin Cancer Res* **15**, 5323-5337 (2009).
80. I. Gavvovidis *et al.*, Targeting Merkel Cell Carcinoma by Engineered T Cells Specific to T-Antigens of Merkel Cell Polyomavirus. *Clin Cancer Res* **24**, 3644-3655 (2018).
81. B. Dasgeb, D. Mehregan, C. Ring, N. Nartker, I. Brownell, Cancer-testis antigens as biomarkers for Merkel cell carcinoma: Pitfalls and opportunities. *J Cutan Pathol* **46**, 748-752 (2019).
82. H. Kurzen, S. Kaul, U. Egner, M. Deichmann, W. Hartschuh, Expression of MUC 1 and Ep-CAM in Merkel cell carcinomas: implications for immunotherapy. *Arch Dermatol Res* **295**, 146-154 (2003).
83. D. Chowell *et al.*, Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* **359**, 582-587 (2018).
84. K. L. Rock, E. Reits, J. Neefjes, Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends Immunol* **37**, 724-737 (2016).
85. S. Shimizu *et al.*, Tumor-infiltrating CD8(+) T-cell density is an independent prognostic marker for oral squamous cell carcinoma. *Cancer Med* **8**, 80-93 (2019).
86. K. G. Paulson *et al.*, CD8+ lymphocyte intratumoral infiltration as a stage-independent predictor of Merkel cell carcinoma survival: a population-based study. *Am J Clin Pathol* **142**, 452-458 (2014).
87. Y. Naito *et al.*, CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. *Cancer Res* **58**, 3491-3494 (1998).
88. J. M. Fletcher, H. G. Prentice, J. E. Grundy, Natural killer cell lysis of cytomegalovirus (CMV)-infected cells correlates with virally induced changes in cell surface lymphocyte function-associated antigen-3 (LFA-3) expression and

not with the CMV-induced down-regulation of cell surface class I HLA. *J Immunol* **161**, 2365-2374 (1998).
89. M. Koutsakos *et al.*, Downregulation of MHC Class I Expression by Influenza A and B Viruses. *Front Immunol* **10**, 1158 (2019).
90. T. W. Phares, S. A. Stohlman, D. R. Hinton, C. C. Bergmann, Enhanced CD8 T-cell anti-viral function and clinical disease in B7-H1-deficient mice requires CD4 T cells during encephalomyelitis. *J Neuroinflammation* **9**, 269 (2012).
91. P. Bhat, G. Leggatt, N. Waterhouse, I. H. Frazer, Interferon-gamma derived from cytotoxic lymphocytes directly enhances their motility and cytotoxicity. *Cell Death Dis* **8**, e2836 (2017).
92. V. Shankaran *et al.*, IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature* **410**, 1107-1111 (2001).
93. P. D. Drew *et al.*, Regulation of MHC class I and beta 2-microglobulin gene expression in human neuronal cells. Factor binding to conserved cis-acting regulatory sequences correlates with expression of the genes. *J Immunol* **150**, 3300-3310 (1993).
94. N. V. Longino *et al.*, Human CD4(+) T Cells Specific for Merkel Cell Polyomavirus Localize to Merkel Cell Carcinomas and Target a Required Oncogenic Domain. *Cancer Immunol Res* **7**, 1727-1739 (2019).
95. S. Kreiter *et al.*, Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* **520**, 692-696 (2015).
96. A. C. Anderson, N. Joller, V. K. Kuchroo, Lag-3, Tim-3, and TIGIT: Co-inhibitory Receptors with Specialized Functions in Immune Regulation. *Immunity* **44**, 989-1004 (2016).
97. L. Golden-Mason *et al.*, Negative immune regulator Tim-3 is overexpressed on T cells in hepatitis C virus infection and its blockade rescues dysfunctional CD4+ and CD8+ T cells. *J Virol* **83**, 9122-9130 (2009).
98. R. H. McMahan *et al.*, Tim-3 expression on PD-1+ HCV-specific human CTLs is associated with viral persistence, and its blockade restores hepatocyte-directed in vitro cytotoxicity. *J Clin Invest* **120**, 4546-4557 (2010).
99. S. Koyama *et al.*, Adaptive resistance to therapeutic PD-1 blockade is associated with upregulation of alternative immune checkpoints. *Nat Commun* **7**, 10501 (2016).
100. M. A. Curran, W. Montalvo, H. Yagita, J. P. Allison, PD-1 and CTLA-4 combination blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors. *Proc Natl Acad Sci U S A* **107**, 4275-4280 (2010).
101. A. Mehta *et al.*, Immunotherapy Resistance by Inflammation-Induced Dedifferentiation. *Cancer Discov* **8**, 935-943 (2018).
102. C. S. Pai *et al.*, Clonal Deletion of Tumor-Specific T Cells by Interferon-gamma Confers Therapeutic Resistance to Combination Immune Checkpoint Blockade. *Immunity* **50**, 477-492 e478 (2019).
103. K. G. Paulson *et al.*, Antibodies to merkel cell polyomavirus T antigen oncoproteins reflect tumor burden in merkel cell carcinoma patients. *Cancer Res* **70**, 8388-8397 (2010).

104. K. G. Paulson *et al.*, Viral oncoprotein antibodies as a marker for recurrence of Merkel cell carcinoma: A prospective validation study. *Cancer* **123**, 1464-1474 (2017).
105. P. R. Spradling *et al.*, Immunity to hepatitis B virus (HBV) infection two decades after implementation of universal infant HBV vaccination: association of detectable residual antibodies and response to a single HBV challenge dose. *Clin Vaccine Immunol* **20**, 559-561 (2013).
106. S. Kikuchi *et al.*, Serum anti-Helicobacter pylori antibody and gastric carcinoma among young adults. Research Group on Prevention of Gastric Carcinoma among Young Adults. *Cancer* **75**, 2789-2793 (1995).
107. K. J. Lastwika *et al.*, Tumor-derived Autoantibodies Identify Malignant Pulmonary Nodules. *Am J Respir Crit Care Med* **199**, 1257-1266 (2019).
108. D. Wu *et al.*, High-throughput sequencing detects minimal residual disease in acute T lymphoblastic leukemia. *Sci Transl Med* **4**, 134ra163 (2012).
109. A. G. Chapuis *et al.*, Tracking the Fate and Origin of Clinically Relevant Adoptively Transferred CD8(+) T Cells In Vivo. *Sci Immunol* **2**, (2017).
110. H. Robins, Immunosequencing: applications of immune repertoire deep sequencing. *Curr Opin Immunol* **25**, 646-652 (2013).
111. A. G. Chapuis *et al.*, T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. *Nat Med* **25**, 1064-1072 (2019).
112. S. A. Grupp *et al.*, Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. *N Engl J Med* **368**, 1509-1518 (2013).
113. A. C. Logan *et al.*, High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. *Proc Natl Acad Sci U S A* **108**, 21194-21199 (2011).

Vita

Miranda Lahman was born and raised in Boulder, Colorado. In 2010, she moved to Bellingham, Washington for her undergraduate studies at Western Washington University. Miranda was a student-athlete where she ran Track and Field and Cross country for the Vikings while also receiving a Bachelor of Science in cellular and molecular biology. During this time, she worked in the laboratory of Dr. Jeff Young in collaboration with Dr. Marion Brodhagen to identify genes that confer resistance to *Aspergillus* infection in a plant model system. After completing her undergraduate degree in 2014 Miranda moved to Seattle, Washington to work for Stratos Genomics. Here she helped develop their DNA sequencing platform, Sequencing by Expansion (SBX^{TM}), that utilizes Stratos' synthetic nucleotides and a proprietary enzymatic amplification process. As a member of the protein engineering team, she worked to engineer novel DNA polymerases and develop a directed evolution platform to advance this technology. In 2017, Miranda began graduate school at the University of Washington in the Molecular Medicine and Mechanisms of disease program. She completed her doctoral dissertation in the laboratory of Dr Aude Chapuis in collaboration with Dr. Roland Strong at the Fred Hutchinson Cancer Center where she studied the role of the proteasome in immunotherapy escape and its contribution the immunopeptidome. Miranda Received her PhD in 2022.