T-Cell Immunogenicity and Dysfunction in Cancer and Viral Diseases

by

Sri Krishna

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2017 by the
Graduate Supervisory Committee:

Karen S. Anderson, Chair
Joshua LaBaer
Bertram L. Jacobs
Douglas F. Lake

ARIZONA STATE UNIVERSITY

December 2017

ABSTRACT

CD8+ T-lymphocytes (CTLs) are central to the immunologic control of infections and are currently at the forefront of strategies that enhance immune based treatment of a variety of tumors. Effective T-cell based vaccines and immunotherapies fundamentally rely on the interaction of CTLs with peptide-human leukocyte antigen class I (HLA-I) complexes on the infected/malignant cell surface. However, how CTLs are able to respond to antigenic peptides with high specificity is largely unknown. Also unknown, are the different mechanisms underlying tumor immune evasion from CTL-mediated cytotoxicity. In this dissertation, I investigate the immunogenicity and dysfunction of CTLs for the development of novel T-cell therapies. Project 1 explores the biochemical hallmarks associated with HLA-I binding peptides that result in a CTL-immune response. The results reveal amino acid hydrophobicity of T-cell receptor (TCR) contact residues within immunogenic CTL-epitopes as a critical parameter for CTL-self/nonself discrimination. Project 2 develops a bioinformatic and experimental methodology for the identification of CTL-epitopes from low frequency T-cells against tumor antigens and chronic viruses. This methodology is employed in Project 3 to identify novel immunogenic CTL-epitopes from human papillomavirus (HPV)-associated head and neck cancer patients. In Project 3, I further study the mechanisms of HPV-specific T-cell dysfunction, and I demonstrate that combination inhibition of Indoleamine 2, 3-dioxygenase (IDO-1) and programmed cell death protein (PD-1) can be a potential immunotherapy against HPV+ head and neck cancers. Lastly, in Project 4, I develop a single-cell assay for high-throughput identification of antigens targeted by CTLs from whole pathogenome libraries. Thus, this dissertation contributes to fundamental T-cell immunobiology by identifying rules of T-cell immunogenicity and dysfunction, as well as to translational immunology by identifying novel CTL-epitopes, and therapeutic targets for T-cell immunotherapy.

For Amma, Appa, and Anna; without all of you, none of this would have been possible

and

For all the giants who stood tall before me

“If I have seen further, it is by standing on the shoulders of giants” – Sir Isaac Newton, 1675
“nanos gigantum humeris insidentes” – Bernard of Chartres, 12 AD

ACKNOWLEDGMENTS

I would first like to acknowledge my dissertation committee, Dr. Karen Anderson, Dr. Joshua LaBaer, Dr. Bertram Jacobs and Dr. Douglas Lake for their support and guidance. In particular, I would like to thank Karen for giving me the opportunity to learn immunology and cancer biology, for allowing me to fail over and over again, for supporting me through all these failures, and for telling me “focus on the science, the rest will come”. I thank Josh for showing me to think in clarity about truly big picture questions and ideas in science. I thank Jake and Doug for their intellectual input and allowing me to use their lab and resources with no restrictions.

This dissertation would not have been possible without help from members of the Anderson Lab, past and present. In particular, I would like to thank Shay Ferdosi, Peaches Ulrich and Eric Wilson for their unflinching friendship, and support inside the lab with experiments, and discussions outside the lab over coffee, and food. I would also like to thank Dr. Diego Chowell, Radwa Ewaisha, Sherie Gwynn, Mark Knappenberger, Dr. Meixuan (Emma) Chen, Marika Hopper, Ting Li, Padhmavathy Yuvaraj, Julia Cheng, Immanuel Purushothaman, Brittany Ulloa, Rizwan Alam and Dr. Benjamin Katchman for their support in lab.

I would like to thank members of the Biodesign Center for Personalized Diagnostics Anasuya Pal, Seron Eaton, Dr. Ji Qiu, Dr. Jin Park, Dr. Vel Murugan, Dr. Stephen Blazie, Dr. Mitch Magee, Dr. Shanshan Yang, Kasuen Kotagama, Heather Geissel, Karie Tepper and Vanessa Baack for their help and support. Special thanks are in order for Dr. Jordan Yaron who has helped me with his career advice, friendship and wisdom over the years.

This research was only made possible with excellent collaborations over the years. In particular, I would like to thank Dr. Joseph Blattman (CIVV, ASU) for his scientific insight, Dr. Linda Klavinskis (Kings College London), Dr. Andrew Sikora and Falguni Parikh (Baylor College of Medicine) and Dr. Marshall Posner and team at Mount Sinai School of Medicine for their great scientific input and collaborations. A special thanks to Dr. Melissa Wilson Sayres (CEM, ASU) for her support and Dr. Pooja Narang from the Sayres Lab for scientific input.

I would also like to acknowledge the student and funding support from the Biological Design Graduate Program faculty past and present; Dr. Tony Garcia, Laura Hawes, Dr. JoAnn

Williams, Dr. Stephen Johnston and Maria Hanlin. The School of Graduate Education and the Graduate and Professional Students Association have been generous to me by supporting me over the years through their travel and research grants and the Dissertation completion Fellowship.

I would like to thank my friends and family outside science, Ramanan, Vinoth, Riyash, Praveen, Swamy, Abhinav, Jake, Alex, my cousins Suby, Karthik and my late pup Denzy for their wonderful friendship and for reminding me of life outside the lab. My special thanks to Samantha Cotsmire for being the source of incredible love, strength, scientific discourse and support past five years. Lastly, I would like to thank my family; my Mom for being the kindest, generous, wonderful person she is and for being a relentless fighter against cancer, and my Dad and Brother for being incredible role models in my life.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

The vertebrate immune system is composed of a constellation of diverse immune cell types, each with an exquisite ability to protect the host from infectious agents, cancers, and other immunopathologies. Central to the success of this elegant complex system is the arm of adaptive immunity, dominated by the effects of T-lymphocytes and B-lymphocytes. In terms of evolutionary history, an immunological “big bang” resulted in the adoption of T/B lymphocyte receptors as dominant anti-pathogen defense system by the vertebrate lineage (Flajnik, 2014; Pancer & Cooper, 2006). The result is the near virtual conservation of the adaptive lymphocyte lineage in jawed vertebrates, underscoring the importance of lymphocyte-mediated immune protection across the tree of life (Pancer & Cooper, 2006). T/B Lymphocytes have three fundamental hallmarks that make them indispensable for host defense: 1) antigen specificity that is critical for self/nonself discrimination limiting off-target effects 2) receptor diversity that ensures sufficient frequency of pre-existing pool of immune cells to effectively combat any unknown pathogen over the course of the host’s lifetime 3) the unique immunologic memory that issues recall responses against a previous pathogen, and forms the basis of all modern vaccines (Abbas, Lichtman, & Pillai, 2014; Moticka, 2016a).

Our understanding of the hallmarks of the lymphocyte immune response and its protective effects dates back to the early days of variolation performed by the Chinese as an effective deterrent against smallpox (Leung, 2011; Moticka, 2016a). It wasn’t until Edward Jenner’s smallpox vaccination efforts in 1796, that a true appreciation for concept of immunological memory impact on modern medicine was truly felt (Stewart & Devlin, 2006). About hundred years later, in 1890, an American pathologist William B Coley observed that injecting sarcoma patients with heat inactivated bacteria into the tumors resulted in occasional prolonged remissions (Decker & Safdar, 2009). It is now believed that a combination of innate and adaptive immune response by trafficking immune cells into the tumor injected with “Coley’s toxins” resulted in tumor remission observed in these patients (Decker & Safdar, 2009). Today, close to 120 years

after Dr. Coley's findings, cancer immune therapies have come to the forefront of therapeutic interventions of cancer (Couzin-Frankel, 2013).

This chapter describes a broad historical account into the discovery of the major players in cell mediated immune response: MHC-restriction and T-cells, along with the foundations of self/nonself discrimination and cancer immunology. Particular focus is given to the last 50 years of research in T-cell immunology and tumor immunology. Finally, relevant unanswered questions specifically relating to this dissertation work, and contributions addressed by the investigations pursued in this dissertation are addressed.

1.1 Discovery, structure and function of MHC

1.1.1. Discovery of MHC

The Major Histocompatibility complex (MHC) genes, also called human leukocyte antigen (HLA), make proteins that form the crux of adaptive cell mediated T-cell response. There is a deep historical relationship between tumor immunology and MHC, going back to early 1900s. During the early years of transplantation and tumor immunology, tumors in inbred and random outbred mice were commonly used to understand immune responses. In 1909, Ernest Tyzzer, an American pathologist, and a founding member of the American association of cancer research (AACR), studied the progression and rejection of tumors transferred between the inbred Japanese waltzing mice and the outbred albino mice (Moticka, 2016c). Tyzzer observed that spontaneous tumors from the Japanese waltzing mice were able to grow when implanted into other Japanese waltzing mice, but were rejected when implanted into the outbred albino mice (Moticka, 2016c). Along with another researcher in his lab, Clarence C. Little, Tyzzer also observed that while F1 hybrids crossed between the Japanese waltzing mice and inbred C57Bl mice accepted the transplanted tumors, almost 98% of the F2 generation hybrids rejected it (Little, 1914, 1920; Moticka, 2016c). These results indicated that there was a genetic basis of tumor rejection in these animals. The data from F2 hybrid mice seemed to indicate a non-mendelian mode inheritance of tumor rejection, which they proposed in a 1914 paper (Little,

1914). Little and Tyzzer later reconciled their data by proposing that perhaps more than one gene (upto 12, according to their calculations) was involved in these tumor rejection experiments (Little, 1914, 1920; Moticka, 2016c).

During World War 2, Peter Groer performed similar experiments to identify genetic traits underlying tumor rejection by outbred mice. Gorer proposed that blood group antigens might be the source of tumor rejection in the different mice receiving the transplantable tumors. He reasoned that if this was the case, they should bear antibodies (which were described years before by Paul Ehrlich) against red blood cells from tumor donor in the recipient (Moticka, 2016c). Gorer was able to show that three types of antibodies indeed developed against blood antigens in the different types of inbred and outbred mice receiving the transplanted tumors (Gorer, 1936). Gorer then collaborated with George D. Snell from Jackson Laboratory who was interested in "histocompatibility genes" by studying tumor rejection in outbred mice. Together, they showed that tumors transplanted from mice with one of the blood antigens Groer identified (antigen II), was immediately rejected by mice which lacked antibodies against antigen II (Gorer, Lyman, & Snell, 1948). Gorer and Snell renamed this gene as Histocompatibility gene 2 (H-2) and proposed rules for tumor rejection during transplantation (Gorer et al., 1948).

The severe need for medical research on transplantation, especially of skin grafts to help soldiers and civilians recover from their severe injuries peaked during World War 2 (Moticka, 2016c). Therefore, under the direction of Medical research Council of UK, Peter Medawar studied transplanting non-malignant tissue in mice models, performing experiments largely similar to Groer and Snell (Moticka, 2016c), and obtained similar results as the MHC-H-2 antigen researchers. Medawar also observed that the severity of rejection of skin grafts transplanted between between animals decreased as the the "relatedness" between the animals increased (Medawar, 1944). Medawar made the very astute observation that graft rejection was therefore an immunological process associated with the adaptive immunity based on 1) the speed of secondary rejection (vastly higher than preliminary rejections) similar to "immune memory" 2) the specificity of graft rejection with limited off-target effects (Medawar, 1944). Medawar and colleagues were further able to demonstrate that it was lymphocytes from draining lymph nodes

in the transplanted animals that were responsible for the vigorous reaction against transplanted skin grafts (Brent, Brown, & Medawar, 1958; Medawar, 1944).

Based on decades of previous research, in 1956, in a landmark study, the first ever successful kidney transplant was performed between monozygotic twins (Merrill, Murray, Harrison, & Guild, 1956). In the following years, Jean Dausset and Baruj Benacerraf identified the human MHC (HLA) protein and it was shown that tumor rejection antigens and transplant rejection antigens were the same, and they were linked to lymphocyte-mediated graft rejection (Benacerraf, 1981; Dausset, 1958). For their discovery of MHC, George Snell, Jean Dausset and Baruch Benacceraf were awarded the Nobel Prize for Physiology or Medicine in 1980 (Moticka, 2016c). Peter Medawar shared the Nobel Prize with Sir Frank Burnett for their contributions to transplantation immunology in 1960 (Moticka, 2016c).

1.1.2. Structure and function of MHC

In the 1970s, there were several labs working on elucidating the structure of the transplant molecule - the MHC. Because Ethan Shevach had previously shown that antibodies raised against MHC could interfere with antigen induced lymphocyte proliferation (Shevach, 1972), Pete Cresswell and colleagues eluted and purified MHC from surface of lymphocytes (Cresswell, Turner, & Strominger, 1973). The results revealed the presence of two polypeptide fragments; a 30 KD fragment and a 11 KD fragment (Cresswell et al., 1973). The smaller 11 KD fragment was later shown by by Howard Grey and Pete Cresswell to be the the Beta-2-microglobulin (β2M) protein (Figure 1B), which non-covalently associates with with MHC-I (Grey, 1973). Finally, in 1987, the crystal structure of human MHC HLA-A*02:01-β2M complex was solved by Pamela J. Björkman, Don C. Wiley and Jack L. Strominger (Bjorkman et al., 1988). Figure 1A shows a 3D reconstructed model of the original crystal structure from Bjorkman et al., which demonstrated a clear cleft formed by the HLA's α1, α2 domains for foreign antigen binding (later shown to be linear peptides). Figure 1B shows the orientation of an antigenic peptide from the Influenza virus A protein M1 forming non-covalent binding in the peptide binding cleft formed by α1-α2 domains. The MHC class II (HLA-II) structure, solved much later again from JL Strominger and DC Wiley's

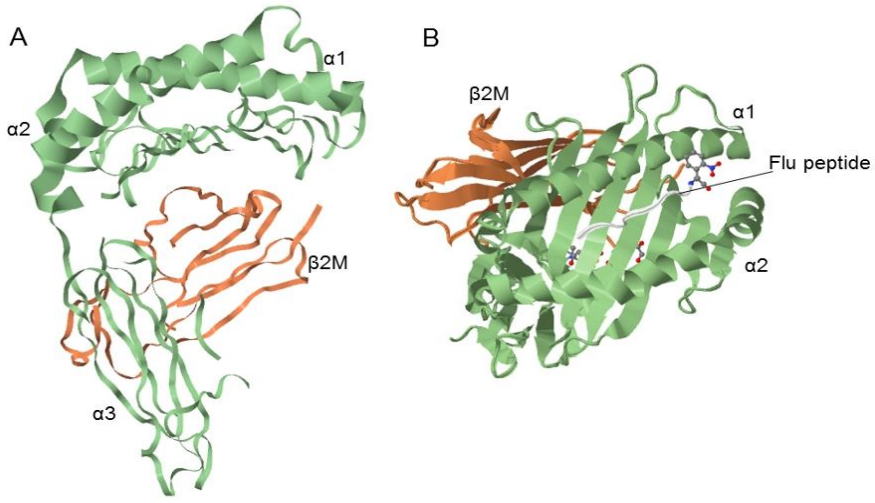

Figure 1-1. 3D protein structure of HLA-A*02:01 derived from Protein Database (PDB). (A) 3D reconstruction of original x-ray crystal structure of the human-MHC HLA-A*02:01 complex (Bjorkman et al., 1988) showing α1-α3 domains of human HLA-A*02:01 heavy chain (green), and β2M (orange) complex derived from PDB (Rose et al., 2017). (B) Top view of the HLA-A*02:01 α1-α2 domains forming the antigen binding cleft in complex with a viral-antigenic peptide (GILGFVFTL in grey) from Influenza A virus matrix protein from (Celie et al., 2009) (PDB 2X70).

groups showed that it was a heterodimer formed by an α chain and a β chain (Brown et al., 1993).

Functionally, MHCs contribute to host immune surveillance by binding and presenting small linear peptides derived from intracellular proteins on the cell surface (Blum, Wearsch, & Cresswell, 2013; Fernando et al., 2008). Encoded in the short arm of human chromosome 6, the MHC-locus encompasses a multi-gene region representing one of the most polymorphic locus of the human genome (Blum et al., 2013; Fernando et al., 2008; Murphy & Weaver, 2016). Variants in this locus contribute to the control and susceptibility to a wide range of immunologic diseases from autoimmunity, infectious diseases to transplantation, and tumor immunity (Blum et al., 2013; Fernando et al., 2008). Classical antigen processing machinery posits that MHC class I (or HLA-I) presents such intracellular peptides to CD8+ T-cells (CTLs) while MHC class II (HLA-II) does the same to CD4+ T-cells (Blum et al., 2013; Fernando et al., 2008). MHC class I is expressed in all nucleated cells except neurons and erythrocytes, while MHC class II proteins are expressed in specialized immune cells such as antigen presenting cells (APCs) (Moticka, 2016c; Murphy &

Weaver, 2016). The peptide products are primarily derived from intracellular proteolysis of cytosolic proteins through the proteasome for the MHC class I pathway, and from exogenous proteins into the lysosome for MHC class II (Blum et al., 2013). In MHC class I pathway, the focus of CTL-recognition in this dissertation, long peptides generated by proteasomal degradation are imported into the endoplasmic reticulum (ER) via the transporter associated with antigen processing (TAP) molecules as shown in Figure 1-2 (Blum et al., 2013). Subsequently, these longer peptides are further trimmed by ER aminopeptidase (ERAAP) to short 9-11 amino acid linear peptides, and the short peptides with binding motifs for the respective MHC subsequently bind with strong affinity (Blum et al., 2013). Peptide binding stabilizes the MHC-β2M complex, and subsequently the pMHC-β2M trimeric complex is exported through the Golgi complex onto the cell surface for CTL- evaluation (Figure 1-2) (Blum et al., 2013). Thus, in classical MHC Class I pathway, the predominant ligands/peptides presented on cell surface via MHC-binding are

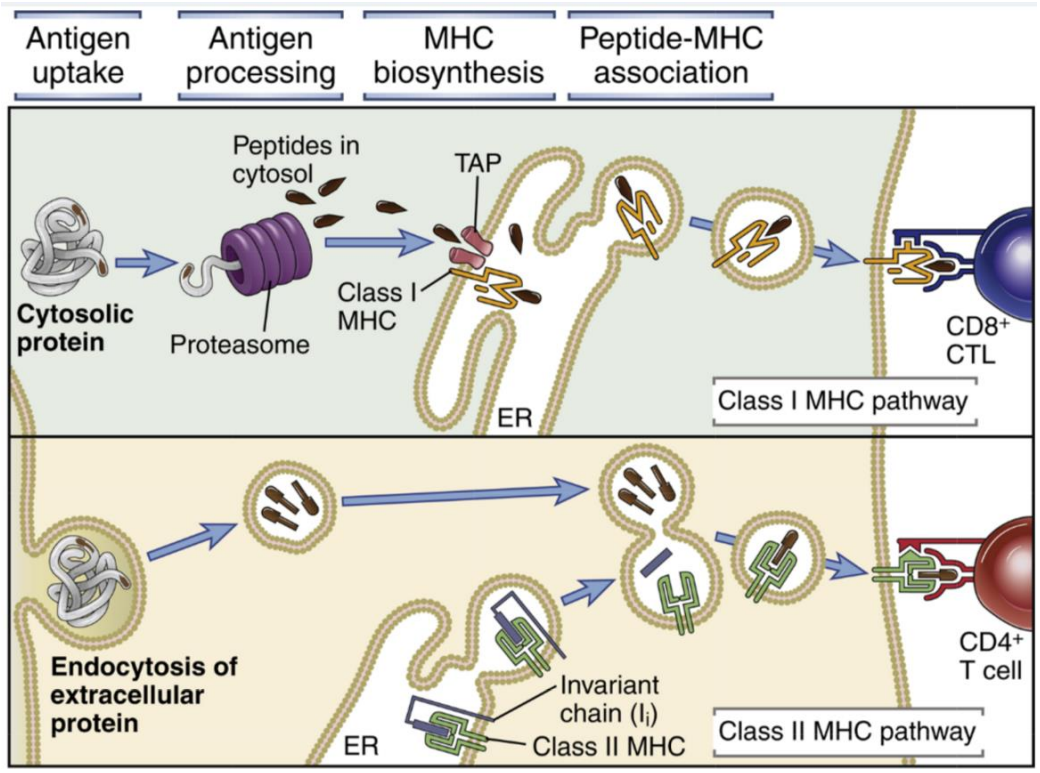

Figure 1-2. Classical MHC class I and II antigen processing pathways. Top: MHC class I pathway. Bottom: MHC Class II pathway. Note that MHC class I is expressed in all nucleated cells, while MHC class II is expressed predominantly in immune cells (APCs). Figure reproduced with permission from Cellular and Molecular Immunology, Chapter 6, pg 127-128, Eighth Edition, Abul K. abbas, Andrew H. Lichtman, Shiv Pillai. Copyright Elsevier 2015 (Appendix C1).

self-peptides. However, during a pathogenic infection for instance, non-self-proteins (as well as autoimmune, tumor associated proteins) are subject to the same rules, resulting in the presentation of non-self antigenic “epitopes” on MHC class I/I. These antigenic peptide-MHC complexes on both class I and class II pathways (pMHC-I, pMHC-II) ultimately leads to T-cell (CD8, CD4 respectively) recognition of the infected cell (Blum et al., 2013; Fernando et al., 2008). This process will be covered in subsequent sections.

1.2 Discovery and function of T-Cells

1.2.1. Discovery of T-Cells

The earliest known indication of cell mediated immunity was observed when Ilya Metchnikov, a Russian zoologist showed that a rose thorn inserted into starfish larvae attracted cells (later shown to be macrophages performing phagocytosis) which surrounded, engulfed and dissolved the thorn (Metchnikoff, 1989; Moticka, 2016e). The relative importance between the types of immune response observed in infections between antibodies (demonstrated by Paul Ehrlich) and cellular responses (Ilya Metchnikov) was the subject of long dispute and debate, despite them sharing the Nobel prize for Physiology and Medicine in 1908 (Metchnikoff, 1989; Moticka, 2016e).

In the early 1940s, Merrill Chase and Karl Landsteiner were researching on skin hypersensitivity and inflammation after synthetic chemical injections (Moticka, 2016e). Chase and Landsteiner showed that serum transfer from guinea pigs sensitized with picryl chloride (a synthetic chemical) injections failed to protect non-sensitized animals indicating that antibodies alone failed to protect against skin hypersensitivity (Landsteiner & Chase, 1942; Moticka, 2016e). However, other immunologists around the same time had shown that skin hypersensitivity matched a number of tenets of immunological memory as proposed by Peter Medawar (section 1.1.1) including immunological memory that mediated a vigorous secondary response to hypersensitivity. After several failures to explain skin hypersensitivity, Chase accidentally used

serum contaminated with cell preparation which elicited skin hypersensitivity in normal guinea pigs (Landsteiner & Chase, 1942). Building on these fortuitous observations, in 1945, Chase demonstrated that peripheral blood cell preparations were able to transfer and initiate skin sensitivity to chemicals and tuberculosis pathogen used in the previous studies (Chase, 1945), demonstrating what is perhaps one of the first attempts at adoptive T-cell transfer therapy.

Around the same time, Peter Medawar was studying the immunologic basis on skin graft rejection (section 1.1.1, (Medawar, 1944)). Because of a substantial amount of antibodies in sera of animals and human undergoing graft rejection, Medawar assumed that serum antibodies were the cause (as opposed to the consequence) of skin graft rejection (Hildemann & Medawar, 1959). Building on Chase's findings, in 1955, Avrion Mitchison, showed that lymphocytes from tumor draining lymph nodes were able to mediate the rejection of tumor grafts in syngeneic animals (Mitchison, 1955). Medawar, in a subsequent seminal study, then reasoned and showed that similar lymphocyte-mediated inflammatory reaction might also occur during non-malignant tissue transplantation using skin grafts (Brent et al., 1958).

1.2.2. Discovery of T-Cell-MHC-restriction

By 1956, it was established that antibodies are synthesized and secreted by cells that are found in the mature Bursa Fabricus of birds (and later in mammalian bone marrow) called B-cells (Glick & Sadler, 1961), but no such equivalent was found for the cell-mediated responses. In 1961, Jacques Miller, an Australian immunologist, discovered the function of the thymus (an organ rich with lymphocytes) by demonstrating that thymectomy dramatically reduced an animal's ability to respond to subsequent antigen challenge (Miller, 1961). Subsequent experiments by James Gowans and colleagues clearly showed antigen-dependent proliferation of small lymphocytes and thymocytes (Gowans, Mcgregor, Cowen, & Ford, 1962). Furthermore, Rupert E. Billingham, a British immunologist in 1968 showed unequivocally that lymphocytes, not antibodies from skin graft recipient mice were able to transfer their reactivity to other untransplanted mice, demonstrating the true importance of lymphocyte mediated histo-incompatibility (Billingham, Silvers, & Wilson, 1963).

In subsequent years several studies indicated the link between T-cells and MHCs (section 1.1.1). For instance, a 1975 study by Peter Doherty and Rolf Zinkernagel demonstrated that activation of cytotoxic T-lymphocytes by LCMV-infected fibroblasts required MHC-compatibility, was contingent on MHC H-2^k interaction and was abrogated in mice lacking MHC-H-2^k molecules (R. M. Zinkernagel, 1975). Similar results of histocompatibility were observed using T-cells and Macrophages, T-cells and B-cells (Katz, 1973). This led Zinkernagel and Doherty to propose that either 1) there were two separate molecules on CTLs (one interacting with virus and other with MHC) - the "intimacy model"; or 2) the virus infection altered the MHC to provide a signal to CTLs-"the altered self" model (Rolf M. Zinkernagel & Doherty, 1974). The duo would later be awarded the Nobel Prize for Physiology or Medicine in 1996.

1.2.3. T-Cell-receptor discovery and structure

Studies on the nature of T-cell receptor (TCR) structure in early 1980s by several groups clearly put to rest the debate of altered self versus intimacy model proposed by Zinkernagel and Doherty (Allison, McIntyre, & Bloch, 1982; Haskins, 1983; Meuer, 1983). The studies indicated that the TCR directly interacts with the MHC-I, and the immunogenic antigen was a linear 8-11 amino acid stretch of peptides derived from the antigen. In essence, the altered self resulted from the presence of the short antigenic epitope from the antigen (virus or tumor-derived) while the TCR itself recognized the MHC-peptide complex. Emil R. Unanue, a Cuban-American immunologist, fluorescently labelled the peptide derived from hen egg lysozyme and visualized on cell surface in complex with MHC-I, and proved the specificity of MHC-I, but not MHC-II (B. P. Babbitt, Matsueda, Haber, Unanue, & Allen, 1986; Bruce P. Babbitt, Allen, Matsueda, Haber, & Unanue, 2005). In 1984, Mark Davis and colleagues, and Tak Wah Mak's group in two successive papers, identified the elusive T-cell receptor gene (TCR from αβ-TCR expressing T-cells) by subtracting T-cell mRNA from that of B-cells (Hedrick, Cohen, Nielsen, & Davis, 1984; Yanagi et al., 1984). Harvey Cantor and colleagues, and a little later, Edgar Engleman and colleagues, showed that MHC-I and MHC-II bound peptides stimulated the two major subpopulations of T-cells, CD4+ T-cells and CD8+ T-cells respectively (Cantor & Boyse, 1975; Engleman, Benike, Grumet, & Evans,

1981). These results were further confirmed by the solved X-ray crystal structure of the MHC-class I peptide complex (section 1.1.2). Lastly, Don Wiley's group in 1996, again succeeded in solving the X-ray crystal structure of TCR in complex with human HLA-A*02:01-Viral peptide complex (Garboczi et al., 2010).

Wiley et al's results largely confirmed decades of previous research. The TCR is a cell surface membrane disulfide linked heterodimer of the α-chain and β-chain belonging to the Immunoglobulin superfamily proteins (Garboczi et al., 2010; Wucherpfennig, Gagnon, Call, Huseby, & Call, 2010). Both chains have a constant (C) region proximal to the cell surface of the T-cell, while the variable (V) region interacts with the MHC-peptide complex (Wucherpfennig et al., 2010). In general, the TCR interacts diagonally with the MHC-peptide complex (Figure 1.3), with different loops within the Vα and Vβ domains making contact with the antigenic peptide, or the MHC, or both (Garboczi et al., 2010; Wucherpfennig et al., 2010). The V and C regions of αβTCRs as well as Igs from B-cells, are generated by the process of somatic recombination, discovered in 1978 by Susumu Tongeawa (Tonegawa, 1983). This process, referred to as V(D)J recombination, results in the generation of the incredible diversity of TCR-repertoire which can in theory, can generate up to 10^{15} to 10^{20} TCR-clonotypes (Laydon, Bangham, & Asquith, 2015).

The diagonal placement of the TCR over pMHC complex ensures that the hypervariable complementarity determining region 3 (CDR3) of Vα and Vβ domains (CDR3α & CDR3β) are responsible for interacting with the center of the antigenic peptide (Wucherpfennig et al., 2010). The CDR1 and CDR2 of Vα and Vβ domains in turn interact with the N or C-terminal of the antigenic peptide, with the CDR2 also thought to be interacting with the presenting MHC, contributing to the TCR-MHC-restriction (Wucherpfennig et al., 2010). Recent studies have now pointed that TCRs although cross-reactive, have very similar recognition motifs across the different antigenic epitopes (Birnbaum et al., 2014; Cole et al., 2016). Despite the generalizable properties of TCR-pMHC interaction, individual residues contributing a TCR's recognition of an antigenic peptide-MHC complex is highly dependent on the specific antigenic peptide-MHC complex, giving high specificity and sensitivity to the sampling methodology for T-cells encountering infected/malignant cells throughout a host's lifetime.

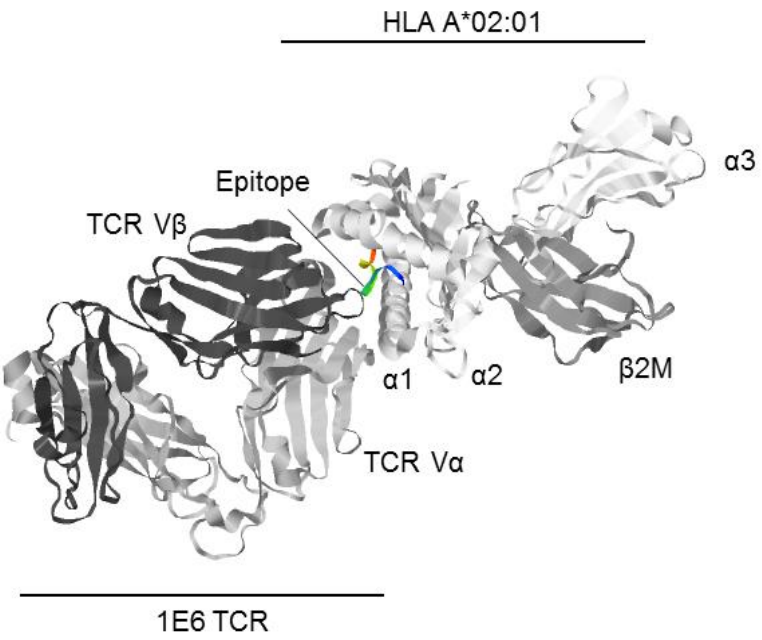

Figure 1-3. 3D protein structure of the 1E6-TCR-pHLA complex. 1E6-TCR in complex with preproinsulin peptide-HLA-A*02:01, derived from a patient with Type 1 diabetes (T1D). 3D reconstruction of X-ray Crystal structure of human-TCR 1E6 binding to the preproinsulin epitope AQWGPDPAAA-HLA-A*02:01 complex (Cole et al., 2016) obtained from PDB. The pMHC complex (right side) and the TCR (left side) are shown, colored by subunits within each protein. The 1E6 TCR also has cross-reactivity with several microbial peptides and is considered to be active in T1D patients by antigen mimicry (Cole et al., 2016; Stadinski, Obst, & Huseby, 2016).

1.2.4. Function of T-Cells

T-cells are the master mediators of the adaptive cell mediated immunity against infected, malignant cells. While antibodies secreted by B-lymphocytes (B-cells) function predominantly to protect an infection from getting established (by “neutralization” and other mechanisms), T-cells function by eliminating infected cells once an infection/malignant cell has already been established (Murphy & Weaver, 2016). When TCR-pMHC interaction is achieved, the membrane proximal Cα and Cβ constant domains region of TCR associate with cell surface CD3 proteins (CD3Ɣ, CD3δ, and CD3ε) (Abbas et al., 2014). Subsequently, a number of signal transducing molecules are recruited to the TCRs, such as the SRC family kinase (SFK) member LCK which phosphorylates immunoreceptor tyrosine-based activation motifs (ITAMs) of the CD3Ɣ, CD3δ,

and CD3ε (Brownlie & Zamoyska, 2013). Following ITAM phosphorylation, a series of phosphorylation events occur via ZAP70 (ζ-chain associated protein kinase of 70 kDa) by LCK, LAT (linker for activation of T cells) by ZAP70, and subsequently assembles a multiprotein complex termed LAT-signalosome (Brownlie & Zamoyska, 2013). The phenotypic outcome of this diverse signaling network is the upregulation of genes necessary for sustained growth, effector functions (e.g. target cell cytolysis by CTLs), secretion of canonical cytokines (e.g. Interleukin-2 by CTLs), cell adhesion molecules (e.g. CD62L for lymph node homing), and other functions associated with the memory state of the T-cell (e.g. Naive vs memory).

Although there are other types of T-cells (e.g. $\boldsymbol{\gamma\delta}$ T-cells, NKT cells etc.), the primary focus of this dissertation will be on CTLs (and to some extent CD4+ T-cells and regulatory T-cells). Functionally the two major subsets of T-cells, CD4+ T-cells and CD8+ T-cells (CTLs) play different yet complementing roles in cell mediated immunity. CD4+ T-cells, termed helper T-cells are stimulated by antigenic peptide-MHC class II complexes are involved in 1) secreting pro-inflammatory cytokines (Th1, Th2 and Th17) that can either directly or indirectly eliminate intracellular, extracellular pathogens; 2) providing T-cell help to B-cells, CTLs to carry out their effector function (Abbas et al., 2014). CTLs, also known as cytotoxic T-cells, are stimulated by antigenic peptide-MHC class I complexes presented by proteasomal processing of antigen by professional APCs (Abbas et al., 2014). Once primed, effector CTLs home to the periphery or the site of infection and directly cause target cell (infected cell or malignant cells) apoptosis (Abbas et al., 2014). Target cell apoptosis by CTLs occur via the effector molecules secreted by CTLs such as Perforin and Granzyme, or via Fas-ligand (FasL) on CTLs binding to Fas on target cells (Abbas et al., 2014). CTLs that recognize target cells via their TCR-pMHC-I interaction, there is massive microtubule reorganization, and secretory granules containing Granzyme B and Perforin traffic towards the immunologic synapse (Voskoboinik, Whisstock, & Trapani, 2015). The secretory granules are then secreted by the CTLs at the immunologic synapse, along with directionally secreted cytokines onto the target cell (Huse, Lillemeier, Kuhns, Chen, & Davis, 2006; Voskoboinik et al., 2015). A second sub-dominant but independent mechanism of CTL-mediated killing involved the ligation of Fas ligand (FasL) on CTLs to Fas protein expressed on

target cells (C. Janeway, 2005). This directional secretion of effector molecules provide high sensitivity and specificity of target cell killing by CTLs, and are exploited in this dissertation (Chapter 5).

1.3 Primer on self/nonself-discrimination

The idea of individuality has been an exercise in intellectual debate beginning with philosophers such as Aristotle, and in the last century has morphed into the phenomenology of biological identity by evolutionary biologists (Clarke, 2010; Thomas Pradeu, 2012; Smith & Szathmary, 1997). At its core, immunology aims to provide mechanistic and conceptual understanding to the biological identity of an organism by its differentiation of self from nonself. Immunologic self/nonself discrimination vigorously protects the biological identity of a living entity by preserving its lifespan, and by extension, its genetic makeup.

The concept of the immunological self pervades the tree of life, from prokaryotes all the way up to vertebrates. The elegant CRISPR system in bacteria and archaea for example, is a prokaryotic adaptive immune mechanism aimed at protecting the bacterial “self” against invading bacteriophages (“nonself”) and other viruses (Barrangou et al., 2007). Regardless of the species and complexity of the immune system, there are commonalities to the conceptual immunological self. First, the immunological self has evolved mechanisms that are almost always harmful for an organism/entity that is perceived as the nonself (Cohn, 2010; Thomas Pradeu, 2012). For instance, in jawed vertebrates, as explained in previous sections, CTLs specifically target and kill infected and malignant cells with great precision (Section 1.2.4). Second, because the mechanism of defection by the immunological self is harmful, there must be a way for the immunological self to prevent itself from its own adverse actions, and reduce “off-target” effects mounted by an immune response against the immunological self to avoid debilitating conditions such as autoimmunity (Cohn, 2010). An example of this limiting self-reactivity criterion can be seen in the vertebrate adaptive immune system, where majority of autoreactive T-cells are eliminated via a process called negative selection during thymic development (Starr, Jameson, & Hogquist, 2003). This results in the dramatic reduction of the diversity in self-reactive TCRs (Yu et

al., 2015). Lastly, the immunological self must maintain cooperation and a status quo with other perceived foreign entities that are not harmful (and are in some cases beneficial) to the host. Examples of this type of interaction are the extension of T-cell tolerance against gut microbiome that is beneficial for humans (Round & Mazmanian, 2009).

In the last century, there have been several attempts to develop unifying immunological theories of self/nonself discrimination. Based on decades of previous research on transplantation and tumor immunology, Frank Burnett, conceptualized one of the earliest known thesis on immunologic self/nonself by proposing that early during development, an organism's immune system acquires the knowledge of self vs. nonself causing immunologic tolerance (F. M. Burnet, 1961; Ebert, 1970; C. A. Janeway Jr, Goodnow, & Medzhitov, 1996). These theories would later spur Peter Medawar's landmark experiments on transplanting donor splenocytes to neonatal mice to acquire immunological tolerance (section 1.1.1) (Billingham, Brent, & Medawar, 1953). This self/nonself theory which dominated the landscape of immune theories over much of the last century proposes that during development, the immune system actively learns to distinguish self and makes an immune response against the nonself foreign entity (C. A. Janeway Jr, 1992; Langman & Cohn, 2000). However, many foreign entities do not garner immune responses (gut microbiota for instance), insofar as there are immune responses often initiated against the self, observed in autoimmunity and immunopathologies (C. A. Janeway Jr et al., 1996). Furthermore, Matzinger and colleagues in 1996 were able to show that neonatal mice can make an immune response to foreign grafts using appropriate signals (derived from dendritic cells) in an apparent contradiction to Medawar's experiments (Ridge, Fuchs, & Matzinger, 1996).

These experiments challenged the notion that tolerance did not depend on the source of the nonself foreign antigen, but rather the context under which the nonself was visible to the self (Ridge et al., 1996). To address these limitations, Polly Matzinger proposed the "danger model" of immune system, which theorized that the immune system does not learn to distinguish self vs. nonself, but rather responds to the perceived presence or absence of danger (Matzinger, 1994, 2002). The danger theory invited much enthusiasm and critique, with respect to the definition of what constitutes molecular danger, and the notion that the immune system is pre-programmed to

respond to pathogen associated molecular patterns (PAMPs) along with the antigenic stimulus (Charles A. Janeway Jr, Goodnow, & Medzhitov, 1996; Thomas Pradeu & Cooper, 2012). Recently, a third theory, put forth by Pradeau and colleagues, tries to unify these theories by proposing that irrespective of the source of immune response (danger vs. self/nonself) it is the breakdown of non-antigenic patterns upon routine surveillance by immune cells such as CTLs, that an immune response can be made (T. Pradeu & Carosella, 2006; Thomas Pradeu & Cooper, 2012). This theory thus abandons the idea of the context of immune response, or the semantical definition of what constitutes self vs. nonself, but rather defines an entity (organism, antigen, or epitope) as immunogenic, if it constitutes a distinct pattern that is normally not encountered by the immune system, innate or adaptive (T. Pradeu & Carosella, 2006).

Regardless of the interpretation of these theories, any definition of what constitutes as nonself, or as danger, or as immunogenic, can only be conceptualized when the underlying biological parameters that make an immune response by the adaptive immune system are comprehensively elucidated. In context of human CTLs, the hallmarks of antigenic epitopes that contribute to a CTL immune response have not been fully discovered. In Chapter 2, I explore the biochemical parameters of epitopes from a variety of sources (pathogenic as well as self) that underlie what T-cells determine as worthy of an immune response. Understanding this phenomenon is particularly relevant when developing T-cell based therapies that target diseases of foreign, as well as of the self (such as autoimmunity, solid tumors).

1.4 Brief history of tumor immunology

Many of the fundamental immunology discoveries on tolerance and autoimmunity and infectious disease response were uncovered based on investigations into tumor immune response (examples in previous sections). This section will describe the history of cancer immunotherapy, major discoveries in the last decade, and current developments in the field.

1.4.1. Coley's toxins

William B. Coley (1862–1936), was an American pathologist and a physician at the Memorial Sloan Kettering cancer center (Moticka, 2016d). During the course of his treatments, Coley noted that occasionally patients would have their tumors regress following a high fever from a bacterial infection (Coley, 1891). Coley through a literature search found that there were close to 47 previous cases of physicians documenting occasional regression of tumors in patients following bacterial fevers (Coley, 1891). Based on these reports, Coley reasoned that the regression was perhaps due to the fever in response to the infection, although the biological mechanism under which that happened remained understandably obscure. Coley started injecting end stage cancer patients (in most cases with inoperable sarcomas) by injecting heat killed *Streptococcus pyogenes* directly into the tumors (Coley, 1898a).

The results of Coley's clinical trials were mixed (Moticka, 2016d). Coley was able to induce tumor regression by his method in several patients, although two patients died of the infection itself (Coley, 1891). Because of the danger of using *S.pyrogenes* strains in cancer patients, Coley also experimented with second less dangerous Enterobacter bacteria, *Serratia marcescens* in combination with *S.pyrogenes*. A subset (~ 10%) of patients in these trials benefitted from regression (Coley, 1895), with some patients succumbing to either the treatment or the disease (Coley, 1898b). However, because Coley did not maintain consistency with respect to preparation of the heat killed bacteria (up to 13 different mixture of strains were used), and administration or the toxins (intravenous vs. intramuscular vs. intratumoral), other physicians had difficulties replicating these trials and treatments (McCarthy, 2006). This resulted in criticism of his experiments, with other physicians and scientists suggesting that the response observed might just be natural regression rates, which prompted Coley to defend his techniques (Coley, 1895). One interesting observation in Coley's response was his remark that some of these patients with inoperable tumors (presumably late-stage metastatic disease) had complete long term benefit (by today's standards) after the administration of his toxins (Coley, 1895). In 1895, Coley wrote "*One of my cases has gone three and one-half years. This was a twice recurrent sarcoma of the neck and tonsil, with the patient in a most desperate condition, with no chance of*

living more than a few months" (Coley, 1895). Based on current knowledge of anti-tumor immunity, it has been proposed that a combination of a strong intratumoral innate inflammatory response mediated by proinflammatory cytokines such as interferon-α (IFNα), tumor necrosis factor α (TNFα), and interleukin 12 (IL-12) (Tsung & Norton, 2006), and cross reactive T-cells targeting neo-epitopes were recruited into these patient tumors causing sustained tumor regression (Snyder et al., 2014).

Despite these occasional but significant response observed in his patients, Coley's toxins were outcompeted by the ever growing prominence and usage of radiation therapy, and subsequently chemotherapy to treat cancers (McCarthy, 2006). In contrast to Coley's toxins which primarily worked best in sarcomas, radiation therapy, using X-rays discovered by Wilhelm Röntgen in 1895, caused therapeutic regression (although short term) of many different types of tumors consistently (Holsti, 1995; McCarthy, 2006). Despite these issues, Coley's studies ushered in early research on factors secreted by immune system in response to pathogenic infections that could have beneficial effects (Kienle, 2012; Tsung & Norton, 2006). For instance, bacterial lipopolysaccharide (LPS) and tumor necrosis factor α (TNFα) were discovered by conducting research on induced sarcoma in mice models (Carswell et al., 1975; Kienle, 2012). Post 1940, Coley's toxins were empirically tested in retrospective analyses, which indicated benefit for up to 57% of the patients treated with heat killed bacterial vaccines (Kienle, 2012). However, these studies did not rigorously undertake spontaneous remission rates to truly delineate the beneficial effects of Coley's toxins. Coley's early immunotherapy was also empirically tested in a randomized clinical trial in 1962, where 20/93 (21%) patients showed regression (Johnston & Novales, 1962). Because of the wide variability in response to heat killed bacterial vaccines, and as mentioned above the difficulties in replicating Coley's techniques, replacement of Coley's toxins with radiation therapy happened for cancer treatment starting in the 1900s (McCarthy, 2006). Post revival of tumor immunology as a viable therapeutic option, there are several clinical trials today evaluating heat killed bacterial vaccines as potential adjuvants for cancer therapy (Karbach et al., 2012).

1.4.2. Early tumor immunology and pre-checkpoint era

As mentioned in earlier sections, many fundamental discoveries made in the last century with respect to histocompatibility and the adaptive immune system were discovered because of studies in tumor rejection in mice (see section 1.1.1 on Tyzzer, Groer and Snell). Paul Ehrlich, known for his Nobel Prize winning contributions to the discovery of antibody, and one of the first adopters of chemotherapy, was also one of the first proponents of the concept that tumors could be recognized by the immune system (Himmelweit, 1958; Moticka, 2016b). Because of large gaps in knowledge with respect to tumorigenesis, as well as on autoimmunity, Ehrlich did not link antitumor immunity as being autoimmune in origin. Instead, Ehrlich proposed the concept of "horror autotoxicus", where immune response against self could not possibly happen because of the disastrous consequences to the organism's integrity (Moticka, 2016b). Later in 1900s, this would be proven otherwise because of the seminal work on tumor and graft transplantation undertaken by George Snell, Peter Medawar and Frank Burnett (sections 1.1.1-1.2.2).

Parallel to the work conducted by early transplantation immunologists, the finding by Paul Uhlenhuth that tissue specific proteins can induce the formation of highly unique and specific antibodies in various animals indicated that tumors could also develop unique tumor-specific antigens (Moticka, 2016d). In 1965, Gold and Freedman showed that extracted human colon cancer cells when injected into mice produced antibodies that reacted with colon and gastrointestinal malignant tissue, proving first experimental evidence of tumor-specific antigens (Gold, 1965; Gold & Freedman, 1965). With the advent of monoclonal antibody production by Georges Kohler and Cesar Milstein in 1965, cancer biology and immunology was given access to unprecedented view of cellular proteins in various model systems for research (Moticka, 2016d). Monoclonal antibody production allowed immunologists in particular to define the many different subsets of the immune system via their cell surface cluster of differentiation (CD) markers (Who Nomenclature Subcommittee, 1984).

In the mid-1970s, Steve Rosenberg and colleagues at the National Cancer Institute, USA pioneered the tumor immunology field with several seminal studies (Moticka, 2016d). First, they showed that interleukin 2 (IL-2) treated mouse splenocytes in *in vitro* cultures were cytotoxic to

mouse tumors when adoptively readministered (Yron, Wood, Spiess, & Rosenberg, 1980). Subsequently, the group demonstrated partial regression (21/55, ~38%) of tumors using autologous IL-2 primed T-cells from the peripheral blood of several metastatic melanoma cancer patients in what would be the first in human adoptively transferred T-cell human clinical trials (S. A. Rosenberg et al., 1985). The Rosenberg group was also the first to demonstrate in mouse models that tumor infiltrating lymphocytes (TILs) were 50-100 times more cytotoxic against autologous tumors, which would later be revealed to be a tumor-specific CTL-memory response (S. Rosenberg, Spiess, & Lafreniere, 1986). In the 1970s, with the discovery of viral mediated tumors in the form of the human papillomavirus (HPV) (zur Hausen, Gissmann, Steiner, Dippold, & Dreger, 1975), Epstein Barr Virus (EBV), and Hepatitis B virus (HBV) (Buynak, 1976), indicated further that the adaptive immune system could be used against treating tumors.

1.4.3. Advances in tumor biology

During 1900s, there were also fundamental breakthroughs made in understanding tumor biology and tumorigenesis as a phenomenon. For instance, Peter Nowell and David Hungerford discovered the Philadelphia chromosome, a major chromosomal abnormality with translocation between chromosomes 9 and 22 in more than 95% of patients with chronic myeloid leukemia (CML) (Koretzky, 2007). PC Nowell would also later propose the seminal hypothesis on the clonal origins of tumorigenesis via acquired somatic mutations (Nowell, 1976). Subsequently, oncogenes (e.g. Her2) (Nowell, 1976; Slamon et al., 1989), tumor suppressor genes TP53 (Oren & Levine, 1983), and the retinoblastoma gene (Rb) (Cavenee et al., 1983) were discovered. All these seminal studies on the nature, the biology and the mutational spectrum of cancer would later be synthesized by Robert Weinberg and Douglas Hanahan into the seminal work of the hallmarks of cancer (D. Hanahan & Weinberg, 2000). However, between 1990-2010, the tumor immunology field would make several fundamental contributions and milestones to the field of both cancer biology and immunology that Weinberg and Hanahan would include Tumor immune evasion and Tumor promoting inflammation as two emerging hallmarks in their revised version of the hallmarks of cancer (Douglas Hanahan & Weinberg, 2011).

1.4.4. Tumor immune editing

In the 1980s, the incidence of acquired immunodeficiency (AIDS) cases caused by the human immunodeficiency virus (HIV) rose to epidemic proportions in the US LGBT community predominantly due to their manifestation as Kaposi's Sarcoma (KS), a viral mediated cancer (Haverkos & Curran, 1982). A major outcome out of the research on HIV-AIDS is the finding that the virus caused massive systemic immune suppression in the infected individuals, which in turn caused the patients to succumb to opportunistic infections such as KS (Haverkos & Curran, 1982). However, HIV-induced KS also highlighted the dominant role of the human immune system to seek and destroy any cancer that could have been formed otherwise. It also implicated that there could be immune evasion mechanisms that are actively employed by viruses and cancers that allow them to persist, and in the subsequent decade these would become the subject of intense investigations.

Frank Burnett in 1957 proposed that tumors might be constantly forming and regressing in an individual because of an immunologic reaction to tumor specific antigens (M. Burnet, 1957; Dunn, Bruce, Ikeda, Old, & Schreiber, 2002). However, because of several failures to adequately test the hypothesis and in lack of an in depth understanding of the adaptive immune system, it would take more than 30 years before several groups revisited the immune surveillance hypothesis. For instance, nude mice lacking a functional immune system were prone to form more tumors when injected with a chemically transformed mouse carcinoma (Engel et al., 1996). The discovery that interferon-γ (IFNγ) produced endogenously or when administered to mice resulted in tumor regression in mice undergoing transplantable tumors suggested a T-cell mediated control of many tumors (Dighe, Richards, Old, & Schreiber, 1994). Similar results were obtained with C57BL/6 mice lacking perforin, effector molecules secreted by CTLs to initiate target cell lysis (van den Broek, 1996), as well as recombination activating gene (RAG-1, RAG2) deficient mice which have profound immunodeficiency (Shinkai et al., 1992). The human relevance of these studies was clearly observable in immunosuppressed HIV-1+ AIDS patients (Haverkos & Curran, 1982), and transplant patients, with both groups being susceptible to higher incidence of many different cancers (Gatti & Good, 1971; Penn & Staezl, 1972). Synthesizing all

these results, Robert Schreiber and colleagues proposed “cancer immunoediting”; the idea that in immunocompetent individuals, tumor evolution is sculpted by the host immune response, giving rise to tumors that are heavily resistant to immune targeting (Dunn et al., 2002).

Cancer immunoediting involves three major phases; 1) Elimination 2) Equilibrium and 3) Escape (Dunn et al., 2002). In the early elimination phase, tumors are immunogenic and are susceptible to host adaptive immune response. The coordinated innate and adaptive response is in part due to tumor intrinsic genomic instability and other hallmarks of the tumor (Dunn et al., 2002). In this phase, the tumors also express tumor-specific/associated antigens, or tumor-specific neoepitopes that are derived from the antigen presentation of non-synonymous mutations on HLA class I molecules. Tumors in this phase can also be susceptible to CD4+ T-cells if they

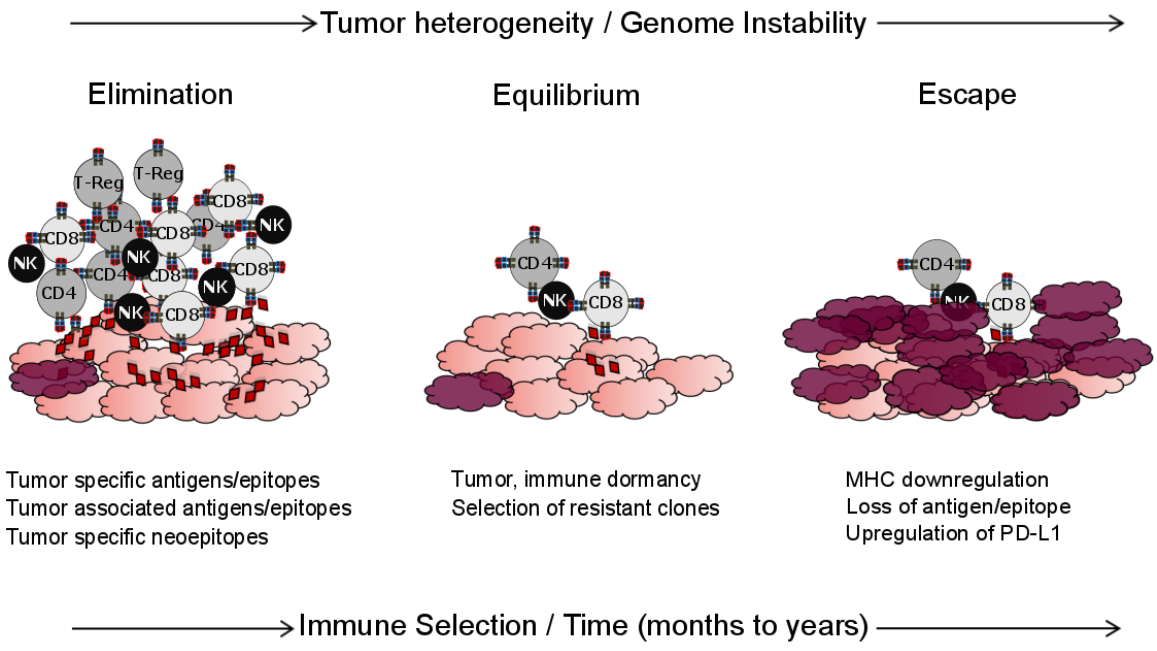

Figure 1-4. The Cancer Immunoediting process. Elimination (left panel) process occurs due to normal immune surveillance where in tumor infiltrating immune cells eliminate immunogenic clones. Effector molecules (e.g. Perforin, IFNγ) are secreted by immune cells. Resistant clones (pre-existing or arising de novo, in magenta) survive and exist in a dynamic equilibrium phase with the immune cells, and strength of immune response decreases. In Escape phase (right panel), occurs with clonal expansion of the resistant clones (select examples of mechanisms are shown). Figure was adapted with modifications from Dunn et al., 2002.

are antigenically processed and presented by APCs, NK cells if they lack antigen presentation, and other various different types of immune-mediated cell death (Dunn et al., 2002). Thus, the elimination phase is a direct consequence of the immune surveillance hypothesis proposed by Burnett and others (Figure 1.4, left panel). In the equilibrium phase, a darwinian selection of immune-resistant clones allows the temporal and spatial survival of the tumor (Figure 1.4, middle panel). Because there is a loss of the immunogenic clones from the tumor due to immune elimination, and a subsequent loss of tumor infiltrating immune cells, there is a dynamic equilibrium that persists over several years between the immune system in the tumor (Dunn et al., 2002). In this phase, it is likely that the resistant clone was pre-existing, or arose *de novo* due to the immune selection pressure (Figure 1.4, middle panel). In the escape phase, variant clones are now resistant and insensitive to immune detection and elimination. This process can involve for instance *de novo* or pre-existing variant clone upregulation of the programmed death ligand 1 (PD-L1) in response to IFNγ secreted by TILs (Tumeh et al., 2014), or loss of the entire antigen/epitope that resulted in the CTL-mediated tumor cell death of the dominant immunogenic clone (Matsushita et al., 2012; Tran et al., 2016). Other mechanisms of immune escape and resistance have been described elsewhere (Schreiber, Old, & Smyth, 2011), and are currently the subject of intense investigation (Pardoll, 2012b; Schreiber et al., 2011). The resistant clones can thus largely avoid immune detection and escape, metastasize, and colonize the body. The cancer immunoediting hypothesis was seminal in that it gave an evolutionary framework to understand cancer immune response, based on clonal origins of tumorigenesis suggested by PC Nowell (section 1.4.3). Today, cancer immunoediting is widely applied to conceptualize and test immune based therapies. In Chapter 4, I apply this framework to investigate the mechanisms of immune resistance employed in response to the HPV-infection in head and neck cancers.

1.4.5 The checkpoint era

The logical extension from studies on tumor immunity, tolerance, and immunodeficiency observed in chronic viruses such as HIV-1, was the idea that there should be immune regulatory molecules

that rein in the efficacy of T-cell mediated immune response. Such biochemical pathways could in be theory leveraged by self-reactive T-cells to avoid causing immunopathology, a major cause for many of the disease symptoms observed in autoimmune conditions.

The first such discovered molecule was cytotoxic T-lymphocyte–associated antigen 4 (CTLA-4) in 1987 by Brunet et al (Brunet et al., 1987). CTLA-4 is expressed on 90% of human CD4+ T-cells and in about 50% of human CD8+ T-cells (Peggs, Quezada, Korman, & Allison, 2006). Because of CTLA-4's homology to CD28 (a co-stimulatory molecule expressed on T-cells required for adequate T-cell stimulation) (Linsley, 1991), CTLA-4's apparently redundant function was under subject of investigation (Pardoll, 2012). However, Jim Allison, Jeffrey Bluestone and colleagues performed *in vitro* antibody cross linking experiments that suggested that CTLA-4 might be a negative regulator of T-cell stimulation (Krummel & Allison, 1995; Walunas et al., 1994). CTLA-4-deficient mice (CTLA-$4^{-/-}$) developed by Arlene Sharpe's group, and independently by Tak Mak and colleagues ultimately confirmed these findings by demonstrating that CTLA-$4^{-/-}$ mice displayed massive inflammatory conditions, characterized by severe lymphadenopathy and lymphoproliferation in all organs (Tivol et al., 1995; Waterhouse et al., 1995). Subsequently, CTLA4 antibodies that can partially block CTLA-4 functions in vivo developed by Allison and colleagues showed that substantial tumor rejection can be achieved in mice models (Leach, Krummel, & Allison, 1996; Sutmuller et al., 2001). The results highlighted the effect and extent of T-cells that are present in the periphery which could be leveraged to treat tumors (Pardoll, 2012).

Functionally, CTLA-4 has been thought to inhibit T-cell function in two ways : 1) by recruiting alternative phosphatases to the TCR attenuating the strength of T-cell activation (Lee et al., 1998); 2) by directly competing for co-stimulatory molecules (CD80, CD86, due to its homology to CD28) on APCs making them unavailable for CD28-ligation (Qureshi et al., 2011). Biochemical and molecular functions of CTLA-4 is still a subject of active investigation. CTLA-4 excited, and revived tumor immunology field spurring investigative humanized antibodies against CTLA-4 that were clinically developed by Medarex in collaboration with Jim Allison (later acquired by Bristol-Myers Squibb) (Pardoll, 2012). The first humanized immune checkpoint blockade antibody Ipilimumab (against CTLA-4) was approved by the FDA for the treatment of recurrent

metastatic melanoma in 2011 (Table 1-1) after a decade of clinical trials (Phan et al., 2003; Robert et al., 2011). In terms of response, Ipilimumab, although only effective in ~20% patients, causes long term complete regression of all metastatic lesions, with some stage III/IV melanoma patients surviving longer than 10 years (Pardoll, 2012).

The second major immune checkpoint molecules to be discovered were programmed death protein 1 (PD-1) and its ligand PD-L1. PD-1 was first discovered as a molecule that was thought to regulate programmed cell death in T-cells (Ishida, Agata, Shibahara, & Honjo, 1992). Although PD-1 knockout mice PD-$1^{-/-}$ did not develop fulminant lymphadenopathy as CTLA-$4^{-/-}$ mice, PD-$1^{-/-}$ mice upon aging would eventually develop strain specific and organ specific autoimmune reactions manifesting largely as Lupus like syndromes (Nishimura et al., 2001; Nishimura, Nose, Hiai, Minato, & Honjo, 1999). These results suggested a more fine-tuned regulatory function to PD-1 expressing T-cells. Subsequently, the ligands for PD-1 expressed on many dendritic cells, PD-L1 was discovered by Lieping Chen and colleagues (H. Dong, Zhu, Tamada, & Chen, 1999). Gordon Freeman and colleagues would then demonstrate that PD-1/PD-L1 interaction delivered inhibitory signals to responding T-cells (Freeman et al., 2000).

In contrast to the CTLA-4 pathway, PD-1/PD-L1 acted on the effector phase of T-cell stimulation, and was shown in several elegant studies by Rafi Ahmed and colleagues that in chronic virus systems, antigen persistence would result in the phenomenon of T-cell exhaustion (Barber et al., 2006; Day et al., 2006). T-cell exhaustion was characterized by the gradual loss of effector functions in antigen-specific CD8+ (and also CD4+) T-cells, with a distinct molecular signature resulting in impaired control of chronic virus infections in human and mice (Wherry et al., 2007; Zajac et al., 1998). Blockade of PD-1 restored (at least partially) functional exhaustion in CTLs and can resume T-cell mediated elimination of infections (Barber et al., 2006). Chen and colleagues would then demonstrate that PD-L1 expressing tumors evade immune response in mice models by causing T-cell apoptosis, establishing the link between PD-1 pathway and tumor immunity (Haidong Dong et al., 2002). As with CTLA-4, investigations into blocking the PD-1 checkpoint pathway was subsequently begun, resulting in several clinical trials again beginning

with melanoma (Brahmer et al., 2010; Topalian et al., 2012). Today, PD-1 antibodies Nivolumab and Pembrolizumab are among the most successful immune checkpoint blockade antibodies

Table 1-1. Selection of FDA approved cancer immune therapies

Name	Target	Type of Cancer	Year
Rituximab	CD20	Non-Hodgkin lymphoma	1997
Trastuzumab	HER2	Breast Cancer	1998
Ibritumomab	CD20	Non-Hodgkin lymphoma	2002
Cetuximab	EGF Receptor	Colorectal Cancer	2004
Bevacizumab	VEGF	Colorectal Cancer	2004
Panitumumab	EGF Receptor	Colorectal Cancer	2004
Ofatumumab	CD20	Chronic lymphocytic leukemia	2009
Ipilimumab	CTLA-4	Metastatic melanoma	2011
Brentuximab	CD30	Hodgkin lymphoma	2011
Pertuzumab	HER2	Breast Cancer	2012
Obinutuzumab	CD20	Chronic lymphocytic leukemia	2013
Ramucirumab	VEGF receptor 2	Gastric cancer	2014
Blinatumomab	CD19 and CD3	ALL, Melanoma	2014
Dinutuximab	GD2	Neuroblastoma	2015
*Pembrolizumab, Nivolumab	PD1	Melanoma NSCLC** Gastric Cancer**, HNSCC** Renal, Bladder Cancers MMR deficient, MSI positive CC Hodgkin lymphoma	2014 2016 2017 2017 2017 2017
Tisagenlecleucel	CAR T-cell therapy	ALL	2017

HER2, human epidermal growth factor receptor; EGF, epidermal growth factor; VEGF, vascular endothelial growth factor; CTLA-4, cytotoxic T lymphocyte antigen 4; PD1, programmed cell death protein 1; GD2,; Non-small cell lung cancer, NSCLC; ALL, Acute Lymphocytic Leukemia; MMR, Mismatch repair; MSI, Microsatellite instability. *Includes combined approvals for both Pembrolizumab and Nivolumab; ** Approved for recurrent metastatic cancers at present.
Data current as of October 2017. Adapted from (Moticka, 2016d).

pursued in cancer immunotherapy with successful FDA approval for more than six cancers (Table 1-1). PD-L1 and other checkpoint molecules are also being pursued for several ongoing clinical trials, as are other immune therapeutic modalities such as CAR T-cell, adoptive T-cell therapies.

The dramatic, persistent, long-term response to CTLA-4 and PD-1 blockade antibodies in the treatment of select human cancers is partly attributed to the memory compartment of T-cell immunity, where recurring immunogenic lesions displaying antigens can be readily recognized and kept under control by a secondary immune response (Pardoll, 2012). However, the evolving fitness landscape of a tumor under cancer immunoediting hypothesis (section 1.4.4.) posits that some cancers will have resistant clones that can successfully thwart this immune assault. Thus, although checkpoint blockade immunotherapies are effective, objective long term clinical responses only occur in a subset of patients (~20% in CTLA-4 for melanoma, between 30%-40% for PD-1 blockade in melanoma and other cancers) (Pardoll, 2012a, 2012b). Thus, it is highly likely that neoantigen and immune landscape of a tumor evolves as a result of specific tumor biology in addition to the immune heterogeneity observed in individuals (Schumacher & Schreiber, 2015). These observations argue for developing tumor-specific and perhaps patient-specific immunotherapy regimens based on the tumor microenvironment and antigenic landscape of each tumor. The emphasis of current cancer immunotherapies has thus been to:

- Understand response to checkpoint blockade immunotherapies to dissect mechanisms of immune dysfunction
- Improve immunogenicity of weakly immunogenic tumors. This involves exploring the antigenic landscape of various tumors to identify epitope/neoepitope vulnerabilities.
- Understand the limits of tolerance and self-reactivity in context of T-cells.
- Understand the tumor microenvironment in context of the local factors influencing and editing the tumor immune response.

In Chapter 4, I explore the landscape of antigenic epitopes and immune dysfunction in tumor microenvironment in the setting of HPV-associated head and neck cancers to develop better immune therapies.

1.5 Key thesis contributions

This dissertation answers several fundamental questions in T-cell immunology, tumor immunology, and HPV-immunology, as well as develops novel techniques for applications in translational immunology. The key findings and contributions are listed as follows:

1. The discovery that immunogenic MHC class I epitopes are characterized by a higher frequency of relatively hydrophobic amino acids at specific TCR-contact residues compared to non-immunogenic self-peptides, leading to a general biochemical parameter for T-cell self/nonself discrimination (Section 1.3) by CTLs (Chapter 2).
2. The demonstration that amino acid biochemical properties in particular hydrophobicity, can be used to enhance the efficiency of prediction of immunogenicity of CTL-epitopes from any given antigen, and can correlate with epitope immunodominance hierarchies (Chapter 2).
3. The discovery that human papillomavirus 16 (HPV16) antigen E2 elicits broad T-cell and B-cell reactivity in HPV+ HNSCC patients and is expressed in a subset of HPV+ HNSCC patients, making HPV16-E2 is a potential immunotherapeutic target for HPV-associated malignancies. (Chapter 4).
4. The first comprehensive experimental definition of the landscape of HPV16 CTL-epitopes from E2, E6 and E7 across 12 different globally frequent HLA class I alleles from HNSCC patients (Chapter 4).
5. The discovery that low immunogenicity of HPV16-E7 may be tied to the relatively higher levels of dysfunctional E7-specific CTLs compared to E2, E6-CTLs observed in HPV+ HNSCC patients (Chapter 4).
6. The computational and experimental demonstration that the immunoregulatory enzyme indoleamine 2,3-dioxygenase (IDO-1) represents an HPV-specific immune evasion mechanism and is highly expressed in HPV-related malignancies (Chapter 4).
7. The first mechanistic and experimental demonstration that IDO-1 inhibition can individually as well synergistically in combination with PD-1 blockade enhance the

cytotoxic potential of HPV-specific CTLs on HPV+ HNSCC cells, representing a new potential immunotherapeutic modality for HPV+HNSCCs (Chapter 4).

8. The development of a novel single cell T-cell assay that can be employed to identify immunogenic T-cell antigens from whole pathogenome cDNA libraries (Chapter 5).

CHAPTER 2

TCR CONTACT RESIDUE HYDROPHOBICITY IS A HALLMARK OF IMMUNOGENIC CD8+ T CELL EPITOPES

This chapter is published:

Diego Chowell[1], Sri Krishna[1], Pablo D. Becker, Clement Cocita, Jack Shu Xuefang Tan, Philip D. Greenberg, Linda S. Klavinskis, Joseph N. Blattman, and Karen S. Anderson. Proceedings of the National Academy of Sciences 112.14 (2015): E1754-E1762

[1]D.C. and S.K. contributed equally to this work

Abstract

Despite the availability of major histocompatibility complex (MHC)-binding peptide prediction algorithms, the development of T-cell vaccines against pathogen and tumor antigens remains challenged by inefficient identification of immunogenic epitopes. CD8+ T cells must distinguish immunogenic epitopes from non-immunogenic self-peptides to respond effectively against an antigen without endangering the viability of the host. Because this discrimination is fundamental to our understanding of immune recognition and critical for rational vaccine design, we interrogated the biochemical properties of 9,888 MHC class I peptides. We identified a strong bias toward hydrophobic amino acids at T-cell receptor contact residues within immunogenic epitopes of MHC allomorphs, which permitted us to develop and train a hydrophobicity-based artificial neural network (ANN-Hydro) to predict immunogenic epitopes. The immunogenicity model was validated in a blinded in vivo overlapping epitope discovery study of 364 peptides from three HIV-1 Gag protein variants. Applying the ANN-Hydro model on existing peptide-MHC algorithms consistently reduced the number of candidate peptides across multiple antigens and may provide a correlate with immunodominance. Hydrophobicity of TCR contact residues is a hallmark of immunogenic epitopes and marks a step toward eliminating the need for empirical epitope testing for vaccine development.

2.1 Significance Statement

The design of effective T-cell vaccines against pathogens and tumor antigens is challenged by the highly inefficient identification of the subset of peptides from a given antigen that effectively stimulate an immune response. Here we report that the relative hydrophobicity of T-cell receptor contact residues is markedly enriched in immunogenic major histocompatibility complex class I epitopes in both human and murine MHCs, and in both self and pathogen-derived immunogenic epitopes. Incorporating hydrophobicity into T-cell epitope prediction models increases the efficiency of epitope identification, which will manifest in the time and cost of T-cell vaccine development. Amino acid hydrophobicity may represent a biochemical basis by which T cells discriminate immunogenic epitopes within the background of self peptides.

2.2 Introduction

The interaction of CD8+ T cells with peptide-MHC complexes (pMHCs) is a key event in the development of cell-mediated immunity (Grakoui, 1999). MHC class I (MHC-I) molecules typically present 8-11 aa peptides derived predominantly from proteasomal degradation of intracellular proteins, either self-peptides or infection-derived antigens (Blum, Wearsch, & Cresswell, 2013). T cell receptors (TCRs) from CD8+ T cells bind antigenic pMHC molecules, triggering a downstream signaling cascade that leads to T cell activation, differentiation, and ultimately to cytolysis of target cells presenting the same epitope (Hennecke & Wiley, 2001). Vaccines and immunotherapies for the treatment of infection and cancer seek to incorporate cytotoxic T cell (CTL) epitopes, but defining such epitopes remains a costly and arduous process (Purcell, McCluskey, & Rossjohn, 2007). Understanding the molecular basis of TCR-pMHC recognition will aid discovery of immunogenic epitopes in infectious and autoimmune disease.

During thymic development, CD8+ T cells undergo both positive and negative selection to acquire the ability to discriminate antigenic peptides from self-peptides (Hogquist et al., 1994). Costimulatory signals can enhance this discrimination (Medzhitov & Janeway, 2002), but a primary event that triggers CD8+ T cell activation is the non-covalent pMHC-TCR interaction. Proteasomal cleavage patterns and binding affinities of peptides to different MHCs have been

extensively studied (Falk, Rötzschke, Stevanović, Jung, & Rammensee, 1991; Kubo et al., 1994; Rammensee et al., 1999). In contrast, the biochemical basis of immunogenic epitopes is less well-defined (van der Merwe & Dushek, 2010). T cell epitope discovery is complicated by the codominance and polymorphism of MHC alleles, diversity of antigens (both infectious and self-antigens), limited mass spectrometry-based confirmation of MHC-bound peptides, and a scarcity of experimentally confirmed immunogenic epitopes within the infectious and self-proteome (Purcell et al., 2007). As a result, T cell epitope prediction algorithms have focused on aa binding affinity for specific MHC-motifs and the proteins proteasomal cleavage pattern to identify candidate T cell epitopes (Honeyman, Brusic, Stone, & Harrison, 1998; Moutaftsi et al., 2006; Nielsen et al., 2007; Tenzer et al., 2005). Although computational tools have improved over the past decade, they have not been trained to predict immunogenicity. The major limitation when using such prediction algorithms is the presence of a significant number of binders from a given antigen that will never lead to an immune response (Newell et al., 2013). Thus, immunogenic CTL epitopes fulfill additional criteria that go beyond antigen processing and MHC-binding.

Here, we sought to identify the biochemical criteria that define immunogenicity within the subset of MHC-I binding peptides. Using a curated repository of MHC-I epitopes from the Immune Epitope Database (IEDB) (Vita et al., 2009), we evaluated the biochemical properties of aas that discriminate between immunogenic epitopes and non-immunogenic self-peptides. We found a strong bias towards hydrophobicity in aa residues of immunogenic CTL epitopes that is highly selective for exposed TCR contact residues. Using these criteria, we trained an artificial neural network (ANN) model to identify immunogenic CTL epitopes from a data set and empirically assessed our prediction model for 3 human immunodeficiency virus 1 (HIV-1) Gag protein variants in a murine model of immunogenicity. We demonstrate the utility of this ANN model, which has the potential to significantly enhance the efficiency of T cell epitope discovery.

2.3 Results

2.3.1. Amino Acid Use Differs Between Immunogenic and non-immunogenic Peptides

CTLs recognize immunogenic epitopes from a background of poorly immunogenic self-peptides. To understand the biochemical basis of differences between these two classes of peptides, we retrieved all known MHC class I-binding peptides reported as T cell reactive (hereafter immunogenic) and self-peptides from MHC-ligand elution experiments with no known immunogenicity (hereafter non-immunogenic) from IEDB. Any eluted peptide that was immunogenic (either pathogen derived or self-antigen derived) was excluded to generate two mutually exclusive datasets that avoid any potential bias. Out of the 34,586 total retrieved peptides from IEDB, 5,035 8-11mer non-redundant peptides were reported to be immunogenic and 4,853 were non-redundant non-immunogenic and were used for further analysis. Frequency distributions of aas in 8-11mer immunogenic and non-immunogenic peptides showed significant variability in aa composition (Fig. 2-1-A).

To identify overrepresentation of certain aa's in immunogenic epitopes, we computed a probability ratio for each aa. We then performed a correlation analysis between the probability ratio of each aa and three major biochemical properties using independent numeric scales: hydrophobicity (Kyte-Doolittle) (Kyte & Doolittle, 1982), polarity (Grantham) (Grantham, 1974), and side chain bulkiness (Zimmerman) (Zimmerman, Eliezer, & Simha, 1968) (Table A-1). We found a strong, statistically significant correlation between probability ratios and hydrophobicity values (Spearman $\rho = 0.71$, $P = 4.24\times10{-}4$; (Fig. 2-2 A). Similarly, we also found a negative correlation between probability ratios and polarity of aas (Spearman $\rho = -0.77$, $P = 6.97\times10{-}5$; (Fig. 2-2 B)), with highly polar aas being underrepresented in immunogenic epitopes. No significant correlation was observed with aa side chain bulkiness (Fig. 2-2 C). Most of the overrepresented and strongly bulky aas were also strongly hydrophobic. Cysteine, a non-polar hydrophobic aa was an outlier in the immunogenic dataset. Two potential sources of bias in our analyses were the variation in peptide length of MHC-I peptides and the dominance of human leukocyte antigen (HLA) A2 epitopes within existing databases. We analyzed on the 9mer epitopes (Fig. 2-1 B) and HLA class I restricted peptides excluding HLA-A2 peptides (Fig. 2-1 C).

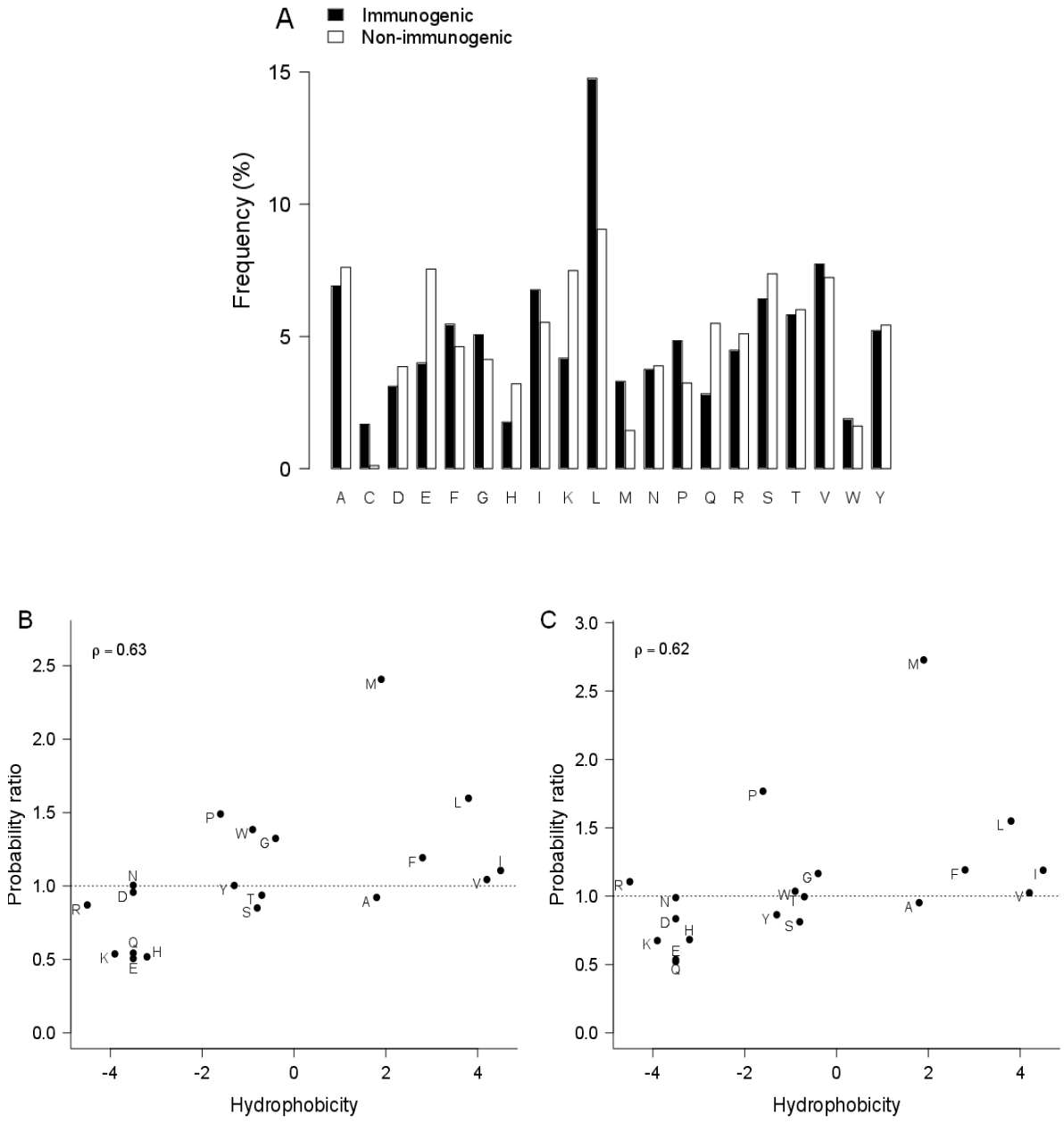

Figure 2-1. Bias in amino acid usage between immunogenic and non-immunogenic MHC-I peptides. (A) Comparison of frequency distributions of amino acids between immunogenic and non-immunogenic datasets. (B) Probability ratio (P(x I immunogenic)/P(x I non-immunogenic)) of each amino acid as a function of its hydrophobicity, analyzed on just 9mer MHC-I peptides. (C) Probability ratio (P(x I immunogenic)/P(x I non-immunogenic)) of each amino acid as a function of its hydrophobicity, analyzed on 9mer HLA-I peptides excluding HLA-A2 restricted peptides.

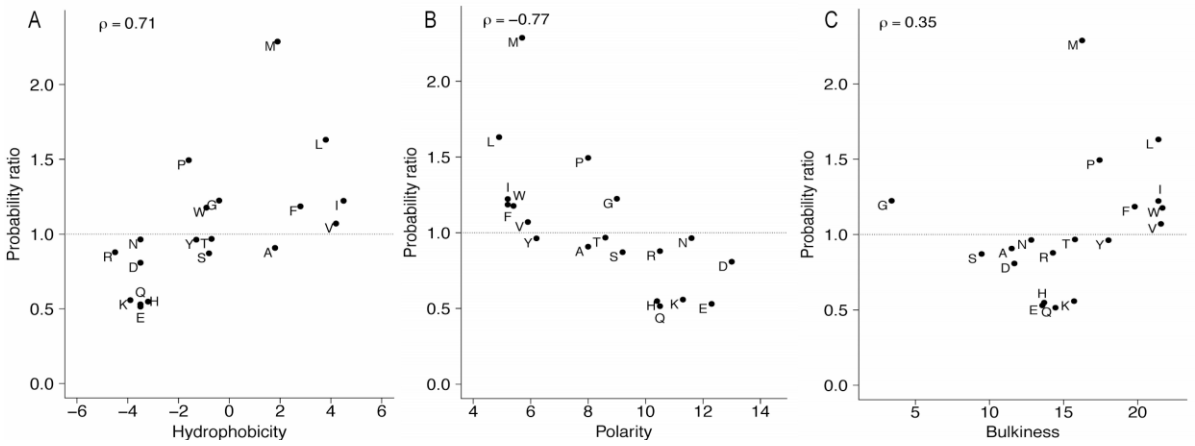

Figure 2-2. Probability ratio of each amino acid as a function of its corresponding biochemical property. Each probability of each amino acid was computed from the frequency distribution of immunogenic epitopes and non-immunogenic peptides. Biochemical properties analyzed were (A) hydrophobicity, (B) polarity, and (C) side-chain bulkiness. A probability ratio >1 indicates overrepresentation of the amino acid in the immunogenic dataset. The overrepresented outlier cysteine (C) was omitted for scale. Spearman correlations coefficients (ρ) are shown.

2.3.1. Hydrophobicity Bias in Selective TCR Contact Residues

We first compared the mean hydrophobicity of each residue between immunogenic and non-immunogenic peptides using the Kyte-Doolittle numeric hydrophobicity scale. Immunogenic 9mer epitopes were significantly more hydrophobic than non-immunogenic 9mer peptides at each residue ($P < 1.6\times10^{-5}$; (Fig. 2-4 A) and (Table A-2)). We observed similar results in 10mer peptides ($P < 2\times10^{-7}$ at every residue; (Fig. 2-3 A)), and within HLA-A2 excluded 9mer and 10mer subsets (Figs. 2-3 B and C). Because the immunogenic dataset is biased to pathogen-derived immunogenic epitopes, we performed similar analyses between immunogenic self-epitopes and non-immunogenic self-peptides ($P < 1\times10^{-4}$ at all residues, except P5 and P6; (Fig. 2-3 D)). We further compared immunogenic HLA-A2 restricted 9mer epitopes derived from pathogens with those derived from self-antigens and observed no significant difference in hydrophobicity ($P > 0.05$ at each aa residue except P6, $P = 0.04$; (Fig. 2-3 G)) revealing that T cells that escape thymic deletion recognize self-peptides with hydrophobicity profile that is virtually the same as that of pathogen-derived epitopes. Lastly, to evaluate if there is potential

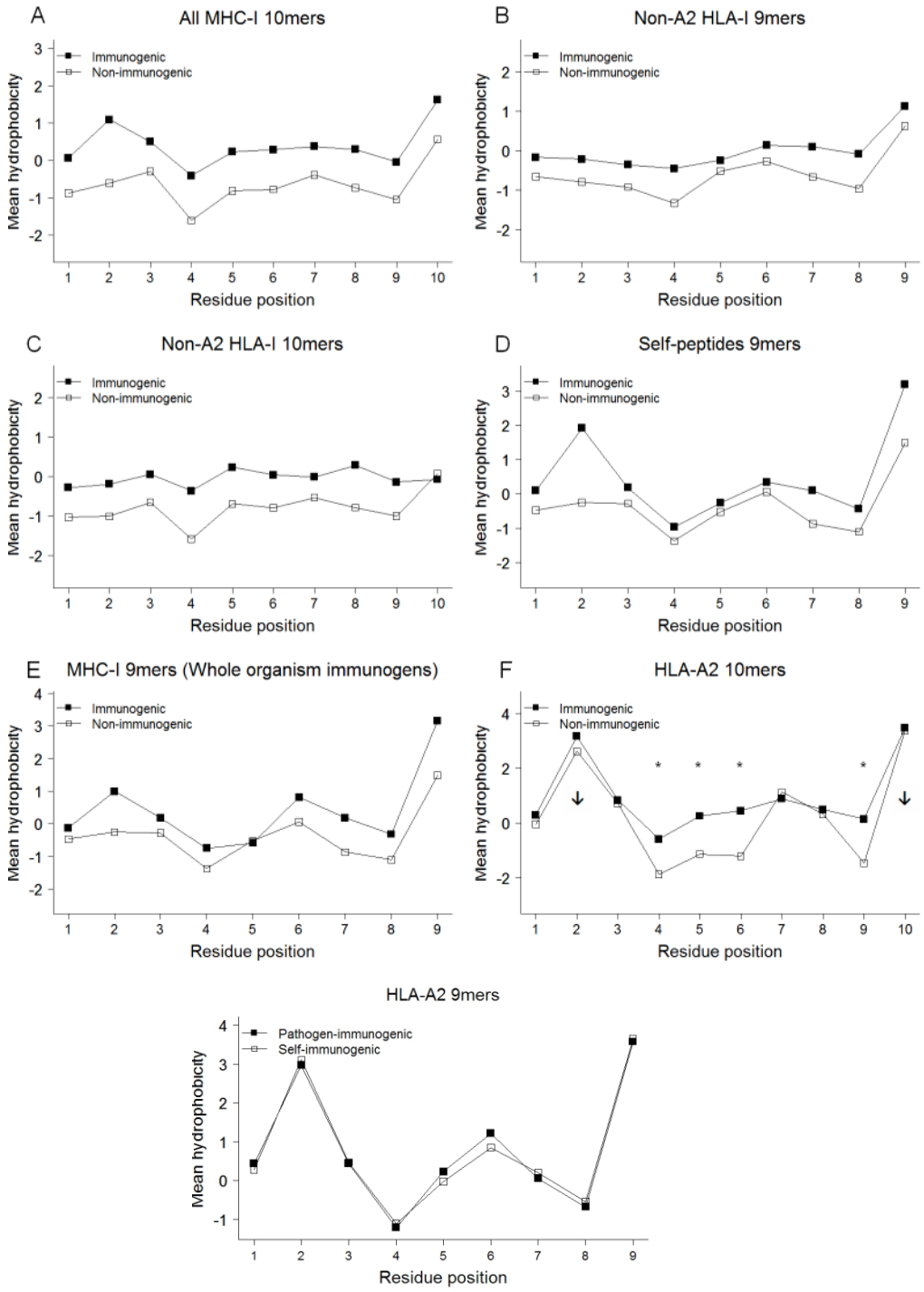

Figure 2-3. Immunogenic pathogen-derived and self-epitope hydrophobicity. Each peptide sequence was transformed into numeric sequence based on hydrophobicity and the mean hydrophobicity at each position was computed. (*A*) Immunogenic and non-immunogenic MHC-I 10mers; all residues $P < 2x10^{-7}$. (*B*) HLA-I immunogenic and non-immunogenic 9mers excluding HLA-A2 epitopes. (*C*) Human HLA-I immunogenic and non-immunogenic 10mers excluding HLA-A2 epitopes. (*D*) Immunogenic and non-immunogenic MHC-I 9mer self-peptides. (*E*) MHC-I 9mers peptides discovered using whole organism as immunogen as opposed to peptide-

immunization experiments (non-immunogenic dataset – same as Fig. 2A) (*F*) Human HLA-A2 restricted immunogenic and non-immunogenic 10mers with arrows indicating anchor residues and stars for $P < 0.005$. (*G*) Human HLA-A2 restricted immunogenic pathogen-derived and immunogenic self 9mer epitopes. *P*-values for each figure were obtained using Wilcoxon rank-sum test and are shown in Table A-2.

bias created by using peptide immunization experiments, we did the same analysis using immunogenic epitopes identified using whole "Organism" as the immunogen ($P < 0.01$ at all residues except P1, P5 (Fig. 2-3 E)). Thus, our results demonstrate a preference for hydrophobicity in immunogenic epitopes across antigenic sources (self and pathogen) and MHC molecules (HLA-A2 and non-HLA-A2).

The locations of anchor residues and TCR contacts have been mapped for many MHC peptides (Rudolph, Stanfield, & Wilson, 2006). If the observed bias toward non-polar hydrophobic aa's within immunogenic epitopes affects TCR affinity, we predicted that it would be selective for TCR contact residues. We analyzed the mean hydrophobicity along the peptide for the highest represented MHC epitopes within the database: HLA-A2 (Fig. 2-4 B), and murine MHC H-2D^b and H-2K^b (Fig. 2-4 C and D). HLA-A2 restricted 9mer peptides are anchored at residues P2 and P9, with P6 as an auxiliary anchor. We observed no statistical difference in hydrophobicity between the anchor residues of immunogenic and non-immunogenic peptides (P2, $P = 0.9$; P9, $P = 0.08$; (Fig. 2-4 B)). The observed difference in hydrophobicity was at specific TCR contact residues P4, P7 and P8 (P4, $P = 6.3\times10{-}12$; P7, $P = 5\times10{-}13$; P8, $P < 2.2\times10{-}16$). In contrast, the auxiliary anchor P6 was more hydrophobic in non-immunogenic peptides ($P = 3.1\times10{-}7$). Similar results were found for HLA-A2 restricted 10mer peptides (Fig. 2-3 F).

To determine if the difference in hydrophobicity was species-specific, we evaluated the subset of known mouse MHC H-2Kb restricted 8mer peptides. Again, we observed a marked increase in relative hydrophobicity for the TCR contact residues P6 and P7 of immunogenic epitopes (P6, $P = 7 \times 10{-}5$; P7, $P = 1.1 \times 10{-}6$) but no difference in anchor residues (P5, $P = 0.67$; P8, $P = 0.15$; (Fig. 2-4 C)). As observed with HLA-A2, the auxiliary anchor residue P3 was more hydrophobic in non-immunogenic peptides ($P = 0.005$). Finally, we analyzed mouse MHC H-2Db restricted 9mer peptides and observed that P7 and P8 TCR contact residues were more

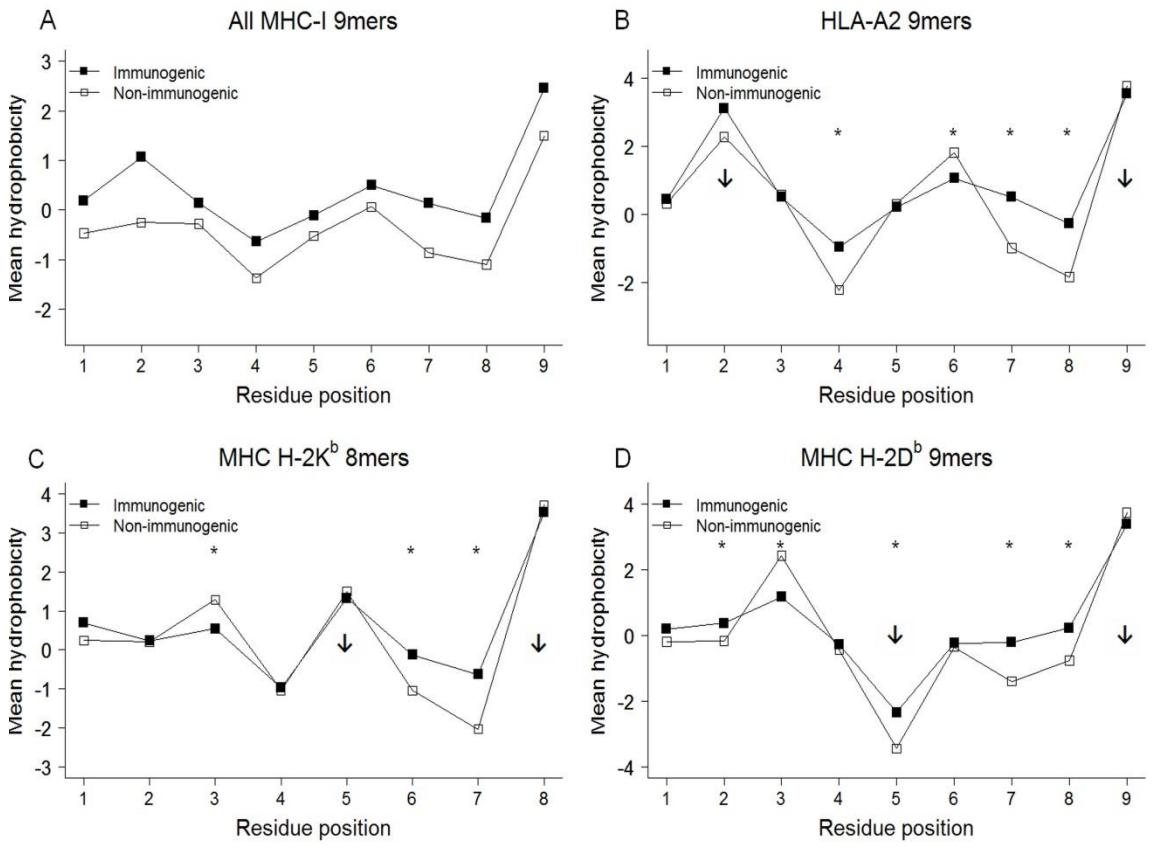

Figure 2-4. Hydrophobicity comparison at each residue position between immunogenic and non-immunogenic MHC-I peptides. Mean hydrophobicity at each position was calculated. (A) All immunogenic and non-immunogenic MHC-I 9-mers; every residue had P < 1.6 X 105. (B) HLA-A2 restricted immunogenic and non-immunogenic 9-mers. (C) Murine MHC H-2Db restricted immunogenic and non-immunogenic 9-mers. (D) Murine MHC H-2Kb restricted immunogenic and non-immunogenic 8-mers. Down-arrows in B-D indicate anchor residues based on specific MHC motifs. ∗P < 0.008 in that residue position. P values are listed in Table A-2.

hydrophobic in immunogenic epitopes (P7, P = 1.1×10−4; P8, P = 0.001; (Fig. 2-4 D)), with no difference in anchor residue P9 (P = 0.127). One exception was the anchor residue P5, which was more hydrophobic in immunogenic epitopes (P = 4.9×10−10). This discrepancy might be due to the presence of other potential anchors at P5 (apart from Asn) within immunogenic dataset. Hence we demonstrate that the observed bias towards relative hydrophobic aas in immunogenic epitopes is selective for TCR contact residues.

2.3.3. Differential Hydrophobicity can Predict Immunogenic CTL Epitopes

While MHC-binding is necessary for antigen presentation, it is not sufficient to stimulate an immune response. We predicted that hydrophobicity could be incorporated into existing binding algorithms to improve prediction of CTL epitopes. To test this hypothesis, we used the IEDB-consensus binding prediction tool to generate peptide predictions for HLA-A2 restricted peptides (9, 10mers) for two viral proteins: polyprotein from dengue virus type 1 (DENV1) and tegument protein pp65 from cytomegalovirus (CMV). Using mean hydrophobicity of aas in TCR contact residues (all residues except anchors: P2, P6, and P9 or P10), each predicted peptide was re-ranked with decreasing TCR contact hydrophobicity values (Fig. 2-5). The rate at which experimentally defined HLA-A2 restricted CTL epitopes (Table A-3) were identified was increased using hydrophobicity-based predictions compared to IEDB-consensus binding predictions (Fig. 2-5 A and B). As a negative control, we performed re-ranking of top predictions from the 2 proteins using mean hydrophobicity of just anchor residues (Fig. 2-5 C-D). The rate of prediction of HLA-A2 restricted CTL epitopes was similar to IEDB-consensus binding predictions, confirming that relative hydrophobicity impacts immunogenicity and not HLA-binding. These results suggest that using TCR contact hydrophobicity could improve prediction of immunogenic epitopes.

2.3.4. Hydrophobicity-based ANN Prediction Model

The relative contribution of each aa residue to immunogenicity varies between MHC allomorph and is motif-dependent (Fig. 2-4) and (Table A-2). Furthermore, the immunogenicity of a peptide might result from nonlinear interactions between different TCR contact residues. Artificial neural networks (ANN) are designed to handle such nonlinearity (Bishop, 2006; Honeyman et al., 1998). Therefore, we developed and trained an ANN-based prediction model of immunogenicity using aa hydrophobicity (ANN-Hydro) with the goal of improving existing CTL epitope prediction algorithms and were used as the trainings sets for the two ANN-Hydro models (Fig. 2-6). An initial assessment of the trained ANN-Hydro model for HLA-A2 assigned a good probability of immunogenicity to 54/64 (> 80%) experimentally defined HLA-A2 restricted epitopes from 3

recent studies (Assarsson et al., 2008; Newell et al., 2013; Weiskopf et al., 2011) ($P < 0.001$, compared to the distribution of probabilities of immunogenicity of 64 randomly generated 9mer

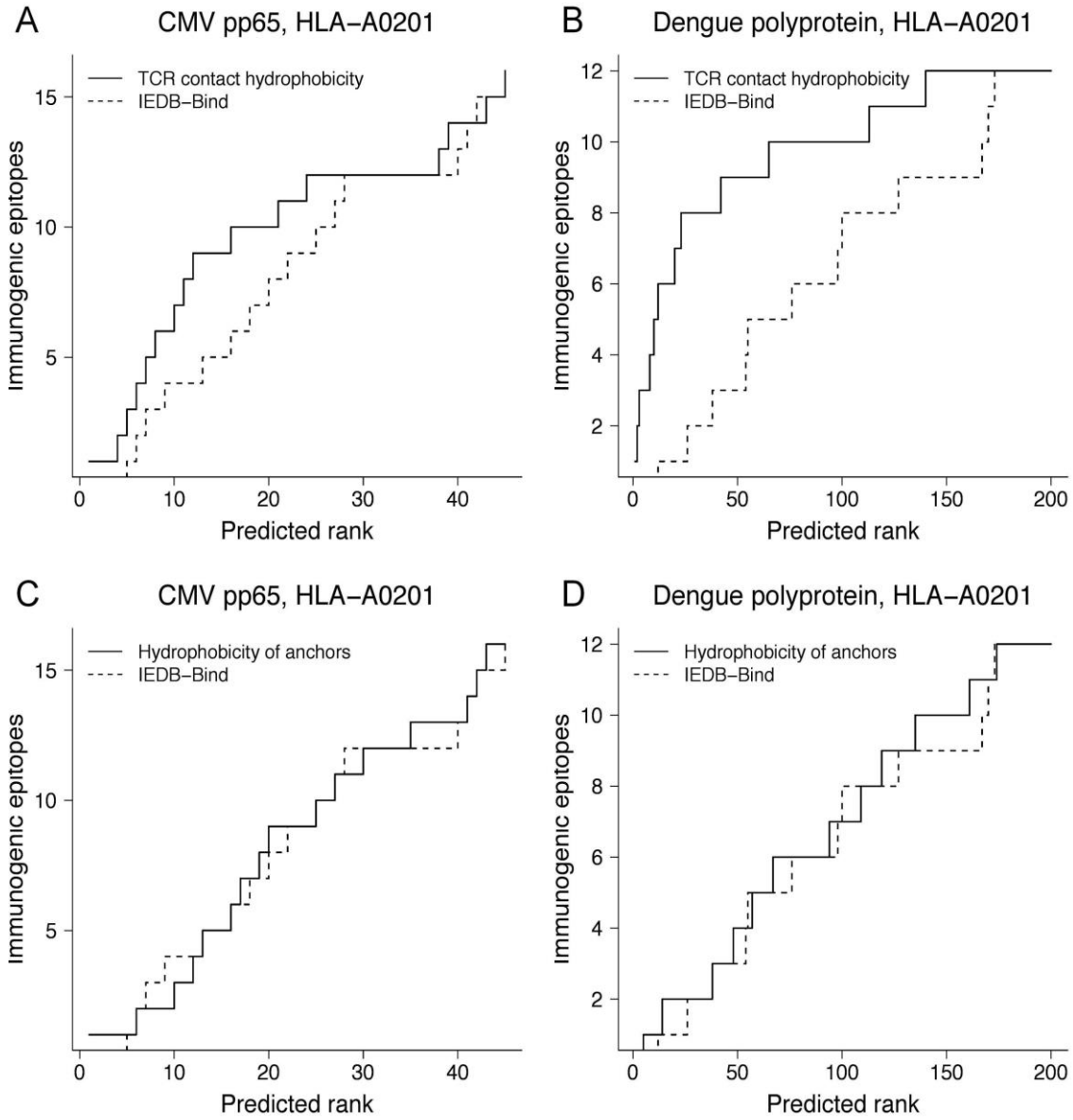

Figure 2-5. Differential Hydrophobicity can predict Immunogenic CTL Epitopes. Efficiency of predicting experimentally defined HLA-A0201 restricted immunogenic epitopes using mean hydrophobicity of TCR contact residues (straight lines) compared to IEDB consensus binding tool (IEDB-Bind) are shown (dashed lines). Tegument protein pp65 from cytomegalovirus (CMV) and polyprotein from dengue virus type 1 were used for predictions. (A-B) Predicted peptides from the IEDB-Bind were re-ranked using mean hydrophobicity of TCR contact residues. (C-D) Predicted peptides from the IEDB-Bind were re-ranked using mean hydrophobicity of anchor residues.

peptides) (Table A-4). We then developed an epitope discovery strategy incorporating the ANN-Hydro model to predict a previous set of experimentally validated H-2D^b and HLA-A2 epitopes from 5 pathogen and 5 tumor antigens (Table A-4). The IEDB-consensus MHC-binding prediction algorithm was used to obtain a list of predicted peptides for each antigen, which were each assigned a normalized binding score (S_B). Since T cell epitopes are a subset of predicted peptides that bind to MHC molecules, normalized scores (S_I) based on probabilities of immunogenicity obtained by ANN-Hydro were assigned to each peptide (Fig. 2-6). We then defined a total score (S) as $S = S_B \,.\, S_I$ for the rate of identifying CTL epitopes from the list of predicted H-2D^b and HLA-A2 peptides from each antigen. The total score is therefore dependent on contribution of both scores, reflecting two critical aspects: binding and immunogenicity (Fig. 2-6). Our strategy of re-ranking by prioritization of high-binding and high-immunogenic peptides over other predicted peptides ((Fig. 2-6), Materials and Methods), scored 42 out of the 43 H-2D^b and HLA-A2 9mer epitopes within the top 20 ranked peptides (Table A-4). Each peptide sequence in the H-2D^b and HLA-A2 datasets was transformed into a corresponding numeric sequence based on the hydrophobicity value of aas prediction algorithms, ranked the same epitopes up to rank 133 (Table A-4). Therefore, the ANN-Hydro model can be used in conjunction with IEDB-consensus to improve the efficiency of prediction of CTL epitopes.

2.3.5. Prediction Validation by in Vivo Discovery of HIV-1 Gag Epitopes

To comprehensively evaluate the predictive capacity of our approach for CTL epitope discovery and to correlate immunodominance, we interrogated 3 HIV-1 Gag variant proteins: Consensus B (ConsB), 96ZM651.8 (ZM96), and 97/CN54 (CN54) (Fig. 3.4). With no prior knowledge of Gag-specific CTL epitopes, our model was used to generate a list of ranked H-2D^b restricted peptides, of which the top 20 predictions for each interrogated Gag sequence are shown (Table B.8). To validate our predicted epitopes in vivo, B6 mice were immunized independently against each of the three different Gag variants and the peptide specificity of effector CD8+ T cell responses analyzed using overlapping peptide pools (Table A-5). Deconvolution and truncation experiments allowed us to define a unique dominant H-2D^b-restricted epitope within each Gag protein (SI9 for

ConsB, QL11 for CN54, RT9 for ZM96), as well as shared subdominant epitopes: Db-restricted RT9, AI9, YI9 and Kb-restricted VL8 (Fig. 2-7). Comparison of empirically defined epitopes to predictions made using ANN-Hydro revealed that H-2D^b restricted 9mer CTL epitopes for HIV-1

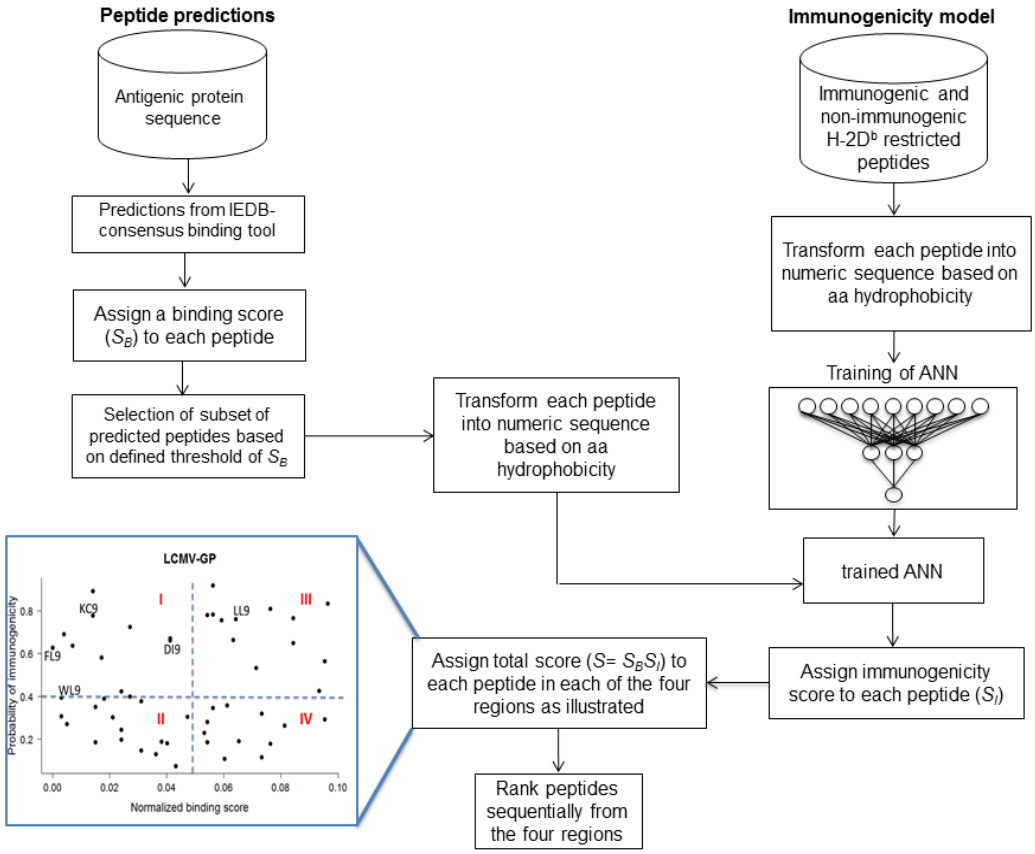

Figure 2-6. Workflow for CTL epitope prediction using the ANN-Hydro model and the MHC-binding prediction tool IEDB-consensus. For training and application of the ANN-Hydro model for immunogenicity scores, each peptide sequence in the HLA-A2 and H-2Db dataset was transformed into a corresponding numeric sequence based on the hydrophobicity value of amino acids. To obtain a list of candidates for MHC-bound peptides from a given antigen, IEDB-consensus binding algorithm was used and a normalized binding score (S_B) was assigned. The trained immunogenicity ANN model was applied on the same list of peptides independently to assign immunogenicity scores (S_I). After the subset of top binding peptides was selected, peptides from each region ranging from high-binding highly-immunogenic peptides to modest-binding low-immunogenic peptides (quadrants 1 through 4 in inset) were re-ranked based on total score $S = S_B . S_I$. An example of epitope prediction is shown in the plot for experimentally defined H-2D^b restricted CTL epitopes from LCMV-GP.

CN54 Gag and ZM96 Gag correlated with ANN-Hydro model epitope sequences predicted within the top 15 ranked peptides; and for ConsB Gag within the top 11 ranked peptides (Table A-5). In striking contrast, prediction of the identified Gag epitopes by individual prediction algorithms was

more varied, with predictions up to rank number 46, depending on the binding or processing algorithm used. Although the IEDB consensus binding and NetMHCpan algorithms predicted the identified Gag epitopes within the top 6 ranked peptides, the performance of these algorithms (unlike the ANN-Hydro model) was highly variant with the antigen selected (variance ranges from66.72 to 220.27; (Table A-5). In sum, the ANN-Hydro model predicted 52 out of 53 experimentally validated H-2D^{b} and HLA-A2 9mer epitopes from 13 different antigens within the

Source	Peptide	Amino Acid Sequence	Immunogenicity	Confirmed Epitope
Cons B	7889	QTGSEELRSLYNTVA	+	H-2D^{b}
Cons B	7890	EELRSLYNTVATLYC	+	Gag76-84
CN54	7080.19	EELRSLFNTVATPYC	+	RSLYNTVAT
CN54	7080.20	SLFNTVATPYCVHTE	+	RSLFNTVAT
ZM96	7116.8	GTEELRSLYNTVATLYCVHE	+	(RT9)
Cons B	7912	EKAFSSPEVIPMFSAL	+	H-2K^{b}
Cons B	7913	SPEVIPMFSALSEGA	+	Gag168-175
CN54	7080.41	F-SPEVIPMFSALSEG	+	VIPMFSAL
CN54	7080.42	VIPMFSALSEGATPQ	+	VIPMFTAL
ZM96	7116.17	EKAF-SPEVIPMFTALSEGAT	+	(VL8)
Cons B	7920	GHQAAMQMLKETINE	+	H-2D^{b}
Cons B	7921	AMQMLKETINEEAAE	+	Gag197-205
CN54	7080.49	AAMQILKDTINEEAA	+	AMQMLKETI
ZM96	7116.2	VGGHQAAMQMLKDTINEEAA	+	AMQILKDTI AMQMLKDTI (AI9)
Cons B	7940	IVRMYSPTSILDIRQ	+	H-2D^{b}
Cons B	7941	YSPTSILDIRQGPKE	+	Gag277-285
CN54	7080.68	KIVRMYSPTSILDIK	+	YSPTSILDI
CN54	7080.69	MYSPTSILDIKQGPK	+	YSPVSILDI
ZM96	7116.28	NKIVRMYSPVSILDIKQGPK	+	(YI9)
Cons B	7949	ASQEVKNWMTETLLV	-	H-2D^{b}
CN54	7080.77	QATQGVKNWMTDTLL	+	Gag310-320
CN54	7080.78	GVKNWMTDTLLVQNA	+	QGVKNWMTDTL
ZM96	7116.32	QEVKNWMTDTLLVQNA	-	(QL11)
Cons B	7963	EAMSQVTNSATIMMQ	+	H-2D^{b}
CN54	7080.91	AEAMSQ-TNSA-ILMQR	-	Gag368-376
ZM96	7116.37	RVLAEAMSQ-TNSVNILMQK	-	SQVTNSATI (SI9)

Figure 2-7. Summary of identified epitopes. Responses to the RT9, VL8, AI9, and YI9 epitopes are observed for all three Gag protein variant peptides. Overlapping sequences of individual peptides are shown. The QL11 epitope was only immunogenic for the CN54 Gag protein, but not ConsB or ZM96 Gag proteins, likely due to the A to E substitution at position 2. The SI9 epitope was only immunogenic in the ConsB Gag protein, as both CN54 and ZM96 had major deletions and substitutions in this sequence. MHC restriction was confirmed using MHC class I tetramer staining, and Gag amino acid positions are in reference to the HXB2 strain.

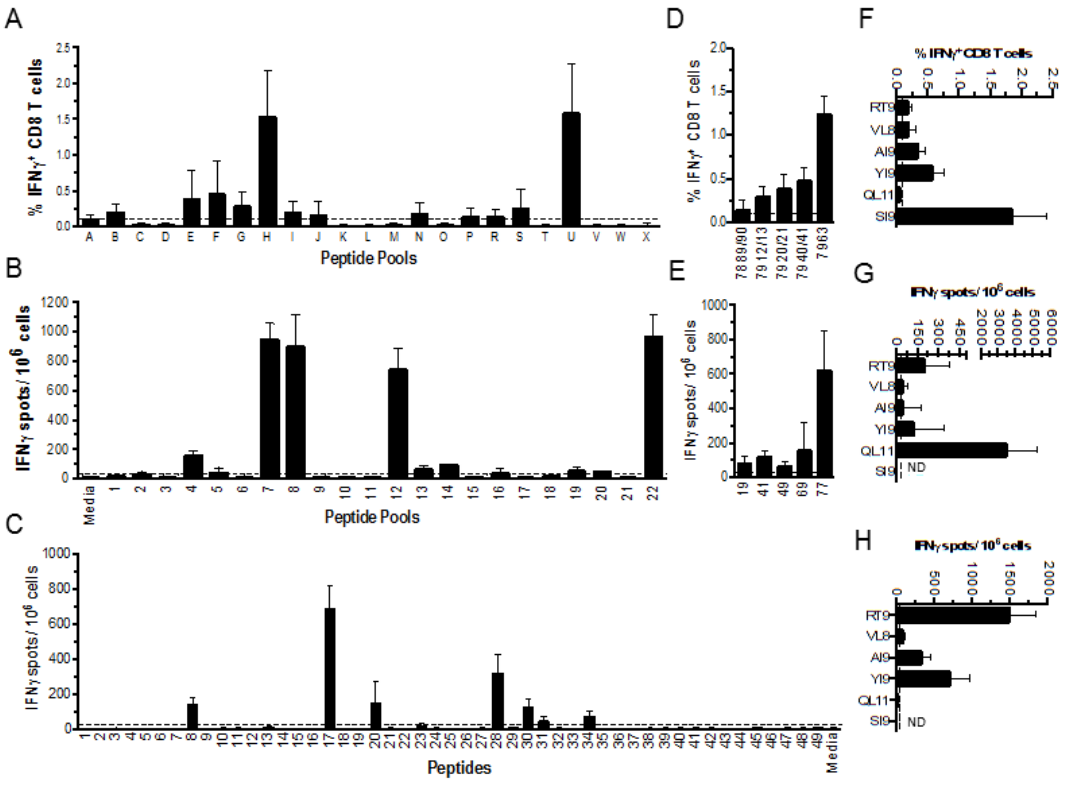

Figure 2-8. ANN-Hydro model prediction validation by in vivo discovery of HIV-1 Gag epitopes. Predictions for H-2D^b epitopes were made for three HIV-1 Gag proteins using the ANN-Hydro model, and then a blinded epitope discovery study was performed in vivo. (A–C) B6 mice were immunized with AdHu5 vaccines expressing the ConsB, CN54, or ZM96 Gag, and CD8+ T-cell responses determined by intracellular IFN-γ or IFN-γ ELISPOT after ex vivo stimulation with peptide pools of 15-mer peptides (overlapping by 11 mer) spanning the entire Gag sequence (ConsB or CN54, A and B) or with a complete set of overlapping 20-mer peptides spanning ZM96 (C). (D–H) Positive responses to pools were deconvoluted by stimulation with individual 15-mer peptides from the positive pools (ConsB or CN54, D and E). Minimal epitopes were identified by stimulation with truncated peptides and are shown (F–H).

top 20 ranked peptides (Fig. 2-9), corresponding to a 98% success rate in identifying immunogenic epitopes. Moreover, this predictive improvement was reflected in lower variability of epitope identification, a variance of 37.72 using ANN-Hydro as opposed to 66.72 by IEDB alone ($P < 0.05$, F-test).

2.3.6. Prediction of Immunodominant Epitopes

The probabilities of immunogenicity assigned by ANN-Hydro were interrogated with respect to epitope immunodominance using three antigens with a clear vertical epitope hierarchy, as

identified by ex vivo experimental data (Oukka et al., 1996; van der Most et al., 1998) and this study). The epitope hierarchy defined experimentally in LCMV-GP, Flu-NP, and ZM96 Gag showed robust correlation with the probabilities of immunogenicity assigned by ANN-Hydro ($r > 0.94, P < 0.05$; (Fig. 2-9 B-D)). In contrast, predicted MHC-binding assigned by IEDB-consensus showed no correlation with epitope immunodominance in LCMV-GP and ZM96-Gag (Fig. 2-9 E and G). Epitope immunodominance in Flu-NP correlated with both predicted ANN-Hydros probability and predicted MHC-binding (Fig. 2-9 C and F). As a further correlate, seven of 13 epitopes predicted in lower rankings by ANN-Hydro along with IEDB-consensus were modest immunogens derived from LCMV-GP, LCMV-NP, ZM96, CN54, and Consensus Gag (Table A-6). Therefore, efficient pMHC-TCR affinity may contribute towards epitope immunodominance. By using ANN Hydro, epitope predictions were consistently less variable, and improved the prediction of immunodominant CTL epitopes.

2.4 Discussion

At present, there is no consensus on the molecular mechanisms by which CD8+ T cells discriminate immunogenic antigens within the background of poorly immunogenic self peptides. Understanding this discrimination has implications in rational vaccine design and the identification of antigenic targets of malignant and autoimmune diseases. While several theories have been proposed to explain the concept of self/non-self discrimination (Pradeu & Carosella, 2006), the present study is the first attempt to provide a biochemical explanation for this fundamental phenomenon. We show that relative aa hydrophobicity within immunogenic epitopes reveal an antigenic pattern that could be recognized by TCRs. We leveraged these findings to design an immunogenicity model, trained and validated using experimentally defined epitopes. ANN-Hydro consistently reduced variable standard prediction outputs across multiple antigens, demonstrating an important step forward in reducing the empirical element of T cell epitope prediction.

The majority of antigens within the immunogenic dataset used in this study are derived from intracellular pathogens, such as viruses, which have been shown to favor a lower G+C genomic content, reflected in their aa usage (Calis, Sanchez-Perez, & Keşmir, 2010). Strongly

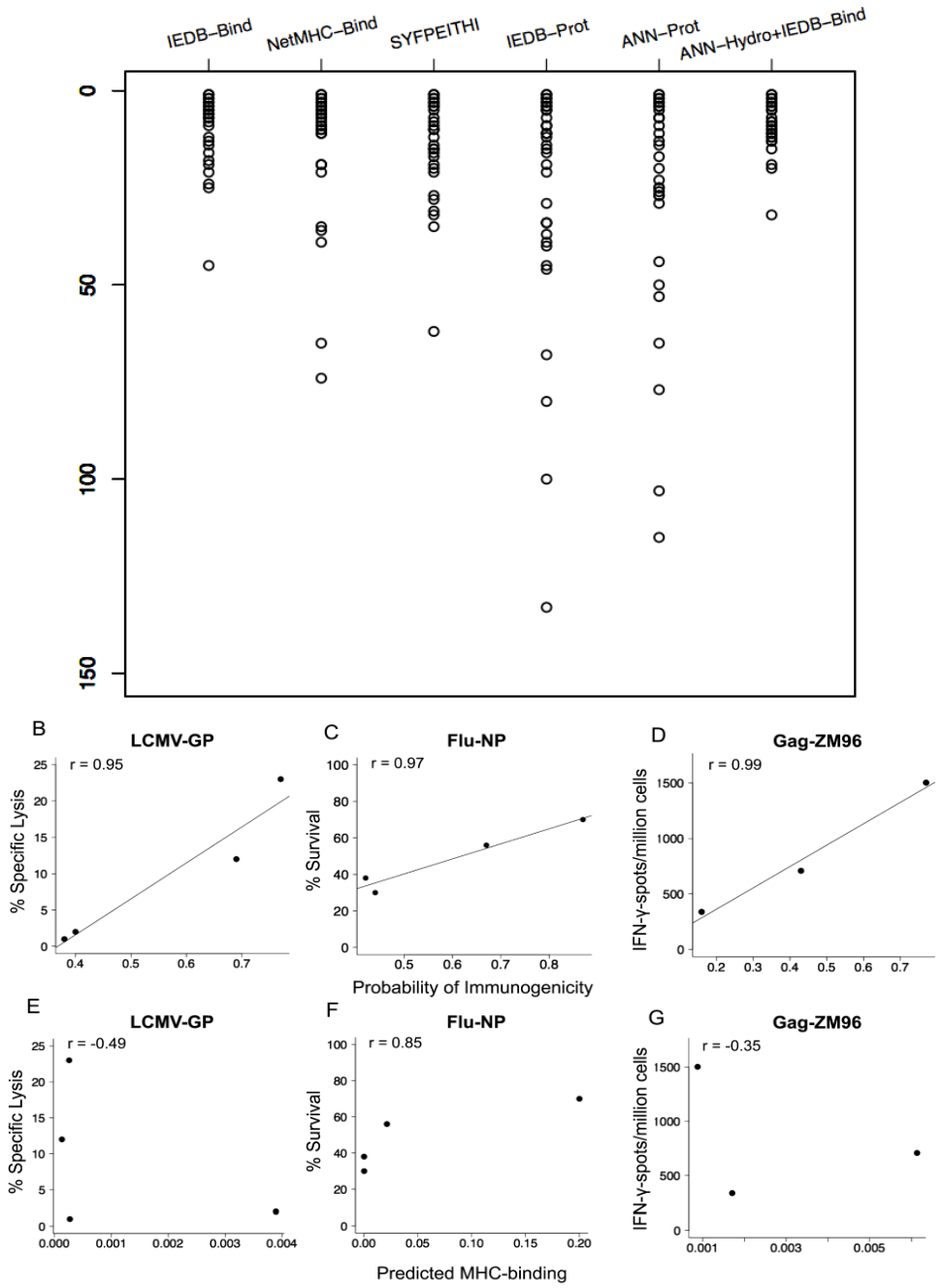

Figure 2-9. Incorporating ANN-Hydro in the IEDB-binding tool improves epitope prediction. (A) Ranked epitopes for 26 H-2D^{b} CTL epitopes from eight well-described antigens and for 27 HLA-A2 CTL epitopes from five tumor antigens (Melan-A, Wt-1, gp100, TRAG-3, and p53), each column is a different prediction algorithm with epitopes and their corresponding predicted ranks shown. (B–D) Epitope immunodominance as a function of probability of immunogenicity for LCMV-GP, Flu-NP, and Gag-ZM96. (E–G) Epitope immunodominance as a function of predicted MHC binding (IEDB consensus) for LCMV-GP, Flu-NP, and Gag-ZM96. Immunodominance was determined from percentage-specific lysis of target cells ex vivo. (B and E) 9-mer versions of SGV11 and CSA10 were used. (C and F) Percent survival of peptide-primed mice on lethal challenge of virus (D and G) IFN-γ spots per million cells on ex vivo peptide stimulation post vaccination with antigen (this study).

hydrophobic aas (e.g. L, I, V, F, M) are characterized by low G+C codons while hydrophilic aas are not (Khrustalev & Barkovsky, 2011). This suggests the possibility that pathogens, in general, have a higher usage of hydrophobic aas that could be exploited for TCR recognition. A second possibility is that antigen presentation inherently favors hydrophobic regions within a protein. A recent study demonstrated that exposing hydrophobic domains significantly enhances the rate of proteasomal degradation and MHC presentation (Huang, Kuhls, & Eisenlohr, 2011). Moreover, immunogenic CTL epitopes are also positionally biased towards the center of their source antigens (Huang et al., 2011; Kim, Yewdell, Sette, & Peters, 2013), consistent with the fact that cytosolic proteins with a central hydrophobic core are the major substrates of proteasomal degradation. Thus, protein hydrophobicity can enhance both antigen presentation and immunogenicity, perhaps an evolutionary adaptation of hydrophobicity driven by damage-associated molecular patterns (Seong & Matzinger, 2004).

TCRs are estimated to recognize on average about 5 non-anchor residues of a presented peptide because of the angle of peptide contact (Burroughs, de Boer, & Keşmir, 2004; Hennecke & Wiley, 2001). For 3 pMHC allomorphs analyzed by hydrophobicity in this study, only 4-5 positions on the peptide were significantly different between immunogenic and non-immunogenic peptides (Fig. 2-4), similar to published pMHC-TCR structures (Rudolph et al., 2006). This hydrophobicity difference is relative, not absolute. Certain aa positions in the peptide may be hydrophilic (e.g. P4 in HLA-A2 9mers, (Fig.2-4 B)). However, even in such inherently hydrophilic residues in the peptide, immunogenic epitopes are less hydrophilic (more hydrophobic). Covering exposed hydrophobic residues on the peptide by a TCR may be a thermodynamically favorable process, facilitating the pMHC-TCR interaction as noted in retrospect, by a recent study (Birnbaum et al., 2014). TCR-engagement of pMHC complexes may be enhanced by water-exclusion from the immunological synapse or by increased K_{on} rates of the TCR-pMHC complex by relatively hydrophobic aas.

In the absence of a good understanding of the biochemical composition of peptide ligands that result in T cell activation, current strategies for epitope discovery either rely on the

unbiased synthesis of a large number of overlapping peptides, or use MHC-binding/antigen processing algorithms to select candidate peptides. While the former is an expensive and laborious process, the latter results in a large number of false positive peptides that are not immunogenic. Advances in the development of combinatorial technologies have allowed the rapid identification and characterization of antigen-specific T cells (Newell & Davis, 2014). However, even such novel technologies rely on binding predictions to create candidate peptides lists that require extensive empirical validation. For instance, 77 candidate good-binders for HLA-A2 from the rotavirus proteome were chosen for recombinant pMHC tetramer production based on their MHC-binding capability, but only 6 (four being 9mer epitopes) were confirmed to be immunogenic epitopes (Newell et al., 2013). Therefore, T cell antigen discovery studies need strategies that improve the efficiency of epitope prediction.

ANN-Hydro assigned high probabilities of immunogenicity to 80% of the HLA-A2 9mer epitopes described in the three proteome wide studies. Of note, three of the four rotavirus 9mer epitopes from the data set scored a probability of immunogenicity greater than 0.8 (Table A-4). In the HIV-1 Gag study, over 364 overlapping peptides were tested in vivo from the Gag variants (length: 500aa) for epitope discovery. Using the ANN-Hydro model combined with SB scores narrows the validation discovery process down to 11-15 peptides per Gag protein to be tested. Similarly, applying ANN-Hydro also improved predictions of immunogenic H-2Db and HLA-A2 epitopes from 10 independent antigens compared to individual prediction algorithms. Thus, models such as ANN-Hydro adds an extra dimension (immunogenicity) to MHC-binding for CTL epitope prediction and could be used to significantly reduce the variability associated with standard prediction algorithms, and the time and cost of experimental validation (Fig. 2-9), (Table A-6). With the advent of tumor exome sequencing in immune therapy settings, we anticipate that immunogenicity models such as ANN-Hydro will be critical in identifying immunogenic neo-antigens for tumor immune therapies (Snyder et al., 2014; Yadav et al., 2014).

The ANN-Hydro model differs from existing MHC-binding/antigen-processing prediction algorithms in two respects: First, the ANN-Hydro was trained on a relative hydrophobicity scale, which facilitates the model to discover complex numeric relationships between different aa

residues. Second, the dataset used for training was immunogenic epitopes and non-immunogenic self-peptides, which do not differ in binding motifs but only in immunogenicity. While some high-binding epitopes (e.g. SI9 from ConsB) are readily predicted by all algorithms, other epitopes (e.g. the immunodominant dominant RT9 from ZM96, LL9 from LCMV-GP) are predicted at variable rankings by different algorithms (Table A-6). In comparison, ANN-Hydro rescued these epitopes by virtue of their probability of immunogenicity. Although ANN-Hydro marks a step forward in efficiently predicting 9mer epitopes, it is currently limited in terms of predicting longer or shorter epitopes, exemplified by the 11mer epitope (QL11) deduced by epitope mapping from the CN54 Gag protein. To improve longer or shorter epitope predictions, larger representative datasets are required for training. Nonetheless, the model predicted a 9mer version of this epitope ranked at 35 and 44, which is consistent with presentation of nested length peptides (Riemer et al., 2010). A second limitation of the current model is its applicability to predict epitopes for other HLA class I alleles. In theory, the ANN-Hydro model could be applied to predict CTL epitopes for any MHC class I allele, but large representative datasets are required for training the model for representative MHC allomorphs. We anticipate that advances in mass spectrometry-based MHC peptide discovery will result in more extensive training databases for predicting longer and shorter epitopes from a broader selection of HLA class I molecules (Riemer et al., 2010; Tan, Croft, Dudek, Williamson, & Purcell, 2011).

While immunogenicity models have been developed by others for prediction of CTL epitopes (Calis et al., 2013; Harndahl et al., 2012), they considered only the impact of pMHC stability and positional significance along the peptide for immunogenicity. In contrast, a crucial feature of our approach is the use of ligand-eluted non-immunogenic self-peptides as the comparator set. Because binding and antigen processing are required for all epitopes, we built upon existing algorithms for immunogenic pMHC predictions. “Layering” the immunogenicity model on top of existing prediction algorithms enabled us to predict epitopes with increased effectiveness than standalone predictions. Importantly, the empirical evaluation of our immunogenicity model and epitope prediction approach without a priori knowledge of the immunodominant HIV-1 Gag epitopes in vivo gives strong support for these results. In summary,

integrating aa hydrophobicity into pMHC prediction algorithms should significantly enhance the success of epitope discovery. The biological mechanism underlying TCR preferences for non-polar hydrophobic residues remains to be evaluated.

2.5 Materials and Methods

For full details of methods and construction of data sets see Appendix A.

2.5.1. Construction of Datasets

All MHC-I peptides used in this study and design of the ANN-Hydro prediction model was retrieved from IEDB (Vita et al., 2010) (www.iedb.org, last accessed: 08/11/2013). Epitopes with a positive T cell response represent the immunogenic epitope group. The non-immunogenic self-peptide group represent cell surface ligand eluted MHC-I self-peptides that have been antigenically processed and MHC-bound. Additional curation and exclusion criteria resulted in a final dataset with 5,035 8-11mer immunogenic epitopes and 4,853 8-11mer non-immunogenic peptides (See Appendix A for further details).

2.5.2. Amino Acid Scales

These were derived from Expasys ProtScale (http://web.expasy.org/protscale/) (Gasteiger et al., 2005), specifically, Hydrophobicity scale (Kyte and Doolittle) (Kyte and Doolittle, 1982), Polarity (Grantham) (Grantham, 1974), and Bulkiness (Zimmerman) (Zimmerman et al., 1968). The scales are relative, e.g., negative to positive values in the hydrophobicity scale correspond to a relative hydrophobicity increase between aas (Table A-1).

2.5.3. Position-based Hydrophobicity Analysis

We transformed our datasets of immunogenic and non immunogenic peptides into numeric arrays using the R statistical software (RDevelopment, 2012). Separate numeric arrays were generated for immunogenic and non-immunogenic 8, 9 and 10mers. Mean hydrophobicity of

immunogenic and non-immunogenic peptides at each position was calculated and were compared residue-by-residue through Wilcoxon rank-sum tests to quantify statistical significance.

2.5.4. Hydrophobicity-based ANN Prediction Model (ANN-Hydro)

The R neuralnet package was used to design and train the two ANN-Hydro models on H-2Db and HLA-A2 restricted 9mer peptides known to be immunogenic (n=204 and n=374, respectively) or non-immunogenic (n=232 and n=201, respectively). Each peptide sequence in the respective H-2Db and HLA-A2 datasets were transformed into a corresponding numeric sequence based on aa hydrophobicity using R statistical software. A three-layer fully connected feed-forward ANN was comprised by nine input neurons, one hidden layer with three neurons, and one output variable (Fig. 2-6).

2.5.5. Application of ANN-Hydro

For each H-2Db and HLA-A2 restricted epitope prediction, we used IEDB-consensus to generate a list of epitope predictions. Each peptide was assigned a normalized binding score (S_B) and a subset of these predicted peptides was then selected by defining a S_B threshold of 0.1 for antigen length >100 aas and a SB-threshold of 0.2 for antigen length ≤ 100 aas. Independently, probabilities of immunogenicity were obtained by applying ANN Hydro to this subset of binding predictions. Normalized scores (S_I) were then assigned based on the probabilities of immunogenicity (Fig. B.3). Predicted peptides were re-ranked based on total score, $S = S_B \cdot S_I$, ranging from lowest score to the highest score. The lower the total score of a predicted peptide, the higher its probability of being an immunogenic epitope. See Appendix A for details.

2.5.6. Vaccines

Recombinant Adenovirus type 5 (rAdHu5) vectors encoding codon optimized HIV-1 Gag from Cons B, strain 96ZM651.8 (ZM96) and strain 97CN54 (CN54) (Bachy et al., 2013) are described in Appendix A.

2.5.7. Immunization of Mice

C57BL/6 mice were immunized with 109 virus particles. All animal studies were conducted in accordance with UK Home Office regulations and Kings College London ethics committee.

2.5.8. Peptides

15mer peptides spanning HIV-1 CN54 Gag and a 20mer set of peptides spanning HIV-1 ZM96 were provided by the UK Centre for AIDS Reagents. 15mer peptides spanning HIV-1 Cons B Gag were provided by the NIH AIDS Reagent Reference Program. Truncated HIV-1 Gag peptides were purchased from Proimmune.

2.5.9. T Cell Epitope Mapping

Spleen cells were re-stimulated either with media alone or with peptides, either in pools or individually (each at 1M final concentration) and IFN- production was detected by intracellular cytokine staining or by ELISPOT assay as previously described (Bachy et al., 2013). Cons B and CN54 Gag epitopes were deconvoluted to individual 15mers from peptide pools, and truncated versions of the 15mer peptides were synthesized and tested. For ZM96 Gag, 49 individual 20mer peptides were tested. Reactive peptide sequences were confirmed against the corresponding 15mer peptide to the reactive sequence and 9mer peptides were synthesized and tested.

CHAPTER 3

T-CELL EPITOPE DISCOVERY FOR THERAPEUTIC CANCER VACCINES

This chapter is published:

Sri Krishna and Karen S. Anderson. Vaccine Design: Methods and Protocols: Volume 1: Vaccines for Human Diseases, 779-796. Springer New York. 2016.

Abstract

The success of recent immune checkpoint blockade trials in solid tumors has demonstrated the tremendous potential of immune-mediated treatment strategies for cancer therapy. These immune therapies activate preexisting cytotoxic CD8+ T cells (CTL) to selectively target and eradicate malignant cells. In vitro models suggest that these therapies may be more effective in combination with priming of CTL using cancer vaccines. CTL-mediated tumor targeting is achieved by its recognition of tumor antigenic epitopes presented on human leukocyte antigen (HLA) class I molecules by tumor cells. Discovering CTL-antigenic epitopes is therefore central to the design of therapeutic T-cell vaccines and immune monitoring of these complex immunotherapies. However, selecting and monitoring T-cell epitopes remains difficult due to the extensive polymorphism of HLA alleles and the presence of confounding non-immunogenic self-peptides. To overcome these challenges, this chapter presents methodologies for the design of CTL-targeted vaccines using selection of target HLA alleles, novel integrated computational strategies to predict HLA-class I CTL epitopes, and epitope validation methods using short-term ex vivo T-cell stimulation. This strategy results in the improved efficiency for selecting antigenic epitopes for CTL-mediated vaccines and for immune monitoring of tumor antigens.

3.1 Introduction

Recent clinical trials of vaccines, checkpoint blockade, and immunotherapy have demonstrated the potential efficacy of harnessing cytotoxic T cells for treatment of many cancers (H.-J. Kim & Cantor, 2014; Mellman, Coukos, & Dranoff, 2011; Trimble & Frazer, 2009). Unlike multimodality

therapy with surgery, radiation, and chemotherapy, immune therapies against tumor-specific or tumor-associated antigens hold great promise for targeted tumor eradication with relatively minimal side effects. Prophylactic subunit vaccines, such as the hepatitis B vaccine (HBV) and the human papillomavirus (HPV) vaccines, stimulate protective antibody responses and have been highly successful with >90 % efficacy (Chang et al., 1997; Giuliano et al., 2011; Mast et al., 2005). However, eradicating preexisting pathogenic infections and malignancies is difficult to achieve by antibody-mediated immunity alone. For instance, the prophylactic HPV VLP vaccine has limited efficacy for the therapeutic treatment of existing lesions (Trimble & Frazer, 2009). Solid tumors, in particular, have a limited number of selective cell surface targets, a striking genomic heterogeneity, and rapid evolution of antigenic escape. Therefore, vaccines that induce T-cell-mediated immunity against established malignancies for therapeutic intervention are needed (H.-J. Kim & Cantor, 2014; Mellman et al., 2011; Trimble & Frazer, 2009).

The primary goal of immune therapies for tumor eradication has been the induction of cytotoxic CD8+ T cells (CTLs). CTLs are activated by their recognition of 8–11 amino acid peptides derived from proteasomal degradation of either pathogen-derived or self-antigens in association with human leukocyte antigen (HLA) class I molecules (Blum, Wearsch, & Cresswell, 2013; Grakoui et al., 1999). The downstream signaling cascade triggered by the binding of T-cell receptors (TCRs) to epitope specific peptide-HLA complex causes antigen-specific effector CTL proliferation and the cytolysis of target cells presenting the epitope (Hennecke & Wiley, 2001). The αβTCR-peptide-HLA interaction is thus a critical event in CTL-mediated immunity and is fundamental for rational vaccine design. These CTL epitopes can be incorporated as a component of the therapeutic vaccine, or they can be useful for immune monitoring post-therapy (Trimble & Frazer, 2009). Identifying immunogenic CTL epitopes remains a major challenge in vaccinology. Three major hurdles impede efficient discovery of CTL epitopes: (1) target antigen selection for vaccine design, (2) the codominance and polymorphism of HLA alleles which vary in populations by ethnicity and geographic location, and (3) the identification of the minimal peptidic sequence that can stimulate antigen-specific effector T-cell responses (Purcell, McCluskey, & Rossjohn, 2007). Identifying antigenic peptides reduces the cost of vaccine manufacture and

limits exposure to competing non-immunogenic peptides within the vaccine formulation (Purcell et al., 2007). Comprehensive T-cell epitope mapping across different HLA alleles is important to identify relevant epitopes that are antigenically processed and presented on the tumor tissue (Riemer et al., 2010).

Antigens for tumor immunotherapy may be derived from mutated, splice-variant, or structurally altered antigens (tumor specific antigens), overexpressed wild-type antigens (tumor-associated antigens), as well as other neo-antigens against which central or peripheral T-cell tolerance has not been established (Ernst & Anderson, 2015; Mellman et al., 2011). Ideal antigens are those that are strongly expressed in tumor tissue and required for tumor pathogenesis. Examples of tumor-specific antigens are the HPV16 viral oncogenes E6 and E7 which are integrated into the host genome in cervical carcinomas and have sustained expression during tumor progression. E6 and E7 are excellent candidates for CTL-mediated recognition of malignant cells harboring these “non - self” antigens (Trimble & Frazer, 2009). Several groups have targeted E6/E7 in therapeutic vaccines (Ma, Xu, Hung, & Wu, 2010), but comprehensive CTL epitope and HLA-restriction mapping of the HPV immunome are still limited (Riemer et al., 2010; Yadav et al., 2014). With recent advances in tumor exome and RNAseq analysis, target antigens are increasingly being discovered using bioinformatics analysis of the tumor genome (Rajasagi et al., 2014), exome (Pulido et al., 2012; H.-G. Rammensee & Singh-Jasuja, 2013; Segal et al., 2008; Yadav et al., 2014), or post-hoc analysis of patients in response to immune therapies . Advances in proteomic tools such as mass spectrometry (MS) are now more routinely used to identify the tumor peptidome for antigen discovery (Rizvi et al., 2015; Snyder et al., 2014). Advances in proteomic tools such as mass spectrometry (MS) are now more routinely used to identify the tumor peptidome for antigen discovery (Fortier et al., 2008; Riemer et al., 2010; Yadav et al., 2014).

The second limitation of CTL epitope discovery is the codominance and polymorphism of HLA alleles (Lund et al., 2004; Purcell et al., 2007). Bioinformatic and sequence analyses have demonstrated that most HLA-class I alleles can be classified into one of the 9–12 common

supertypes of HLA alleles, providing a population coverage of over 90 % within the HLA supertypes (Lund et al., 2004; Sette & Sidney, 1999). HLA alleles and their supertypes can be obtained using bioinformatic analyses from the global HLA database (http://www.ebi.ac.uk/ipd/imgt/hla/). For targeting HLA alleles from each supertype with maximal population coverage for tumor antigen discovery, we have chosen ten common HLA alleles corresponding to a 60 % of the HLA-A locus and a >35 % HLA-B locus-specific population coverage (HLA-A alleles: A*0101, A*0201, A*0301, A*1101, A*2402; HLA-B alleles: B*0702, B*0801, B*2705, B*3501, B*5701). Using these HLA supertypes, the selected alleles represent the most common representatives of HLA supertypes for a CTL vaccine targeting over 90% of the global population coverage, according to Lund et al. (Lund et al., 2004).

There are several methods to define antigenic epitopes for T-cell immunotherapy (Fig. 3-1). Conventional discovery of CTL epitopes has relied on in vitro or in vivo testing of overlapping peptides spanning the entire target antigen length, followed by peptide deconvolution and serial truncation to identify the minimal immunogenic epitope(s). Alternatively, the protein sequence can be scanned for potential HLA-binding motifs based on known amino acid preferences of different HLAs for peptide binding (Vonderheide, Anderson, et al., 2001). Potential peptides can be tested for HLA-binding affinity on cell lines with defective antigen processing such as T2 (Riemer et al., 2010; Vonderheide, Anderson, et al., 2001). Recently, computational tools developed over the past decade have become increasingly reliable for predicting HLA-peptide affinity (Honeyman, Brusic, Stone, & Harrison, 1998; Moutaftsi et al., 2006; Tenzer et al., 2005). These computational prediction tools leverage large experimentally derived datasets on peptide-HLA binding for training Markov models or neural networks and have now been expanded to additionally include antigen-processing elements such as proteasomal cleavage patterns (Honeyman et al., 1998; Moutaftsi et al., 2006; Tenzer et al., 2005). Additionally, we and others have developed computational models which predict HLA-binding peptide immunogenicity (Calis et al., 2013; Chowell et al., 2015). These immunogenicity models can be used in conjunction with existing prediction algorithms to further improve efficiency of CTL epitope predictions (Chowell et al., 2015). Additional algorithms for HLA-class II peptide predictions reviewed in Nielsen, &

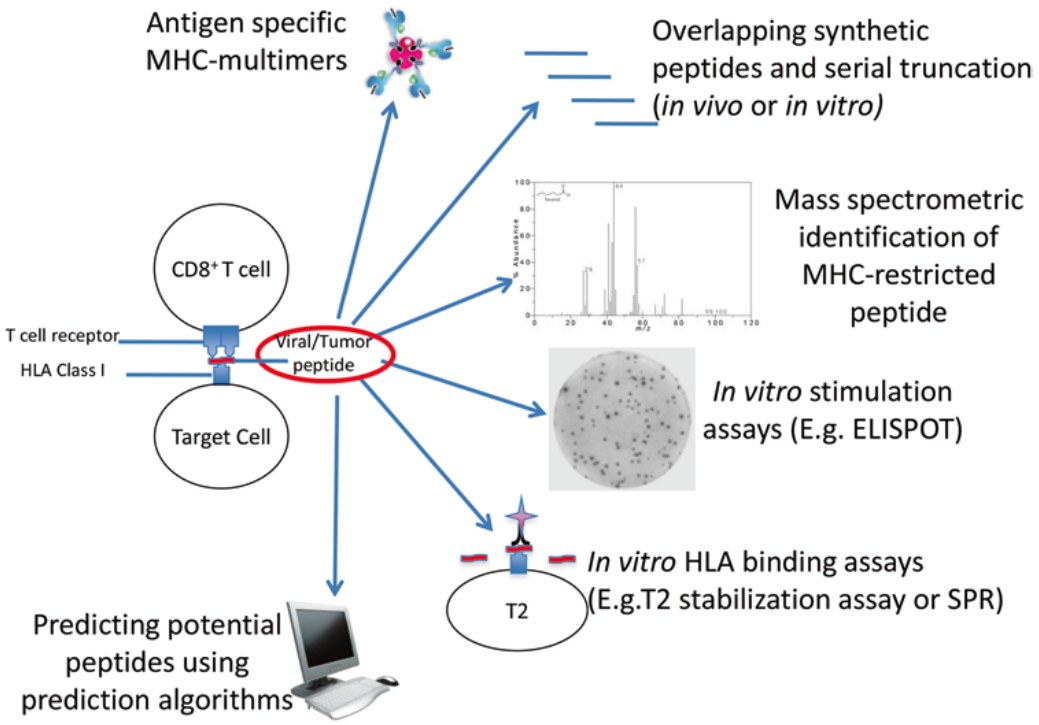

Figure 3-1. Techniques to identify T-cell epitopes. Note that computationally predicted peptides need validation by one or more experimental techniques.

Lundegaard, 2010 can be used concurrently with class I predictions to improve T-cell vaccine targets for tumor antigens.

Predicted CTL epitopes are conventionally tested for HLA binding using recombinant HLA proteins or cellular assays, T-cell stimulation with Elispot assays (Vonderheide, Schultze, et al., 2001), antigen processing using mass spectrometry (Fortier et al., 2008; Hirano, 2006), or HLA multimers to determine the frequencies of antigen-specific T cells in the peripheral blood (Fig. 3-1). Despite these major advances in computational immunology, there is huge variability that exists between the different prediction algorithms and results in a significant number of non-immunogenic false-positive epitopes from a given antigen. Here, we will focus on a reverse immunology CTL epitope discovery strategy that improves the efficiency of epitope prediction and experimental validation by short-term ex vivo T-cell cultures (Newell et al., 2013; Newell & Davis, 2014). Despite these major advances in computational immunology, there is huge variability that exists between the different prediction algorithms (Lin, Ray, Tongchusak, Reinherz, & Brusic,

2008) and results in a significant number of non-immunogenic false-positive epitopes from a given antigen (Newell et al., 2013). Here, we will focus on a reverse immunology CTL epitope discovery strategy that improves the efficiency of epitope prediction and experimental validation by short-term ex vivo T-cell cultures.

3.2 Comparative CTL epitope prediction strategy

CTL epitope identification strategies for tumor and pathogen derived antigens are predominantly limited to the well-represented HLA-A2 allele. There is a need to define CTL epitopes for other non-A2 major HLA supertypes in order to develop globally relevant immune therapies. A number of open-access prediction algorithms are available for peptide-MHC binding and antigen processing (Honeyman et al., 1998; Nielsen et al., 2007). However, a recent study showed that there is huge variability associated with the use of these prediction algorithms depending on the HLA type and antigen chosen (Lin et al., 2008). To counter this variation in performance and scores, we employ a strategy that makes use of commonly used algorithms (three HLA-binding tools and two antigen processing). This strategy of pooling multiple epitope prediction algorithms increases the likelihood of obtaining a true positive epitope. Potential HLA binders for the desired antigen are predicted for the five HLA-A alleles (A*0101, A*0201, A*0301, A*1101, A*2402) and five HLA-B alleles (B*0702, B*0801, B*2705, B*3501, B*5701). Five prediction algorithms are used to predict candidate peptides per antigen per HLA. Three of these algorithms (IEDB-consensus (Moutaftsi et al., 2006), NetMHCpan (Nielsen et al., 2007), and Syfpeithi (H. Rammensee, Bachmann, Emmerich, Bachor, & Stevanović, 1999)) predict HLA binding, while the other two algorithms (IEDB recommended (Tenzer et al., 2005) and SMMPMBEC (Y. Kim, Sidney, Pinilla, Sette, & Peters, 2009)) predicted candidate peptides based on antigen processing. A common pool of top-ranked peptides from each algorithm is re-ranked using a normalization score from three binding algorithms, and the top candidate peptides are selected.

3.3 Pooled Epitope Screen-1

Of the potential peptidome from a target antigen, only those peptides that can stimulate CD8+ T-cell response in tumor samples will be useful targets for immunotherapy. Both epitope targets and HLA restriction of tumor-reactive T cells are largely unknown for both cancers and pathogens. Long overlapping peptides from target antigens have been widely used in T-cell assays, but serial truncation of positive peptides is still required for epitope identification. This is a labor-intensive and expensive process, limited by the number of samples and more difficult for large antigens (Purcell et al., 2007). The low frequencies of precursor tumor antigen-specific CTLs can be amplified by expansion of antigen-specific CTLs ex vivo in a 10-day stimulation protocol using autologous peripheral blood mononuclear cells (PBMCs) as antigen presenting cells (Fig. 2) (Hida et al., 2002; Parikh et al., 2014). Because of emerging evidence of the role of PD-1/PD-L1 checkpoint blockade to activate a potent antitumor immunity in HPV-associated as well as other tumors (Binder et al., 2013; Lyford-Pike et al., 2013), anti-PD-1 antibody is used on day 1 of our culture protocol to inhibit antigen-specific T-cell suppression. Since candidate peptides are predicted for several HLAs, they are pooled into separate 8–10 peptide pools. Peptide pools are designed to limit intra-pool binding competition by different peptides to the same HLA. Activation of antigen-specific CTLs is identified by standard interferon gamma (IFNγ) enzyme-linked immunospot (Elispot) assay. The mean spot-forming units (SFUs) from each peptide pool are assessed in triplicate. The mean SFU of any peptide pool greater than twice the mean SFU of negative control (PBS-DMSO or irrelevant peptide pool) with statistical significance ($P < 0.05$ by two-sample T-test) is considered as a positive response.

3.4 Peptide Pool Deconvolution Screen-2

Once peptide pool(s) that has a positive signal from several patient PBMCs is identified by the primary screen culture protocol, the minimal peptide(s) responsible for CTL stimulation is identified by deconvoluting the peptide pool (Fig. 3-2). The same ex vivo short-term culture protocol (including anti-PD-1) is repeated now with individual peptides from the positive peptide pool from the initial screen (screen-2). The reactive parent peptide pools from screen-1 are

included as a biological replicate. Concurrently, the positive responder's HLA-class I type is identified either by HLA-specific monoclonal antibodies using flow cytometry (low-resolution HLA typing) or by commercial HLA-typing (high-resolution) services such as the type HLA (ProImmune, Oxford, UK). Low-resolution HLA typing by flow cytometry is performed on PBMCs set aside during screen-1 during any of the 2–3 days following day-1 stimulation. High resolution commercial HLA typing requires genomic DNA (2µg total) isolated from the PBMCs and sent out to commercial services. Because it requires more material, it is usually done on the last day of the Elispot screen. Cells are collected from the Elispot plate and washed once, and genomic DNA is isolated.

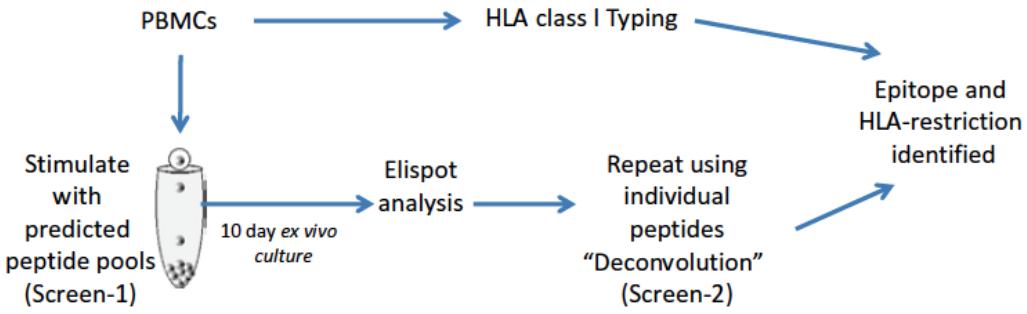

Figure 3-2. Ex vivo short term cultures, epitope deconvolution, and HLA-restriction identification.

If there are limitations on sample availability, only those peptides that correspond to the patient's HLA-class I types are tested. In such cases, a small number (~1 million cells) of donor PBMCs is set aside during screen-1, and HLA typing is performed (low and/or high resolution). Once screen-1 is complete, only those candidate peptides from a pool that are predicted to bind the donor's HLA type are tested in the subsequent screen-2. This minimizes the amount of sample and the number of peptides to be tested. However, this approach will not successfully identify cross- reactive promiscuous HLA binders, which can be lost when focusing on donor-HLA-specific candidate peptides.

3.5 Materials

3.5.1. Comparative CTL Epitope Prediction Strategy

1. Antigenic protein sequence. Usually obtained from literature or from National Center for Biotechnology (NCBI) RefSeq database (Pruitt et al., 2014) and UniProt servers (UniProt Consortium, 2015).

2. Immune epitope database (IEDB) MHC-peptide binding algorithms: NetMHCpan, IEDB consensus binding prediction tool, both accessible at http://tools.immuneepitope.org/main/tcell/.

3. SYFPEITHI algorithm: http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm.

4. IEDB antigen-processing: IEDB-consensus antigen-processing tool, SMMPMBEC antigen-processing algorithms http://tools.immuneepitope.org/main/tcell/

5. Microsoft Excel

6. R statistical software.

3.5.2. Assessing peptide immunogenicity by short term ex vivo cultures (Screen-1)

1. 1x tissue-culture grade phosphate buffered saline (PBS) (Cellgro, Mediatech, VA, USA).

2. Dimethyl sulfoxide (DMSO) (Sigma, St. Louis, MO, USA).

3. Acetic acid (Amresco, Solon, OH, USA).

4. T-cell culture media: To RPMI-1640 (ATCC, Manassas, VA, USA), add 100 U(µg)/mL penicillin-streptomycin (Gibco, Grand Island, NY, USA), 10 mM HEPES (Gibco, Grand Island, NY, USA), 2 mM l-glutamine (Gibco, Grand Island, NY, USA), and 10 % human serum (Gemcell, USA). Filter through 0.22µm Corning sterile filter, IL-2, IL-7. Store at 4℃.

6. Recombinant human IL-2 (R&D Systems, MN, USA): Reconstitute at 100µg/mL (=1,640,000 U/mL) in 100mM sterile acetic acid. Add 100 µL acetic acid into 16.6 mL water to make 100 mM acetic acid; filter through 0.22µm Corning sterile filter. Store at -80°C. Working solution is 2000 U/mL in sterile PBS. Final concentration in culture is 20 U/mL and can be stored at 4°C.

6. Recombinant human IL-7 System Systems, MN, USA): Reconstitute at 50µg/mL in sterile PBS. Keep at −80℃. Working solution at 1µg/mL in sterile 1x PBS, stored at 4℃ until use. Final concentration in culture at 10ng/mL and can be stored at 4℃.

7. Peptide pools (ProImmune, Oxford, UK, other commercial sources are also available): Peptide purity >70%. Reconstitute all stock peptides at 20mg/mL in DMSO or according to manufacturer specifications (for certain peptides) and store at −20°C. To create a working peptide pool, add each peptide corresponding to 1 mg/mL final concentration and make up the rest of the volume in sterile 1x PBS to 1mL. Make smaller 100µL aliquots and store at –20°C. Working peptide pool tube can be stored at 4°C for about 2 months. Final concentration in culture is 10µg/mL.

8. CEF-peptide pool (ProImmune, UK): Reconstitute at 20mg/mL in DMSO. Follow manufacturer's instructions and make a 1 mg/mL stock using sterile 1x PBS. Store in −20°C and an aliquot in 4°C. Final concentration in culture is 1µg/mL.

9. Phytohemagglutinin M form (PHA-M), for positive stimulation (Gibco, NY, USA).

10. DMSO in 1x sterile PBS can be used as a negative control.

11. Anti-PD-1 antibody: Antihuman CD279 (PD-1) purified, clone eBioJ105 (eBioscience, CA, USA). Store at 4°C. Stock 0.5mg/mL. Working concentration is 1µg/mL. Store at 4°C.

12. Levy counting chamber (Hausser Scientific, USA).

13. Centrifuge (Beckman Coulter, Pasadena, CA, USA).

14. Antihuman IFNγ monoclonal antibody, 1-D1K (1 mg/ mL) (Mabtech, OH, USA). Store at 4℃.

15. Biotinylated antihuman IFNγ monoclonal antibody, 7-B6-1 biotin (1 mg/mL) (Mabtech, OH, USA). Store at 4°C.

16. BCIP/NBT Color Development Substrate (Promega, Madison, WI, USA). Store at −20℃. 17. Ethanol (Avantor Performance Materials, Center Valley, PA, USA).

18. Fetal bovine serum (FBS). Store at −20℃ after heat inactivation for 20 min at 56℃.

19. Tris (Sigma, St. Louis, MO, USA). Store at room temperature.

20. Magnesium chloride (MgCl2) (Sigma, St. Louis, MO, USA). Store at room temperature.

21. Sodium chloride (NaCl) (Sigma, St. Louis, MO, USA). Store at room temperature.

22. Multiscreen filter plate, 2EM004M9 or MSIPS4W10 (Millipore, Billerica, MA, USA).

23. AID or other ELISPOT reader (Autoimmun Diagnostika, Strassberg, Germany).

24. 1x PBS 0.5 % FBS buffer for ELISPOT washes (0.5 % FBS wash solution).

3.5.3. HLA-Typing and Positive Peptide Pool Deconvolution (Screen-2)

1. Materials 1 through 23 from Subheading 2.2.

2. Individual peptides from the immunogenic peptide pool identified through screen-1. Reconstitute all stock peptides at 20 mg/mL in DMSO or according to manufacturer specifications (for certain peptides) and store at −20℃. To create a working peptide solution, make a 1mg/mL final concentration and make up the rest of the volume in sterile 1x PBS to 1mL. Make smaller 100µL aliquots and store at 20℃. Working peptide tubes can be stored at 4℃ for about 2 months. Final concentration in culture is 10µg/mL.

3. Fluorescently conjugated monoclonal antibodies (mAbs) for 10 HLA alleles e.g., HLA-A2 mAb clone BB7.2-PE conjugated; from BD Pharmingen, San Jose, CA, USA).

4. Fluorescently conjugated isotype controls (e.g., mouse IgG2a-PE for HLA-A2 staining, BD Pharmingen, San Jose, CA, USA).

5. Attune or a similar flow cytometer (Life Technologies, Grand Island, NY, USA).

6. Staining buffer. 1x PBS (Cellgro, VA, USA) with 1% bovine serum albumin (BSA, Sigma, USA).

7. DNAzol Reagent (Life Technologies, NY, USA) for genomic DNA isolation.

8. Sodium hydroxide monobasic (Sigma, St. Louis, MO, USA) used with DNAzol.

9. 100 % ethanol (Avantor Performance Materials, Center Valley, PA).

10. Nuclease-free water (Hyclone, Thermo Fisher Scientific, Waltham, MA, USA).

11. Nanodrop 2000 C (Thermo Fisher Scientific, Waltham, MA, USA).

3.6 Methods

3.6.1. Comparative CTL Epitope Prediction Strategy

1. Access IEDB prediction server: (http://tools.immuneepitope.org/main/tcell/).

2. Select “Peptide binding to MHC class I molecules” link.

3. Enter the antigen protein sequence in FASTA format.

4. Select “Consensus” or “NetMHCpan” as the prediction method.

5. Select “human” as the MHC source species.

6. Select the desired “HLA allele” (each of HLA-A’s and HLA-B’s listed in Subheading 1.3).

7. Select “all lengths” as the predicted peptide length.

8. Choose “Percentile rank” for IEDB-consensus and “Predicted IC50” for NetMHCpan as the output type to sort the peptides.

9. Choose “XHTML table” as the output format.

10. Download the results as an excel file.

11. Select “Proteasomal cleavage/TAP transport/MHC class I combined predictor” IEDB server.

12. Enter the antigen protein sequence in FASTA format.

13. Select “IEDB recommended” or “SMMPMBEC” as the prediction method.

14. Repeat steps 5–7 for the different HLA alleles.

15. Sort the peptide list by “Total score” for both the algorithms (should be in a decreasing order).

16. Repeat steps 9 and 10.

17. Access SYFPEITHI: http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm.

18. Select the “MHC type” for each of the HLA alleles.

19. Choose “All mers” for peptide length.

20. Paste the antigen protein sequence in non-FASTA format. Click Run.

21. Save the output by copying all the predicted peptides into MS Excel and removing blank spaces between different lengths.

22. From each prediction algorithm’s output per HLA, select the top two-thirds (66 %) of all predicted peptides (i.e., for 100 predicted peptides, select the top 66). Do this for each prediction algorithm’s output in MS Excel or database program.

23. By using a local script on R, create a separate list of peptides common between the five prediction algorithms.

24. Using the output scores as described above normalize the common pool of predicted peptides as follows: (a) For algorithms displaying results ranging from low score to high score

(IEDB-consensus binding and NetMHCpan), $\overline{\delta_i} = (\delta_i - \delta_{\min})/(\delta_{\max} - \delta_{\min})$, where $\overline{\delta_i}$ represents the normalized score of the peptide; δ_i, the assigned output score by the prediction algorithm; $\delta_{\min}$, the minimum score assigned in prediction output; and $\delta_{\max}$, the maximum score assigned in the entire prediction output.

25. Calculate the average binding score S_B from the three normalized binding scores.

26. Using the "sort" function in MS Excel, sort the S_B score list ranging from the lowest to the highest. Lower the S_B score, higher its probability of being a candidate peptide.

27. Select the top candidate peptides including promiscuous binders (usually up to 5 or 10, depending on feasibility) and order synthetic peptides for immunogenicity assessment. Greater than 70% purity is sufficient for T-cell assays.

28. Once synthesized peptides are received, reconstitute the peptides according to manufacturer's instructions.

29. Create the peptide pools by having between 1 and 3 binders per HLA allele within each pool. This limits intra-pool binding competition by different peptides for the same HLA. Store the peptide pools at 4℃.

3.6.2. Peptide pool immunogenicity assessment by short-term ex vivo cultures

1. Obtain a frozen PBMC cryovial from liquid nitrogen and rapidly thaw in 37℃ water bath.

2. Pipette the cells drop by drop into a 15 mL tube with pre-warmed T-cell culture media.

3. Centrifuge at 300g for 5 minutes, remove supernatant for wash.

4. Resuspend cells gently in T-cell culture media. Count cells using a hemacytometer. Set aside 500,000 PBMCs for low-resolution HLA typing (Subheading 3.3).

5. Prepare cell suspension at 1 x 10^6 cells/mL. Seed cells in a round bottom 96-well plate, 200 µL/well (=200,000 cells/well).

6. On day 1, stimulate with 20U/mL IL-2, 10ng/mL IL-7 (1µg/mL Anti PD-1), and 10µg/mL peptide pool or CEF-peptide pool. Use 1 % PHA-M for positive control stimulation (2µL/well).

7. On day 5, remove 100µL media from each well and add 100µL/well fresh T-cell culture media. Add IL-2 for final concentration at 20 U/mL and peptide pool at concentration 10µg/well.

8. On day 7, pre-wet an ELISPOT plate with 35% ethanol at room temperature, 35µL/well.

9. Wash with sterile water for five times, 200µL/well.

10. Coat with capture antibody 1-D1K, 100µL/well. Dilute capture antibody in 1x PBS at final concentration 5µg/mL.

11. Seal the plate and incubate at 4℃ overnight.

12. On day 8, wash with RPMI 1640 (serum free is fine) for five times, 200µL /well.

13. Add T-cell culture media, 200µL/well. Keep at room temperature in the hood for 30 minutes.

14. On day 8, remove 100µL media from each well and add 100µL/well fresh T-cell culture media. Add IL-2 for final concentration at 20U/mL, and peptide pool at concentration

10 µg/well. Add IL-2 for final concentration at 20U/mL.

15. Add cells from each of the wells (200µL/well) from the culture plate into a corresponding well on the ELISPOT plate, triplicate for assays. Stimulate with antigenic peptide at final concentration 10µg/mL (2µg/200 µL). The same quantity of CEF-peptide pool as peptide controls. Use 1 % PHA-M (2µL/well) for positive control.

16. Cover with plate lid and incubate in CO2 incubator for 48 hours (days 8–10).

17. On day 10, wash with sterile 1x PBS for three times, 200µL/well. Save cells for genomic DNA isolation (high-resolution HLA typing, Subheading 3.3).

18. Wash with 0.5% FBS wash solution three times, 200µL/well with intervals between washes.

19. Add detection antibody 7-B6-1 Biotin, 100µL/well. Dilute detection antibody in 0.5% FBS wash solution at final concentration 1µg/mL.

20. Seal the plate and incubate at room temperature for 2 hours in the dark. (Leave the sealed plate in the hood if preferred.)

21. Wash with 1x PBS for three times, 150µL/well.

22. Wash with 0.5 % FBS wash solution for three times, 150µL/well, 5 minutes interval washes.

23. Add streptavidin-ALP, 100µL/well. Dilute streptavidin-ALP in 0.5 % FBS wash solution at final concentration 1:1000.

24. Incubate at room temperature (preferably dark) for 1 hour.

25. Wash with 1x PBS for three times, 150µL/well.

26. Wash with 0.5 % FBS wash solution for three times150µL/well.

27. Add BCIP/NBT mixed solution, 100µL well.

28. Watch for spot development.

29. Stop reaction by rinsing with cool tap water.

30. Dry the plate in the dark at room temperature overnight (leave the lid partially open).

31. Read the ELISPOT plate after 24–48 h on the AID ELISPOT reader.

32. Calculate mean Spot Forming Units for each peptide pool.

3.6.3. HLA-typing and positive peptide pool Deconvolution (Screen-2)

1. Low-resolution HLA typing is performed using the 500,000 cells set aside on day 1 of screen-1.

2. Pool cells in media into a single tube and wash once with staining buffer (500g, 3 min).

3. Divide the washed PBMCs into appropriate number of HLAs and isotype controls to be tested, in 100µL staining buffer per tube.

4. Add appropriate mAbs and isotype controls to be analyzed.

5. Incubate on ice in the dark for 30 min.

6. Wash twice with staining buffer. Resuspend in a final volume of 200–300µL 1x PBS for flow cytometry analysis.

7. Run isotype controls, gate on live lymphocyte population, and collect 10,000 events/gate.

8. Run the individual samples stained with HLA-mAbs and analyze.

9. Genomic DNA (gDNA) extraction for high-resolution commercial HLA typing is performed on the last day (day 10) of screen-1; when incubating, PBMCs are washed from the Elispot plate.

10. Collect cells from the Elispot plate (contains media) in a reservoir.

11. Wash the Elispot plate with 1x PBS and collect remaining cells in the reservoir.

12. Pool cells from reservoir into a 15 mL conical tube.

13. Centrifuge the cells, 500g 3 minutes. Remove supernatant.

14. Wash the pellet once with 5 mL 1x sterile PBS, centrifuge and remove the supernatant.

15. Add 500 µL DNAzol directly to the pellet.

16. Follow manufacturer's protocols and isolate gDNA.

17. Determine the concentration of gDNA using nanodrop. 2µg gDNA is shipped to the company for high-resolution HLA typing.

18. Once screen-1 (Subheading 3.2) is completed, perform data analysis to identify the peptide pool with statistically significant higher mean SFUs compared to the negative control

19. Repeat the ex vivo short-term culture protocol (steps 1 through 32) using individual peptides from the reactive peptide pool from screen-1.

20. At the end of the short-term culture, analyze mean SFU from individual peptides as described before and correlate with donor-HLA type to identify reactive CTL epitope and its HLA restriction.

CHAPTER 4

IMMUNOGENIC AND DYSFUNCTIONAL CD8+ T-CELLS IN HPV+HEAD AND NECK CANCER

Abstract

Human papillomavirus subtype 16 (HPV16) is the primary cause of an increasing number of head and neck squamous cell carcinomas (HNSCCs), providing strong rationale for T-cell immune therapies against HPV+ HNSCC. Here, we sought to assess immunogenicity of HPV16-specific CD8+ T-cells (CTLs), and to understand HPV-specific mechanisms of T-cell dysfunction. We experimentally identified 16 strong and 29 moderately immunogenic CTL-epitopes from HPV16 E2, E6 and E7 antigens restricted by 12 common HLA class I alleles. Relative to E6 and E7, E2-specific CTL-reactivity is higher in HPV+ HNSCC patients than in healthy controls (>3-fold, P = 0.012). Upon antigen re-exposure, E7/E2-CTL-dysfunction phenotype was observed in more patients than E6-CTL-dysfunction, indicating inter-antigenic heterogeneity of HPV-CTL-exhaustion. Immunogenomic analyses of 119 HNSCC transcriptomes revealed high T-cell infiltration and dysfunction in HPV+ HNSCCs, and correlation of HPV-antigen expression with T-cell exhaustion gene signatures. We found that Indoleamine 2,3-dioxygenase (IDO-1), is strongly expressed in HPV+ HNSCCs compared to HPV- HNSCCs (P = 0.001), and correlates with E7-expression in vitro (R^2 = 0.84, P = 0.033). Combination treatment with PD-1 blockade and IDO-1 inhibition overcomes profound CTL-dysfunction, enhancing HPV+ HNSCC sensitivity to CTL-cytotoxicity in vitro (up to 10-fold in E7-CTLs, P = 0.011). Our findings implicate mechanisms of T-cell escape in HPV+ HNSCC, wherein high tumoral HPV-antigen load results in high expression of immune dysfunction genes on tumor cells (e.g. IDO-1), and dysfunction of HPV-specific CTLs (e.g. E7,E2-CTLs). The HPV16 CTL-epitopes identified in this study, in synergy with combination blockade of HPV+ HNSCC-specific checkpoints may be useful for targeted immunotherapy.

4.1 Significance

Strategies enhancing immune based targeting of tumor cells have come to the forefront of cancer treatments. A subset of head and neck cancers (HNSCCs), caused by human papillomavirus 16 (HPV16), is an ideal candidate for T-cell cancer immunotherapies. Here, we identified immunogenic CD8+ T-cell (CTL) epitopes from 3 HPV16-antigens, and studied T-cell dysfunction mechanisms in HPV+HNSCC. We detected several novel CTL-epitopes from HPV16-genes E2, E6 and E7 across multiple HLA-alleles in peripheral blood CTLs of HPV+ HNSCC patients. We identified that tumoral viral load largely drives T-cell infiltration and subsequent CTL-exhaustion observed in HPV+HNSCC. Our study underscores the importance of host immune control of HPV, and identifies combination PD-1/IDO-1 inhibition as a novel strategy to enhance CTL-targeting of HPV+HNSCC.

4.2 Introduction

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer worldwide with close to 600,000 cases diagnosed annually (Torre et al., 2015). A subset of HNSCCs are caused by the human papillomavirus (HPV), (HPV+ HNSCCs) (Marur, D'Souza, Westra, & Forastiere, 2010; Parfenov et al., 2014a), which are molecularly and clinically distinct from non-HPV associated HNSCCs (HPV- HNSCCs)(The Cancer Genome Atlas Network, 2015). In the United States, 70-80% of HPV+ HNSCCs are caused by the oncogenic HPV type 16 (HPV16) (Chaturvedi et al., 2013; Dahlstrom, Anderson, & Sturgis, 2017; Giordano & Macaluso, 2016). Incidence of HPV+ HNSCC increased 225% from 1984-2004 and has now surpassed the incidence of cervical cancer (Maura L. Gillison et al., 2012; Jemal et al., 2013). Although HPV vaccines effectively prevent HPV-related cancers, the impact of vaccination on HNSCC incidence may not occur until 2060, likely due in part to slow vaccine uptake, and the decades between infection and clinical HPV+ HNSCC diagnosis (M. L. Gillison, Chaturvedi, Anderson, & Fakhry, 2015; Williams et al., 2017). As a result, over 600,000 cases are predicted in the interim, providing a strong rationale for the development of novel therapeutic strategies against HPV+ HNSCC.

The recent development of clinically effective tumor immunotherapies, such as checkpoint blockade (CKB) using PD-1/PD-L1 inhibitors (Herbst et al., 2014; Pardoll, 2012; Robert et al., 2015), has led to FDA approval of nivolumab and pembrolizumab for HNSCC (Chow et al., 2016; Ferris et al., 2016). It is now established that the clinical response to CKB is correlated with tumor neo-epitope load. Tumor-specific neo-epitopes have been directly targeted using therapeutic vaccines and/or adoptive T-cell therapy (ACT), and have been shown to enhance cytotoxic T-cell targeting of multiple solid tumors (Ott et al., 2017; Schumacher & Schreiber, 2015; Tran et al., 2016). Thus, there is a renewed interest in defining the human leukocyte antigen (HLA) restricted antigenic repertoire of tumor infiltrating lymphocytes (TILs) to develop targeted therapeutic vaccines (Kenter et al., 2009), to identify T-cell receptors (TCR) for ACT (Schumacher & Schreiber, 2015), and to monitor clinical responses to complex cancer immunotherapies (Rizvi et al., 2015; Rosenblatt et al., 2016).

HPV-associated cancers express multiple viral neo-antigens. HPV integration into host genome in cervical cancer results in derepression of the oncogenic drivers E6 and E7 (Woodman, Collins, & Young, 2007). Immune therapies targeting E6 and E7 have thus been developed, including peptides, DNA, and ACT therapies (Draper et al., 2015; Kenter et al., 2009; Trimble et al., 2015). However, in comparison to cervical cancer, HPV+ HNSCCs have both lower rates of genome integration, and less interruption of the viral transcriptional regulatory gene E2 (Akagi et al., 2014; Parfenov et al., 2014b; Rusan, Li, & Hammerman, 2015). Thus, subsets of HPV+ HNSCCs also express E2 in addition to E6 and E7 (Parfenov et al., 2014b; Zhang et al., 2016).

We and others have previously shown that high titers of serum antibodies against HPV16-E2, E6 and E7 are detectable in most HPV+ HNSCC patients, indicating immunogenicity and persistence of these antigens (Anderson et al., 2015; D'Souza et al., 2007). We thus hypothesized that patients with HPV+ HNSCC would have pre-existing HPV-specific CTLs, and that HPV-antigen expression levels would influence CTL-dysregulation in tumor microenvironment. We identified the T-cell antigenic landscape of globally frequent HLA class I alleles from HPV16 E2, E6 and E7. By phenotyping HPV-specific CTLs from HPV+ HNSCC

patients, and analyzing the immune transcriptomes of 119 HNSCCs, we demonstrate intratumoral and peripheral CTL-dysfunction in HPV+ HNSCC. We show that this CTL-dysfunction can be reversed using targeted HPV-specific T-cell expansion, and synergistic inhibition of IDO-1 and PD-1. These results have implications for the development of effective T-cell therapies for HPV+ HNSCC.

4.3 Results

4.3.1. Frequency and specificity of HPV16 E2, E6 and E7-specific CTLs in HPV+ HNSCC

We performed a systematic analysis of potential CTL-epitopes from HPV16 E2, E6 and E7 antigens restricted by 15 globally frequent HLA class I alleles representative of major HLA supertypes (Lund et al., 2004) (Fig 4-1 A-D). We used a comprehensive CTL-epitope prediction strategy we previously developed by incorporating stringent selection criteria (Appendix B, Methods) to control for inter-algorithmic variations (Chowell et al., 2015; Krishna & Anderson, 2016). Fifty-nine candidate peptides (24 from E2, 20 from E6 and 15 from E7) were selected covering 13 of the 15 common HLA class-I alleles as candidate HPV-CTL peptides based on predicted HLA-affinity and antigen processing percentile scores (Fig. 4-1 A, 4-1 E-F). Several previously described HLA-A*02:01-restricted HPV16 E6 and E7 epitopes were predicted with high scores (e.g. E6-KLP epitope, total percentile 94.6, Table B1) confirming our prediction strategy. Within the 59 candidate HPV16-peptides, E2 had the lowest number of previously defined CTL-epitopes (3/24, 12%), while E6 and E7 had higher number of previously described CTL-epitopes (35% and 46% respectively). The number of predicted HPV16-peptides ranged from 15 peptides (A*02:01), to 0 peptides (B*40:01, B*44:02) among the selected HLA-alleles (Fig 4-1 A). To determine if lack of HLA-binding motifs in the 3 HPV-antigens can poise specific HLA-alleles as risk-factors for HPV+ HNSCCs, we calculated the odds-ratio of HLA-allele frequencies in HPV+ HNSCCs (N=77), compared to HPV- HNSCCs (N=64) (Appendix B). HLA B*40:01, which had no predicted HPV16-peptides for E2, E6 and E7 had an odds-ratio of 7.48 compared to HPV-HNSCCs (Fig. 1*A-B*, $P = 0.059$), and had poor-binding peptides for all HPV16-antigens (bottom 20^{th} percentile compared to other HLAs, Appendix Fig. B1). HLA-alleles A*24:02, B*07:02, and

B*51:01 were also overrepresented (OR>=2) in HPV+ HNSCCs, although they were not statistically significant. Of note, HLA-B*07:02 (OR = 2, Fig. 4-1 B), has been previously reported to be associated with poor clinical outcome in cervical cancer and escape HPV-specific T-cell (HPV-CTL) recognition (Ellis et al., 1995). These results point to the importance of CTL-mediated control of HPV16 malignancies.

Because HPV-CTLs in peripheral blood mononuclear cells (PBMCs) are not abundant (Badoual et al., 2013; Riemer et al., 2010), we used PBMCs stimulated for 10 days with candidate peptides and CKB antibodies αCTLA4 and αPD-1 to enhance HPV-CTL reactivity (Appendix B, Fig. B-2). We compared HPV-CTL frequency in PBMCs between HPV+ HNSCC patients (N=18) and healthy controls (N=14) by interferon gamma (IFNγ) Elispots using antigen specific peptide pools (Fig. 4-2 A-B, Appendix B Table B1) (Parikh et al., 2014a). The HLA frequency distribution of this cohort largely mirrored median HLA-frequency distribution in the USA (Fig. 4-1 D). IFNγ responses against HPV16-E2 were substantially more common (>3-fold higher) in HPV+ HNSCC PBMCs compared to healthy control PBMCs (Unpaired Welch's T-Test, *P=0.012,* 4-2 A). Moderate to high E6-reactivity was observed in HPV+ HNSCC patients (1.5 fold higher in HPV+ HNSCCs), while E7-reactivity was generally low (Fig. 4-2 B-C). To determine if PBMC T-cell reactivity correlates with B-cell immunity, we measured IgG serologic responses to the E2, E6, and E7 antigens in the 18 patients. E2 and E7-specific serum IgG titers were higher relative to E6 (>2-fold, $P < 0.05$) in the patients (Fig. 4-2 C). The majority of patients who had IgG to E2, E6 and E7 also had a measurable CTL response (E2=72%, E6=60%, E7=70%, respectively, Fig. 4-2 D). There was strong concordance between seroreactivity and T-cell reactivity within same antigens (Chi-squared independence test, $P = 0.03$). These results indicate E2 and E6 antigens are more CTL-reactive than E7 in patients, and HPV-CTL response enhanced by CKB antibodies.

4.3.2. Mapping immunodominant epitopes of HPV16 E2, E6 and E7 in HPV+ HNSCCs

To identify novel CTL epitopes from E2, E6 and E7, we performed a second IFNγ Elispot analysis using individual predicted HPV16-peptides against patient-specific HLA-alleles (Fig. 4-3). 51 out

of 59 predicted peptides elicited a T-cell response in at least one patient, indicating a high degree of success (86%) of our prediction-validation strategy (Fig. 4-3, Table B1). Consistent with pooled

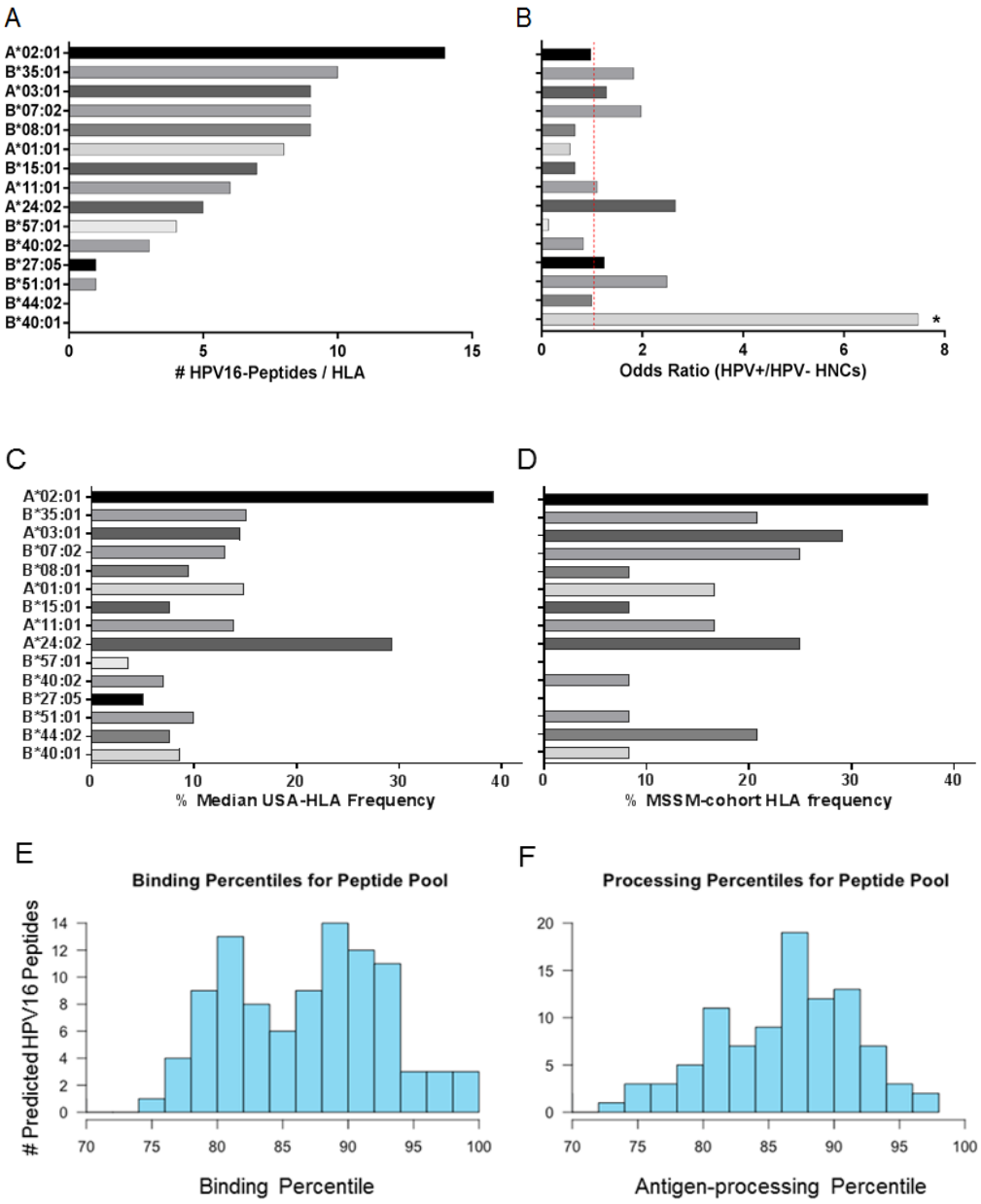

Figure 4-1. HPV16 E2, E6 & E7 predicted epitope distribution and risk factor for HPV+ HNSCCs. (A) Distribution of 59 predicted HPV16-peptides by each HLA-allele ranked from highest to lowest. (B) Odds ratio of HLA-allele frequency in HPV+ HNSCC patients compared to HPV- HNSCC patients **P = 0.059* (C) Median USA HLA-allele frequencies obtained from Allele frequency net database. (D) HLA-frequency distributions of MSSM-cohort. Distribution of 59 predicted HPV16-peptides binned according to (E) total binding percentile scores (F) total antigen-processing percentile scores.

antigen Elispot results (Fig. 4-2 A-B), we observed sub-dominant E7-specific CTL-reactivity relative to E2, E6 epitope-specific CTL-responses (Fig. 4-3). Sixteen epitopes had an average response frequency of >=100 mean spot forming units (SFUs)/10^6 PBMCs and were classified as strongly immunogenic (representative e.g. Fig. 4-4, Table B1). Twenty nine CTL-epitopes had an average response frequency between 10-100 SFU/10^6 PBMCs (moderately immunogenic), while 6 epitopes had an average response frequency <10 SFU/10^6 PBMCs (low immunogenic). The majority of moderate to highly immunogenic epitopes (77%) was novel, or had not been described with the observed HLA restriction (Figs 4-3, 4-4, Table B1). We also observed 16 unique epitopes that elicited a cross-reactive response to other alleles within the same supertype supporting the strategy for HLA-supertype based epitope prediction (Table B1). Figure 4-4 D shows a representative example, where an HLA-A*11:01 restricted E2-peptide had strong predicted binding affinity and elicited strong CTL-reactivity to HLA-A*68:01 (A3 supertype).

Within the E2 antigen, most CTL-epitopes (52%) were clustered within the trans-activating DNA-binding domain, 23% in the hinge region, and 24% in the DNA-binding domain (Fig. 4-5 A). Within E6, immunodominant regions (70% of epitopes) encompassed AAs 37-109 with 40% of epitopes arising in the first zinc finger domain (Fig. 4-5 B). Interestingly, the zinc finger domain of E7 also had 42% of the CTL-epitopes (Fig. 4-5 C). Thus, we have defined immunodominant regions and enhanced the landscape of E2, E6 and E7 CTL-epitopes for future studies.

4.3.3. HPV-specific T-cells exhibit dysfunctional phenotype in HPV+ HNSCC patients

While CTL dysfunction in chronic viral infections and cancers has been described (Barber et al., 2006), few studies have focused on the extent of T-cell exhaustion in HPV+ HNSCC patients because of the difficulties in studying low-frequency HPV-CTLs (Badoual et al., 2013). We assessed HPV-CTL dysfunction in HPV+ HNSCC patients after ex vivo stimulation by autologous antigen-presenting cells (APCs) presenting cognate HPV16-antigen in the absence of CKB antibodies. Our rationale here was that activated HPV-specific PD1+ CTLs that are poised towards the exhaustion spectrum, will become further dysfunctional after APC-stimulation and

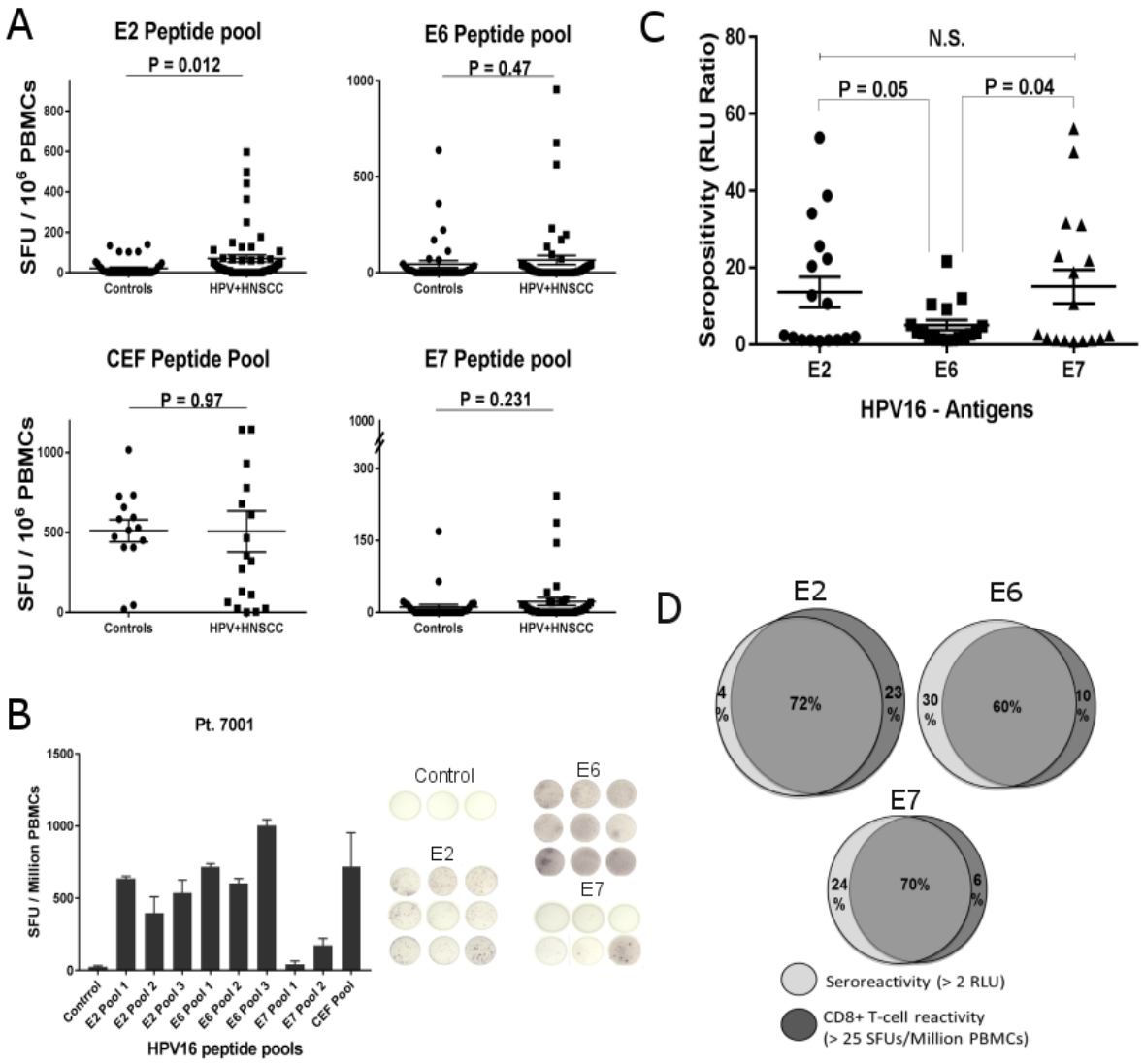

Figure 4-2. HPV16 E2, E6 & E7 T-cell and B-cell immunogenicity in HPV+ HNSCCs. (A) Summary of CTL-reactivity. Predicted HPV16-peptides were pooled according to antigen (Table S1) and tested for CTL-reactivity by IFNγ Elispot. *P*-values from Unpaired Welch's T-test are shown. (B) Representative example of CTL-reactivity from one HPV+ HNSCC patient PBMC with SFUs after background subtraction (left panel) and images from each pool in triplicate (right panel). (C) Seroreactivity of HPV16-E2, E6, E7 antigens in HPV+ HNSCC MSSM patients screened for CTL-responses by Rapid-ELISA *(SI Methods*) (D) Seroreactivity and CTL-reactivity concordance for each HPV-antigen in responding HPV+ HNSCC patients. Circles are proportional to number of responding HPV+ HNSCC patients for each antigen. % represent patient responses.

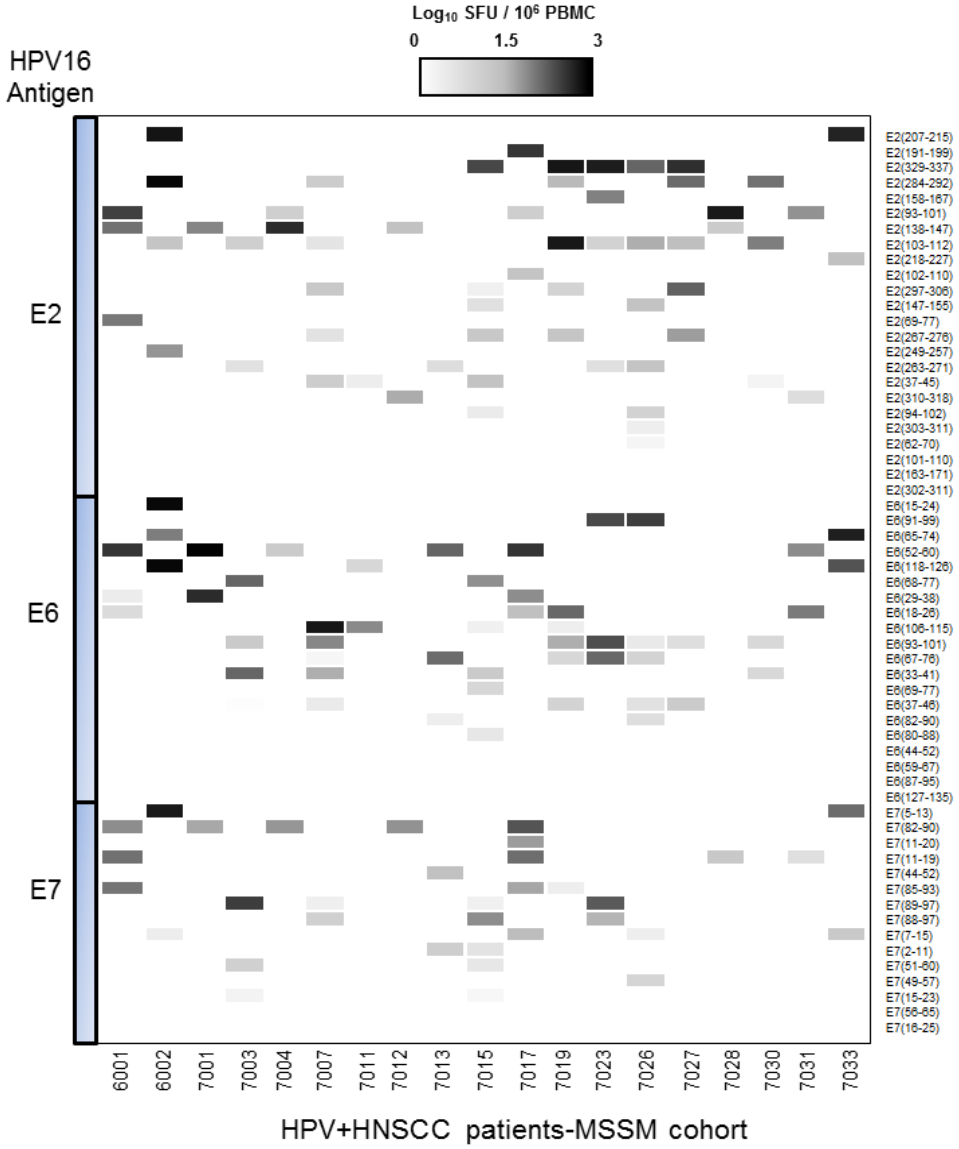

Figure 4-3. Landscape of CTL-epitopes from HPV16 E2, E6 and E7 in HPV+ HNSCCs. HPV+ Summary of Elispot epitope deconvolution screen showing all responding HPV+ HNSCC patients (each column) against tested HPV16-peptides (each row) in log scale. Within each antigen, peptides are ranked from most number of CTL responses (top) to the least (bottom).

acquire additional inhibitory markers characteristic of profound dysfunction, such as CD39 and TIM-3 (Gupta et al., 2015; Wherry, John Wherry, & Kurachi, 2015). In 4 HLA-A*02:01+ patients with HPV-specific CTLs detectable by antigen-specific tetramers (Fig. 4-6), and one HLA-A*68:01+ patient with CD137-positivity defining HPV-CTLs, we phenotyped total CD8+ and E2,

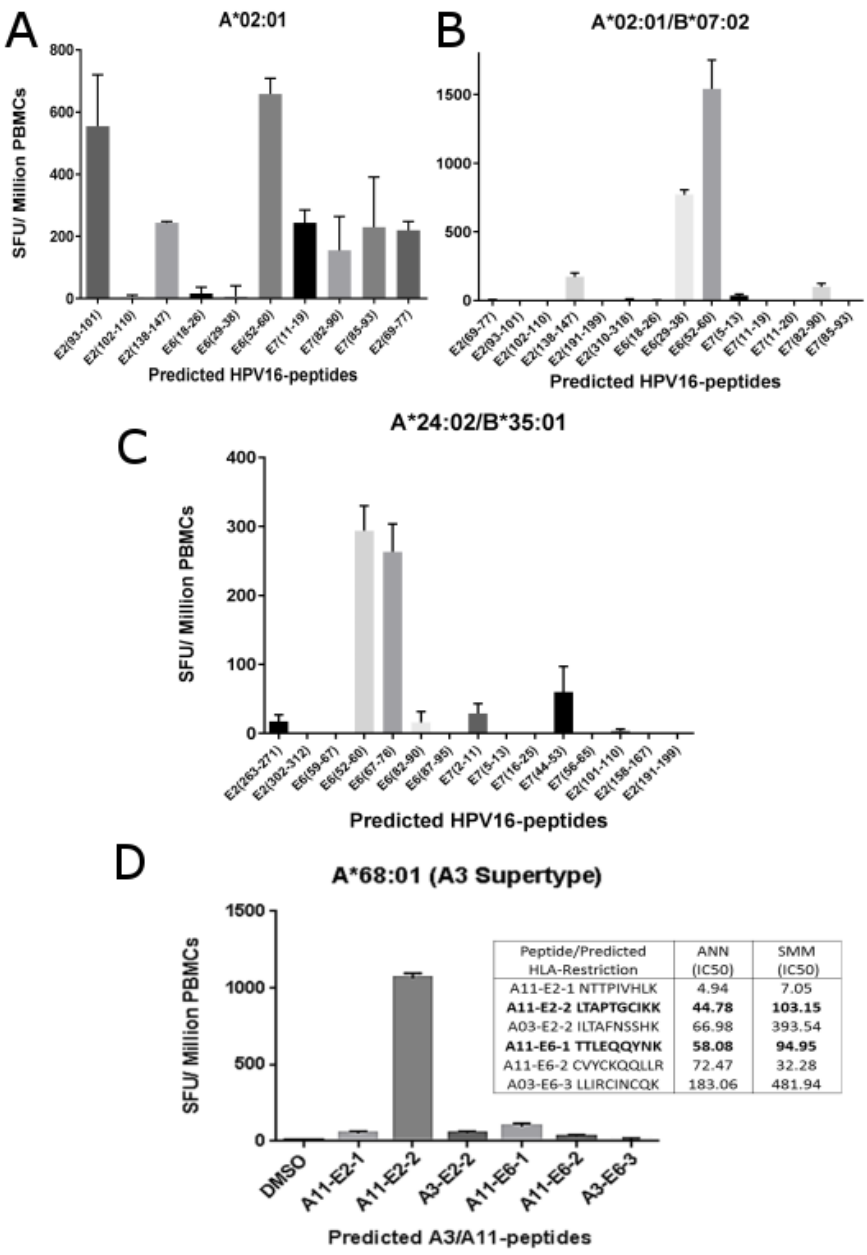

Figure 4-4. Identifying antigen specific epitopes from HPV16 E2, E6 and E7 in HPV+ HNSCCs. Examples of individual responding patients after background subtraction. (A) HLA-A*02:01+ patient; (B) HLA-A*02:01/B*07:02+ patient; (C) HLA-A*24:02/B*35:01+ patient. (D) HPV16-peptides predicted for the HLA-A3-supertype can stimulate a CTL response to representative allele (A*68:01). Inset shows binding affinities for predicted peptides for A*68:01, peptide labels shows HLA-allele the peptide was originally predicted. Positive responders are shown in bold. *$P < 0.1$, **$P < 0.01$, Unpaired two-tailed Welch's T-test.

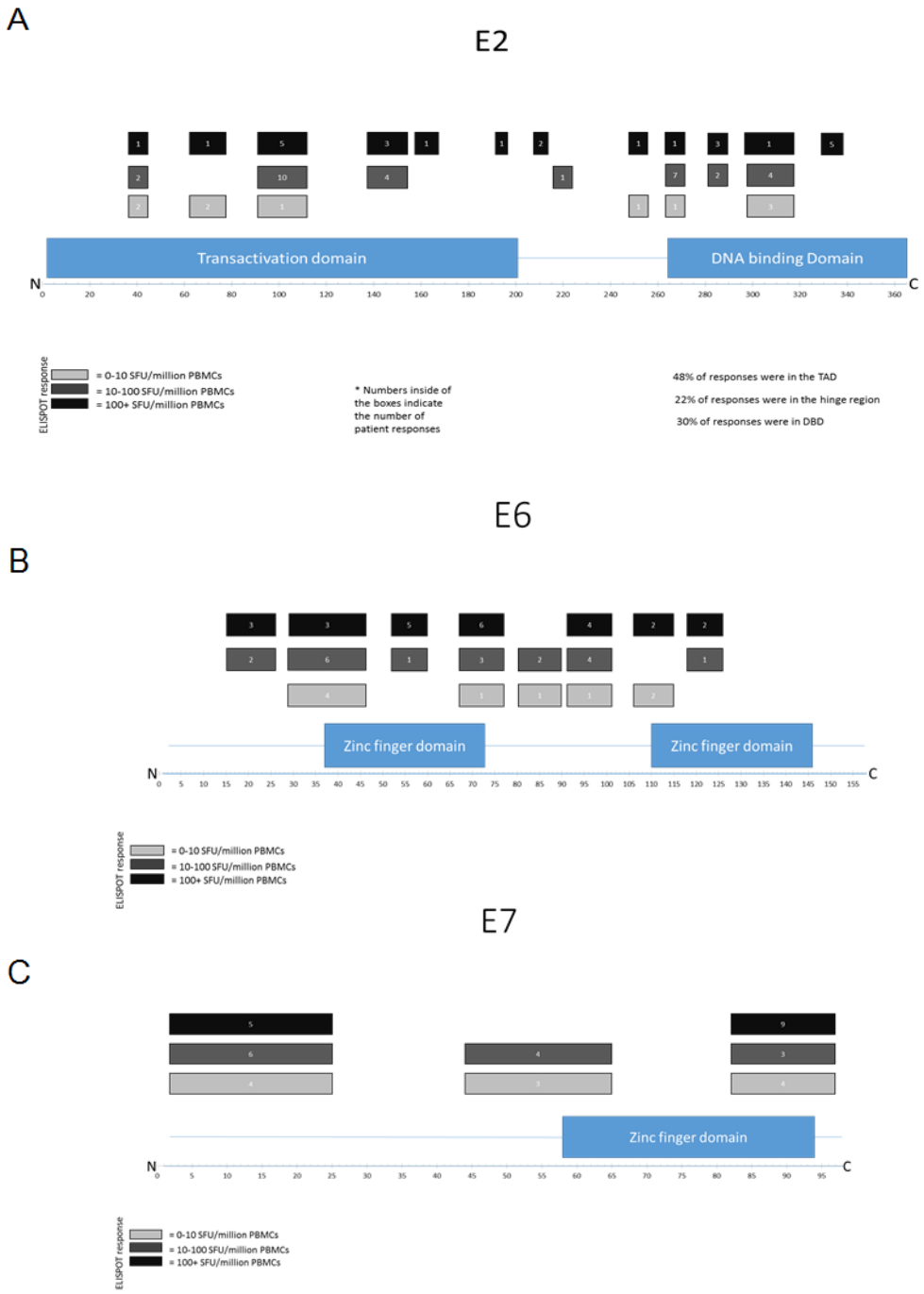

Figure 4-5. Immunodominant regions of HPV16-E2, E6 and E7. All immunogenic CTL-epitopes identified in this study mapped onto the 3 HPV16-antigens. Protein domain information was obtained from PAVE. Strength of immune response in the regions are indicated by the shaded boxes encompassing all peptides from the region, numbers indicate the number of unique HPV+HNSCC MSSM patient-specific responses.

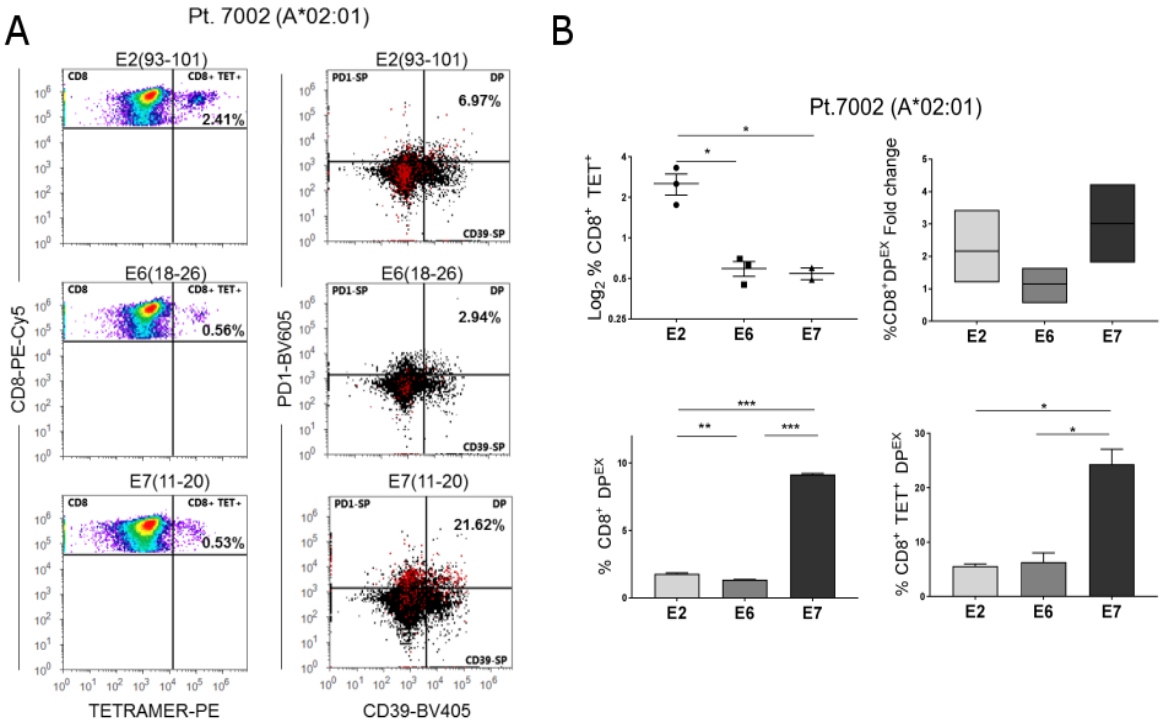

Figure 4-6. HPV16-specific T-cells acquire dysfunctional phenotype upon ex vivo stimulation. (A) Representative flow cytometry plots from an HLA-A*02:01 HPV+ HNSCC patient CTLs stimulated with autologous APCs transfected with cognate antigen. Left: HPV16-Tetramer+ CD8+ T-cells one example for each HPV16-antigen; labels correspond to HPV16-epitope, % indicate tetramer+ events within CD8-gate; Right: CD8+PD1+CD39+ (in black) or CD8+Tetramer+PD1+CD39+ (DP^{Ex}-phenotype, back gated in red). % - back gated CD8+Tetramer+DP^{Ex}. (B) Pt.7002 HPV-CTL dysfunction, Top panel: CD8+ Tetramer+ events from the 3 HPV16-antigens (left), each point is an HPV16-epitope-tetramer. Fold change in total CD8+DP^{Ex} % after CTL-stimulation with HPV16-antigen transfected APCs, compared to mock transfected APCs (right). Bottom panel: % Total CD8+DP^{Ex} (left), % CD8+Tetramer+DP^{Ex} (right). $P < 0.1$, $^{**}P < 0.01$, $^{***}P < 0.001$, Unpaired two-tailed Welch's T-test.

E6 and E7-specific CTLs exhibiting CD8+CD39+ PD-1+ or CD8+TIM-3+ PD-1+ phenotype (DP^{Ex}-phenotype) indicating substantial exhaustion. As shown in the representative example Pt.7002, two weeks after stimulation, E2-CTLs were higher in frequency than E6 and E7-CTLs (Fig. 4-6 A). Within the HPV16-antigens in Pt.7002, CTLs stimulated with E7-transfected APCs exhibited the highest levels of CD8+DP^{Ex} fold-change relative to mock-antigen transfected (2-4 fold,Fig. 4-6 B) followed by E2 (1-3 fold, Fig. 4-6 C) and E6 respectively.

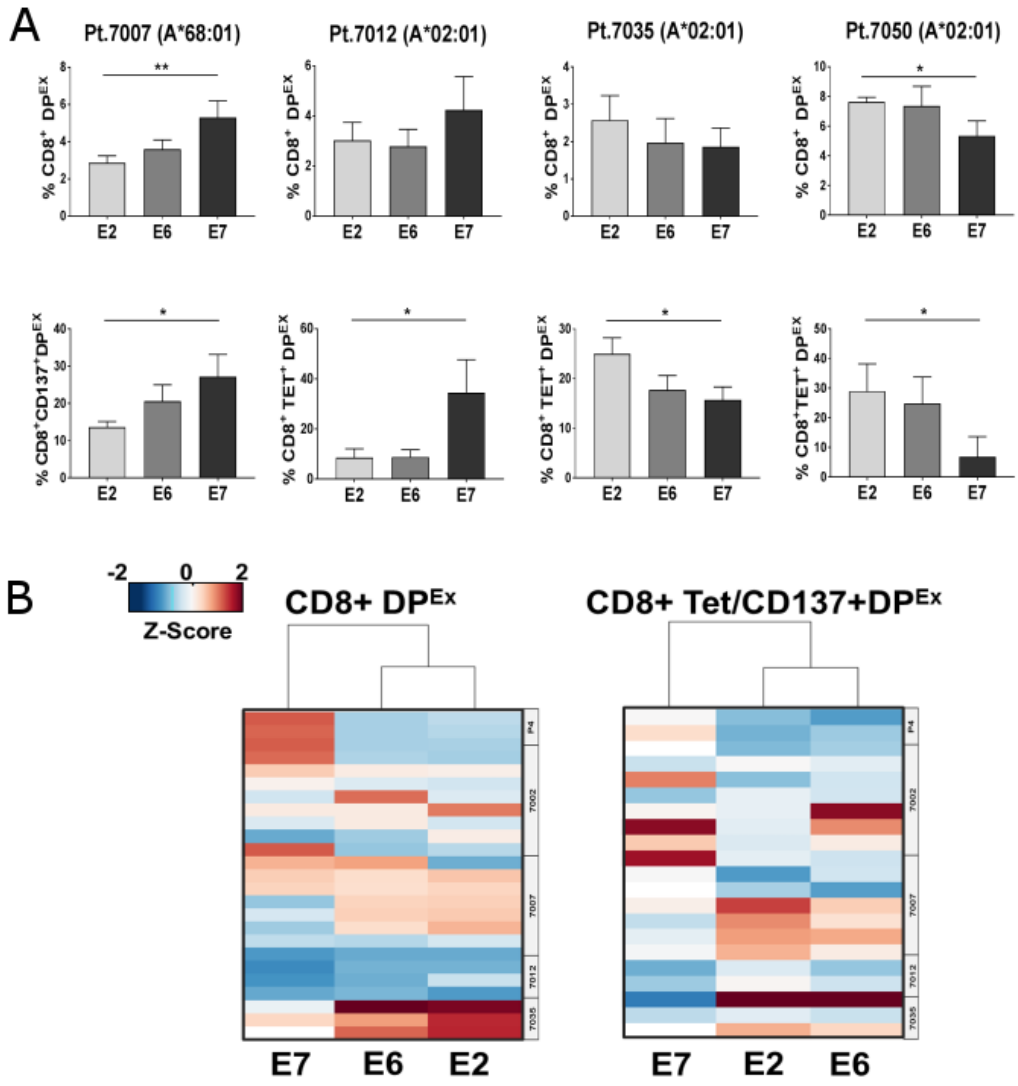

Figure 4-7. Dysfunctional phenotype heterogeneity upon ex vivo stimulation. (A) Summary of dysfunction experiments in 4 other HPV+ HNSCC patients. Top: % Total CD8+DP^{Ex}, Bottom: % CD8+Tetramer+DP^{Ex} for all except Pt. 7007 (% CD8+CD137+DP^{Ex}). (B) Unsupervised hierarchical clustering of %DP^{Ex} from E2, E6 and E7 total CTLs (left) and HPV-specific CTLs (HPV-Tetramer+, CD137+, right) from patients analyzed in 3A-C. Each row is an epitope specific DP^{Ex} response per patient. *P < 0.1, **P < 0.01, ***P < 0.001, Unpaired two-tailed Welch's T-test.

In 3/5 HPV+ HNSCC patients, total CD8+ DP^{Ex} and HPV-specific CD8+DP^{Ex} cells were higher in CTLs stimulated with E7-antigen (between 2-10 fold) relative to E2/E6 antigen-stimulated CTLs (Pts. 7002, 7007, 7012, figs. 4-6 B, 4-7 A, Unpaired Welch's T-test, $P < 0.1$), independent of HLA-status (e.g. HLA-A*68:01+ Pt. 7007) (Fig. 4-7 A). In the other two patients,

E2-CD8+DP^{Ex} was higher than E7-CD8+DP^{Ex} (2-3 fold, Pts. 7035, 7050, Fig. 4-7 A, Unpaired Welch's T-test, $P < 0.1$), indicating heterogeneity in HPV-specific CTL-dysfunction in HPV+ HNSCC. Interestingly, compared to E2 and E7-CTLs, E6-CD8+DP^{Ex} remained relatively low in most patients, and there was an inverse relationship within patients between E7-CD8+DP^{Ex} and E2/E6-CD8+DP^{Ex} (Fig. 4-7 A). Unsupervised hierarchical clustering of DP^{Ex}-frequencies of total CD8+ as well as HPV-specific CD8+ T-cells revealed this trend where high E2-CD8+DP^{Ex} and E6-CD8+DP^{Ex} co-occurred in patients who had relatively lower E7-CD8+DP^{Ex} and vice versa (Tukey's multiple comparisons test, E2 vs. E7, P = 0.014, E6 vs. E7, P = 0.084, E2 vs. E6, P = NS) (Fig. 4-7 B). These results indicate that in most HPV+ HNSCC patients, E7-CTL dysfunction is distinct and might be either temporally or mechanistically unrelated to E2/E6 CTL-dysfunction.

4.3.4. HPV16-antigen load correlates with T-cell exhaustion

To provide a broader analysis of immune dysfunction from the tumor side in HPV+ HNSCCs, we performed an immune signature analysis of publicly available HNSCC transcriptomes (TCGA, UM-cohorts N =119, 51 HPV+, 68 HPV-)(The Cancer Genome Atlas Network, 2015; Zhang et al., 2016). We used previously-validated immune signatures representing tumor infiltrating immune cell subsets and performed single-sample gene set enrichment analysis (ssGSEA, Methods, Table B2 lists the gene signatures) to score the HPV+ and HPV- subsets (De Simone et al., 2016; Mandal et al., 2016a; Şenbabaoğlu et al., 2016a). HPV+ HNSCC patients in general had higher immune infiltration scores compared to HPV- HNSCCs (Fig. 4-8), with 36/51 (70%) of HPV+ HNSCC samples represented in the T-cell-high gene cluster, and few HPV+ HNSCC samples (17%) with very low immune cell infiltration scores, confirming and expanding the findings in previous studies (Mandal et al., 2016a). To assess the impact of HPV gene expression on immune cell infiltration, we calculated the Spearman correlation coefficients among ssGSEA scores for the entire gene sets across all patients, including HPV16 genes and performed an unsupervised clustering on the correlation matrix (Fig. 4-9, Table B3).

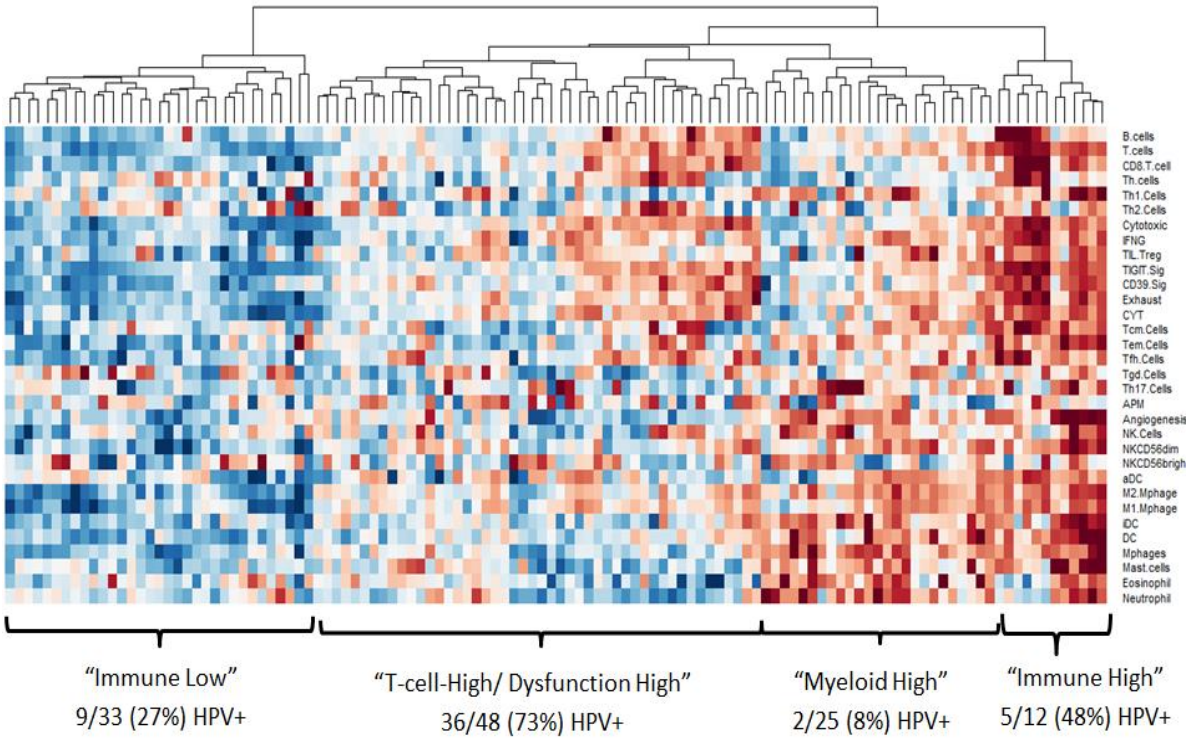

Figure 4-8. T-cell infiltration and dysfunction signatures are enriched in HPV+HNSCCs. Unsupervised hierarchical clustering of validated normalized immune signatures (Table S3), in 119 HNSCC transcriptomes (TCGA+UM cohorts) by ssGSEA (SI Methods). Each column represents one HNSCC patient tumor ssGSEA scores. Four clusters are revealed and their names are indicated in the bottom. Number and percentages of HPV+HNSCCs in each cluster are indicated.

Interestingly, HPV-gene signatures also negatively correlated with neutrophils and other myeloid gene signatures, indicating that lymphocytes dominate the immune landscape of HPV+ HNSCCs (Fig. 4-9). We also observed a previously described exhaustion gene set correlating with HPV-gene sets in this module (Spearman ρ = 0.33, Fig. 4-9). Unsupervised hierarchical clustering on expression levels of constituent genes within the exhaustion gene set revealed three main groups with low (L), moderate (M), and high (H) expression of immune regulatory gene expression within HPV16+HNSCC patients (Fig. 4-10 A). We then analyzed individual HPV16-gene expression in the HPV16+HNSCC tumors (N=40) stratified into Exhaustion-high (EX-H) and Exhaustion moderate/ low tumors (Fig. 4-10 B). EX-High HPV16+HNSCC tumors had higher gene expression of E1, E2, E4, E6 and L2 genes compared to the EX-ML subset (Fig. 4-10 B, Unpaired Welch's T-test; E1, E2, $P < 0.01$; E4, E6, L2, $P < 0.1$). E7-expression was

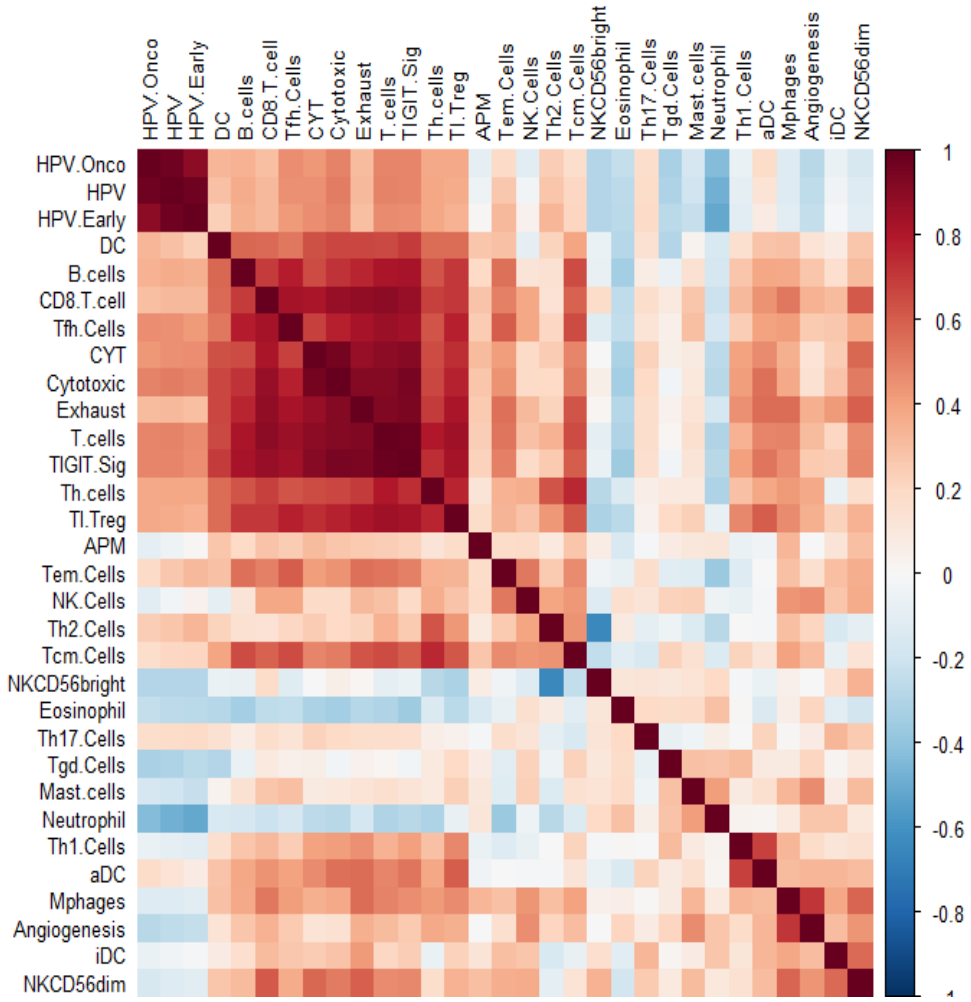

Figure 4-9. T-cell exhaustion signatures correlate with HPV16-antigen expression. 119 HNSCC transcriptomes (68 HPV- HNSCCs and 51 HPV+ HNSCCs), were analyzed for immune cell infiltration by ssGSEA (Appendix B, Methods). Clustered correlation matrix of immune signatures with HPV-gene sets (HPV - All 8 HPV genes, HPV.Early - E1, E2, E4, E5, HPV.Onco - E6, E7). All gene sets are listed in Appendix B Table B-2. Gene set correlations were clustered by hierarchical clustering creating distinct modules (Table B3).

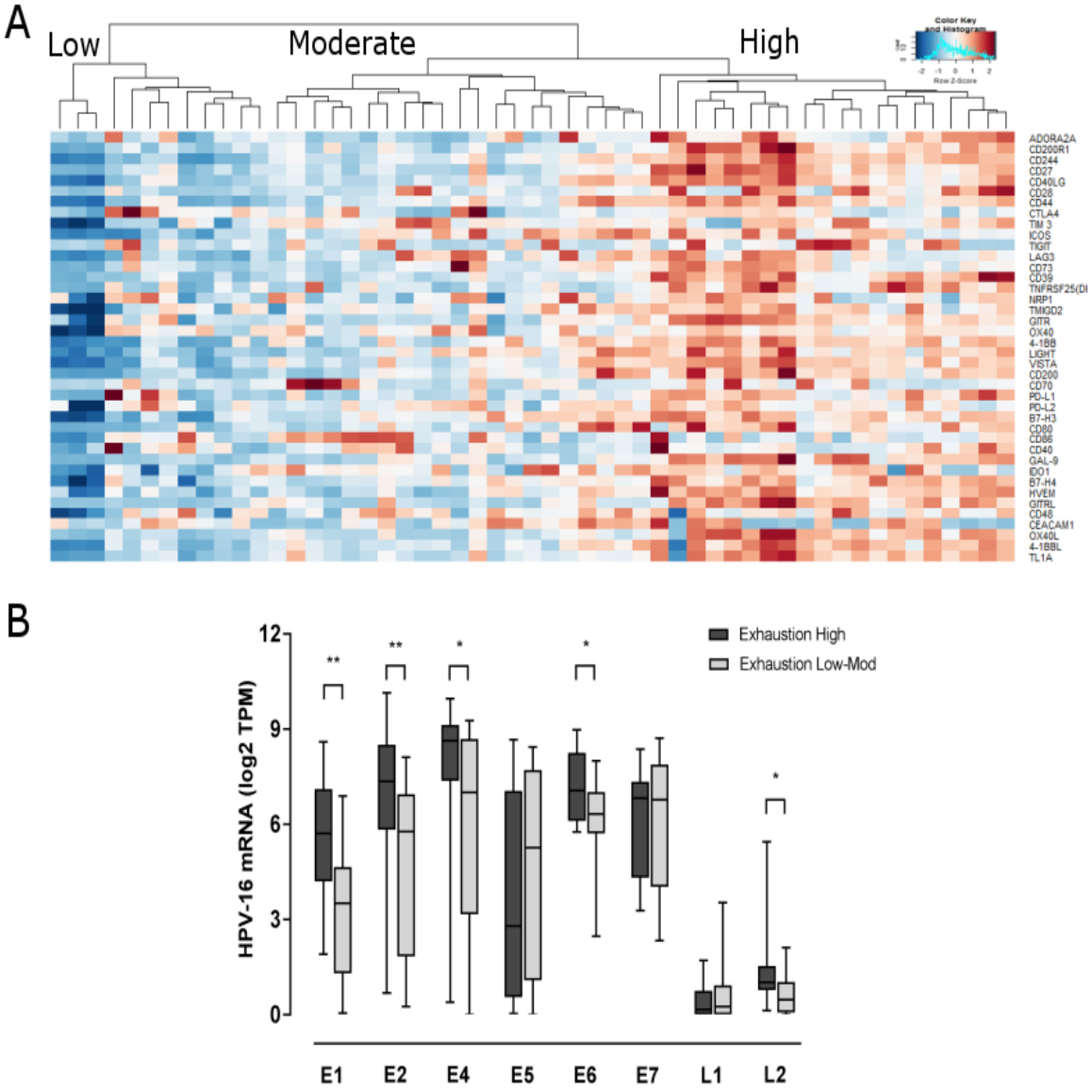

Figure 4-10. HPV16-antigen load likely drives T-cell infiltration and dysfunction. (A) Unsupervised hierarchical clustering of 49 genes in the exhaustion gene set (De Simone et al., 2016) in HPV+ HNSCC. Patient clusters: low exhaustion (L), moderate (M) and high exhaustion (H). (B) HPV gene levels (Log_2-TPM+1) in 40 HPV16+HNSCCs from 4-8 A classified into exhaustion high (EX-H in main text, N=15) and exhaustion low+moderate subsets (EX-ML in main text, N=25). *P < 0.1, **P < 0.01, Unpaired two-tailed Welch's T-test.

comparably high in both the subsets, while E5 known to downregulate MHC-class I expression (Campo et al., 2010), remained low in the EX-H subset (Fig. 4-10 B). These computational analyses along with previous experiments (Fig. 4-7) suggest that HPV-specific CTLs have T-cell exhaustion at tumor sites, driven by intra-tumoral HPV-antigen expression.

4.3.5. IDO-1 represents a novel HPV+ HNSCC specific immune target

We analyzed differential expression of constituent genes within the exhaustion gene set between HPV+ HNSCCs and HPV- HNSCCs (Fig. 4-11). We observed several well-known T-cell regulatory genes such as LAG-3, GAL-9, CEACAM-1 and CTLA-4 overexpressed in HPV+ HNSCCs compared to HPV- HNSCCs, consistent with high T-cell infiltration and dysfunction as observed in our results (Figs. 4-8, 4-9), and other studies (Mandal et al., 2016a). Interestingly, we also observed exhaustion genes NRP1, CD39 and CD73 that were selectively upregulated in HPV- HNSCCs compared to HPV+ HNSCCs (Fig. 4-11), indicating distinct types of T-cell dysfunction between the two HNSCC subtypes. Indoleamine 2,3-dioxygenase (IDO-1), an L-Tryptophan catabolizing enzyme was one of the highest differentially expressed gene (based on To validate the immune signature analysis indicating that HPV-antigen expression can impact immune regulatory gene expression such as IDO-1 (Fig. 4-12 A), we performed immunoblotting for IDO-1 expression in a panel of HPV16+ cancer cell lines (2 cervical, 4 HPV+ HNSCC). These showed variability in HPV16-E7 protein expression (Fig. 4-12 B). IDO-1 expression followed a striking correlation with E7-protein expression in the same cell lines (R^2 = 0.84, P = 0.033, Fig. 5B). Transfection of the 3-HPV16-antigens into a non-HPV cell line (HEK-293-T) did not alter IDO-1 protein expression (Appendix B-3A). A cross-cancer (N = 30 types, 45708 total tumors) gene expression analysis from cBioPortal (Cerami et al., 2012), revealed that IDO-1 is also highly expressed in the HPV-malignancy cervical cancer (Fig. 4-12 C). These results suggest that tumor IDO-1 expression is linked to immune selection pressure from TILs rather than a direct molecular/biochemical consequence of the HPV-life cycle. Of note, PD-L1 protein expression on the same cell lines did not correlate with E7-antigen expression (Spearman ρ = 0.17, P = NS, Appendix B-3B).

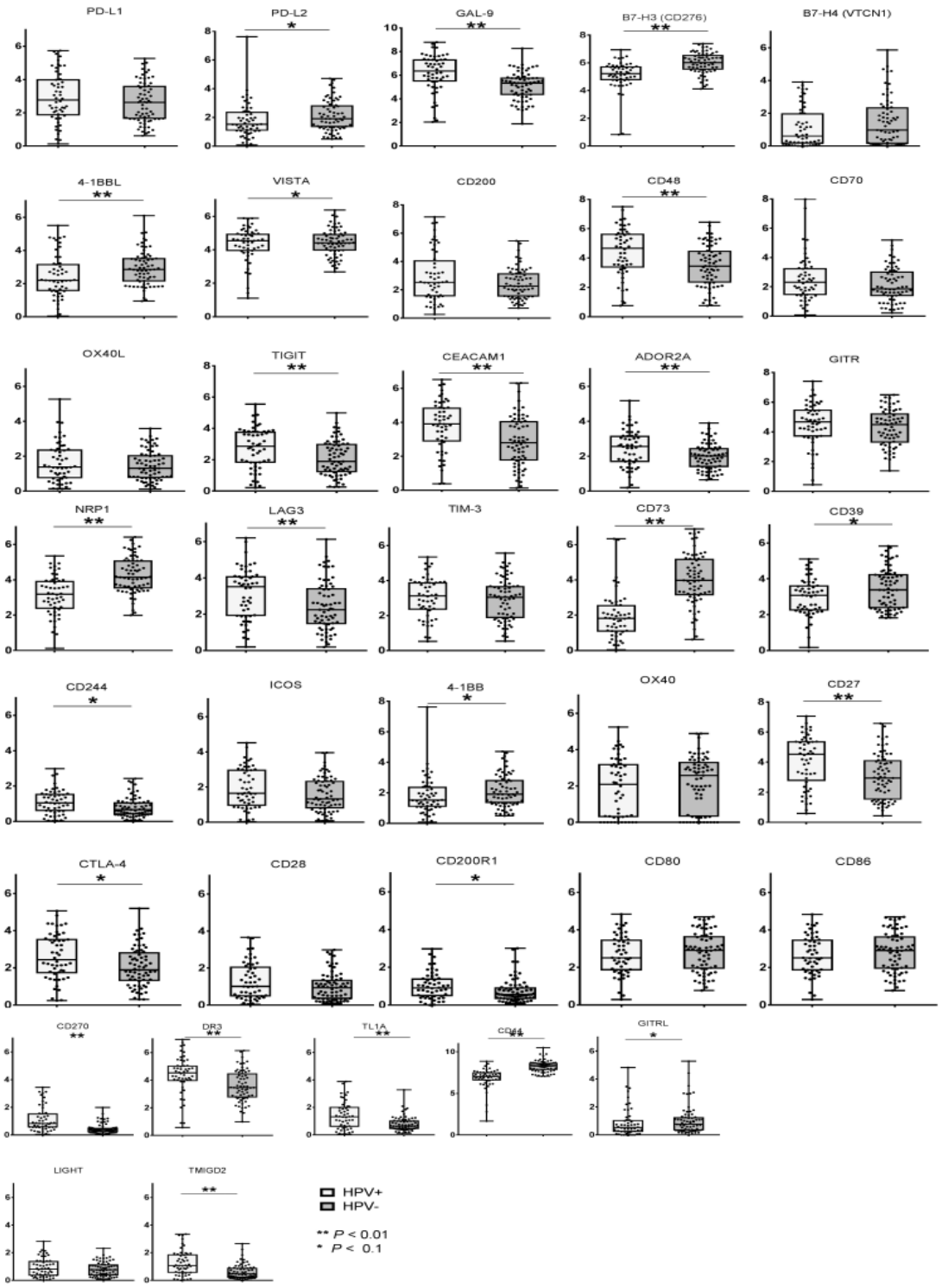

Figure 4-11. Differential expression of exhaustion genes in HPV+ vs. HPV- HNSCCs. Box plots representing Log2 TPM from HPV+, HPV- HNSCC patients showing median from exhaustion gene set (Table B2), whiskers indicate 10-90th percentile. *P*<0.1, ***P*<0.01 ranked *P*-value) from exhaustion set in HPV+ HNSCCs compared to HPV- HNSCCs (Fig. 4-12A; Mann-Whitney test, *P* = 0.0012).

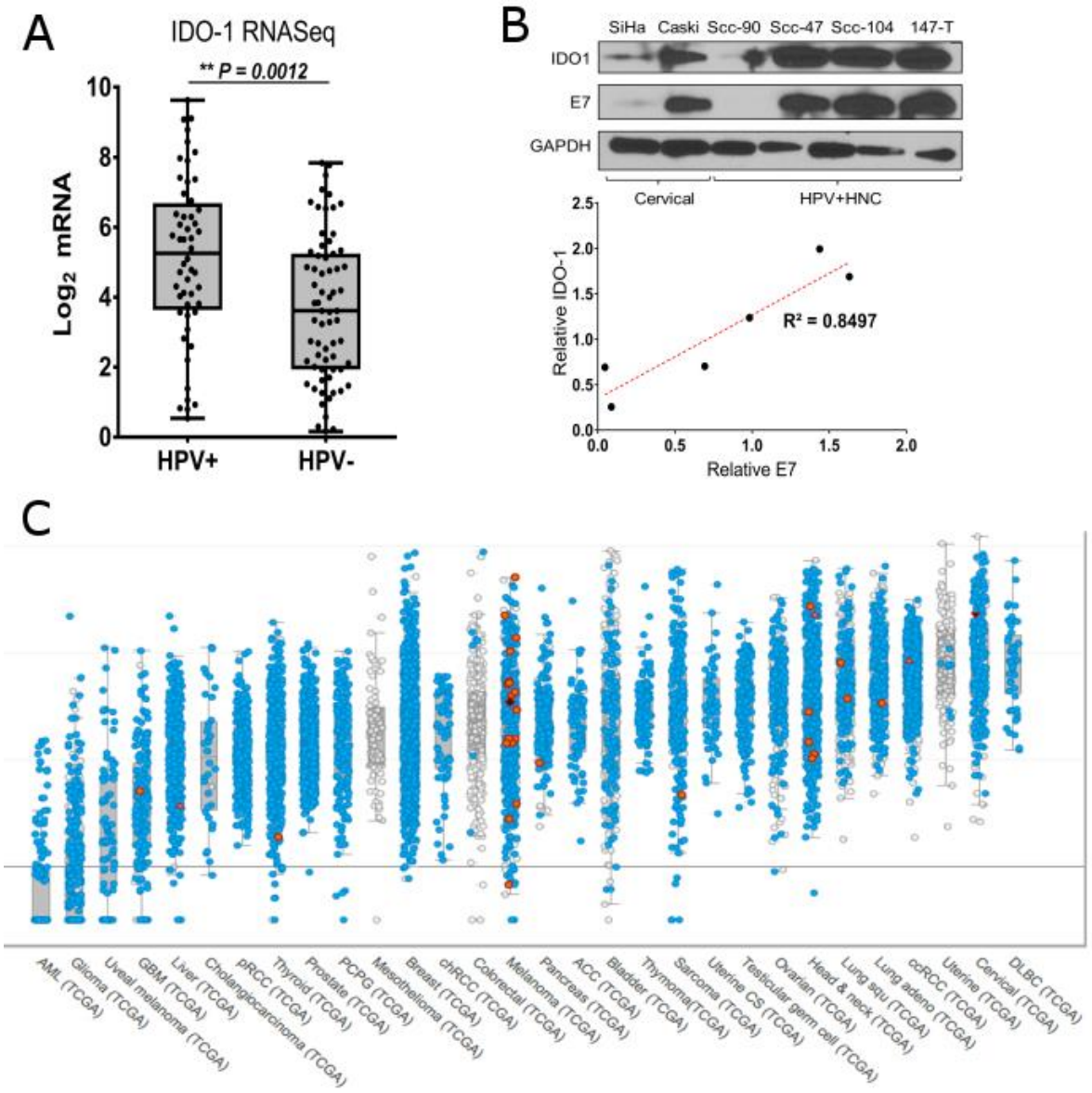

Figure 4-12. IDO-1 expression is an HPV-specific immune regulatory gene. (A) Box plots representing Log2 TPM from HPV+, HPV- HNSCC patients showing median from exhaustion gene set (Table S3), whiskers indicate 10-90th percentile. Each data point represents one patient. $P < 0.1$, $^{**}P < 0.01$ by Mann-Whitney test. (B) Correlation of IDO-1 protein levels to HPV16-E7 antigen expression in 6 HPV16 cell lines (2 cervical cancer, 4 HPV16+ HNSCC). Top pane: Immunoblot, Bottom panel: correlation of IDO-1 and E7 protein levels normalized to GAPDH ($R^2 = 0.84$, $P = 0.033$).(C) Log2 mRNA TPM levels were obtained from cBioportal, for 30 different types of cancers from 45708 total tumors from TCGA. Each blue data point represents expression from one patient tumor for respective cancer, red representing a mutation or other alteration and grey represents unsequenced tumors.

4.3.6. IDO-1 inhibition enhances T-cell targeting of HPV+ HNSCCs

IDO-1 inhibitors are being evaluated in pre-clinical and clinical settings to enhance tumor immunity (Gangadhar et al., 2015; Sheridan, 2015). We therefore explored the possibility of exploiting IDO-1 inhibition to overcome HPV-CTL dysfunction, especially using E7-CTLs. We fluorescently labelled the HLA-A*02:01+ HPV+ HNSCC cell line UM-SCC-104 (SCC-104), which has high IDO-1, E7 and PD-L1 expression (Fig. 4-12 B, Appendix B-3B). SCC-104 cells were previously reported to have a distinct hierarchy of HPV16-antigen expression where E7 > E6 > E2 (Olthof et al., 2015). We assessed HPV-CTL mediated cytotoxicity on SCC-104 cells, after co-incubation with ex vivo expanded HPV-CTLs from an HLA-A*02:01+ HPV+ HNSCC patient in the presence of either anti-PD-1 antibody (αPD-1+DMSO) or IDO-1 inhibitor Epacadostat (Ig+IDO-1i), or both (αPD-1+IDO-1i). Within E7-CTLs (Fig. 4-13 A, 4-13 B right side), single-agent treatment with either αPD-1 or IDO-1i individually resulted in a 3-5 fold increase in sensitivity of SCC-104 to E7-CTL mediated cytotoxicity compared to mock (Ig+DMSO) treatment (Fig. 4-13 A-B, αPD-1 vs. mock, $P = 0.024$, IDO-1i vs. mock, $P = 0.064$). In contrast, combination blockade with both αPD-1+IDO-1i resulted in a 10-fold increase in tumor cytotoxicity compared to mock treatment (4-13 A-B, $P = 0.011$), and a 2-3 fold increase in tumor cytotoxicity compared to the single-agent treatments 4-13 A-B, $P = 0.04$ compared to IDO-1i, $P = 0.013$ compared to αPD-1). Similar results were obtained with αPD-1+IDO-1i combination therapy on E2 and E6 CTL cytotoxicity although to a lesser extent (3-fold increase for E2, and 5-fold increase for E6 compared to mock treatment, $P < 0.01$, Fig. 4-13 B) likely reflecting the lower expression of these antigens in SCC-104 cell line (Olthof et al., 2015). These results demonstrate that IDO-1 is a novel HPV+ HNSCC specific checkpoint correlating with HPV-antigen expression, and combination inhibition of PD-1 and IDO-1 can sensitize HPV+ HNSCCs to HPV-CTL mediated cytotoxicity.

4.4 Discussion

HPV-driven malignancies remain an ideal model system for cancer immunotherapy, due to 1) a long lead time from infection to malignancy, 2) emerging immune and viral biomarkers for early

detection, 3) the persistent tissue expression of viral oncogenes and 4) epidemiologic evidence of the central role of T-cell control of viral persistence. However, the dynamics of viral persistence within immunocompetent individuals and the mechanisms of tumor immune escape remain

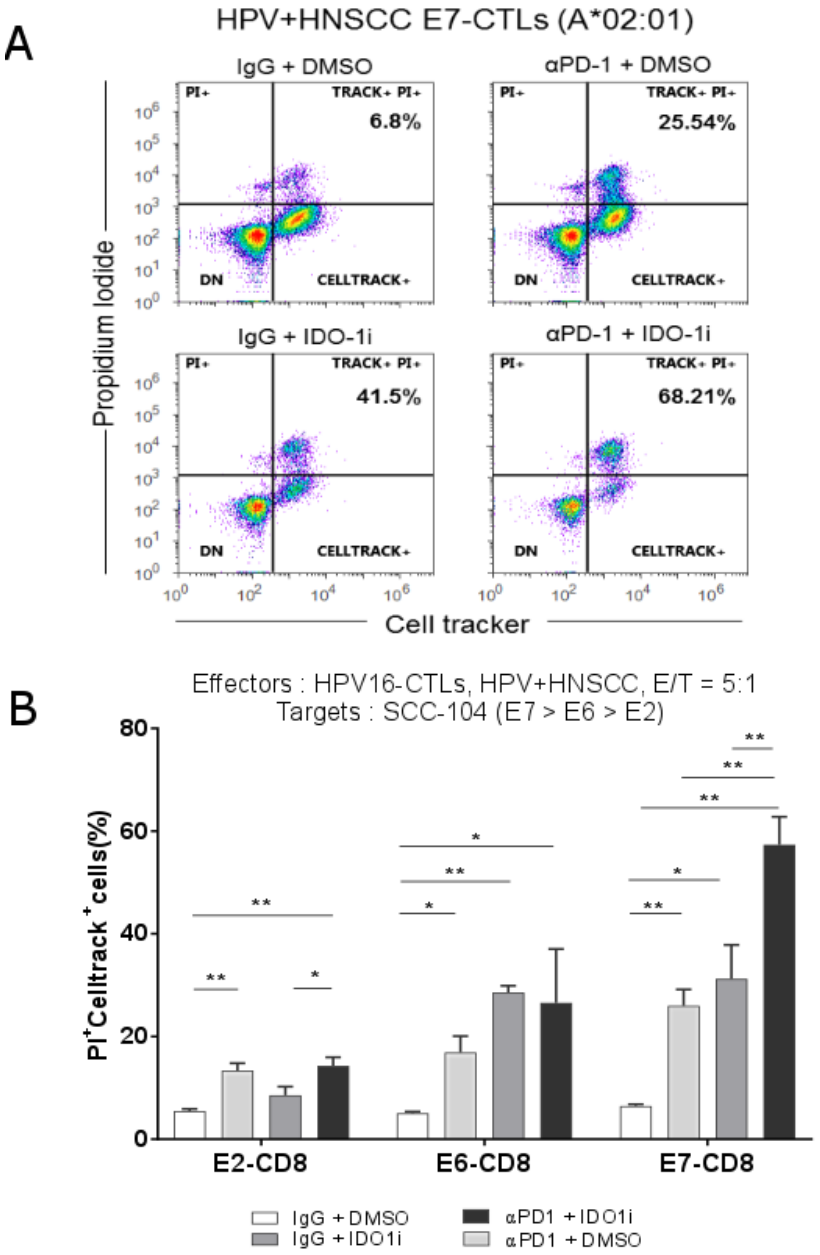

Figure 4-13. IDO-1 inhibition enhances T-cell targeting of HPV+ HNSCCs. Celltracker labeled SCC-104 cells were co-incubated with polyclonal E2, E6 and E7-specific T-cells at an effector/target ratio of 5:1 for 48 hours, and assessed for cell death (TRACK+ PI+ events) under αPD-1, IDO-1 single or dual-inhibition conditions. (A) Representative flow cytometry plot from E7-specific T-cell mediated cytotoxicity on SCC-104 cells. % indicates celltracker labeled dead SCC-104 cells. (B) Summary of SCC-104 cytotoxicity (3 biological replicates) by E2, E6 and E7-CTLs under treatments indicated. *$P < 0.1$, **$P < 0.01$, Unpaired two-tailed Welch's T-test.

largely unknown, in particular for HPV+ HNSCC. The emerging epidemic of HPV+ HNSCC and lack of screening modalities represents a major clinical challenge and opportunity for targeted T-cell immunotherapy.

In this study, we have expanded the spectrum of HPV+ HNSCC-specific immune therapeutic targets at the CTL-epitope level and at the target tumor cell-modulatory level. We chose E2, E6 and E7-antigens, as they induce strong B-cell immunity, have been detected in pre-invasive and/or invasive cervical cancer, and we confirmed viral antigen expression in HPV16+ HNSCC transcriptomes. Most studies that have attempted to define CTL-immunogenicity from HPV16 have primarily focused on a limited number of HLA-alleles (e.g. A*02:01) and peptides from E6 and E7 (Table B1), with limited data on immunogenic targets in in HPV+ HNSCC (Ressing et al., 1995; Riemer et al., 2010; Rudolf, Man, Melief, Sette, & Kast, 2001). The 15 HLA alleles chosen for this study are predicted to include 10/12 of HLA supertypes and over 95% of the global population (Lund et al., 2004; Riemer et al., 2010). Of the 15 HLA alleles, we failed to identify peptides for HLA-B*40:01 and HLA-B*4402. HLA B*40:01 is significantly overrepresented in the HPV+ HNSCC cohort compared with HPV- HNSCC (Fig. 4-1 B), but these data remain to be confirmed in larger datasets and association studies. Viral immune escape by altering HLA-binding CTL epitopes has been documented in HIV-1 and HCV infections (Petrovic, Dempsey, Doherty, Kelleher, & Long, 2012; Price et al., 1997), but not as well for DNA viruses such as HPV, where the mutation rates are markedly lower.

We identified several immunogenic CTL-epitopes from the 3 HPV16-antigens (Fig.4-3, Table B1). In our experiments, addition of PD-1 and CTLA-4 CKB antibodies aided our ability amplify and detect low-frequency HPV-CTL-response in both healthy and HPV+ PBMCs ex vivo (Figs.4-1-4.4, B1). Our results indicate that HPV16-E2 and E6 induce more CTL responses than HPV16-E7 (Figs. 4-1, 4-2). HPV16-E6 and E7 have been the dominant targets for T-cell based immune therapies against HPV thus far Draper et al., 2015; Kenter et al., 2009). In contrast E2-specific CTL-reactivity has been unexplored as an immunotherapeutic target in HPV+ HNSCCs due to the assumption that E2-locus is interrupted by viral integration, similar to that observed in cervical cancer (Woodman et al., 2007). However, several recent whole genome studies in

HNSCCs have indicated that viral breakpoints in HPV+ HNSCCs are distributed throughout the genome, with preferential integration in the E1 region (Akagi et al., 2014; Parfenov et al., 2014b). E2 is also a larger antigen >3-times the size of E6, E7, possibly explaining the bigger spectrum of CTL-epitopes from the protein. These results, taken along with the high proportion of episomal full length HPV DNA in HNSCC lesions (Olthof et al., 2014), and our data demonstrating strong E2-specific T-cell and B-cell reactivity (Figs. 4-1, 4-2) warrant further investigation of E2 as a T-cell therapeutic target in addition to E6 and E7 in HPV+ HNSCCs.

Despite the addition of CKB antibodies in the ex vivo T-cell stimulation protocol, we detected low levels of E7-CTLs compared to E2 and E6-CTLs in this study. This can be due to 1) inaccurate prediction of CTL-epitopes, 2) inherently low immunogenicity of E7-antigen, 3) low antigen load in patients, or 4) higher levels of dysfunctional E7-specific CTLs. Our ability to accurately predict previously described epitopes from E7 and the successful identification of novel CTL-epitopes from E2 and E6 across various HLA-alleles (Fig. 4-3), argues against a sub-optimal prediction strategy. The presence of high levels of serum titers against E7 in HPV+ HNSCC patients indicates that the antigen is presented and is immunogenic at least in context of B-cell immunity (Fig. 4-2C-D). Gene expression analysis of HPV+ HNSCC tissue and cell lines showed that E7-antigen load is high in patient tumors, consistent with several other studies (Olthof et al., 2014; Zhang et al., 2016). However, E7-CTLs tended to exhibit higher levels of PD1+CD39+ or PD1+TIM3+ DP^{Ex}-phenotypes compared to E2 and E6-CTLs after ex vivo stimulation in 3/5 independent HPV+ HNSCC patients (Figs. 4-6, 4-7). In particular, E7-CTL dysfunction rarely occurred in concert with E2/E6 CTL dysfunction within the same patient (Fig. 4-7). Since antigen persistence and subsequent magnitude of CTL-response are major factors in chronic viral T-cell exhaustion (Mueller & Ahmed, 2009), we speculate that in each of these patients, the variable HPV-dysfunctional CTLs might reflect temporal tumor HPV-load or tumor heterogeneity, although this remains to be elucidated. Future studies that compare and correlate tumor antigen load in vivo with the dynamics of E2, E6 and E7-CTL phenotypes in peripheral blood and tumor will be needed to verify in vitro CTL-dysfunction findings from our study.

While other studies have shown high levels of immune infiltration in HPV+ HNSCCs (Mandal et al., 2016a; Zhang et al., 2016), we show here that HPV-antigen load likely drives high CTL-infiltration and CTL-dysfunction (Fig. 4-9), arguing for better response to immune CKB therapies. Indeed, preliminary data from ongoing HNSCC clinical trials targeting the PD-1/PD-L1 axis indicate more benefit for HPV+ HNSCCs compared to HPV- HNSCCs (Chow et al., 2016; Ferris et al., 2016). Our data thus provides mechanistic insights into this clinical response, wherein, high HPV-antigen load likely drives T-cell infiltration into HPV+ HNSCCs causing immune selection pressure for HPV-CTL dysfunction (in particular E7-CTLs), and that immune checkpoint blockade can at least partially reverse this effect in HPV+ HNSCCs (~32% ORR in (Chow et al., 2016)).

In addition to the PD-1 checkpoint, we demonstrate both by computational and experimental in vitro assays that IDO-1 is highly expressed in HPV+ HNSCC and other HPV-driven malignancies (Figs. 4-10, 4-12). In preclinical murine melanoma models, TILs increase tumor IDO-1 expression (Spranger et al., 2013) and IDO-1 inhibitors are showing promise in clinical trials in particular with PD-1/PD-L1 blockade (Sheridan, 2015). HPV+ HNSCCs expressing IDO-1 might similarly be driven by HPV-specific-CTL infiltration in response to high tumoral HPV antigen load (Fig. 4-12). In vitro, this resistance to CTL-targeting by HPV+ HNSCCs is apparent in the absence of PD-1/IDO-1 inhibition, where only ~5% of SCC-104 cells were sensitive to CTL-cytotoxicity regardless of the antigen-specificity of the HPV-CTLs (Fig. 4-13). In contrast, inhibition of IDO-1 alone or in combination with PD-1 blockade significantly enhances tumor cell cytotoxicity of E7-CTL (and, to a lesser extent, E2 and E6-CTL) derived from patients with HPV+ HNSCCs (Fig. 4-13). These data suggest that IDO-1/PD-1 blockade may have a significant effect to activate pre-existing HPV-specific CTL in the majority of HPV+ HNSCCs.

Sixty years after the discovery of HPV and 10 years after FDA approval of the first HPV vaccine, HPV-associated malignancies remain a major public threat, with an estimated 14 million new HPV-infections occurring every year (Dunne et al., 2014). The presence of highly-expressed viral antigens makes HPV+ HNSCCs a promising setting for targeted immunotherapies. We propose that vaccination or adoptive T-cell therapy to HPV16-specific CTL epitopes from E2, E6

and E7, and targeted immune modulation with a combination of PD-1 and IDO-1 inhibition warrant further evaluation in HPV+ HNSCC and other HPV-associated malignancies.

4.5 Materials and Methods

For full details of methods and construction of data sets see Appendix B.

4.5.1. HPV16 candidate CTL-epitope prediction

HPV16-candidate CTL epitopes were predicted using previously described prediction strategies developed by us (Chowell et al., 2015; Krishna & Anderson, 2016), except for the incorporation of immunogenicity scores. For the 15 HLA-class I alleles, 9-mer and 10-mer candidate epitopes derived from the HPV16 proteins E2, E6, and E7 were predicted from 5 independent prediction algorithms and normalized. Top 4-5 candidate peptides / HLA-allele were used for experiments. Further details are provided in Appendix B. For full list of peptides see Appendix Table B1.

4.5.2. Epitope mapping from HPV+ HNSCC PBMCs

PBMCs were obtained from stage III or stage IV HPV+ HNSCC patients (MSSM cohort). Patient characteristics are described in(Parikh et al., 2014b). All HPV16-peptides (> 80% purity) were synthesized by Proimmune, UK. HPV+ HNSCC PBMCs were thawed, rested with 1µg/mL of CKB antibodies anti-PD1 (eBosciences, USA) , anti-CTLA4 (eBosciences, USA) for 1 hour at 37 C. HPV16-peptides in pool or individually were added subsequently in biological triplicates, along with recombinant human IL-2 (20U/mL), human IL-7 (5ng/mL). On day 5, half the media was removed and replaced with fresh IL-2 and peptide pool. On day 8, half the media was removed and fresh media, IL-2 and peptide was added to the cells and replated into a 96-well multiscreen elispot plate for Elispot detection. Elispot detection is described in Appendix B.

4.5.3. HPV-CTL stimulation for phenotyping

HPV-specific T-cells were generated by stimulating autologous HPV+ HNSCC patient B-cell APCs (Appendix B). APCs were either peptide pulsed with HPV16-epitopes, or transfected with

whole HPV-antigen encoded in mammalian expression plasmid pCDNA3.2 (Invitrogen, CA, USA). APCs were washed and incubated with thawed whole HPV+ HNSCC PBMCs at a ratio of 1:2 (200,000 APCs : 400,000 PBMCs) supplemented with 20U/mL recombinant human IL-2 (R&D Systems, MN, USA), 5ng/mL IL-7 (R&D Systems, MN, USA). On day 5, partial media exchange was performed. On day 10, expanded HPV-CTLs were restimulated with peptide-pulsed, transfected APCs similar to day 1. HPV-CTLs were used for cytolytic assays or immunophenotyped after day 20.

4.5.4. Tetramer staining, HPV-CTL and HPV+cell line phenotyping

HPV16-tetramers were obtained from NIH Tetramer Core Facility at Emory University. For tetramer staining, cells were re-suspended in 100µL staining buffer with 5% human serum and 1mM Dasatanib (ThermoFisher Scientific, MA, USA), and each tetramer was added at concentration of 1:100 for 30 minutes at room temperature. Cells were washed twice and restained with anti-CD8-PC5, anti-CD4-FITC, anti-CD14-FITC and anti-CD19-FITC for exclusion gates, and either a combination of anti-PD1-BV605 and anti-CD39-BV-405 or anti-PD1-BV-605 and anti-TIM3-BV-405 for 30 minutes on ice. HPV+ cell line PD-L1 staining was done for 30 minutes on ice. Samples were then washed twice in 1x PBS, and analyzed by Attune flow cytometer (ThermoFisher Scientific, MA, USA). Tetramer and antibody details are provided in Appendix B.

4.5.5. HPV-CTL cytotoxicity assays

HLA-A*02:01+ HPV+ HNSCC+ SCC-104 cells were pre-labelled with 0.5µM CellTracker Green CMFDA (ThermoFisher Scientific, MA, USA) for 1 hour and washed. HPV-specific CTLs were pooled by HPV-antigen, washed and resuspended in media supplemented with 20U/mL IL-2, with 1µg/mL isotype IgG or anti-PD1 antibody, DMSO, and 1µM IDO-1 inhibitor Epacadostat (Selleck Chemicals, MA, USA) in various combinations as described. HPV-CTLs were added at ratio of 5:1 to SCC-104 cells and incubated for 48 hours at 37 C, 5% CO_2. Cocultured cells were harvested, neutralized with media supernatant from each well containing dead cells and

centrifuged for 850g, 10 minutes. Cell pellets were washed twice with sterile 1X PBS, resuspended with 1mL 1X PBS, and 2uL Propidium Iodide (ThermoFisher Scientific, MA, USA) and cell death was assessed by flow cytometry.

4.5.6. RNASeq data alignment

RNA-Seq reads for each sample were quality checked using FastQC (version 0.10.1, Babraham bioinformatics, Babraham Institute, Cambridge, UK) and aligned to the human genome build 38 (GRCh38) primary assembly and HPV16 genome (GCF_000863945.1) simultaneously using STAR (version 2.5.2B). After alignment, variants were discovered following GATK Best Practices workflow for RNAseq. Raw RNAseq reads were pre-processed by adding read groups, indexing, marking duplicates and sorting, Split'N'Trim, reassigning mapping quality and base recalibration.

4.5.7. ssGSEA analysis of HPV and immune gene signatures

Log transformed transcripts per million (Log_2 TPM+1) from each HNSCC sample, after subtraction of low expression genes was used for ssGSEA as previously described in Şenbabaoğlu et al (Şenbabaoğlu et al., 2016b). Pre-defined immune signatures (Appendix B Table B2), have been extensively validated in Şenbabaoğlu et al (Şenbabaoğlu et al., 2016b) and Mandal et al(Mandal et al., 2016b). ssGSEA scores were computed for each tumor sample using the R package GSVA, and Z-transformed across the cohort prior to analysis. To assess impact of HPV-gene expression on immune signatures, a correlation matrix was built using the R-library Corrplot with the Z-transformed ssGSEA scores and were displayed by hierarchical clustering of correlations (Appendix B Table B3). Individual gene expression analysis was performed by unsupervised hierarchical clustering methods and were used for heatmap analysis.

4.5.8 Statistical Analysis

Categorical variables, such as Elispot data, and Flow cytometric data were summarized as SFUs, and percentages. Continuous variables (RNAseq data) were presented with mean with standard error of mean (SEM). Unpaired T-test with Welch's correction was used for all categorical variable

analyses, and for continuous variable analyses non-parametric Mann-Whitney's test was used. For heatmaps of T-cell frequencies and ssGSEA RNASeq analyses, Z-transformation was performed to normalize the data across the cohorts. R statistical software V3.4.0 and Prism software (GraphPad Software) were used for data managements and statistical analyses. Significance levels were set at 0.1 (*), and P-values of 0.01 (**) or 0.001 (***) for all tests are indicated.

CHAPTER 5

T-CELL ANTIGEN DISCOVERY BY SINGLE CELL CYTOKINE CAPTURE

Abstract

Peripheral CD8+ T-lymphocytes (CTL) are critical components of the human immune system performing constant immune surveillance and elimination of infected, malignant cells thereby protecting the host against a multitude of pathogenic infections. Despite major advances in technologies for T-cell epitope discovery, CTL antigen and epitope identification from large complex genomes remain a major challenge. Here, we develop a novel single cell assay using autologous CTLs and antigen presenting cells (APCs) to enable identification of immunogenic antigens from a cDNA library encoding whole genomes without the need for HLA-typing. The assay relies on the capture of interferon gamma secreted by effector CTLs on the surface of autologous APCs encoding an immunogenic antigen. We empirically show the this assay using peptide pulsed as well as transfected APCs as well as with a limit of detection of up to 10% of antigenic epitope pulsed APCs. We validated the IFNɣ APC capture assay by magnetic sorting of the IFNɣ + APCs followed by PCR amplification of an antigen cassette. We anticipate that this technology will enable the accurate identification of immunogenic antigens and epitopes relevant for T-cell vaccine and immunotherapy design.

5.1 Introduction

CD8+ T-cell (CTL) recognition of an antigenically processed intracellular epitope is critical for the adaptive immune based clearance of many pathogenic infections. CTLs use their T-cell receptor (TCR) to interact with linear immunogenic epitopes in complex with the human leukocyte antigen (HLA) class I molecules of infected cell surface (Grakoui et al., 1999). In the lymph node, professional antigen presenting cells (APCs) such as dendritic cells (Mildner & Jung, 2014), or activated B-cells (Coughlin, Vance, Grupp, & Vonderheide, 2004), can activate CTLs by integrating TCR-HLA-peptide complex (Signal 1) and costimulatory molecules CD80/86 (Signal 2)

to prime naive CTLs as well as boost memory CTLs (Pollizzi & Powell, 2014). This complex cellular interaction ultimately leads activated CTLs to seek infected, malignant cells expressing the same epitope, CTL-release of many proinflammatory cytokines such as interleukin-2 (IL-2), interferon gamma (IFNY), and finally release of effector molecules causing target cell lysis (Hennecke & Wiley, 2001). Thus, the identification of CTL-epitopes and antigens is a major effort in translational immunology which can aid in the design of effective T-cell vaccines, immunotherapies, and immune monitoring of many malignancies (Ott et al., 2017; Schumacher & Schreiber, 2015; Tran et al., 2016).

The immense diversity of T-cell repertoire, the universe of pathogenic epitopes, and the polymorphism, codominance of HLA-alleles makes the identification of antigenic T-cell repertoire of any pathogen daunting. Several advances in genomics, proteomics and computational techniques have resulted in the development of novel approaches for identifying the HLA class I-restricted antigenic repertoire of CTLs in the past decade. For instance, biochemical methods such as mass spectrometry (MS) can perform unbiased identification of the HLA-ligandome of any cell type (Kowalewski et al., 2015; Shao et al., 2017). However MS-identification while being sensitive, does not identify immunogenic epitopes that result in T-cell activation, and generally requires large amounts of sample. Fluorescent HLA-multimers in combination with multiparameter flow cytometry can be used identify immune reactive CTLs, and have recently been combined with genomics to identify immunogenic epitopes from complex libraries (Bentzen et al., 2016). However, such methodologies still require knowledge of the donor's HLA-type, the predicted CTL-epitope, and its HLA-restriction. More recently, machine learning computational algorithms have been developed to predict HLA-binding candidate peptides from an antigen (Moutaftsi et al., 2006), and more recently, immunogenicity of HLA-binding CTL-epitopes (Chowell et al., 2015). However, computational predictions have an inherent error in their accuracy, and will need to be experimentally validated. Experimental techniques involve arduous low throughput immunological techniques such as Elispot, or flow cytometric evaluation (Miyahira et al., 1995; Zaritskaya, Shurin, Sayers, & Malyguine, 2010).

Our goal here was to develop a sensitive and inexpensive immunologic assay that overcomes many of the limitations inherent to experimental T-cell antigen discovery approaches. We hypothesized that directional secretion of cytokines by activated T-cells into the immunologic synapse can enable the detection of antigenic APCs from a cDNA library. Our assay employs two color flow cytometry combined with magnetic or fluorescence based identification with deep sequencing allowed us to correctly identify immunogenic antigen from a pool of control genes, with high sensitivity (upto ~10% antigen encoding APCs), specificity at low effector:target ratios (~1:1) without the need of knowledge of donor's HLA type.

5.2 Results

5.2.1. Development of autologous APC system and APC capture assay

Interaction between a T-cell and an APC presenting a cognate antigenic epitope causes polarization of the T-cell cytoskeleton resulting in the directional secretion of specific cytokines such as IL-2 and IFNγ preferentially into the immunologic synapse secreted directionally into the immunological synapse into the target cell (Huse, Lillemeier, Kuhns, Chen, & Davis, 2006; Xie, Tato, & Davis, 2012). The directional CTL cytokine secretion is temporally dependent on early time points after TCR-HLA-epitope interaction (Han et al., 2011; Sanderson et al., 2012; von Bergwelt-Baildon et al., 2002). We leverage this dynamic paracrine cytokine secretion by capturing IFNγ secreted by CTLs on the immunologic synapse of APCs presenting cognate antigen via a commercially available bispecific IFNγ catch antibody (Brosterhus et al., 1999; Campbell, 2003) (Fig. 5-1 B-C). We utilize CD40L activated B-cells (CD40L.APCs) expanded from T-cell depleted fraction, and purified CD8+ T-cells from subject blood mononuclear cells (PBMCs). CD40L.APCs have been previously shown to be a rapid source of autologous APCs, have abundant expression of the target CD45 protein for bispecific catch antibody, and are highly programmable by transfection methods (von Bergwelt-Baildon et al., 2002). As shown in Fig. 5-1 A top panel, after two weeks of CD40L feeder cell stimulation, >90% of cells are CD19+ CD86+

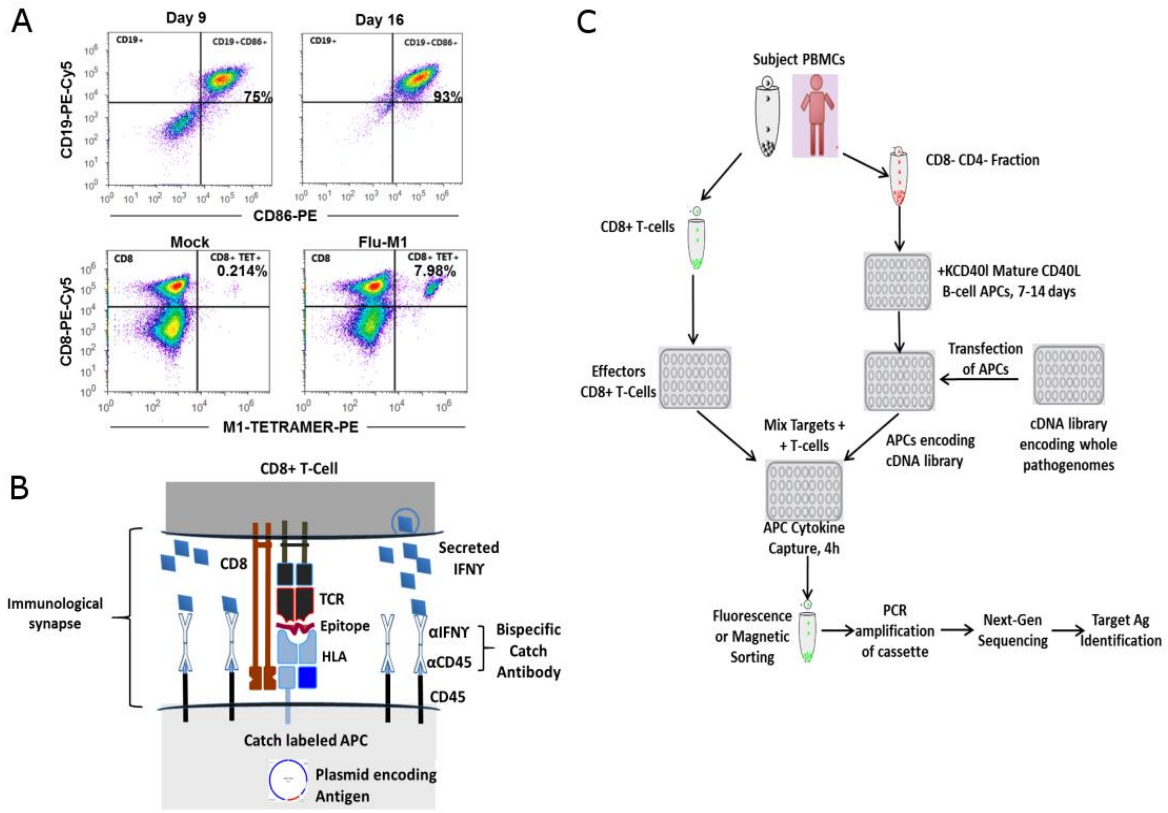

Figure 5-1. Development of CD40L-APC system and workflow of APC-cytokine capture assay. (A) Autologous CD40L.APCs can be expanded and matured (CD19+ CD86+ fraction) in two weeks (top panel); CD40L.APCs are efficient in inducing CTL-memory responses (FLUM1 epitope from Influenza A virus) in control peptide compared to M1-epitope pulsed APCs (bottom panel) (B) IFNϒ captured by bispecific catch antibody on immunological synapse of APCs. (C) Workflow of the APC-capture assay followed next generation sequencing of antigen identification.

expressing B-cells, indicating a pure mature APC fraction. CD40L.APCs are also very efficient in stimulating virus-specific CTLs as demonstrated in Fig. 5-1 A, where an HLA-A*02:01+ healthy donor had a 37-fold increase in FLUM1-epitope specific CTLs in two weeks compared to baseline frequencies (Fig. 5-1A, bottom panel). We developed a workflow based on this assay where mature CD40L.APCs are transfected with antigens derived from a cDNA library (encoding the whole pathogenome) and are then subject to APC-cytokine capture assay (Methods, Fig. 5-1C). APCs with the captured cell surface IFNϒ are subject to secondary antibody staining, followed by standard fluorescent or magnetic cell sorting assays. Following cell sorting, the target antigen-

expressing plasmid responsible for T-cell stimulation is isolated, amplified by universal primers, and identified by deep sequencing (Fig. 5-1 C).

5.2.2. APC-capture assay methodology

We first tested the feasibility of our APC-capture assay by peptide pulsing APCs overnight with pre-mixed virus specific peptide pools (CEF), labeled the APCs next day with the bispecific-IFNϒ antibody, and co-cultured them with purified CTLs for 4.5 hours at 37°C. A representative figure is shown in Fig. 5-2 A, where the CD8- fraction representing APCs have a significant increase (~ 5-fold) in cell surface IFNϒ+ APCs compared to DMSO-pulsed APCs or APCs alone. We then tested the limits of the assay using peptide pulsed autologous CD40L.APCs at various CTL:APC ratios, and time points of APC maturity (Fig. 5-2). Although the assay worked well across all the conditions, we observed that an APC:CTL ratio of 2.5:1 was sufficient to distinguish CTL-targeted CEF-pulsed APCs from control-APCs, and increasing the number of CTLs led to slight increase in

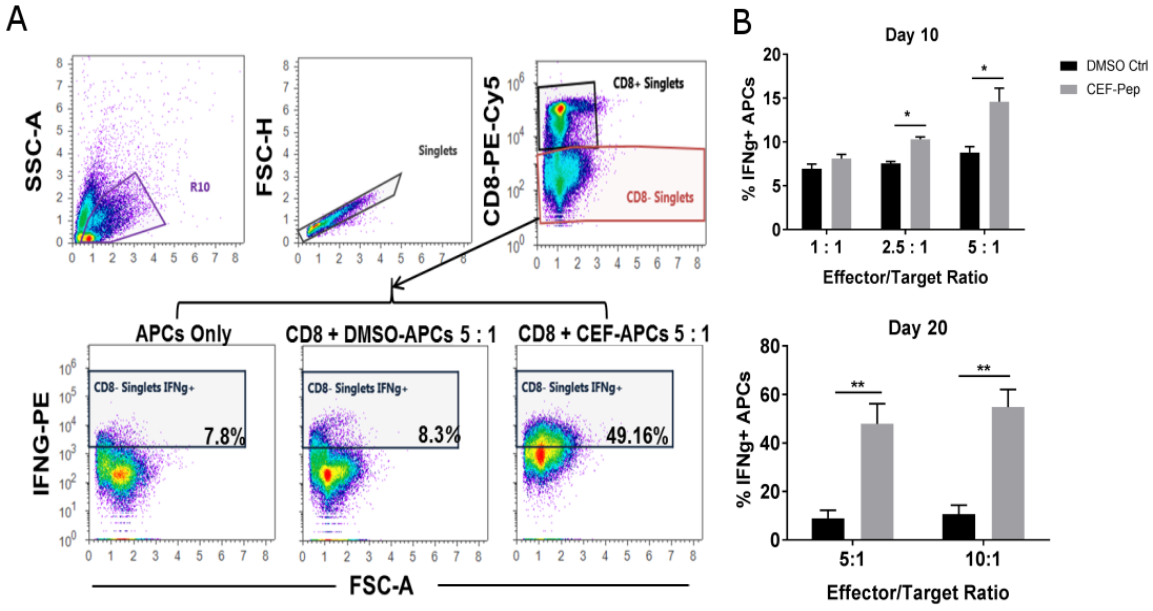

Figure 5-2. Experimental testing of APC-cytokine capture assay. (A) Representative example and gating strategy of APC-IFNϒ capture on peptide pulsed APCs E:T = 5:1. % indicate APC IFNϒ+ events within CD8- gate. (B) Quantification of APC-capture assay using different E:T (CTL:APC) ratios on day 10 of mature CD40L.APCs (top panel), and on day 20 of mature CD40L.APCs (bottom panel). CEF-Pep represents the pre-mixed immunogenic epitopes from CMV, EBV and FLU. Data represents mean of 3 biological replicates. *P < 0.05, **P < 0.01, Unpaired two-tailed Welch's T-test.

background (5:1, Fig. 5-2 B top panel), and there was no benefit to increase CTL:APC ratio beyond 5:1 (Fig. 5-2 B bottom panel). We also observed a higher proportion of IFNɣ+ APCs when APCs were used by day 20, compared to less mature APCs from day 10 (>3-fold for 5:1 E:T, 45% IFNɣ+ APCs compared to 15%), consistent with a highly pure CD19+ CD86+ faction observed in second week of CD40L.APC maturation (Fig. 5-1 A). We used an CTL:APC ratio of 2:1, in week 3 of APC maturation for subsequent experiments.

5.2.3. Specificity optimization of the APC-capture assay

Because this workflow depends on capturing IFNɣ on single APCs encoding cognate antigen, we were concerned about "off-target" effects in which IFNɣ secreted by CTLs are captured on

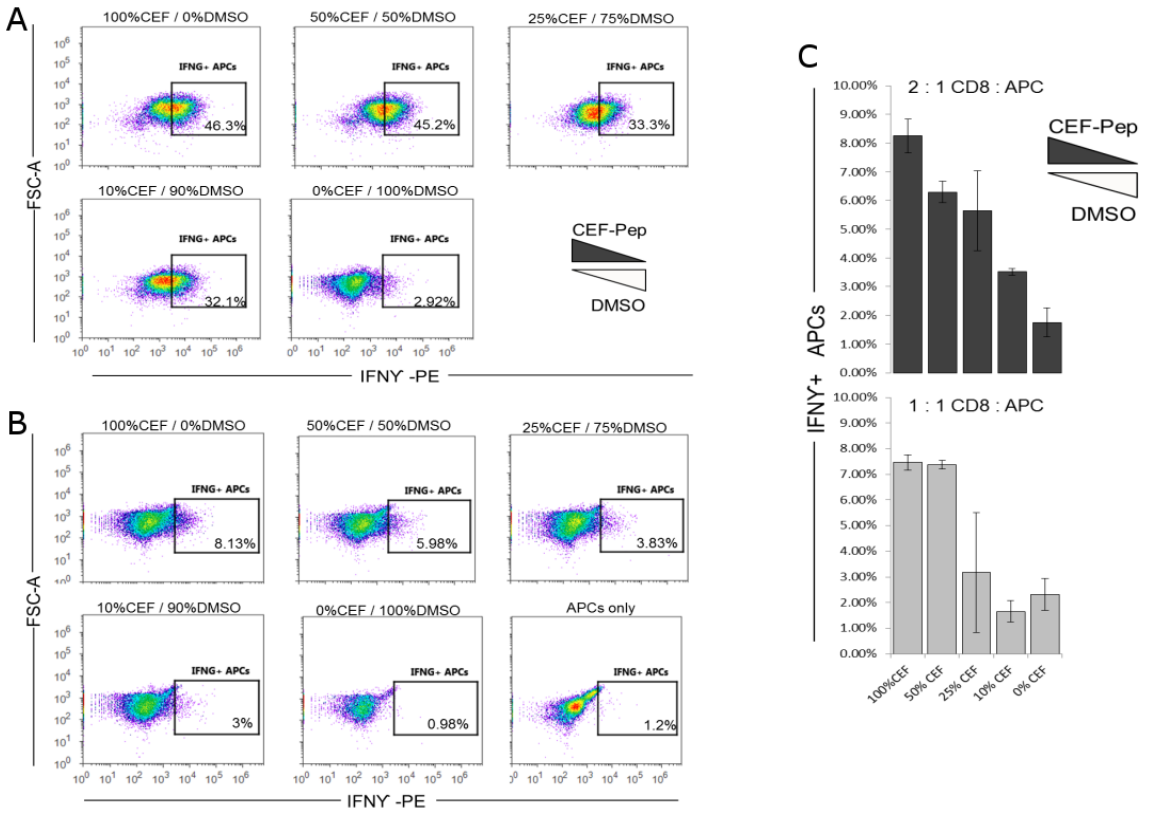

Figure 5-3. IFNɣ APC capture assay specificity optimization. (A) Representative flow cytometry plots from a donor with off-target effects showing no decrease in IFNɣ+ APCs with decreasing proportions of antigenic APCs. (B) Representative flow cytometry plots from same donor with reduced off-target effects after optimization and dose-dependence (C) Quantification of biological triplicates in 5-3 B, with two different CTL:APC ratios.

bystander APCs without a cognate immunogen, leading to false positives. To test the specificity and sensitivity of this assay, we assessed CTL targeting of APCs with cognate antigen (CEF-pulsed) mixed in varying proportions with control DMSO-pulsed “cold target” APCs. Our rationale here was that if there was off-target effects, bystander cold target APCs would also be labeled with IFNγ regardless of the number of true antigen-specific APCs present in the pool. This assessment would allow us to estimate the specificity of the assay, especially for true target APCs present at low frequencies in the population.

Our first attempts at this specificity experiments indicated that there were indeed off target effects, because IFNγ+ APCs did not sufficiently decrease with decreasing proportion of CEF-pulsed APCs (Fig. 5-3 A). We further noted that the percentage of IFNγ+ APCs in some cases far exceeded the total percent of immunogenic APCs (e.g. 25%, 10% CEF.APCs in Fig. 5-3 A). However, because there was a general trend of decrease in overall IFNγ+ APCs (100% CEF.APCs versus 10% CEF.APCs), we reasoned that IFNγ secreted by CTLs diffused and labelled cold target APCs when the samples were mixed to transfer and perform antibody staining in the plate. To overcome these issues, we stopped CTL-IFNγ secretion by incubating samples on ice for 30 minutes, followed by careful removal of any cytokine containing supernatant without mixing. These optimizations substantially decreased off-target issues, when the same donor showed substantially lower percentage of IFNγ+ APCs even at 100% CEF.APCs (8.13% in Fig. 5-3 B compared to 46.3% in Fig. 5-3 A). More importantly, there was a dose-dependent decrease of IFNγ+ APCs following the trend of decreasing true antigenic APCs (Fig. 5-3 B-C), and was overall applicable to varying CTL:APC ratios (Fig. 5-3 C). These results indicated that the optimized APC capture assay was able to specifically label antigenic APCs, and had a sensitive detection limit of up to 10% target APCs (Fig. 5-3B).

5.2.4. APC-capture assay and antigen recovery using transfected APCs

We then tested the applicability of the assay to antigen transfected APCs in order to recover the target antigen responsible for CTL immunogenicity. We first assessed transfection efficiency by GFP expression 24 hours after transfection of APCs and observed that close to 60% of APCs

were positive (Fig. 5-4 A). We then repeated the cold targets experiment, where we mixed APCs transfected with either FLUM1 antigen (from Influenza A virus) or the control self-antigen GAPDH

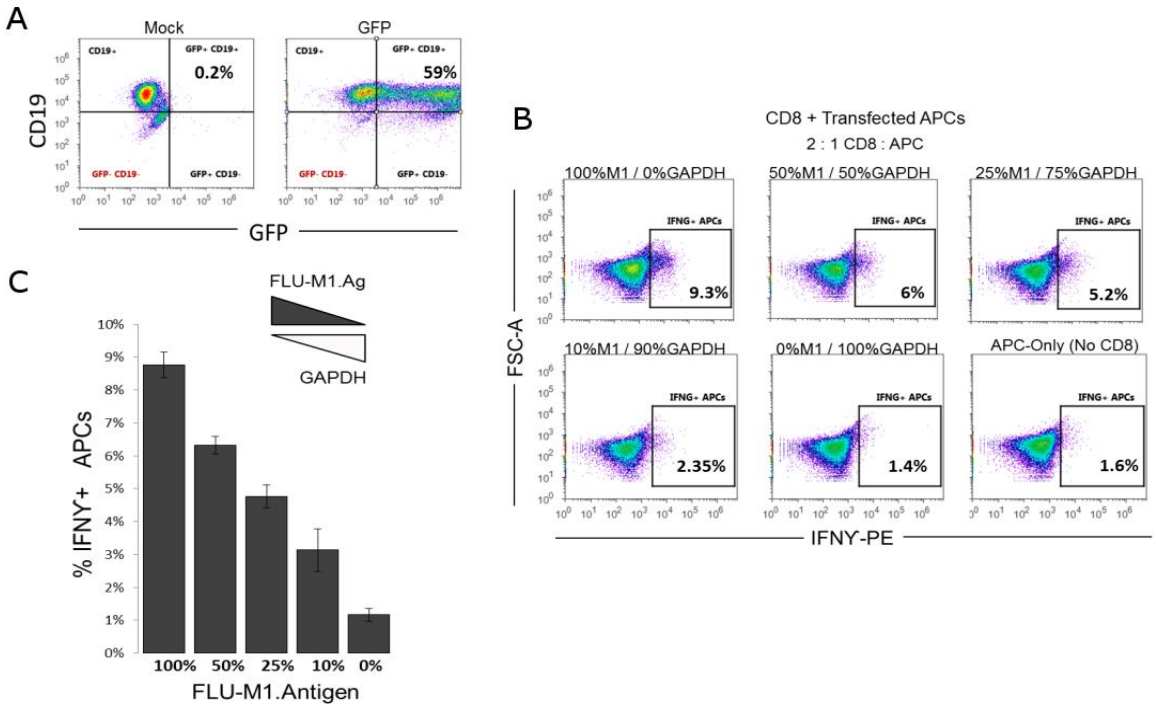

Figure 5-4. IFNɣ APC capture assay on transfected APCs. (A) Transfection efficiency of APCs assessed with GFP 24 hours post nucleofection (B) Representative flow cytometry plots from same donor with transfected cold targets showing dose dependent decrease in IFNɣ+ APCs (C) Quantification of biological triplicates in 5-4 B.

in varying proportions and co-incubated with autologous CTLs (Fig. 5-4 B-C). Similar to the peptide pulse experiments, we observed a clear hierarchy of dose dependence from 9.5% mean IFNɣ+ APCs for 100% FLUM1.APcs down to 3% mean IFNɣ+ APCs with varying decreasing proportions in between (Fig. 5-4 C). We further noted that there was more consistency in hierarchy of dose dependence for antigenic IFNɣ+ APCs in transfected APCs compared to peptide pulsed APCs (Fig. 5-3). Similar results were obtained with other independent donors (data not shown). This might be due to a reduced antigen load in the APCs as opposed to saturating levels of epitopes loaded on HLA-I molecules when synthetic peptides are being used to load APCs.

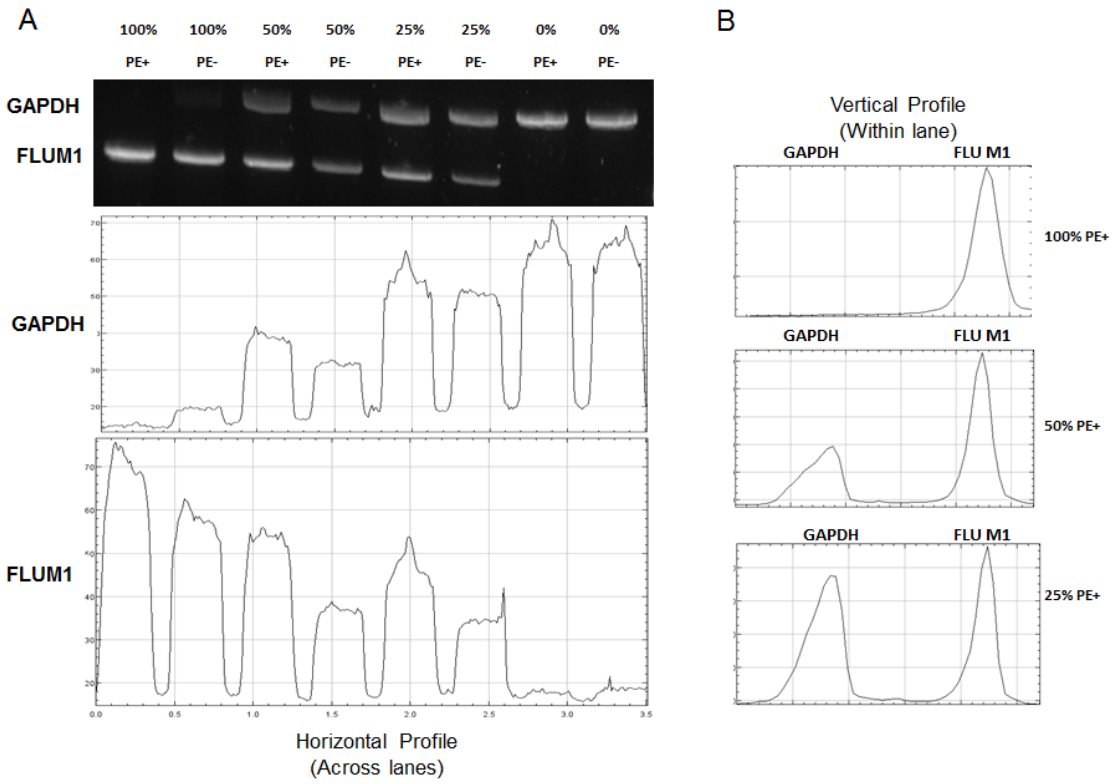

Figure 5-5. Antigen recovery of IFNɣ+ APCs. (A) PCR amplification of the antigen cassette after MACS sorting of PE+ and PE- fractions (top panel); Gel intensity across lanes (horizontal profile) for GAPDH (middle panel), and FLUM1 antigen (bottom panel). (B) Gel intensity of PCR cassette within the sample (vertical profile) indicating GAPDH and FLUM1 peaks.

We then tested assessed the efficiency of antigen recovery from the transfected immunogenic APCs for applicability in the workflow described in Fig. 5-1 C. We labelled the pooled APCs and CTLs with IFNɣ-phycoerythrin (PE) conjugated secondary antibody, followed by magnetic associated cell sorting (MACS) with anti-PE conjugated magnetic beads (Methods). Following MACS isolation, cells were subject to plasmid extraction to recover the antigen, PCR amplified and run on an agarose gel to visualize the recovered antigen (Fig, 5-5). We compared the band intensities of the PCR-amplified cassette for both the IFNɣ+ fraction (PE+) and IFNɣ- unbound flow through fraction (PE-) (Fig. 5-5). We observed a consistently higher FLUM1 band intensity in the PE+ fraction compared to the PE- fraction irrespective of the number of initial FLUM1.APCs (Fig. 5-5 A, bottom panel) indicating successful antigen recovery and that FLUM1.specific APCs are being enriched by the IFNɣ APC capture assay (Fig. 5-5 A-B).

Comparing the PE+ fractions alone, we observed a proportional decrease in FLUM1 bands with decreasing FLUM1-specific initial APCs (mirroring the flow cytometry results (Fig. 5-4). GAPDH bands on the other hand showed an increasing trend with decreasing FLUM1.APCs and were comparably represented in both the PE+ and PE- fractions in 100% GAPDH.APCs (Fig. 5-5, middle panel, last two samples). Lastly, within the sample, we observed higher intensity of the FLUM1 band relative to GAPDH band (Fig. 5-5 B) indicating specificity, although the specificity decreased with lower proportion of antigenic FLUM1.APCs (25%). These results indicated that while antigen specificity was significant by flow cytometry methods, magnetic sorting might be also selecting for non-antigenic targets via non-specific binding to cells (see Discussion).

5.3 Discussion

There is considerable interest in identifying the antigenic repertoire of CD8+ T-cells in a given individual. Knowledge of such antigens are critical to the development of novel T-cell based vaccines against pathogens (Robinson & Amara, 2005), cancer immunotherapies targeting tumor specific antigens, neoantigens (Carreno et al., 2015; van Rooij et al., 2013; Verdegaal et al., 2016), identifying the T-cell receptor repertoire specific for antigens (Glanville et al., 2017; Strønen et al., 2016), and immune monitoring of various malignancies (Jäger et al., 2000; Snyder et al., 2014).

In this study, we aimed to develop an immunologic assay that when combined with next generation sequencing could yield a high throughput overview of immunogenic antigens from a given pathogenome. To this end, we leveraged directional IFNγ secretion by CTLs to develop an IFNγ APC capture assay presenting cognate immunogenic T-cell activating antigen. We optimized and tested the limits of this assay first by peptide pulsed APCs, then through antigen transfected APCs via flow cytometry and MACS sorting.

Our preliminary results indicate that IFNγ APC capture can be used to identify immunogenic antigens and peptides with specificity and sensitivity to up to 10% APCs presenting cognate antigen (Figs. 5-3, 5-4). These results support other studies which have shown that even low frequency antigens (upto one epitope-HLA complex per cell) can stimulate and activate pre-

existing T-cells (Huang et al., 2013; Sykulev, Joo, Vturina, Tsomides, & Eisen, 1996). As such, this simple flow cytometry based assay is performed using low sample numbers (200,000 CTLs : 100,000 APCs per well), using frozen CTLs within a day (4-5 hours incubation). Thus, without invoking sequencing techniques, one can envision that differential labeling of antigens can enhance the throughput of this flow cytometric assay.

However, in large pathogenomes (e.g. CMV ~ 200 genes) methods beyond multi-parameter flow cytometry are required for high throughput antigen discovery. Therefore, we tested IFNɣ APC capture in transfected APCs and showed that the assay is consistent in transfected APCs, with results largely similar to peptide-pulsed APCs. After controlling for IFNɣ diffusion and off target effects (Fig. 5-3 A-B), transfected APC experiments showed that there is a clear dose dependent decrease in IFNɣ+ FLUM1.APCs dependent on starting proportion of either FLUM1 or control APCs (Fig. 5-4) which is critical for a highly sensitive single-cell assay. Furthermore, we were successful in isolating the FLUM1 amplicon from IFNɣ+ APCs and we demonstrated enrichment of the antigen in samples with varying proportions of the antigenic APCs (Fig. 5-5). These experiments will enable us to employ the assay to identify T-cell immunogenic antigens from complex pooled antigen libraries in future studies.

The current experiments have also revealed caveats in our methodology. Although, flow cytometry experiments revealed high sensitivity and specificity across multiple donors, at low frequency antigenic APCs there were discrepancies when flow cytometry results were compared to MACS based antigen recovery. For instance, at 25% FLUM1.APC frequency, there was a five fold higher percentage of IFNɣ+APCs compared to control APCs (0% FLUM1.APCs) by flow cytometry (Fig. 5-4 B). However, MACS isolation antigen recovery and PCR amplification of the exact same sample showed only a modest enrichment of FLUM1 band relative to the control GAPDH band (Fig. 5-5 B, bottom panel). Based on these results, I speculate that at low antigen APC frequencies MACS sorting can lead to non-specific isolation of bystander cells and might therefore not be ideal for this assay in complex libraries. We therefore propose either further optimization of MACS sorting methodology or employ fluorescent activated cell sorting (FACS) to isolate highly pure enriched populations of antigenic IFNɣ+APCs for future experiments.

5.4 Materials and Methods

5.4.1. Donor PBMC processing and cell separation

Peripheral blood was obtained from healthy donors at Arizona State University under IRB protocol #MOD00006783. All subject blood samples were obtained under informed consent in accordance with ASU's human subject research policies. For some samples, trima leuko reduction chamber residuals from Blood Centers of the Pacific blood collection center, CA, USA. PBMCs were extracted from both these source samples using Sepmate PBMC isolation tubes (StemCell technologies, MA, USA) according to manufacturer's instructions. Isolated PBMCs were counted and CD8+ T-cells were isolated using human CD8+ T-cell isolation kit (Miltenyi Bitoec, Germany) according to manufacturer's instructions, and purified CTLs were frozen until use. CD8- fraction was further subject to CD4 fraction removal using CD4 microbeads (Miltenyi Biotec, Germany) and frozen. CD4- CD8- fractions were counted and either frozen or used to generate CD40L activated APCs. Purity of each fraction was assessed by flow cytometry prior to being frozen.

5.4.2. Autologous APC generation

Autologous APCs were generated by methods described in Chapter 5 (Appendix B5). Briefly, CD40L-activated B-cell APCs were generated from donor CD4- CD8- PBMC fractions by incubating them with irradiated (32Gy) K562-cell line expressing human CD40L (KCD40L) at a ratio of 4:1 (800,000 CD4- CD8- to 200,000 irradiated KCD40Ls) in each well. The cells were maintained in BCM consisting of IMDM (Gibco, USA), 10% heat inactivated human serum (Gemini Bio Products, CA, USA), Antibiotic-Antimycotic (Anti-Anti, Gibco, USA). BCM was supplemented with 10 ng/mL recombinant human IL-4 (R&D Systems, MN, USA), 2µg/mL Cyclosporin A (Sigma-Aldrich, CA, USA), 1X insulin transferrin supplement (ITES, Lonza, MD, USA). APCs were re-stimulated with fresh irradiated KCD40Ls on days 5 and 10, after washing with 1X PBS and expanding into a whole 24-well plate. After two weeks, APC purity was assessed by CD19+ CD86+ expressing cells by flow cytometry, and were generally used for T-

cell stimulation after >90% purity. APCs were either restimulated upto 4 weeks or frozen and re-expanded as necessary.

5.4.3. IFNγ APC capture assay

APCs were used on days as indicated for the APC capture assay. For peptide experiments, peptide pulsing of APCs were done under BCM 5% human serum, with recombinant IL-4 overnight. Transfection of APCs were done using the Lonza 4D Nucleofector, primary P3 buffer, program EO117 (Lonza, MD, USA) and incubated in BCM-10% human serum, IL-4 without any Anti-Anti. Twenty four hours later APCs were washed resuspended in MACS staining buffer, counted, and 8uL of IFNγ catch reagent (IFN-γ Secretion Assay - Cell Enrichment and Detection Kit, Miltenyi Biotec, Germany) for 20 minutes on ice. Purified CTLs were thawed either the night before or on the day of the assay, and resuspended at desired E:T concentrations as indicated in the experiment in BCM-5% human serum. After catch reagent labelling, APCs were washed three times with BCM-5% human serum, resuspended in desired E:T ratios as necessary for the experiment, and co-cultured with pure CTLs in a 96 well U-bottomed plate (Costar, USA), total volume of 200µL. The plate was centrifuged at 400g, 1 minute, without brakes to promote conjugate formation, and incubated at 37C for 4-5 hours. For the optimization experiments, after the incubation time, the plate was chilled for 30 minutes on ice to stop IFNγ secretion, and supernatant media was carefully removed. For initial washing of the cells, cold MACS buffer was added drop by drop to minimize IFNγ diffusion. The plate was centrifuged at 500g, 5 minutes, followed by careful removal of supernatant and same wash procedure was repeated. After centrifugation, secondary detection IFNγ-PE antibody(IFN-γ Secretion Assay - Cell Enrichment and Detection Kit, Miltenyi Biotec, Germany) in 2µL/100µL/well MACS buffer was added and incubated for 20 minutes on ice. For flow cytometry analysis, additional CD8 antibody anti-CD8-PC5 (clone B9.11; Beckman Coulter 1:100) was added. After incubation the plates were washed twice with MACS buffer, resuspended in 1x sterile PBS and analyzed by Attune flow cytometer (ThermoFisher Scientific, MA, USA). For MACS sorting, after IFNγ-PE antibody staining, cells were pelleted and 10µL anti-PE microbeads were added to the pellet in 90µL MACS buffer for 15

minutes. Cells were then washed once, isolated using MACS MS columns (Miltenyi Biotec, Germany), and flow through and eluate were subject to Hirt plasmid extraction.

5.4.4. Tetramer staining

FluM1-tetramer (corresponding to the HLA-A*02:01-restricted epitope GILGFVFTL) was obtained from Proimmune, UK. For tetramer staining, cells were re-suspended in 100µL staining buffer with 5% human serum and 1mM Dasatanib (ThermoFisher Scientific, MA, USA), and each tetramer was added at concentration of 10uL for 30 minutes at room temperature. Cells were washed twice and restained with anti-CD8-PC5, anti-CD4-FITC, anti-CD14-FITC and anti-CD19-FITC for exclusion gates for 30 minutes on ice. Samples were then washed twice in 1x PBS, and analyzed by flow cytometry.

5.4.5. Antigen recovery by Hirt plasmid extraction

MACS sorted PE+/PE- CD8- cells were first pelleted by centrifugation at 500xg for 10 minutes. The supernatant was carefully aspirated and the cells re-suspended gently in 250 uL of chilled Hirt Buffer #1 (50 mM Tris, 10mM EDTA, 50ug/mL RNase A). Following resuspension 250 uL of Lysis Buffer #2 (1.2% SDS in Water) was added to each sample and the tubes were slowly and fully inverted 3 times and incubated for 5 minutes at room temperature. Cell lysis was stopped via the addition of 350 uL of Hirt Precipitation Buffer #3 (3M CsCl, 1M Potassium Acetate, 0.67M Acetic Acid) to each sample. After incubation the samples were quickly loaded into a microcentrifuge chilled to 4 degrees C and centrifuged for 30 minutes at 17,300xg. The supernatant was loaded onto Qiagen QIAprep Spin Miniprep Columns by spinning the columns at 17,300xg for 1 minute. The column was washed twice with 700 uL of Hirt Column Wash #4 (60% EtOH, 10mM Tris, 50uM EDTA, 80 mM Potassium Acetate) and the flow through was discarded. Plasmid DNA was eluted from the column in nuclease-Free Water to the membrane and heating the column at 70 degrees C for 1 minute, followed by spinning the warmed column at 17,300 xg for 1 minute. The eluted DNA concentration was quantified by Nanodrop.

5.4.6. PCR amplification and gel analysis

For each reaction, 25ng of plasmid DNA added to a mastermix solution containing 10uL SapphireAmp PCR Mix, 1 uL of each primer at 10 uM, with water added up to 20 uL for the total reaction volume. After a 5 minute initial denaturation and activation of the hot-start polymerase at 95°C, the samples were cycled 27 times. Each cycle consisted of a 30 second denaturation step at 95°C, a 50 second annealing step at 57C, and a 3 minute extension step at 72°C. Finally, after 27 cycles, the samples were extended once more for 7 minutes at 72°C, followed by cooling to 4 C for overnight storage. 2 uL of each PCR reaction was loaded onto a 2% agarose gel prestained with 1uL GelRed / 2mL 1X TAE and visualized.

CHAPTER 6

CONCLUSIONS AND FUTURE OUTLOOK

CD8+ T-cells form a critical part of a complex network of cell subsets in the human immune system causing immune mediated the elimination of many pathologies and malignancies. Despite major advances in the understanding of the molecular mechanisms underlying T-cell immunobiology, the core driving features of T-cell immunogenicity and tumor-specific mechanisms that cause T-cell dysfunction remain largely unknown. This dissertation is an attempt to address these questions experimentally using classical and novel immunological techniques, and computationally using data mining, statistical and machine learning tools.

In Chapter 2, I studied the biochemical hallmarks underlying T-cell immunogenicity in context of self/nonself discrimination. We identified an over representation of relatively hydrophobic amino acids at specific TCR contact residues of immunogenic CD8+ T-cell epitopes. We leveraged these findings to develop a computational immunogenicity model which can consistently improve prediction of immunogenic CTL epitopes. We further validated these findings by *in vivo* experiments in a murine HIV-1 infection model (Chowell et al., 2015).

Interestingly, strong support for our study came from a subsequent 2016 study by Stadinski et al. (Stadinski et al., 2016), describing importance of hydrophobic amino acids at positions 6 and 7 of TCR CDR3β in the development of self-reactive T cells. One implication of Stadinski et al's findings is that hydrophobic amino acids in the CDR3β region may de facto be crucial for T cell activation via recognition of self or foreign antigen as alluded in their study. Since negative selection eliminates the majority of self-reactive TCRs, current thinking would anticipate that the biochemical composition of TCRs in the periphery would be skewed towards amino acids that are not strongly hydrophobic. Nonetheless, for a selected TCR to be able to recognize a foreign antigen it must bind to the epitope with very high affinity, which can be effectively accomplished if the epitope is also composed of hydrophobic amino acids residues. This interpretation is consistent with our findings from Chapter 2 (Chowell et al., 2015), where we

reported that the hydrophobic composition of MHC class I epitopes determine their immunogenicity. Notably, we found that the relative differential hydrophobicity of TCR-contact residues (positions 4 to 7) is a hallmark of immunogenic pathogen-derived and self-reactive epitopes, in both human and mouse MHC-allomorphs (Chapter 2). This would explain that, both non-self-antigen and self-antigen reactive TCRs with hydrophobic CDR3β residues are perhaps not completely eliminated due to lack of antigenic gene expression in the thymus or inefficiencies in negative selection, respectively (Yu et al., 2015).

Together, our study and that of Stadinski et al, shed light on understanding of the biochemical basis of self and non-self discrimination by T cells. They further underscore the predictive value of incorporating amino acid biochemical properties in identifying immunogenic epitopes and their corresponding TCRs. Several important unresolved questions remain with respect to the role of hydrophobic amino acid residues in promoting T-cell activation. For instance, is TCR-pMHC interfacial hydrophobicity a general rule that applies across different MHC class I and II alleles and epitope lengths? Do TCRs tread a threshold value of affinity between their CDR3β and self-pMHC hydrophobic contact residues to avoid negative selection? Lastly, is the T cell receptor a hydrophobicity-driven pattern recognition receptor (Seong & Matzinger, 2004)? If so, the evolutionary basis of TCR recognition of amino acid hydrophobicity will be interesting to explore. Future studies will address these fascinating questions.

In Chapter 3, I developed an experimental and bioinformatic methodology for the amplification and detection of low frequency T-cells in a short term ex vivo stimulation protocol (Krishna & Anderson, 2016). In Chapter 4, I used this methodology to define several novel CTL epitopes from the HPV16 antigens E2, E6 and E7. Although previous studies have attempted to define HPV16 CTL-epitopes (Ressing et al., 1995b; Riemer et al., 2010d; Rudolf, Man, Melief, Sette, & Kast, 2001b), to my knowledge this is one of the first comprehensive epitope, immunodominant domain mapping studies from E2, E6 and E7. Knowledge of these epitopes and regions from the antigens can aid the design and development of HPV-specific immunotherapies. My results also indicate that, in addition to E6 and E7, HPV16-E2 is an immunogenic antigen with

broad T-cell and B-cell reactivity and is expressed in a subset of HPV+ HNSCCs. My findings of E2-immunogenicity, supported by other recent studies indicating lower rates of genome disruption at the E2-locus (Parfenov et al, PNAS, 2014), warrants further investigation of E2 in immune therapeutic modalities in addition to E6 and E7.

In Chapter 4, I also studied the different modes of HPV-specific CTL dysfunction in HPV+ HNSCC. I compared and contrasted dysfunctional phenotypes of HPV16-specific E2, E6 and E7 CD8+ T-cells after *ex vivo* stimulation in HPV+ HNSCC patients. Most patients had higher levels of E7 or E2-specific T-cell dysfunctional phenotypes, compared to E6-T-cells, indicating a predictable but heterogeneous mode of HPV-specific T-cell dysfunction in HPV+ HNSCC. HPV-T-cell dysfunction in HNSCCs has not been described before, and my findings aids better understanding of HPV-immunobiology. I also found by transcriptomic analysis that high HPV-antigen load is strongly correlated with, and likely drives, T-cell infiltration and subsequent T-cell dysfunction observed in these tumors. I further show computationally, and then experimentally, that the immune modulatory indoleamine 2,3-dioxygenase (IDO-1) expression correlates with HPV-antigen load and is highly expressed in HPV+ cancers. Inhibiting the IDO-1 pathway can synergize PD-1 blockade by enhancing cytotoxic potential of E2, E6 and E7-specific CTLs *in vitro*. This result is particularly significant considering ongoing clinical trials that are exploring PD-1 blockade with IDO-1 inhibition as immune therapies against HNSCCs (Ferris et al, NEJM, 2016, Gangadhar et al, JITC, 2015). To my knowledge, this is the first description of a mechanistic explanation of HPV-specific CTL dysfunction and immunogenicity in HNSCC patients. Future studies that compare the dynamics and interplay of HPV-CTLs from tumor, periphery with tumor HPV load will be necessary to verify the HPV T-cell dysfunction findings from this study.

In the last chapter (Chapter 5), I developed an in vitro immunologic assay that can enable the identification of CTL-immunogenic antigens by flow cytometry. The IFNɣ APC capture assay was antigen specific, was sensitive to upto 10% of antigen encoded APCs, and was highly reproducible in multiple donors. Furthermore, the IFNɣ APC capture assay is inexpensive, is performed within one day, and requires minimal sample (CTL:APC ratios of 2:1). I also

demonstrated efficient antigen recovery from antigen transfected IFNY+ APCs. However, the assay needs to be optimized for isolating and antigen recovery of low frequency by magnetic or fluorescence based sorting. This is especially important when antigenic APC frequencies are low for high throughput antigen identification. To this end, future experiments that assess and compare the efficiency of FACS and MACS sorting methods for antigen recovery after the IFNY APC capture assay. After optimization of antigen recovery, the assay will be employed to identify immunogenic antigens from three common human viruses CMV, EBV and Flu.

Thus, this dissertation makes a step towards understanding fundamental CD8+ T-cell immunobiology in context of cancer and viral infections, and develops methodologies for the design of effective T-cell immunotherapies. As such, I believe this dissertation will be of interest and of relevance to the fields of tumor immunology, virology, vaccinology and immunogenomics.

REFERENCES

Abbas, A. K., Lichtman, A. H. H., & Pillai, S. (2014). *Cellular and Molecular Immunology*. Elsevier Health Sciences.

Akagi, K., Li, J., Broutian, T. R., Padilla-Nash, H., Xiao, W., Jiang, B., … Gillison, M. L. (2014). Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Research*, *24*(2), 185–199.

Allison, J. P., McIntyre, B. W., & Bloch, D. (1982). Tumor-specific antigen of murine T-lymphoma defined with monoclonal antibody. *Journal of Immunology* , *129*(5), 2293–2300.

Anderson, K. S., Dahlstrom, K. R., Cheng, J. N., Alam, R., Li, G., Wei, Q., … Sturgis, E. M. (2015a). HPV16 antibodies as risk factors for oropharyngeal cancer and their association with tumor HPV and smoking status. *Oral Oncology*, *51*(7), 662–667.

Assarsson, E., Bui, H.-H., Sidney, J., Zhang, Q., Glenn, J., Oseroff, C., … Sette, A. (2008). Immunomic analysis of the repertoire of T-cell specificities for influenza A virus in humans. *Journal of Virology*, *82*(24), 12241–12251.

Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E., & Unanue, E. R. (2005). Binding of immunogenic peptides to Ia histocompatibility molecules. 1985. *Journal of Immunology* , *175*(7), 4163–4165.

Babbitt, B. P., Matsueda, G., Haber, E., Unanue, E. R., & Allen, P. M. (1986). Antigenic competition at the level of peptide-Ia binding. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(12), 4509–4513.

Bachy, V., Hervouet, C., Becker, P. D., Chorro, L., Carlin, L. M., Herath, S., … Klavinskis, L. S. (2013). Langerin negative dendritic cells promote potent CD8 T-cell priming by skin delivery of live adenovirus vaccine microneedle arrays. *Proceedings of the National Academy of Sciences*, *110*(8), 3041–3046.

Badoual, C., Hans, S., Merillon, N., Van Ryswick, C., Ravel, P., Benhamouda, N., … Tartour, E. (2013). PD-1-expressing tumor-infiltrating T cells are a favorable prognostic biomarker in HPV-associated head and neck cancer. *Cancer Research*, *73*(1), 128–138.

Bai, Y., Ni, M., Cooper, B., Wei, Y., & Fury, W. (2014a). Inference of high resolution HLA types using genome-wide RNA or DNA sequencing reads. *BMC Genomics*, *15*(1), 325.

Barber, D. L., Wherry, E. J., Masopust, D., Zhu, B., Allison, J. P., Sharpe, A. H., … Ahmed, R. (2006a). Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature*, *439*(7077), 682–687.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S.,… Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709–1712.

Benacerraf, B. (1981). Role of MHC gene products in immune regulation. *Science*, *212*(4500), 1229–1238.

Bentzen, A. K., Marquard, A. M., Lyngaa, R., Saini, S. K., Ramskov, S., Donia, M., … Hadrup, S. R. (2016). Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nature Biotechnology*, *34*(10), 1037–1045.

Billingham, R. E., Brent, L., & Medawar, P. B. (1953). “Actively Acquired Tolerance” of Foreign Cells. *Nature*, *172*(4379), 603–606.

Billingham, R. E., Silvers, W. K., & Wilson, D. B. (1963). Further studies on adoptive transfer of sensitivity to skin homografts. *The Journal of Experimental Medicine*, *118*, 397–420.

Binder, D. C., Engels, B., Arina, A., Yu, P., Slauch, J. M., Fu, Y.-X., … Schreiber, H. (2013). Antigen-specific bacterial vaccine combined with anti-PD-L1 rescues dysfunctional endogenous T cells to reject long-established cancer. *Cancer Immunology Research*, *1*(2), 123–133.

Birnbaum, M. E., Mendoza, J. L., Sethi, D. K., Dong, S., Glanville, J., Dobbins, J., … Garcia, K. C. (2014a). Deconstructing the peptide-MHC specificity of T cell recognition. *Cell*, *157*(5), 1073–1087.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. *Springer Verlag.*

Abbas, A. K., Lichtman, A. H. H., & Pillai, S. (2014). Cellular and Molecular Immunology. Elsevier Health Sciences.

Akagi, K., Li, J., Broutian, T. R., Padilla-Nash, H., Xiao, W., Jiang, B., … Gillison, M. L. (2014). Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. Genome Research, 24(2), 185–199.

Allison, J. P., McIntyre, B. W., & Bloch, D. (1982). Tumor-specific antigen of murine T-lymphoma defined with monoclonal antibody. Journal of Immunology , 129(5), 2293–2300.

Anderson, K. S., Dahlstrom, K. R., Cheng, J. N., Alam, R., Li, G., Wei, Q., … Sturgis, E. M. (2015a). HPV16 antibodies as risk factors for oropharyngeal cancer and their association with tumor HPV and smoking status. Oral Oncology, 51(7), 662–667.

Assarsson, E., Bui, H.-H., Sidney, J., Zhang, Q., Glenn, J., Oseroff, C., … Sette, A. (2008). Immunomic analysis of the repertoire of T-cell specificities for influenza A virus in humans. Journal of Virology, 82(24), 12241–12251.

Babbitt, B. P., Allen, P. M., Matsueda, G., Haber, E., & Unanue, E. R. (2005). Binding of immunogenic peptides to Ia histocompatibility molecules. 1985. Journal of Immunology , 175(7), 4163–4165.

Babbitt, B. P., Matsueda, G., Haber, E., Unanue, E. R., & Allen, P. M. (1986). Antigenic competition at the level of peptide-Ia binding. Proceedings of the National Academy of Sciences of the United States of America, 83(12), 4509–4513.

Bachy, V., Hervouet, C., Becker, P. D., Chorro, L., Carlin, L. M., Herath, S., … Klavinskis, L. S. (2013). Langerin negative dendritic cells promote potent CD8 T-cell priming by skin delivery of live adenovirus vaccine microneedle arrays. Proceedings of the National Academy of Sciences, 110(8), 3041–3046.

Badoual, C., Hans, S., Merillon, N., Van Ryswick, C., Ravel, P., Benhamouda, N., … Tartour, E. (2013). PD-1-expressing tumor-infiltrating T cells are a favorable prognostic biomarker in HPV-associated head and neck cancer. Cancer Research, 73(1), 128–138.

Bai, Y., Ni, M., Cooper, B., Wei, Y., & Fury, W. (2014a). Inference of high resolution HLA types using genome-wide RNA or DNA sequencing reads. BMC Genomics, 15(1), 325.

Barber, D. L., Wherry, E. J., Masopust, D., Zhu, B., Allison, J. P., Sharpe, A. H., … Ahmed, R.

(2006a). Restoring function in exhausted CD8 T cells during chronic viral infection. Nature, 439(7077), 682–687.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S.,… Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science, 315(5819), 1709–1712.

Benacerraf, B. (1981). Role of MHC gene products in immune regulation. Science, 212(4500), 1229–1238.

Bentzen, A. K., Marquard, A. M., Lyngaa, R., Saini, S. K., Ramskov, S., Donia, M., … Hadrup, S. R. (2016). Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. Nature Biotechnology, 34(10), 1037–1045.

Billingham, R. E., Brent, L., & Medawar, P. B. (1953). “Actively Acquired Tolerance” of Foreign Cells. Nature, 172(4379), 603–606.

Billingham, R. E., Silvers, W. K., & Wilson, D. B. (1963). Further studies on adoptive transfer of sensitivity to skin homografts. The Journal of Experimental Medicine, 118, 397–420.

Binder, D. C., Engels, B., Arina, A., Yu, P., Slauch, J. M., Fu, Y.-X., … Schreiber, H. (2013). Antigen-specific bacterial vaccine combined with anti-PD-L1 rescues dysfunctional endogenous T cells to reject long-established cancer. Cancer Immunology Research, 1(2), 123–133.

Birnbaum, M. E., Mendoza, J. L., Sethi, D. K., Dong, S., Glanville, J., Dobbins, J., … Garcia, K. C. (2014a). Deconstructing the peptide-MHC specificity of T cell recognition. Cell, 157(5), 1073–1087.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer Verlag.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1988). Structure of the human class i histocompatibility antigen, HLA-A2. https://doi.org/10.2210/pdb1hla/pdb

Blum, J. S., Wearsch, P. A., & Cresswell, P. (2013a). Pathways of Antigen Processing. Annual Review of Immunology, 31(1), 443–473.

Brahmer, J. R., Drake, C. G., Wollner, I., Powderly, J. D., Picus, J., Sharfman, W. H., … Topalian, S. L. (2010). Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 28(19), 3167–3175.

Brent, L., Brown, J., & Medawar, P. B. (1958). Skin transplantation immunity in relation to hypersensitivity. The Lancet, 2(7046), 561–564.

Brosterhus, H., Brings, S., Leyendeckers, H., Manz, R. A., Miltenyi, S., Radbruch, A., … Schmitz, J. (1999). Enrichment and detection of live antigen-specific CD4(+) and CD8(+) T cells based on cytokine secretion. European Journal of Immunology, 29(12), 4053–4059.

Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., & Wiley, D. C. (1993). Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature, 364(6432), 33–39.

Brownlie, R. J., & Zamoyska, R. (2013). T cell receptor signalling networks: branched, diversified

and bounded. Nature Reviews. Immunology, 13(4), 257–269.

Brunet, J.-F, Denizot, F, Luciani, M.-F, Roux-Dosseto, M., Suzan, M, Mattei, M.-G., & Golstein, P. (1987). A new member of the immunoglobulin superfamily—CTLA-4. Nature, 328(6127),

Burnet, F. M. (1961). Immunological recognition of self. Science, 133(3449), 307–311.

Burnet, M. (1957). Cancer--A Biological Approach: I. The Processes Of Control. II. The Significance of Somatic Mutation. BMJ , 1(5022), 779–786.

Burroughs, N. J., de Boer, R. J., & Keşmir, C. (2004). Discriminating self from nonself with short peptides from large proteomes. Immunogenetics, 56(5), 311–320.

Buynak, E. B. (1976). Vaccine against human hepatitis B. JAMA: The Journal of the American Medical Association, 235(26), 2832–2834.

Calis, J. J. A., Maybeno, M., Greenbaum, J. A., Weiskopf, D., De Silva, A. D., Sette, A., … Peters, B. (2013a). Properties of MHC class I presented peptides that enhance immunogenicity. PLoS Computational Biology, 9(10), e1003266.

Calis, J. J. A., Sanchez-Perez, G. F., & Keşmir, C. (2010). MHC class I molecules exploit the low G C content of pathogen genomes for enhanced presentation. European Journal of Immunology, 40(10), 2699–2709.

Campbell, J. D. M. (2003). Detection and enrichment of antigen-specific CD4+ and CD8+ T cells based on cytokine secretion. Methods , 31(2), 150–159.

Campo, M. S., Graham, S. V., Cortese, M. S., Ashrafi, G. H., Araibi, E. H., Dornan, E. S., … Man, S. (2010). HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells. Virology, 407(1), 137–142.

Cantor, H., & Boyse, E. A. (1975). Functional subclasses of T lymphocytes bearing different Ly antigens. II. Cooperation between subclasses of Ly+ cells in the generation of killer activity. The Journal of Experimental Medicine, 141(6), 1390–1399.

Carreno, B. M., Magrini, V., Becker-Hapak, M., Kaabinejadian, S., Hundal, J., Petti, A. A., … Linette, G. P. (2015). Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. Science, 348(6236), 803–808.

Carswell, E. A., Old, L. J., Kassel, R. L., Green, S., Fiore, N., & Williamson, B. (1975). An endotoxin-induced serum factor that causes necrosis of tumors. Proceedings of the National Academy of Sciences of the United States of America, 72(9), 3666–3670.

Cavenee, W. K., Dryja, T. P., Phillips, R. A., Benedict, W. F., Godbout, R., Gallie, B. L., … White, R. L. (1983). Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature, 305(5937), 779–784.

Celie, P. H. N., Toebes, M., Rodenko, B., Ovaa, H., Perrakis, A., & Schumacher, T. N. M. (2009). UV-induced ligand exchange in MHC class I protein crystals. Journal of the American Chemical Society, 131(34), 12298–12304.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., … Schultz, N. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. Cancer Discovery, 2(5), 401–404.

Chang, M.-H., Chen, C.-J., Lai, M.-S., Hsu, H.-M., Wu, T.-C., Kong, M.-S., … Chen, D.-S. (1997). Universal Hepatitis B Vaccination in Taiwan and the Incidence of Hepatocellular Carcinoma in Children. The New England Journal of Medicine, 336(26), 1855–1859.

Chase, M. W. (1945). The Cellular Transfer of Cutaneous Hypersensitivity to Tuberculin. Experimental Biology and Medicine, 59(2), 134–135.

Chaturvedi, A. K., Anderson, W. F., Lortet-Tieulent, J., Curado, M. P., Ferlay, J., Franceschi, S., Gillison, M. L. (2013). Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 31(36), 4550–4559.

Chowell, D., Krishna, S., Becker, P. D., Cocita, C., Shu, J., Tan, X., … Anderson, K. S. (2015a). TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. Proceedings of the National Academy of Sciences of the United States of America, 112(14),

Chow, L. Q. M., Haddad, R., Gupta, S., Mahipal, A., Mehra, R., Tahara, M., … Seiwert, T. Y. (2016). 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. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology. https://doi.org/10.1200/JCO.2016.68.1478

Clarke, E. (2010). The Problem of Biological Individuality. Biological Theory, 5(4), 312–325.

Cohn, M. (2010). The evolutionary context for a self-nonself discrimination. Cellular and Molecular Life Sciences: CMLS, 67(17), 2851–2862.

Cole, D. K., Bulek, A. M., Dolton, G., Schauenberg, A. J., Szomolay, B., Rittase, W., … Sewell, A. K. (2016). Hotspot autoimmune T cell receptor binding underlies pathogen and insulin peptide cross-reactivity. The Journal of Clinical Investigation, 126(6), 2191–2204.

Coley, W. B. (1891). Contribution to the knowledge of sarcoma. Annals of Surgery, 14, 199–220.

Coley, W. B. (1895). The Erysipelas Toxins—A Reply to the Editorial of Dec. 15, 1894. JAMA: The Journal of the American Medical Association, XXIV(1), 28.

Coley, W. B. (1898a). The treatment of inoperable sarcoma with the 'mixed toxins of erysipelas and bacillus prodigiosus. Journal of the American Medical Association, XXXI(9), 456.

Coughlin, C. M., Vance, B. A., Grupp, S. A., & Vonderheide, R. H. (2004). RNA-transfected CD40-activated B cells induce functional T-cell responses against viral and tumor antigen targets: implications for pediatric immunotherapy. Blood, 103(6), 2046–2054.

Couzin-Frankel, J. (2013). Breakthrough of the year 2013. Cancer immunotherapy. Science, 342(6165), 1432–1433.

Cresswell, P., Turner, M. J., & Strominger, J. L. (1973). Papain-solubilized HL-A antigens from cultured human lymphocytes contain two peptide fragments. Proceedings of the National Academy of Sciences of the United States of America, 70(5), 1603–1607.

Dahlstrom, K. R., Anderson, K. S., & Sturgis, E. M. (2017). Human Papillomavirus–Associated Oropharyngeal Cancer. JAMA Oncology, 3(2), 161.

Dausset, J. (1958). [Iso-leuko-antibodies]. Acta haematologica, 20(1-4), 156–166.

Day, C. L., Kaufmann, D. E., Kiepiela, P., Brown, J. A., Moodley, E. S., Reddy, S., … Walker, B. D. (2006). PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. Nature, 443(7109), 350–354.

Decker, W. K, & Safdar, A. (2009). Bioimmunoadjuvants for the treatment of neoplastic and infectious disease: Coley's legacy revisited. Cytokine & Growth Factor Reviews, 20(4) 271

De Simone, M., De Simone, M., Arrigoni, A., Rossetti, G., Gruarin, P., Ranzani, V., … Pagani, M. (2016a). Transcriptional Landscape of Human Tissue Lymphocytes Unveils Uniqueness of Tumor-Infiltrating T Regulatory Cells. Immunity, 45(5), 1135–1147.

Dighe, A. S., Richards, E., Old, L. J., & Schreiber, R. D. (1994). Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFN gamma receptors. Immunity, 1(6), 447–456.

Dong, H., Strome, S. E., Salomao, D. R., Tamura, H., Hirano, F., Flies, D. B., … Chen, L. (2002). Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. Nature Medicine, 8(8), 793–800.

Dong, H., Zhu, G., Tamada, K., & Chen, L. (1999). B7-H1, a third member of the B7 family, co-stimulates T-cell proliferation and interleukin-10 secretion. Nat Medicine, 5(12), 1365–1369.

Draper, L. M., Kwong, M. L. M., Gros, A., Stevanović, S., Tran, E., Kerkar, S., … Hinrichs, C. S. (2015). Targeting of HPV-16+ Epithelial Cancer Cells by TCR Gene Engineered T Cells Directed against E6. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 21(19), 4431–4439.

D'Souza, G., Kreimer, A. R., Viscidi, R., Pawlita, M., Fakhry, C., Koch, W. M., … Gillison, M. L. (2007). Case–Control Study of Human Papillomavirus and Oropharyngeal Cancer. The New England Journal of Medicine, 356(19), 1944–1956.

Dunne, E. F., Markowitz, L. E., Saraiya, M., Stokley, S., Middleman, A., Unger, E. R., … Centers for Disease Control and Prevention (CDC). (2014). Reducing the burden of HPV-associated cancer and disease. MMWR. Morbidity and Mortality Weekly Report, 63(4), 69–72.

Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J., & Schreiber, R. D. (2002). Cancer immunoediting: from immunosurveillance to tumor escape. Nature Immunology, 3(11), 991–998.

Ebert, J. D. (1970). Self and Not-Self (Cellular Immunology, Book 1). Macfarlane Burnet. Melbourne University Press, Carlton, Victoria, Australia; Cambridge University Press, New York, 1969. viii 320 pp. plates. Cloth, 8.50; paper, 2.95. Science, 167(3925), 1608–1608.

Ellis, J. R., Keating, P. J., Baird, J., Hounsell, E. F., Renouf, D. V., Rowe, M., … Young, L. S. (1995). The association of an HPV16 oncogene variant with HLA-B7 has implications for vaccine design in cervical cancer. Nature Medicine, 1(5), 464–470.

Engel, A. M., Svane, I. M., Mouritsen, S., Rygaard, J., Clausen, J., & Werdelin, O. (1996). Methylcholanthrene-induced sarcomas in nude mice have short induction times and relatively low levels of surface MHC class I expression. APMIS: Acta Pathologica, Microbiologica, et Immunologica Scandinavica, 104(9), 629–639.

Engleman, E. G., Benike, C. J., Grumet, F. C., & Evans, R. L. (1981). Activation of human T lymphocyte subsets: helper and suppressor/cytotoxic T cells recognize and respond to distinct histocompatibility antigens. Journal of Immunology , 127(5), 2124–2129.

Ernst, B., & Anderson, K. S. (2015). Immunotherapy for the Treatment of Breast Cancer. Current Oncology Reports, 17(2). https://doi.org/10.1007/s11912-014-0426-9

Falk, K., Rötzschke, O., Stevanović, S., Jung, G., & Rammensee, H. G. (1991). Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature, 351(6324), 290–296.

Fernando, M. M. A., Stevens, C. R., Walsh, E. C., De Jager, P. L., Goyette, P., Plenge, R. M., … Rioux, J. D. (2008). Defining the role of the MHC in autoimmunity: a review and pooled analysis. PLoS Genetics, 4(4), e1000024.

Ferris, R. L., Blumenschein, G., Jr, Fayette, J., Guigay, J., Colevas, A. D., Licitra, L., … Gillison, M. L. (2016). Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. The New England Journal of Medicine, 375(19), 1856–1867.

Flajnik, M. F. (2014). Re-evaluation of the immunological Big Bang. Current Biology: CB, 24(21),

Fortier, M.-H., Caron, E., Hardy, M.-P., Voisin, G., Lemieux, S., Perreault, C., & Thibault, P. (2008). The MHC class I peptide repertoire is molded by the transcriptome. The Journal of Experimental Medicine, 205(3), 595–610.

Freeman, G. J., Long, A. J., Iwai, Y., Bourque, K., Chernova, T., Nishimura, H., … Honjo, T. (2000). Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. The Journal of Experimental Medicine,

Gangadhar, T. C., Hamid, O., Smith, D. C., Bauer, T. M., Wasser, J. S., Luke, J. J., … Gajewski, T. F. (2015). Preliminary results from a Phase I/II study of epacadostat (incb024360) in combination with pembrolizumab in patients with selected advanced cancers. Journal for ImmunoTherapy of Cancer, 3(Suppl 2), O7.

Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E., & Wiley, D. C. (2010). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature. 1996. 384: 134-141. Journal of Immunology , 185(11), 6394–6401.

Gatti, R. A., & Good, R. A. (1971). Occurrence of malignancy in immunodeficiency diseases: A literature review. Cancer, 28(1), 89–98.

Gillison, M. L., Broutian, T., Pickard, R. K. L., Tong, Z.-Y., Xiao, W., Kahle, L., … Chaturvedi, A. K. (2012). Prevalence of Oral HPV Infection in the United States, 2009-2010. JAMA: The Journal of the American Medical Association, 307(7), 693.

Gillison, M. L., Chaturvedi, A. K., Anderson, W. F., & Fakhry, C. (2015). Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 33(29), 3235–3242.

Giordano, A., & Macaluso, M. (2016). Gynecological Cancers: Genetic and Epigenetic Targets and Drug Development. Springer.

Giuliano, A. R., Palefsky, J. M., Goldstone, S., Moreira, E. D., Penny, M. E., Aranda, C., … Guris, D. (2011). Efficacy of Quadrivalent HPV Vaccine against HPV Infection and Disease in Males. The New England Journal of Medicine, 364(5), 401–411.

Glanville, J., Huang, H., Nau, A., Hatton, O., Wagar, L. E., Rubelt, F., … Davis, M. M. (2017). Identifying specificity groups in the T cell receptor repertoire. Nature, 547(7661), 94–98.

Glick, B., & Sadler, C. R. (1961). The Elimination of the Bursa of Fabricius and Reduction of Antibody Production in Birds from Eggs Dipped in Hormone Solutions. Poultry Science, 40(1), 185–189.

Gold, P. (1965). Specific carcinoembryonic antigens of the human digestive system. The Journal of Experimental Medicine, 122(3), 467–481.

Gold, P., & Freedman, S. O. (1965). Demonstration of tumor-specific antigens in human colonic carcinomata by immunological tolerance and absorption techniques. The Journal of Experimental Medicine, 121, 439–462.

Gorer, P. A. (1936). The detection of a hereditary antigenic difference in the blood of mice by means of human group a serum. Journal of Genetics, 32(1), 17–31.

Gorer, P. A., Lyman, S., & Snell, G. D. (1948). Studies on the Genetic and Antigenic Basis of Tumour Transplantation. Linkage between a Histocompatibility Gene and "Fused" in Mice. Proceedings of the Royal Society B: Biological Sciences, 135(881), 499–505.

Gowans, J. L., Mcgregor, D. D., Cowen, D. M., & Ford, C. E. (1962). Initiation of Immune Responses by Small Lymphocytes. Nature, 196(4855), 651–655.

Grakoui, A., Bromley, S. K., Sumen, C., Davis, M. M., Shaw, A. S., Allen, P. M., & Dustin, M. L. (1999a). The immunological synapse: a molecular machine controlling T cell activation. Science, 285(5425), 221–227.

Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. Science, 185(4154), 862–864.

Grey, H. M. (1973). The small subunit of hl-a antigens is 2-microglobulin. The Journal of Experimental Medicine, 138(6), 1608–1612.

Gupta, P.K., Godec, J., Wolski, D., Adland,E., Yates, K., Pauken, K. E., … Haining, W. N. (2015). CD39 Expression Identifies Terminally Exhausted CD8+ T Cells. PLoS Pathogens 11(10)

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. Cell, 100(1), 57–70.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. Cell, 646–674.

Han, Q., Bagheri, N., Bradshaw, E. M., Hafler, D. A., Lauffenburger, D. A., & Love, J. C. (2011). Polyfunctional responses by human T cells result from sequential release of cytokines. Proceedings of the National Academy of Sciences, 109(5), 1607–1612.

Harndahl, M., Rasmussen, M., Roder, G., Pedersen, I. D., Sørensen, M., Nielsen, M., & Buus, S. (2012). Peptide-MHC class I stability is a better predictor than peptide affinity of CTL immunogenicity. European Journal of Immunology, 42(6), 1405–1416.

Haskins, K. (1983). The major histocompatibility complex-restricted antigen receptor on T cells. I. Isolation with a monoclonal antibody. Journal of Experimental Medicine, 157(4), 1149–1169.

Haverkos, H. W., & Curran, J. W. (1982). The current outbreak of Kaposi's sarcoma and opportunistic infections. CA: A Cancer Journal for Clinicians, 32(6), 330–339.

Hedrick, S. M., Cohen, D. I., Nielsen, E. A., & Davis, M. M. (1984). Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature, 308(5955), 149–153.

Hennecke, J., & Wiley, D. C. (2001a). T cell receptor-MHC interactions up close. Cell, 104(1),1–4.

Herbst, R. S., Soria, J.-C., Kowanetz, M., Fine, G. D., Hamid, O., Gordon, M. S., … Hodi, F. S. (2014). Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature, 515(7528), 563–567.

Hida, N., Maeda, Y., Katagiri, K., Takasu, H., Harada, M., & Itoh, K. (2002). A simple culture protocol to detect peptide-specific cytotoxic T lymphocyte precursors in the circulation. Cancer Immunology, Immunotherapy: CII, 51(4), 219–228.

Hildemann, W. H., & Medawar, P. B. (1959). Relationship between skin transplantation immunity and the formation of humoral isoantibodies in mice. Immunology, 2(1), 44–52.

Himmelweit,F.(1958). Collected papers of paul ehrlich, vol.1. histology, biochemistry and pathology. American Journal of the Medical Sciences, 235

Hirano, N. (2006). Efficient Presentation of Naturally Processed HLA Class I Peptides by Artificial Antigen-Presenting Cells for the Generation of Effective Antitumor Responses. Clinical Cancer Research, 12(10), 2967–2975.

Hogquist, K. A., Jameson, S. C., Heath, W. R., Howard, J. L., Bevan, M. J., & Carbone, F. R. (1994). T cell receptor antagonist peptides induce positive selection. Cell, 76(1), 17–27.

Holsti, L. R. (1995). Development of clinical radiotherapy since 1896. Acta Oncologica , 995–

Honeyman, M. C., Brusic, V., Stone, N. L., & Harrison, L. C. (1998a). Neural network-based prediction of candidate T-cell epitopes. Nature Biotechnology, 16(10), 966–969.

Hoof, I., Peters, B., Sidney, J., Pedersen, L. E., Sette, A., Lund, O., … Nielsen, M. (2009). NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics, 61(1), 1–13.

Huang, J., Brameshuber, M., Zeng, X., Xie, J., Li, Q.-J., Chien, Y.-H., … Davis, M. M. (2013). A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4(+) T cells. Immunity, 39(5), 846–857.

Huang, L., Kuhls, M. C., & Eisenlohr, L. C. (2011). Hydrophobicity as a driver of MHC class I antigen processing. The EMBO Journal, 30(8), 1634–1644.

Huse, M., Lillemeier, B. F., Kuhns, M. S., Chen, D. S., & Davis, M. M. (2006a). T cells use two directionally distinct pathways for cytokine secretion. Nature Immunology, 7(3), 247–255.

Ishida, Y., Agata, Y., Shibahara, K., & Honjo, T. (1992). Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. The EMBO Journal, 11(11), 3887–3895.

Jäger, E., Nagata, Y., Gnjatic, S., Wada, H., Stockert, E., Karbach, J., … Knuth, A. (2000). Monitoring CD8 T cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. Proceedings of the National Academy of Sciences of the United States of America, 97(9), 4760–4765.

Janeway, C. (2005). Immunobiology: The Immune System In Health And Disease. Harcourt Health Sciences.

Janeway, C. A., Jr. (1992). The immune system evolved to discriminate infectious nonself from

noninfectious self. Immunology Today, 13(1), 11–16.

Janeway, C. A., Jr, Goodnow, C. C., & Medzhitov, R. (1996). Danger - pathogen on the premises! Immunological tolerance. Current Biology: CB, 6(5), 519–522.

Jemal, A., Simard, E. P., Dorell, C., Noone, A.-M., Markowitz, L. E., Kohler, B., … Edwards, B. K. (2013). Annual Report to the Nation on the Status of Cancer, 1975–2009, Featuring the Burden and Trends in Human Papillomavirus–Associated Cancers and HPV Vaccination Coverage Levels. JNCI: Journal of the National Cancer Institute, 105(3), 175–201.

Johnston, B. J., & Novales, E. T. (1962). Clinical effect of Coley's toxin. II. A seven-year study. Cancer Chemotherapy Reports. Part 1, 21, 43–68.

Karbach, J., Neumann, A., Brand, K., Wahle, C., Siegel, E., Maeurer, M., … Jäger, E. (2012). Phase I clinical trial of mixed bacterial vaccine (Coley's toxins) in patients with NY-ESO-1 expressing cancers: immunological effects and clinical activity. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 18(19), 5449–5459.

Katz, D. H. (1973). Cell interactions between histoincompatible t and b lymphocytes: ii. failure of physiologic cooperative interactions between t and b lymphocytes from allogeneic donor strains in humoral response to hapten-protein conjugates. The Journal of Experimental Medicine, 137(6), 1405–1418.

Kenter, G. G., Welters, M. J. P., Valentijn, A. R. P. M., Lowik, M. J. G., Berends-van der Meer, D. M. A., Vloon, A. P. G., … Melief, C. J. M. (2009). Vaccination against HPV-16 oncoproteins for vulvar intraepithelial neoplasia. The New England Journal of Medicine, 361(19), 1838–1847.

Khrustalev, V. V., & Barkovsky, E. V. (2011). Percent of highly immunogenic amino acid residues forming B-cell epitopes is higher in homologous proteins encoded by GC-rich genes. Journal of Theoretical Biology, 282(1), 71–79.

Kienle, G. S. (2012). Fever in Cancer Treatment: Coley's Therapy and Epidemiologic Observations. Global Advances in Health and Medicine, 1(1), 92–100.

Kim, H.-J., & Cantor, H. (2014). The Path to Reactivation of Antitumor Immunity and Checkpoint Immunotherapy. Cancer Immunology Research, 2(10), 926–936.

Kim, Y., Sidney, J., Pinilla, C., Sette, A., & Peters, B. (2009). Derivation of an amino acid similarity matrix for peptide: MHC binding and its application as a Bayesian prior. BMC Bioinformatics, 10, 394.

Kim, Y., Yewdell, J. W., Sette, A., & Peters, B. (2013). Positional bias of MHC class I restricted T-cell epitopes in viral antigens is likely due to a bias in conservation. PLoS Computational Biology, 9(1), e1002884.

Koretzky, G. A. (2007). The legacy of the Philadelphia chromosome. The Journal of Clinical Investigation, 117(8), 2030–2032.

Kowalewski, D. J., Schuster, H., Backert, L., Berlin, C., Kahn, S., Kanz, L., … Stickel, J. S. (2015). HLA ligandome analysis identifies the underlying specificities of spontaneous antileukemia immune responses in chronic lymphocytic leukemia (CLL). Proceedings of the National Academy of Sciences of the United States of America, 112(2), E166–75.

Krishna, S., & Anderson, K. S. (2016). T-Cell Epitope Discovery for Therapeutic Cancer Vaccines. Methods in Molecular Biology , 1403, 779–796.

Krummel, M. F., & Allison, J. P. (1995). CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation. The Journal of Experimental Medicine, 182(2), 459–465.

Kubo, R. T., Sette, A., Grey, H. M., Appella, E., Sakaguchi, K., Zhu, N. Z., … Michel, H. (1994). Definition of specific peptide motifs for four major HLA-A alleles. Journal of Immunology , 152(8), 3913–3924.

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology, 157(1), 105–132.

Landsteiner, K., & Chase, M. W. (1942). Experiments on Transfer of Cutaneous Sensitivity to Simple Compounds. Experimental Biology and Medicine, 49(4), 688–690.

Langman, R. E., & Cohn, M. (2000). A minimal model for the self-nonself discrimination: a return to the basics. Seminars in Immunology, 12(3), 189–95; discussion 257–344.

Laydon, D. J., Bangham, C. R. M., & Asquith, B. (2015). Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 370(1675). https://doi.org/10.1098/rstb.2014.0291

Leach, D. R., Krummel, M. F., & Allison, J. P. (1996). Enhancement of antitumor immunity by CTLA-4 blockade. Science, 271(5256), 1734–1736.

Lee, K. M., Chuang, E., Griffin, M., Khattri, R., Hong, D. K., Zhang, W., … Bluestone, J. A. (1998). Molecular basis of T cell inactivation by CTLA-4. Science, 282(5397), 2263–2266.

Leung, A. K. C. (2011). “Variolation” and Vaccination in Late Imperial China, Ca 1570–1911. In History of Vaccine Development (pp. 5–12).

Lin, H. H., Ray, S., Tongchusak, S., Reinherz, E. L., & Brusic, V. (2008). Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research. BMC Immunology, 9, 8.

Linsley, P. S. (1991). Binding of the B cell activation antigen B7 to CD28 costimulates T cell proliferation and interleukin 2 mRNA accumulation. The Journal of Experimental Medicine, 173(3), 721–730.

Little, C. C. (1914). A possible mendelian explanation for a type of inheritance apparently non-mendelian in nature. Science, 40(1042), 904–906.

Little, C. C. (1920). The heredity of susceptibility to a transplantable sarcoma (j. w. b.) of the japanese waltzing mouse. Science, 51(1323), 467–468.

Lund, O., Nielsen, M., Kesmir, C., Petersen, A. G., Lundegaard, C., Worning, P., … Brunak, S. (2004a). Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics, 55(12), 797–810.

Lyford-Pike, S., Peng, S., Young, G. D., Taube, J. M., Westra, W. H., Akpeng, B., … Pai, S. I. (2013). Evidence for a role of the PD-1:PD-L1 pathway in immune resistance of HPV-associated head and neck squamous cell carcinoma. Cancer Research, 73(6), 1733–1741.

Ma, B., Xu, Y., Hung, C.-F., & Wu, T.-C. (2010). HPV and Therapeutic Vaccines: Where are We in 2010? Current Cancer Therapy Reviews, 6(2), 81–103.

Mandal, R., Şenbabaoğlu, Y., Desrichard, A., Havel, J. J., Dalin, M. G., Riaz, N., … Morris, L. G. T. (2016a). The head and neck cancer immune landscape and its immunotherapeutic implications. JCI Insight, 1(17), e89829.

Marur, S., D'Souza, G., Westra, W. H., & Forastiere, A. A. (2010). HPV-associated head and neck cancer: a virus-related cancer epidemic. The Lancet Oncology, 11(8), 781–789.

Mast, E. E., Margolis, H. S., Fiore, A. E., Brink, E. W., Goldstein, S. T., Wang, S. A., … Advisory Committee on Immunization Practices (ACIP). (2005). A comprehensive immunization strategy to eliminate transmission of hepatitis B virus infection in the United States: recommendations of the Advisory Committee on Immunization Practices: immunization of infants, children, and adolescents. MMWR. Recommendations and Reports: Morbidity and Mortality Weekly Report. Recommendations and Reports / Centers for Disease Control, 54(RR-16), 1–31.

Matsushita, H., Vesely, M. D., Koboldt, D. C., Rickert, C. G., Uppaluri, R., Magrini, V. J., … Schreiber, R. D. (2012). Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. Nature, 482(7385), 400–404.

Matzinger, P. (1994). Tolerance, Danger and the Extended Family. Annual Review of Immunology, 12(1), 991–1045.

Matzinger, P. (2002). The Danger Model: A Renewed Sense of Self. Science, 296(5566), 301

McCarthy, E. F. (2006). The toxins of William B. Coley and the treatment of bone and soft-tissue sarcomas. The Iowa Orthopaedic Journal, 26, 154–158.

Medawar, P. B. (1944). The behaviour and fate of skin autografts and skin homografts in rabbits: A report to the War Wounds Committee of the Medical Research Council. Journal of Anatomy, 78(Pt 5), 176–199.

Medzhitov, R., & Janeway, C. A., Jr. (2002). Decoding the patterns of self and nonself by the innate immune system. Science, 296(5566), 298–300.

Mellman, I., Coukos, G., & Dranoff, G. (2011). Cancer immunotherapy comes of age. Nature, 480(7378), 480–489.

Merrill, J. P., Murray, J. E., Harrison, J. H., & Guild, W. R. (1956). Successful homotransplantation of the human kidney between identical twins. Journal of the American Medical Association, 160(4), 277–282.

Metchnikoff, E. (1989). On the Present State of the Question of Immunity in Infectious Diseases. Scandinavian Journal of Immunology, 30(4), 387–398.

Meuer, S. C. (1983). Clonotypic structures involved in antigen-specific human T cell function. Relationship to the T3 molecular complex. The Journal of Experimental Medicine, 157(2), 705–719.

Mildner, A., & Jung, S. (2014). Development and Function of Dendritic Cell Subsets. Immunity, 40(5), 642–656.

Miller, J. (1961). Immunological function of the thymus. The Lancet, 278(7205), 748–749.

Mitchison, N. A. (1955). Studies on the immunological response to foreign tumor transplants in the mouse: ii. the relation between hemagglutinating antibody and graft resistance in the normal mouse and mice pretreated with tissue preparations. The Journal of Experimental Medicine,

102(2), 179–197.

Miyahira, Y., Murata, K., Rodriguez, D., Rodriguez, J. R., Esteban, M., Rodrigues, M. M., & Zavala, F. (1995). Quantification of antigen specific CD8 T cells using an ELISPOT assay. Journal of Immunological Methods, 181(1), 45–54.

Moticka, E. J. (2016a). Hallmarks of the Adaptive Immune Responses. In A Historical Perspective on Evidence-Based Immunology (pp. 9–19).

Moticka, E. J. (2016b). Immune Responses Directed Against Self. In A Historical Perspective on Evidence-Based Immunology (pp. 299–308).

Moticka, E. J. (2016c). The Adaptive Immune Response and Histocompatibility Genes. In A Historical Perspective on Evidence-Based Immunology (pp. 129–139).

Moticka, E. J. (2016d). Tumor Immunology. In A Historical Perspective on Evidence-Based Immunology (pp. 329–339).

Moticka, E. J. (2016e). Two Effector Mechanisms of the Adaptive Immune Response. In A Historical Perspective on Evidence-Based Immunology (pp. 21–29).

Moutaftsi, M., Peters, B., Pasquetto, V., Tscharke, D. C., Sidney, J., Bui, H.-H., … Sette, A. (2006a). A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. Nature Biotechnology, 24(7), 817–819.

Mueller, S. N., & Ahmed, R. (2009). High antigen levels are the cause of T cell exhaustion during chronic viral infection. Proceedings of the National Academy of Sciences of the United States of America, 106(21), 8623–8628.

Murphy, K., & Weaver, C. (2016). Janeway's Immunobiology, 9th edition. Garland Science.

Newell, E. W., & Davis, M. M. (2014a). Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells. Nature Biotechnology, 32(2), 149–157.
Newell, E. W., Sigal, N., Nair, N., Kidd, B. A., Greenberg, H. B., & Davis, M. M. (2013a). Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. Nature Biotechnology, 31(7), 623–629.

Nielsen, M., Lundegaard, C., Blicher, T., Lamberth, K., Harndahl, M., Justesen, S., … Buus, S. (2007a). NetMHCpan, a Method for Quantitative Predictions of Peptide Binding to Any HLA-A and -B Locus Protein of Known Sequence. PloS One, 2(8), e796.

Nielsen, M., Lund, O., Buus, S., & Lundegaard, C. (2010). MHC class II epitope predictive algorithms. Immunology, 130(3), 319–328.

Nishimura, H., Nose, M., Hiai, H., Minato, N., & Honjo, T. (1999). Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. Immunity, 11(2), 141–151.

Nishimura, H., Okazaki, T., Tanaka, Y., Nakatani, K., Hara, M., Matsumori, A., … Honjo, T. (2001). Autoimmune dilated cardiomyopathy in PD-1 receptor-deficient mice. Science, 291(5502), 319–322.

Nowell, P. (1976). The clonal evolution of tumor cell populations. Science, 194(4260), 23–28.

Olthof, N. C., Huebbers, C. U., Kolligs, J., Henfling, M., Ramaekers, F. C. S., Cornet, I., … Speel,

E.-J. M. (2015). Viral load, gene expression and mapping of viral integration sites in HPV16-associated HNSCC cell lines. International Journal of Cancer. Journal International Du Cancer, 136(5), E207–18.

Olthof, N. C., Speel, E.-J. M., Kolligs, J., Haesevoets, A., Henfling, M., Ramaekers, F. C. S., … Huebbers, C. U. (2014). Comprehensive analysis of HPV16 integration in OSCC reveals no significant impact of physical status on viral oncogene and virally disrupted human gene expression. PloS One, 9(2), e88718.

Oren, M., & Levine, A. J. (1983). Molecular cloning of a cDNA specific for the murine p53 cellular tumor antigen. Proceedings of the National Academy of Sciences of the United States of America, 80(1), 56–59.

Ott, P. A., Hu, Z., Keskin, D. B., Shukla, S. A., Sun, J., Bozym, D. J., … Wu, C. J. (2017a). An immunogenic personal neoantigen vaccine for patients with melanoma. Nature, 547(7662), 217.

Oukka, M., Manuguerra, J. C., Livaditis, N., Tourdot, S., Riche, N., Vergnon, I., … Kosmatopoulos, K. (1996). Protection against lethal viral infection by vaccination with nonimmunodominant peptides. Journal of Immunology , 157(7), 3039–3045.

Pancer, Z., & Cooper, M. D. (2006). The evolution of adaptive immunity. Annual Review of Immunology, 24(1), 497–518.

Pardoll, D. M. (2012). Immunology beats cancer: a blueprint for successful translation. Nature Immunology, 13(12), 1129–1132.

Parfenov, M., Pedamallu, C. S., Gehlenborg, N., Freeman, S. S., Danilova, L., Bristow, C. A., … Cancer Genome Atlas Network. (2014a). Characterization of HPV and host genome interactions in primary head and neck cancers. Proceedings of the National Academy of Sciences of the United States of America, 111(43), 15544–15549.

Parikh, F., Duluc, D., Imai, N., Clark, A., Misiukiewicz, K., Bonomi, M., … Sikora, A. G. (2014a). Chemoradiotherapy-induced upregulation of PD-1 antagonizes immunity to HPV-related oropharyngeal cancer. Cancer Research, 74(24), 7205–7216.

Peggs, K. S., Quezada, S. A., Korman, A. J., & Allison, J. P. (2006). Principles and use of anti-CTLA4 antibody in human cancer immunotherapy. Current Opinion in Immunology, 18(2), 206–213.

Penn, I., & Staezl, T. E. (1972). Malignant tumors arising de novo in immunosuppressed organ transplant recipients. Transplantation, 14(4), 407–417.

Petrovic, D., Dempsey, E., Doherty, D. G., Kelleher, D., & Long, A. (2012). Hepatitis C virus--T-cell responses and viral escape mutations. European Journal of Immunology, 42(1), 17–26.

Phan, G. Q., Yang, J. C., Sherry, R. M., Hwu, P., Topalian, S. L., Schwartzentruber, D. J., … Rosenberg, S. A. (2003). Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma. Proceedings of the National Academy of Sciences of the United States of America, 100(14), 8372–8377.

Pollizzi, K. N., & Powell, J. D. (2014). Integrating canonical and metabolic signalling programmes in the regulation of T cell responses. Nature Reviews. Immunology, 14(7), 435–446.

Pradeu, T. (2012). The Limits of the Self: Immunology and Biological Identity. OUP USA.

Pradeu, T., & Carosella, E. D. (2006a). On the definition of a criterion of immunogenicity. Proceedings of the National Academy of Sciences, 103(47), 17858–17861.

Pradeu, T., & Cooper, E. L. (2012). The danger theory: 20 years later. Frontiers in Immunology 3,

Price, D. A., Goulder, P. J., Klenerman, P., Sewell, A. K., Easterbrook, P. J., Troop, M., … Phillips, R. E. (1997). Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proceedings of the National Academy of Sciences of the United States of America, 94(5), 1890–1895.

Pruitt, K. D., Brown, G. R., Hiatt, S. M., Thibaud-Nissen, F., Astashyn, A., Ermolaeva, O., … Ostell, J. M. (2014). RefSeq: an update on mammalian reference sequences. Nucleic Acids Research, 42(Database issue), D756–63.

Pulido, J., Kottke, T., Thompson, J., Galivo, F., Wongthida, P., Diaz, R. M., … Vile, R. (2012). Using virally expressed melanoma cDNA libraries to identify tumor-associated antigens that cure melanoma. Nature Biotechnology, 30(4), 337–343.

Purcell, A. W., McCluskey, J., & Rossjohn, J. (2007a). More than one reason to rethink the use of peptides in vaccine design. Nature Reviews. Drug Discovery, 6(5), 404–414.

Qureshi, O. S., Zheng, Y., Nakamura, K., Attridge, K., Manzotti, C., Schmidt, E. M., … Sansom, D. M. (2011). Trans-endocytosis of CD80 and CD86: a molecular basis for the cell-extrinsic function of CTLA-4. Science, 332(6029), 600–603.

Rajasagi, M., Shukla, S. A., Fritsch, E. F., Keskin, D. B., DeLuca, D., Carmona, E., … Wu, C. J. (2014). Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. Blood, 124(3), 453–462.

Rammensee, H., Bachmann, J., Emmerich, N. P., Bachor, O. A., & Stevanović, S. (1999). SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics, 50(3-4), 213.

Rammensee, H.-G., & Singh-Jasuja, H. (2013). HLA ligandome tumor antigen discovery for personalized vaccine approach. Expert Review of Vaccines, 12(10), 1211–1217.

Ressing, M. E., Sette, A., Brandt, R. M., Ruppert, J., Wentworth, P. A., Hartman, M., … Kast, W. M. (1995a). Human CTL epitopes encoded by human papillomavirus type 16 E6 and E7 identified through in vivo and in vitro immunogenicity studies of HLA-A*0201-binding peptides. Journal of Immunology , 154(11), 5934–5943.

Ridge, J. P., Fuchs, E. J., & Matzinger, P. (1996). Neonatal tolerance revisited: turning on newborn T cells with dendritic cells. Science, 271(5256), 1723–1726.

Riemer, A. B., Keskin, D. B., Zhang, G., Handley, M., Anderson, K. S., Brusic, V., … Reinherz, E. L. (2010a). A conserved E7-derived cytotoxic T lymphocyte epitope expressed on human papillomavirus 16-transformed HLA-A2+ epithelial cancers. The Journal of Biological Chemistry, 285(38), 29608–29622.

Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., … Chan, T. A. (2015a). Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science, 348(6230), 124–128.

Robert, C., Schachter, J., Long, G. V., Arance, A., Grob, J. J., Mortier, L., … KEYNOTE-006 investigators. (2015). Pembrolizumab versus Ipilimumab in Advanced Melanoma. The New England Journal of Medicine, 372(26), 2521–2532.

Robert, C., Thomas, L., Bondarenko, I., O'Day, S., Weber, J., Garbe, C., … Wolchok, J. D. (2011). Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. The New England Journal of Medicine, 364(26), 2517–2526.

Robinson, H. L., & Amara, R. R. (2005). T cell vaccines for microbial infections. Nature Medicine, 11(4s), S25–S32.

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

Rosenberg, S. A., Lotze, M. T., Muul, L. M., Leitman, S., Chang, A. E., Ettinghausen, S. E., … Vetto, J. T. (1985). Observations on the systemic administration of autologous lymphokine-activated killer cells and recombinant interleukin-2 to patients with metastatic cancer. The New England Journal of Medicine, 313(23), 1485–1492.

Rosenberg, S., Spiess, P., & Lafreniere, R. (1986). A new approach to the adoptive immunotherapy of cancer with tumor-infiltrating lymphocytes. Science, 233(4770), 1318.

Rosenblatt, J., Stone, R. M., Uhl, L., Neuberg, D., Joyce, R., Levine, J. D., … Avigan, D. (2016). Individualized vaccination of AML patients in remission is associated with induction of antileukemia immunity and prolonged remissions. Science Translational Medicine, 8(368),

Rose, P. W., Prlić, A., Altunkaya, A., Bi, C., Bradley, A. R., Christie, C. H., … Burley, S. K. (2017). The RCSB protein data bank: integrative view of protein, gene and 3D structural information. Nucleic Acids Research, 45(D1), D271–D281.

Round, J. L., & Mazmanian, S. K. (2009). The gut microbiota shapes intestinal immune responses during health and disease. Nature Reviews. Immunology, 9(8), 600–600.

Rudolf, M. P., Man, S., Melief, C. J., Sette, A., & Kast, W. M. (2001a). Human T-cell responses to HLA-A-restricted high binding affinity peptides of human papillomavirus type 18 proteins E6 and E7. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 7(3 Suppl), 788s–795s.

Rudolph, M. G., Stanfield, R. L., & Wilson, I. A. (2006). How TCRs bind MHCs, peptides, and coreceptors. Annual Review of Immunology, 24, 419–466.

Rusan, M., Li, Y. Y., & Hammerman, P. S. (2015). Genomic landscape of human papillomavirus-associated cancers. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 21(9), 2009–2019.

Sanderson, N. S. R., Puntel, M., Kroeger, K. M., Bondale, N. S., Swerdlow, M., Iranmanesh, N., …Lowenstein, P. R. (2012). Cytotoxic immunological synapses do not restrict the action of interferon- to antigenic target cells. Proceedings of the National Academy of Sciences, 109(20), 7835–7840.

Schreiber, R. D., Old, L. J., & Smyth, M. J. (2011). Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science, 331(6024), 1565–1570.

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

Segal, N. H., Parsons, D. W., Peggs, K. S., Velculescu, V., Kinzler, K. W., Vogelstein, B., & Allison, J. P. (2008). Epitope landscape in breast and colorectal cancer. Cancer Research

Şenbabaoğlu, Y., Gejman, R. S., Winer, A. G., Liu, M., Van Allen, E. M., de Velasco, G., … Hakimi, A. A. (2016a). Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. Genome Biology, 17(1), 231.

Seong, S.-Y., & Matzinger, P. (2004a). Hydrophobicity: an ancient damage-associated molecular pattern that initiates innate immune responses. Nature Reviews. Immunology, 4(6), 468

Sette, A., & Sidney, J. (1999). Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics, 50(3-4), 201–212.

Shao, W., Pedrioli, P. G. A., Wolski, W., Scurtescu, C., Schmid, E., Vizcaíno, J. A., … Caron, E. (2017). The SysteMHC Atlas project. Nucleic Acids Research

Sheridan, C. (2015). IDO inhibitors move center stage in immuno-oncology. Nature Biotechnology, 33(4), 321–322.

Shevach, E. M. (1972). Histocompatibility-linked immune response gene function in guinea pigs: specific inhibition of antigen-induced lymphocyte proliferation by alloantisera. The Journal of Experimental Medicine, 136(5), 1207–1221.

Shinkai, Y., Rathbun, G., Lam, K. P., Oltz, E. M., Stewart, V., Mendelsohn, M., … Stall, A. M. (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell, 68(5), 855–867.

Slamon, D., Godolphin, W., Jones, L., Holt, J., Wong, S., Keith, D., … Et, A. (1989). Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science, 244(4905), 707–712.

Smith, J. M., & Szathmary, E. (1997). The Major Transitions in Evolution. Oxford University Press.

Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J. M., Desrichard, A., … Chan, T. A. (2014a). Genetic basis for clinical response to CTLA-4 blockade in melanoma. The New England Journal of Medicine, 371(23), 2189–2199.

Spranger, S., Spaapen, R. M., Zha, Y., Williams, J., Meng, Y., Ha, T. T., & Gajewski, T. F. (2013). Up-Regulation of PD-L1, IDO, and Tregs in the Melanoma Tumor Microenvironment Is Driven by CD8 T Cells. Science Translational Medicine, 5(200), 200ra116–200ra116.

Stadinski, B. D., Obst, R., & Huseby, E. S. (2016). A “hotspot” for autoimmune T cells in type 1 diabetes. The Journal of Clinical Investigation, 126(6), 2040–2042.

Stadinski, B. D., Shekhar, K., Gómez-Touriño, I., Jung, J., Sasaki, K., Sewell, A. K., … Huseby, E. S. (2016). Hydrophobic CDR3 residues promote the development of self-reactive T cells. Nature Immunology, 17(8), 946–955.

Starr, T. K., Jameson, S. C., & Hogquist, K. A. (2003). Positive and negative selection of T cells. Annual Review of Immunology, 21, 139–176.

Stewart, A. J., & Devlin, P. M. (2006). The history of the smallpox vaccine. The Journal of Infection, 52(5), 329–334.

Strønen, E., Toebes, M., Kelderman, S., van Buuren, M. M., Yang, W., van Rooij, N., …

Schumacher, T. N. (2016). Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. Science, 352(6291), 1337–1341.

Sutmuller, R. P., van Duivenvoorde, L. M., van Elsas, A., Schumacher, T. N., Wildenberg, M. E., Allison, J. P., … Melief, C. J. (2001). Synergism of cytotoxic T lymphocyte-associated antigen 4 blockade and depletion of CD25(+) regulatory T cells in antitumor therapy reveals alternative pathways for suppression of autoreactive cytotoxic T lymphocyte responses. The Journal of Experimental Medicine, 194(6), 823–832.

Sykulev, Y., Joo, M., Vturina, I., Tsomides, T. J., & Eisen, H. N. (1996). Evidence that a Single Peptide–MHC Complex on a Target Cell Can Elicit a Cytolytic T Cell Response. Immunity, 4(6), 565–571.

Tan, C. T., Croft, N. P., Dudek, N. L., Williamson, N. A., & Purcell, A. W. (2011). Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring. Proteomics, 11(11), 2336–2340.

Tenzer, S., Peters, B., Bulik, S., Schoor, O., Lemmel, C., Schatz, M. M., … Holzhütter, H.-G. (2005a). Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. Cellular and Molecular Life Sciences: CMLS, 62(9), 1025–1037.

The Cancer Genome Atlas Network. (2015). Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature, 517(7536), 576–582.

Tivol, E. A., Borriello, F., Schweitzer, A. N., Lynch, W. P., Bluestone, J. A., & Sharpe, A. H. (1995). Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. Immunity, 3(5), 541–547.

Tonegawa, S. (1983). Somatic generation of antibody diversity. Nature, 302(5909), 575–581.

Topalian, S. L., Hodi, F. S., Brahmer, J. R., Gettinger, S. N., Smith, D. C., McDermott, D. F., … Sznol, M. (2012). Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. The New England Journal of Medicine, 366(26), 2443–2454.

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. CA: A Cancer Journal for Clinicians, 65(2), 87–108.

Tran, E., Robbins, P. F., Lu, Y.-C., Prickett, T. D., Gartner, J. J., Jia, L., … Rosenberg, S. A. (2016a). T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer. The New England Journal of Medicine, 375(23), 2255–2262.

Trimble, C. L., & Frazer, I. H. (2009). Development of therapeutic HPV vaccines. The Lancet Oncology, 10(10), 975–980.

Trimble, C. L., Morrow, M. P., Kraynyak, K. A., Shen, X., Dallas, M., Yan, J., … Bagarazzi, M. L. (2015). Safety, efficacy, and immunogenicity of VGX-3100, a therapeutic synthetic DNA vaccine targeting human papillomavirus 16 and 18 E6 and E7 proteins for cervical intraepithelial neoplasia 2/3: a randomised, double-blind, placebo-controlled phase 2b trial. The Lancet, 386(10008), 2078–2088.

Tsung, K., & Norton, J. A. (2006). Lessons from Coley's Toxin. Surgical Oncology, 15(1), 25–28.

Tumeh, P. C., Harview, C. L., Yearley, J. H., Shintaku, I. P., Taylor, E. J. M., Robert, L., … Ribas, A. (2014). PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature,

515(7528), 568–571.

UniProt Consortium. (2015). UniProt: a hub for protein information. Nucleic Acids Research, 43(Database issue), D204–12.

van den Broek, M. E. (1996). Decreased tumor surveillance in perforin-deficient mice. The Journal of Experimental Medicine, 184(5), 1781–1790.

van der Merwe, P. A., & Dushek, O. (2010). Mechanisms for T cell receptor triggering. Nature Reviews. Immunology, 11(1), 47–55.

van der Most, R. G., Murali-Krishna, K., Whitton, J. L., Oseroff, C., Alexander, J., Southwood, S., … Ahmed, R. (1998). Identification of Db- and Kb-restricted subdominant cytotoxic T-cell responses in lymphocytic choriomeningitis virus-infected mice. Virology, 240(1), 158–167.

van Rooij, N., van Buuren, M. M., Philips, D., Velds, A., Toebes, M., Heemskerk, B., … Schumacher, T. N. (2013). Tumor Exome Analysis Reveals Neoantigen-Specific T-Cell Reactivity in an Ipilimumab-Responsive Melanoma. Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology, 31(32), e439–e442.

Verdegaal, E. M. E., de Miranda, N. F. C. C., Visser, M., Harryvan, T., van Buuren, M. M., Andersen, R. S., … van der Burg, S. H. (2016). Neoantigen landscape dynamics during human melanoma–T cell interactions. Nature, 536(7614), 91–95.

Vita, R., Overton, J. A., Greenbaum, J. A., Ponomarenko, J., Clark, J. D., Cantrell, J. R., … Peters, B. (2014). The immune epitope database (IEDB) 3.0. Nucleic Acids Research,

von Bergwelt-Baildon, M. S., Vonderheide, R. H., Maecker, B., Hirano, N., Anderson, K. S., Butler, M. O., … Schultze, J. L. (2002). Human primary and memory cytotoxic T lymphocyte responses are efficiently induced by means of CD40-activated B cells as antigen-presenting cells: potential for clinical application. Blood, 99(9), 3319–3325.

Vonderheide, R. H., Anderson, K. S., Hahn, W. C., Butler, M. O., Schultze, J. L., & Nadler, L. M. (2001). Characterization of HLA-A3-restricted cytotoxic T lymphocytes reactive against the widely expressed tumor antigen telomerase. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 7(11), 3343–3348.

Vonderheide, R. H., Schultze, J. L., Anderson, K. S., Maecker, B., Butler, M. O., Xia, Z., … Nadler, L. M. (2001). Equivalent induction of telomerase-specific cytotoxic T lymphocytes from tumor-bearing patients and healthy individuals. Cancer Research, 61(23), 8366–8370.

Voskoboinik, I., Whisstock, J. C., & Trapani, J. A. (2015). Perforin and granzymes: function, dysfunction and human pathology. Nature Reviews. Immunology, 15(6), 388–400.

Walunas, T. L., Lenschow, D. J., Bakker, C. Y., Linsley, P. S., Freeman, G. J., Green, J. M., … Bluestone, J. A. (1994). CTLA-4 can function as a negative regulator of T cell activation. Immunity, 1(5), 405–413.

Waterhouse, P., Penninger, J. M., Timms, E., Wakeham, A., Shahinian, A., Lee, K. P., … Mak, T. W. (1995). Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. Science, 270(5238), 985–988.

Weiskopf, D., Yauch, L. E., Angelo, M. A., John, D. V., Greenbaum, J. A., Sidney, J., … Sette, A. (2011). Insights into HLA-Restricted T Cell Responses in a Novel Mouse Model of Dengue Virus Infection Point toward New Implications for Vaccine Design. The Journal of Immunology, 187(8),

4268–4279.

Wherry, E. J., Ha, S.-J., Kaech, S. M., Haining, W. N., Sarkar, S., Kalia, V.… Ahmed, R. (2007). Molecular signature of CD8+ T cell exhaustion during chronic viral infection. Immunity, 27(4),

Wherry, E. J., John Wherry, E., & Kurachi, M. (2015). Molecular and cellular insights into T cell exhaustion. Nature Reviews. Immunology, 15(8), 486–499.

Who Nomenclature Subcommittee, W. N. S. (1984). Nomenclature for clusters of differentiation (CD) of antigens defined on human leukocyte populations (Vol. 62, pp. 809–815). Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6334575

Williams, W. W., Lu, P.-J., O'Halloran, A., Kim, D. K., Grohskopf, L. A., Pilishvili, T., … Fiebelkorn, A. P. (2017). Surveillance of Vaccination Coverage among Adult Populations — United States, 2015. Morbidity and Mortality Weekly Report. Surveillance Summaries , 66(11), 1–28.

Woodman, C. B. J., Collins, S. I., & Young, L. S. (2007). The natural history of cervical HPV infection: unresolved issues. Nature Reviews. Cancer, 7(1), 11–22.
Wucherpfennig, K. W., Gagnon, E., Call, M. J., Huseby, E. S., & Call, M. E. (2010). Structural biology of the T-cell receptor: insights into receptor assembly, ligand recognition, and initiation of signaling. Cold Spring Harbor Perspectives in Biology, 2(4), a005140.

Xie, J., Tato, C. M., & Davis, M. M. (2012). How the immune system talks to itself: the varied role of synapses. Immunological Reviews, 251(1), 65–79.

Yadav, M., Jhunjhunwala, S., Phung, Q. T., Lupardus, P., Tanguay, J., Bumbaca, S., … Delamarre, L. (2014a). Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. Nature, 515(7528), 572–576.
Yanagi, Y., Yoshikai, Y., Leggett, K., Clark, S. P., Aleksander, I., & Mak, T. W. (1984). A human T cell-specific cDNA clone encodes a protein having extensive homology to immunoglobulin chains. Nature, 308(5955), 145–149.

Yron, I., Wood, T. A., Jr, Spiess, P. J., & Rosenberg, S. A. (1980). In vitro growth of murine T cells. V. The isolation and growth of lymphoid cells infiltrating syngeneic solid tumors. Journal of Immunology , 125(1), 238–245.

Yu, W., Jiang, N., Ebert, P. J. R., Kidd, B. A., Müller, S., Lund, P. J., … Davis, M. M. (2015a). Clonal Deletion Prunes but Does Not Eliminate Self-Specific αβ CD8 T Lymphocytes. Immunity, 42(5), 929–941.

Zajac, A. J., Blattman, J. N., Murali-Krishna, K., Sourdive, D. J., Suresh, M., Altman, J. D., & Ahmed, R. (1998). Viral immune evasion due to persistence of activated T cells without effector function. The Journal of Experimental Medicine, 188(12), 2205–2213.

Zaritskaya, L., Shurin, M. R., Sayers, T. J., & Malyguine, A. M. (2010). New flow cytometric assays for monitoring cell-mediated cytotoxicity. Expert Review of Vaccines, 9(6), 601–616.
Zhang, Y., Koneva, L. A., Virani, S., Arthur, A. E., Virani, A., Hall, P. B., … Sartor, M. A. (2016a). Subtypes of HPV-Positive Head and Neck Cancers Are Associated with HPV Characteristics, Copy Number Alterations, PIK3CA Mutation, and Pathway Signatures. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research, 22(18), 4735–4745.

Zimmerman, J. M., Eliezer, N., & Simha, R. (1968). The characterization of amino acid sequences in proteins by statistical methods. Journal of Theoretical Biology, 21(2), 170–201.

Zinkernagel, R. M. (1975). H-2 compatability requirement for T-cell-mediated lysis of target cells infected with lymphocytic choriomeningitis virus. Different cytotoxic T- cell specificities are associated with structures coded for in H-2K or H-2D; The Journal of Experimental Medicine, 141(6), 1427–1436.

Zinkernagel, R. M., & Doherty, P. C. (1974). Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytes choriomeningitis. Nature, 251(5475), 547–548.

zur Hausen, H., Gissmann, L., Steiner, W., Dippold, W., & Dreger, I. (1975). Human papilloma viruses and cancer. Bibliotheca Haematologica, (43), 569–571.

Blum, J. S., Wearsch, P. A., & Cresswell, P. (2013a). Pathways of Antigen Processing. *Annual Review of Immunology*, *31*(1), 443–473.

APPENDIX A

TCR CONTACT RESIDUE HYDROPHOBICITY IS A HALLMARK OF IMMUNOGENIC CD8+ T CELL EPITOPES

A.1 Construction of datasets

All MHC class I peptides used in this study for analyses and the design of ANN-Hydro prediction model were retrieved from IEDB (www.iedb.org, last accessed: 08/11/2014). The IEDB is the largest curated dataset of MHC-I peptides identified from different primary research studies from over 334 different source organisms. We set the “Immune recognition context” as T cell response and selected “MHC class I” as the criteria for data retrieval. In total, there were 28,444 T cell epitopes reported to be immunogenic by T cell assays, including self and pathogenic epitopes and 6,142 peptides were reported to be positive by ligand elution analysis (either mass spectrometry or HPLC). To avoid redundancy and overrepresentation bias, we excluded all duplicate peptides, so that each peptide is present only once in the dataset. Positive CTL epitopes represent the immunogenic epitope group. Ligand eluted MHC-I self-peptides are generally eluted from cell surface and therefore they have been antigenically processed and MHC-bound. A vast majority of eluted self-peptides are derived from endogenous proteins. To completely separate immunogenic and non-immunogenic datasets, any immunogenic eluted self-peptide associated with autoimmunity or cancer was excluded. The remaining peptides were used as the non-immunogenic peptide dataset for our analyses. Additionally, we removed any pathogen derived non-self- eluted peptides from the eluted peptide dataset to generate mutually exclusive datasets. These unique peptides were annotated for antigen name, peptide starting position, peptide ending position, and MHC restriction, which were required for inclusion. Peptides with “undetermined class I alleles” were also excluded. These filtering criteria resulted in a final dataset of 5,035 8-11mer immunogenic and 4,853 8-11mer non-immunogenic peptides.

A.2 Position-based hydrophobicity analysis

We transformed our datasets of immunogenic and non-immunogenic peptides into numeric arrays using the R statistical software. Separate numeric arrays were generated for immunogenic and non-immunogenic 8, 9 and 10mers. Mean hydrophobicity of immunogenic and non-

immunogenic peptides at each position was calculated and were compared residue-by-residue through Wilcoxon rank-sum tests to quantify statistical significance.

A.3 Rate analysis of predicted peptides

An efficient prediction algorithm identifies consistently all possible CTL epitopes from a given protein in the fewest number of "hits" consistently. For each test protein, we created a subset with unique CTL epitopes retrieved either from the IEDB database. Each predicted peptide starting from rank one was queried for an exact match in the dataset of CTL epitopes. When there was an exact match, a positive hit was recorded. Graphical representations comparing the rate of predictions by the IEDB-consensus binding prediction algorithm and hydrophobicity-based predictions were generated (Fig 2-5).

A.4 Hydrophobicity-based ANN prediction model (ANN-Hydro)

The R neuralnet package was used to design and train the two ANN-Hydro models on H-2D^b and HLA-A2 restricted 9mer peptides known to be immunogenic (n=204 and n=374, respectively) or non-immunogenic (n=232 and n=201, respectively)**.** Each peptide sequence in the respective H-2D^b and HLA-A2 datasets were transformed into a corresponding numeric sequence based on the hydrophobicity value of amino acids. Training peptides were derived from IEDB and SYFPEITHI's epitope database. A three-layer fully connected feed-forward ANN was comprised by nine input neurons, one hidden layer with three neurons, and one output variable (Fig 2-6).

Our ANN-Hydro prediction model is given by the following mathematical framework:

$$y(H) = f\left(w_0 + \sum_{i=1}^{3} w_i \cdot f(w_{0i} + W_i^T H) \right),$$

Where w_0 denotes the intercept of the output neuron and w_{0i} the intercept of the i^{th} hidden neuron. Additionally, w_i denotes the synaptic weight corresponding to the synapse starting at the i^{th} hidden neuron and leading to the output neuron. $W_i = (w_{1i}, w_{2i}, ..., w_{9i})$ is the vector of all synaptic weights corresponding to the synapse leading to the i^{th} hidden neuron, and

$H = (h(R_1), h(R_2), ..., h(R_9))$ the vector of all inputs, which corresponds to the numeric hydrophobicity representation of a 9mer peptide, where $h(R_i)$ is the hydrophobicity value of the amino acid . Finally, the output variable $y(H)$ denotes the probability of a peptide being immunogenic (p-ANN-Hydro). Since the starting values for the weights are drawn from the standard normal distribution, the outputs were averaged over 60 realizations. The activation function $f(v)$ was chosen to be the sigmoid function $f(v) = 1/(1+e^{-v})$, and the sum of squared errors was used for the error function. The learning procedure was the resilient back-propagation with learning rate set to 0.01; a threshold set to 0.01 was defined for the partial derivatives of the error function.

A.4 Application of ANN-Hydro

For each H-2D^b and HLA-A2 restricted epitope prediction, we used the MHC-binding prediction tool IEDB-consensus to generate a list of epitope predictions on which the immunogenicity model could be applied. We normalized prediction binding scores (percentile rank) using the expression $S_{B_i} = (\delta_i - \delta_{\min})/(\delta_{\max} - \delta_{\min})$ where S_{B_i} represents the normalized score of a given peptide; δ_i, the assigned output score by IEDB-consensus; $\delta_{\min}$, the minimum score assigned in prediction output by IEDB; and $\delta_{\max}$, the maximum score assigned in the entire prediction output by IEDB. To remove poor binding peptides from the list, a subset of predicted peptides was selected by defining a S_B -threshold of 0.2 for antigen length <= 100 aa's and a S_B -threshold of 0.1 (10th percentile of predicted binders) for antigen length >100 aa's. Independently, probabilities of immunogenicity were obtained by applying the ANN-Hydro model to this subset of binding predictions. Normalized scores (S_I) were then assigned based on these probabilities of immunogenicity. Within the spectrum of predicted binders, we prioritized epitope re-ranking based on both S_B and S_I scores with first priority given to high-immunogenicity high-binders (probability of immunogenicity >= 0.4 and S_B <=0.05; region I in Fig. S3), followed by modest-immunogenicity high-binders (probability of immunogenicity < 0.4 and S_B <=0.05; region II in Fig. S3), then high-

immunogenicity modest-binders (probability of immunogenicity >= 0.4 and S_B > 0.05; region III in Fig. S3), and modest-immunogenicity modest-binders (probability of immunogenicity < 0.4 and S_B > 0.05; region IV in Fig. S3). For the antigens with length <= 100 aa's, the S_B cutoff for the four regions was set to 0.1 and probability of immunogenicity threshold remained at 0.4. Predicted peptides in each section were re-ranked based on a total score defined as $S = S_B S_I$. Final ranked list was obtained by sequential appending of the re-ranked peptides from each region. The list of predicted peptides was ranked based on this total score ranging from lowest score to the highest score. The lower the total score of a predicted peptide, the higher its probability of being an immunogenic epitope. Workflow of the prediction strategy is shown in Fig. 2-6.

A.5 Statistical analysis of predicted CTL epitopes

We used the F-test to quantify statistical significance (P < 0.05) of the variation of predicted rankings of T cell epitopes across different antigens between ANN-Hydro together with IEDB-consensus and IEDB-consensus alone.

A.6 *In vivo* discovery of HIV-1 Gag epitopes

Mice. C57BL/6 mice were obtained from Harlan Laboratories. All mice used were between 6 and 8 weeks of age. All animal study protocols were conducted in accordance with guidelines approved by the Institutional Animal Care and Use Committee at Kings College London and in full compliance with UK Home Office regulations under a project license to L.S.K.

Vaccine immunization. Codon optimized HIV-1 gag plasmid DNA ZM96 from strain 96ZM651.8 (provided by B Hahn, through the Centre for AIDS Reagents [CFAR] UK) and codon optimized HIV-1 gag Consensus B plasmid DNA (provided by D Garber, Emory University, USA) were used to construct and propagate replication defective (E1, E3 deleted) recombinant Adenovirus type 5 (rAdHu5) vectors as described previously for the HIV-1 gag strain 97CN54(4). Animals were

immunized with 10^9 virus particles (vp) as determined by the DNA Pico-Green assay (Invitrogen) and administered either i.m. in the quadricep muscle (rAdHu5 Consensus B gag) or i.d. at the base of the tail (rAdHu5 ZM96 and rAdHu5 CN54).

Peptides. 15mer peptides with an 11 amino acid overlap spanning the HIV-1 CN54 Gag protein and a 20mer set of peptides with 10 amino acid overlap spanning HIV-1 ZM96 were provided by CFAR, a set of 15mers with an 11 amino acid overlap spanning the HIV-1 Consensus subtype B Gag protein were provided from the NIH AIDS Reagent and Reference Program. 'Optimal' 9mer or 11mer peptides from HIV-1 CN54 Gag, ZM96 Gag and HIV-1 Consensus B were purchased from Proimmune.

T cell epitope mapping by intracellular interferon gamma staining. Spleens were harvested 14 days after immunization, homogenized to single-cell suspensions, and RBCs were lysed using ACK lysis buffer (Lonza). Splenocytes were then used for *in vitro* re-stimulation, where 10^6 cells were incubated for 6 h at 37°C with anti-CD28 (2mg/ml; BD Pharmingen), either alone (unstimulated control) or with peptides, either in pools or individually (each at 1mg peptide/ml), derived from Consensus B Gag. Brefeldin A (10mg/ml, Sigma-Aldrich) was added for the last 5 h of culture. After washing, cells were stained with anti-CD8 (clone 37.51, BD Biosciences) for 20 min, then fixed and permeabilized with the BD Cytofiix/Cytoperm Kit according to the manufacturers' instructions, and then stained 30 min with anti-IFN-g (clone XMG1.2, eBiosciences), washed and analyzed by flow cytometry. Consensus B epitopes were deconvoluted to individual 15mers from peptide pools, where each peptide is present in two independent pools within the matrix and reactive peptides confirmed in the second round against the 15mer peptide. Finally, based on the sequence of the reactive 15mer peptide, truncated versions of the 15mer peptides were synthesized and tested.

T cell epitope mapping by ELISPOT assay. 14 days after immunization, splenocytes prepared (as detailed above) were re-stimulated *in vitro* with media alone, or with peptides, either in pools

or individually (each at 1mM final concentration) derived from CN54 or ZM96 Gag on mouse anti-INF-g antibody coated 96 well plates (U-Cytech) and incubated for 16 h at 37°C, 5% $CO_{2.}$ IFN-g production was revealed according to the manufacturer's instructions and IFN-g spot forming cells (SFCs) enumerated using an immunospot image analyser (Bioreader 5000). In the first round, CN54 Gag epitopes were deconvoluted to individual 15mers from peptide pools, where each peptide is present in two independent pools within the matrix and reactive peptides confirmed in the second round against the 15mer peptide. Finally, based on the sequence of the reactive 15mer peptide, 9mer peptides were synthesized and tested. For ZM96 (due to the absence of a complete set of overlapping 15mer peptides), 49 individual 20mer peptides were tested. The reactive peptide sequences were confirmed against the corresponding 15mer peptide (data not shown) to the reactive sequence and then 9mer peptides synthesized and tested.

Table A-1. Amino acid property scales used for analyses. Hydrophobicity scale (Kyte-Doolittle) (5), Polarity (Grantham) (6), and Bulkiness (Zimmerman).

Amino acid	Hydrophobicity	Bulkiness	Polarity
Alanine (A)	1.8	11.5	8
Cysteine (C)	2.5	13.46	5.5
Aspartic acid (D)	-3.5	11.68	13
Glutamic acid (E)	-3.5	13.57	12.3
Phenylalanine (F)	2.8	19.8	5.2
Glycine (G)	-0.4	3.4	9
Histidine (H)	-3.2	13.69	10.4
Isoleucine (I)	4.5	21.4	5.2
Lysine (K)	-3.9	15.71	11.3
Leucine (L)	3.8	21.4	4.9
Methionine (M)	1.9	16.25	5.7
Asparagine (N)	-3.5	12.82	11.6
Proline (P)	-1.6	17.43	8
Glutamine (Q)	-3.5	14.45	10.5
Arginine(R)	-4.5	14.28	10.5
Serine (S)	-0.8	9.47	9.2
Threonine (T)	-0.7	15.77	8.6
Valine (V)	4.2	21.57	5.9
Tryptophan (W)	-0.9	21.67	5.4
Tyrosine (Y)	-1.3	18.03	6.2

Table A-2. Hydrophobicity comparison between immunogenic and non-immunogenic MHC class I peptides. *P*-values were calculated using Wilcoxon sum-ranked test using different datasets as described in main text.

Dataset	Sample size		Residue position in peptide									
	Immunogenic	Non Immunogenic	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
All MHC-I 9mers	2841	2660	6.6×10^{-15}	$< 2.2\times10^{-16}$	1.7×10^{-8}	$< 2.2\times10^{-16}$	4.6×10^{-5}	1.6×10^{-5}	$< 2.2\times10^{-16}$	$< 2.2\times10^{-1}$	$< 2.2\times10^{-16}$	-
All MHC-I 10mers	1007	1013	1.4×10^{-13}	$< 2.2\times10^{-16}$	1.7×10^{-9}	$< 2.2\times10^{-16}$	1.3×10^{-13}	1.7×10^{-15}	2×10^{-7}	3.3×10^{-12}	1.2×10^{-15}	$< 2.2\times10^{-16}$
NonA2 HLA-I 9mers	1318	1897	1.4×10^{-6}	7.8×10^{-3}	7.6×10^{-3}	$< 2.2\times10^{-16}$	0.1561	3×10^{-4}	2.6×10^{-13}	$< 2.2\times10^{-1}$	3.5×10^{-14}	-
NonA2 HLA-I 10mers	518	859	3.4×10^{-7}	1.4×10^{-3}	1.4×10^{-3}	4.6×10^{-13}	8.8×10^{-3}	1.4×10^{-3}	0.0023	1.5×10^{-3}	2.9×10^{-3}	0.2293
Self-peptides 9mers	565	2660	2.1×10^{-4}	$< 2.2\times10^{-16}$	1×10^{-4}	7.8×10^{-5}	0.292	0.1423	1.2×10^{-11}	2.2×10^{-7}	$< 2.2\times10^{-16}$	-
MHC-I 9mers (Immunogen=Organism)	323	2660	0.06781	6.5×10^{-9}	0.0071	1×10^{-5}	0.6156	1.3×10^{-4}	1.5×10^{-3}	3.2×10^{-5}	$< 2.2\times10^{-16}$	-
HLA A2 9mers	936	321	0.6384	0.8929	0.8487	6.3×10^{-12}	0.2917	3.1×10^{-7}	5×10^{-13}	$< 2.2\times10^{-16}$	0.08106	-
HLA A2 10mers	288	77	0.4029	0.2247	0.5054	0.00094	0.00087	1×10^{-5}	0.3368	0.9912	5.5×10^{-6}	0.09147
MHC H-2Kb 8mers	306	315	0.06678	0.9329	0.0055	0.2288	0.6742	7×10^{-5}	1.1×10^{-6}	0.1513	-	-
MHC H-2Db 9mers	194	228	0.1825	0.0087	2.1×10^{-4}	0.3298	4.9×10^{-10}	0.7829	1.1×10^{-4}	0.0013	0.1273	-
	Immunogenic pathogenic	Immunogenic self										
HLA A2 9mers	1508	244	0.389	0.385	0.877	0.97	0.13	0.04	0.776	0.35	0.63	

MHC motif-specific anchor residues

Known TCR contact residues

Table A-3. HLA-A2 restricted CTL epitopes for dengue virus 1 polyprotein and cytomegalovirus pp65 used in the rate analysis of predicted epitopes as shown in (Fig.2-5). All epitopes were retrieved from IEDB Vita et al, 2010.

Antigen	Epitope	Epitope length
CMV-pp65	NLVPMVATV	9
	MLNIPSINV	9
	VLGPISGHV	9
	RLLQTGIHV	9
	LMNGQQIFL	9
	ILARNLVPM	9
	SLILVSQYT	9
	SIYVYALPL	9
	VIGDQYVKV	9
	YLESFCEDV	9
	AMAGASTSA	9
	KYQEFFWDA	9
	GLSISGNLL	9
	RQYDPVAAL	9
	VAALFFFDI	9
	ALFFFDIDL	9
	KISHIMLDVA	10
	SDNEIHNPAV	10
	FTWPPWQAGI	10
	LLCPKSIPGL	10

Antigen	Epitope	Epitope length
Dengue-Polyprotein	VLMLVAHYA	9
	ILLMRTTWA	9
	MLLALIAVL	9
	TLYAVATTI	9
	QEGAMHTAL	9
	LPAIVREAI	9
	SRNSTHEMY	9
	AIVREAIKR	9
	YLPAIVREA	9
	TLLCLIPTV	9
	VLNPYMPSV	9
	LMMMLPATL	9
	VTYECPLLV	9
	MMMLPATLA	9
	IILEFFLMV	9
	KTDFGFYQV	9
	VQADMGCVV	9
	GLLFMILTV	9
	QLWAALLSL	9
	LLMRTTWAL	9
	CLMMMLPATL	10
	ELMRRGDLPV	10
	MLLILCVTQV	10
	FLMVLLIPEP	10
	TLMLLALIAV	10
	LMLLALIAVL	10
	IILEFFLMVL	10
	TLTAAVLLLV	10
	VLLLVTHYAI	10
	ITLLCLIPTV	10
	KVLNPYMPSV	10
	HQLWATLLSL	10
	YTPEGIIPTL	10
	SIILEFFLMV	10
	LSMGLITIAV	10
	NQLIYVILTI	10
	LMMMLPATLA	10
	TLMAMDLGEL	10
	FTMGVLCLAI	10

Table A-4: Probabilities assigned by the ANN-Hydro A2-model for HLA-A2 restricted 9mer CTL epitopes. Three recent epitope discovery studies Assarsson et al. (2008); Weiskopf et al. (2011); Newell et al. (2013) that were based on a proteome-wide screen of various viral antigens and self-epitopes were chosen for assessment of the predictive capacity of the 'A2-model'. Neoepitopes were obtained from rotavirus Newell et al. (2013) dengue virus Weiskopf et al. (2011) and influenza A Assarsson et al. (2008) and other positive control epitopes from several antigens (pathogenic and self) were obtained from Newell et al. Newell et al. (2013). Any epitope that was present in the training set for ANN-Hydro was removed. A cutoff probability (p-ANN-Hydro) of 0.4 was set for a positive "hit".

Source	Epitope	Antigen	p-ANN-Hydro	Reference
Neo-epitopes				
Rotavirus	SLISGMWLL	Rota-VP2_4	0.89	Newell *et* al.
	TLLANVTAV	Rota-VP6_4	0.87	
	FLDSEPHLL	Rota-NSP1_2	0.84	
	LLNYILKSV	Rota-VP7_1	0.37	
Influenza-A (FluA)	QIAILVTTV	NA	0.90	Assarsson *et* al.
	GLIYNRMGA	M1	0.89	
	GILGFVFTL	Flu_1	0.80	
	FVEALARSI	PB1	0.58	
	VMNILLQYL	GAD	0.57	
	FVANFSMEL	PB1	0.54	
	TTYQRTRAL	NP	0.47	
	GLADQLIHL	HIV_7	0.47	
Dengue-Virus 2 (DENV-2)	GLLTVCYVL	NS2B	0.88	Weiskopf *et* al.
	RLITVNPIV	E	0.88	
	IMAVGMVSI	NS2B	0.85	
	IILEFFLIV	NS4A	0.79	
	ALSELPETL	NS4A	0.61	
	YLPAIVREA	NS3	0.50	
	KLAEAIFKL	NS5	0.44	
	AAAWYLWEV	NS3	0.27	

Table A-4 continued.

Source	Epitope	Antigen	p-ANN-Hydro	Reference
Positive Control epitopes				
Human herpesvirus 5 (CMV)	ALFFFDIDL	CMV_5	0.84	Newell *et* al.
	LMNGQQIFL	CMV_18	0.81	
	RIFAELEGV	CMV_22	0.81	
	QMWQARLTV	CMV_21	0.78	
	NLVPMVATV	CMV_1	0.75	
	VLEETSVML	CMV-IE1	0.72	
	FLMEHTMPV	CMV_8	0.61	
	IIYTRNHEV	CMV_13	0.54	
	SLLSEFCRV	CMV_23	0.52	
	ILSPLTKGI	CMV_15	0.46	
	VLAELVKQI	CMV_2	0.24	
Human herpesvirus 4 (EBV)	GLCTLVAML	EBV_2	0.87	Newell *et* al.
	YVLDHLIVV	EBV_1	0.84	
	YLQQNWWTL	EBV_5	0.79	
	CLGGLLTMV	EBV_4	0.79	
	YLLEMLWRL	EBV_3	0.28	
Influenza-A (FluA)	FLDIWTYNA	Flu_4	0.85	
	NMLSTVLGV	Flu_14	0.78	
	LLIDGTASL	Flu_12	0.68	
	FMYSDFHFI	Flu_5	0.63	
	MMMGMFNML	Flu_13	0.48	
	GMFNMLSTV	Flu_7	0.29	
Human Immunodeficiency virus (HIV-1)	NVWATHACV	HIV_1	0.94	Newell *et* al.
	TLNAWVKVV	HIV_2	0.86	
	KLTPLCVTL	HIV_4	0.84	
	SLYNTVATL	HIV_5	0.61	
	ALVEMGHHA	HIV_8	0.44	
	ILKEPVHGV	HIV_9	0.38	
	LTFGWCFKL	HIV_6	0.36	
Mycobacterium *tuberculosis*	GLPVEYLQV	TB_1	0.85	
	KLIANNTRV	TB	0.79	
Plasmodium *falciparum*	YLNKIQNSL	CSP	0.79	
LCMV	YLVSIFLHL	LCMV	0.77	
Herpes simplex virus (HSV-1)	SLPITVYYA	HSV1/2	0.90	
RSV	KMLKEMGEV	RSV	0.29	
Self Antigens	ALWMRLLPL	pp-Insulin	0.83	
	YMCSFLFNL	EZH2	0.51	
	YMDGTMSQV	Tyrosinase	0.49	

Table A-5. Schematic of ConsB (top) and CN54 (bottom) 15mer peptide pools. Peptides were combined at 1mM/each peptide such that each peptide occurs in only two pools numbered 7872 –7994 for ConsB (top) or 7080.01-7080.121 for CN54 (indicated by 1–121, bottom). Yellow highlight indicates positive response to peptide pool. Green highlight indicates positive response to individual 15mer peptide, and red indicates negative response to individual 15mer peptide.

Consensus B Gag peptide pools

	A	B	C	D	E	F	G	H	I	J	K	L
M	7872	7873	7874	7875	7876	7877	7878	7879	7880	7881	7882	7883
N	7884	7885	7886	7887	7888	7889	7890	7891	7892	7893	7894	7895
O	7896	7897	7898	7899	7900	7901	7902	7903	7904	7905	7906	7907
P	7908	7909	7910	7911	7912	7913	7914	7915	7916	7917	7918	7919
R	7920	7921	7922	7923	7924	7925	7926	7927	7928	7929	7930	7931
S	7932	7933	7934	7935	7936	7937	7938	7939	7940	7941	7942	7943
T	7944	7945	7946	7947	7948	7949	7950	7951	7952	7953	7954	7955
U	7956	7957	7958	7959	7960	7961	7962	7963	7964	7965	7966	7967
V	7968	7969	7970	7971	7972	7973	7974	7975	7976	7977	7978	7979
W	7980	7981	7982	7983	7984	7985	7986	7987	7988	7989	7990	7991
X	7992	7993	7994									

CN54 Gag peptide pools

	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22
P1	1	2	3	4	5	6	7	8	9	10	11
P2	12	13	14	15	16	17	18	19	20	21	22
P3	23	24	25	26	27	28	29	30	31	32	33
P4	34	35	36	37	38	39	40	41	42	43	44
P5	45	46	47	48	49	50	51	52	53	54	55
P6	56	57	58	59	60	61	62	63	64	65	66
P7	67	68	69	70	71	72	73	74	75	76	77
P8	78	79	80	81	82	83	84	85	86	87	88
P9	89	90	91	92	93	94	95	96	97	98	99
P10	100	101	102	103	104	105	106	107	108	109	110
P11	111	112	113	114	115	116	117	118	119	120	121

Table A-6: Ranking comparison of all the predicted epitopes (Prevalidation and *in vivo* validation) used in this study. The predictions used are as follows: ANN-Hydro - ANN-hydrophobicity prediction model combined with normalized binding scores from prediction algorithms, IEDB-Bind -IEDB consensus binding tool, NetMHC-Bind - NetMHCpan binding tool, SYFPEITHI - SYFPEITHI epitope prediction tool, IEDB-Prot-IEDB recommended processing prediction, ANN-Prot - IEDB processing predictions using ANN. p-ANN-Hydro – Probability of immunogenicity assigned by the corresponding ANN-Hydro immunogenicity model.

Antigen	Epitope	IEDB.bind	Syfpeithi	NetMHC.bind	IEDB.prot	ANN.prot	IEDB*Si	p-ANN-Hydro
LCMV_glyco	FALISFLLL	1	10	1	3	1	1	0.62
LCMV_glyco	WLVTNGSYL	3	1	2	5	4	11	0.4
LCMV_glyco	LIDYNKAAL	45	12	39	68	77	32	0.8
LCMV_glyco	KAVYNFATC	8	9	10	39	53	5	0.77
LCMV_glyco	DEVINIVII	24	4	74	133	115	10	0.66
LCMV_NP	FQPQNGQFI	1	1	1	2	1	1	0.69
LCMV_NP	SEVSNVQRI	7	2	7	37	50	12	0.14
Ad.v.T.antigen	VNIRNCCYI	1	1	1	1	1	13	0.31
Ad.v.T.antigen	CSDGNCHLL	21	4	9	11	44	20	0.8
Flu_NP	ASNENMETM	1	1	1	1	2	1	0.87
Flu_NP	RLIQNSLTI	3	3	2	3	3	2	0.67
Flu_NP	GERQNATEI	18	2	36	100	103	8	0.42
Flu_NP	YRRVNGKWM	19	4	35	80	65	9	0.44
FluA-Neuraminidase	FCGVNSDTV	3	2	4	11	3	13	0.35
FluA-Neuraminidase	ITYKNSTWV	4	8	3	4	2	1	0.52
FluA-Neuraminidase	YRYGNGVWI	5	7	11	29	7	4	0.45
Consensus Gag	SQVTNSATI	1	1	1	1	1	7	0.2
Consensus Gag	AMQMLKETI	4	9	6	4	2	9	0.19
Consensus Gag	YSPTSILDI	6	19	4	11	3	11	0.39
Consensus Gag	RSLYNTVAT	3	32	5	45	29	1	0.82
ZM96 Gag	AMQMLKDTI	1	4	4	3	2	13	0.17
ZM96 Gag	YSPVSILDI	5	12	3	9	1	3	0.43
ZM96 Gag	RSLYNTVAT	4	28	5	46	27	2	0.77
97CN54 Gag	AMQILKDTI	1	4	4	3	2	13	0.16
97CN54 Gag	YSPTSILDI	5	19	2	9	3	15	0.36
97CN54 Gag	RSLFNTVAT	2	35	3	40	23	2	0.76
Melan-A	ALMDKSLHV	1	3	1	1	1	1	0.65

Melan-A	GILTVILGV	2	1	2	2	2	2	0.86
Melan-A	ILTVILGVL	7	5	5	4	5	8	0.9
Melan-A	AAGIGILTV	9	4	7	19	17	9	0.86
Wt-1	SLGEQQYSV	1	3	3	3	3	1	0.79
Wt-1	RMFPNAPYL	2	10	2	2	2	2	0.53
Wt-1	ALLPAVPSL	3	1	1	1	1	3	0.44
Wt-1	DLNALLPAV	6	2	8	16	27	4	0.9
Wt-1	VLDFAPPGA	7	31	7	15	13	8	0.87
Wt-1	KLGAAEASA	9	16	9	21	26	5	0.94
Wt-1	NLGATLKGV	12	4	10	12	14	12	0.76
Wt-1	CMTWNQMNL	13	27	11	7	9	13	0.69
Wt-1	RVPGVAPTL	25	17	21	11	11	19	0.81
gp100	RLMKQDFSV	1	21	1	2	2	1	0.83
gp100	MLGTHTMEV	2	15	2	3	1	2	0.65
gp100	KTWGQYWQV	5	62	3	4	3	5	0.5
gp100	YLEPGPVTA	16	20	19	34	25	11	0.93
TRAG-3	GLIQLVEGV	1	1	1	1	1	1	0.43
TRAG-3	HACWPAFTV	9	10	6	9	20	9	0.82
TRAG-3	SILLRDAGL	6	2	8	5	7	5	0.92
TRAG-3	ILLRDAGLV	3	3	2	4	4	3	0.7
TRAG-3	ALSKFPRQL	4	5	4	3	2	4	0.34
p53	RMPEAAPPV	1	1	8	2	2	1	0.51
p53	LLGRNSFEV	2	2	4	1	4	2	0.46
p53	VVPCEPPEV	13	14	21	14	11	9	0.77
p53	YQGSYGFRL	5	5	65	3	3	10	0.41
p53	KTCPVQLWV	14	15	19	34	25	11	0.77

Table A-7: Ranked list of top 20 predicted peptides for each of the Gag variant using the ANN-Hydro combined with normalized binding scores from predictions. SB) - Binding score, p-ANN-Hydro - probability of immunogenicity obtained by applying ANN-Hydro model to each peptide, (S) - Total score. This list was ranked based on total score S ranging from lowest score to the highest score within each section (I through IV) classified based on p-ANN-Hydro and SB (see section Application of ANN-Hydro).

ConsB Gag predictions				
Rank	Epitope	Binding (S_B)	p-ANN-Hydro	Total score (S)
1	**RSLYNTVAT**	**0.006**	**0.82**	**0.001**
2	ATPQDLNTM	0.033	0.77	0.008
3	QVSQNYPIV	0.049	0.83	0.008
4	RFAVNPGLL	0.025	0.66	0.008
5	RMYSPTSIL	0.039	0.72	0.011
6	KARVLAEAM	0.021	0.4	0.013
7	**SQVTNSATI**	**0**	**0.2**	**0**
8	SQVSQNYPI	0.004	0.12	0.004
9	**AMQMLKETI**	**0.008**	**0.19**	**0.006**
10	GWMTNNPPI	0.008	0.14	0.007
11	**YSPTSILDI**	**0.02**	**0.39**	**0.012**
12	RSLFGNDPS	0.023	0.27	0.017
13	ASVLSGGEL	0.022	0.07	0.02
14	KALGPAATL	0.036	0.36	0.023
15	AAMQMLKET	0.039	0.33	0.026
16	VQNANPDCK	0.053	0.84	0.009
17	SALSEGATP	0.054	0.84	0.009
18	LLVQNANPD	0.051	0.76	0.012
19	ASLRSLFGN	0.096	0.79	0.02
20	SLYNTVATL	0.061	0.64	0.022

ZM96 Gag predictions				
Rank	Epitope	Binding (S_B)	p-ANN-Hydro	Total score
1	MSQTNSVNI	0	0.64	0
2	**RSLYNTVAT**	**0.001**	**0.78**	**0**
3	**YSPVSILDI**	**0.004**	**0.43**	**0.002**
4	YMIKHLVWA	0.016	0.77	0.004
5	KVSQNYPIV	0.032	0.83	0.005
6	ATPQDLNTM	0.028	0.75	0.007
7	RMYSPVSIL	0.04	0.8	0.008
8	VQNANPDCK	0.048	0.83	0.008
9	LLVQNANPD	0.046	0.81	0.009
10	RFALNPGLL	0.027	0.67	0.009
11	KARVLAEAM	0.016	0.42	0.009
12	NFLQNRPEP	0.042	0.61	0.017
13	**AMQMLKDTI**	**0**	**0.18**	**0**
14	KSLFGSDPL	0	0.08	0
15	KALGPGATL	0.026	0.36	0.017
16	KIVRMYSPV	0.026	0.18	0.021
17	IMKQLQPAL	0.039	0.34	0.026
18	VKNWMTDTL	0.033	0.18	0.027
19	AWMTSNPPI	0.033	0.1	0.03
20	WMTSNPPIP	0.092	0.92	0.007

CN54 Gag predictions				
Rank	Epitope	Binding (SB)	p-ANN-Hydro	Total score
1	MSQTNSAIL	0.002	0.85	0.0003
2	**RSLFNTVAT**	**0.002**	**0.76**	**0.0004**
3	YMLKHLVWA	0.018	0.72	0.005
4	KVSQNYPIV	0.032	0.83	0.005
5	ATPQDLNTM	0.028	0.78	0.006
6	SALSEGATP	0.049	0.84	0.008
7	RFALNPGLL	0.027	0.7	0.008
8	VQNANPDCK	0.048	0.82	0.008
9	LLVQNANPD	0.046	0.78	0.01
10	RMYSPTSIL	0.034	0.7	0.01
11	NFLQNRPEP	0.042	0.65	0.015
12	SALQTGTEE	0.042	0.57	0.018
13	**AMQILKDTI**	**0**	**0.16**	**0**
14	KAKVLAEAM	0.012	0.35	0.008
15	**YSPTSILDI**	**0.015**	**0.36**	**0.01**
16	RALGPGASI	0.02	0.24	0.015
17	KSLFGNDPS	0.025	0.28	0.018
18	IMKQLQSAL	0.037	0.33	0.025
19	VKNWMTDTL	0.033	0.2	0.027
20	WMTSNPPVP	0.077	0.92	0.006

APPENDIX B

IMMUNOGENIC AND DYSFUNCTIONAL CD8+ T-CELLS IN HPV+HEAD AND NECK CANCER

B.1 HPV16 candidate CTL-epitope prediction

HPV16-candidate CTL epitopes were predicted using previously described prediction strategies developed by us (Chowell et al., 2015a; Krishna & Anderson, 2016a), except for the incorporation of immunogenicity scores. Briefly, we predicted HLA-class I restricted 9-mer and 10-mer candidate epitopes derived from the HPV16 proteins E2, E6, and E7 were predicted for the HLAs A*01:01, A*02:01, A*03:01, A*11:01, A*24:02, B*07:02, B*08:01, B*15:01, B*27:05, B*35:01, B*40:01, B*40:02, B*44:02 B*51:01, and B*57:01. The protein reference sequences for each of HPV16 proteins were obtained from Papillomavirus Episteme (PAVE) and were then entered into 5 different prediction algorithms; 3 MHC-binding : IEDB-consensus binding (Moutaftsi et al., 2006a), NetMHCpan binding (Hoof et al., 2009), Syfpeithi (Rammensee et al., 1999a) and 2 antigen-processing algorithms: IEDB-consensus processing, ANN processing (Rammensee et al., 1999b). The individual scores from each of the prediction algorithms were then normalized within the pool of predicted peptides after removal of poor-binders as described in (Chowell et al., 2015b; Krishna & Anderson, 2016b), and the mean normalized binding scores were used to re-rank the candidate peptides. Top 4-5 candidate peptides satisfying binding percentile scores >80% were chosen per antigen per HLA-allele for experimental testing. Predicted candidate HPV16-peptides, individual normalized and total binding percentile scores are listed in Table B1.

B.2 Ex vivo stimulation and epitope mapping of HPV+HNSCC PBMCs

PBMCs were obtained from stage III or stage IV HPV+HNSCC patients (MSSM cohort) as described previously (Parikh et al., 2014a). All human PBMCs were obtained using informed consent under clinical protocol HSM 10-00585. Patient characteristics are listed in (Parikh et al., 2014b). PBMCs were stimulated as previously described (Krishna & Anderson, 2016a). Briefly, HPV16-peptide pools shown in Fig.4-1 (and listed in Table B1) were designed to have equal representation of peptides predicted for each HLA-allele to prevent intra-pool peptide competition for binding to the same HLA (Table B1). All peptides (> 80% purity) were synthesized by Proimmune, UK. The HPV-peptide pools were created by mixing 7-8 HPV16 candidate peptides by antigen, each at a concentration of 1 mg/mL per peptide in sterile 1X PBS. For individual

peptides, each peptide was reconstituted at 1mg/mL in sterile 1X PBS. Frozen PBMCs were thawed rapidly and stimulated with 10µg/mL pre-mixed HIV-negative control peptide pool, HPV16-peptide pools or pre-mixed CEF-positive control pool (all from ProImmune, UK) in biological triplicates in a sterile 96-well U-bottomed plate(Costar, Washington DC, USA). Recombinant human IL-2 (20U/mL), human IL-7 (5ng/mL) and 1µg/mL of checkpoint blockade antibodies anti-PD1 (clone J105, eBosciences, USA) , anti-CTLA4 (clone 14D3, eBosciences, USA) were added and cells were rested for two hours at 37C prior to peptide stimulation. On day 5, half the media was removed and replaced with fresh IL-2 and peptide pool. On day 8, half the media was removed and fresh media, IL-2 and peptide was added to the cells and replated into a 96-well multiscreen elispot plate for Elispot detection. Same procedure was repeated for individual epitope mapping and deconvolution using selected candidate epitopes as per the patient's HLA-restriction (Fig. 4-3).

B.3 Elispot detection of IFNɣ secretion

Elispot detection assay was performed as previously described (Krishna & Anderson, 2016c). Briefly, sterile multiscreen Elispot plates, (Merck Millipore, Billerica, MA, USA) precoated overnight with 5µg/well anti-IFNg capture antibody (clone D1K, Mabtech, USA) in sterile 1X PBS. Eight days after stimulation, HPV+HNSCC PBMCs were subject to media change and IL-2, peptide (pools or individual) were added. Cells in each well were transferred to the Elispot plate and incubated at 37C 5% CO_2 incubator for 48 hours. Plates were washed with elispot buffer (PBS + 0.5% FBS) and incubated with 1µg/mL anti-IFNg secondary detection antibody (clone 7-B6-1, Mabtech, USA) for 2 hours at room temperature, washed and reincubated with 1µg/mL Streptavidin ALP conjugate for 1 hour at room temperature. The wells were washed again with elispot buffer and spots were developed by incubating for 8-10 minutes with detection buffer (33µL NBT, 16.5µL BCIP, in 100mM Tris-HCl pH 9, 1mM $MgCl_2$, 150mM NaCl). Plates were dried for 2 days and spots were read using the AID Elispot reader (Autoimmun Diagnostika GmbH, Germany). Average number of spot forming units for the triplicates were calculated for each test

peptide/pool and subtracted from background (either HIV-control peptide pool or PBS-DMSO controls).

B.4 RAPID-ELISA for E2, E6 and E7 seroreactivity in HPV+HNSCC patients

RAPID-ELISA was performed as described previously (Anderson et al., 2015). Briefly, patient sera were diluted 1:100 and blocked with *E. coli* lysate. Each antigen was expressed from template cDNA and captured onto 96-well plates coated with anti-GST Ab (GE Healthcare, Piscataway, NJ) in duplicates. Horseradish peroxidase (HRP) anti-human IgG Abs (Jackson ImmunoResearch Laboratories, West Grove, PA) were added at 1:10,000, and detected using Supersignal ELISA Femto Chemiluminescent substrate (Thermo Scientific). Luminescence was detected as relative light units (RLU) on a Glomax 96 Microplate Luminometer (Promega, Madison, WI) at 425 nm. To control for nonspecific and GST-specific antibodies, the ratio of RLU for individual HPV-specific Abs to the RLU for the control GST-antigen was measured.

B.5 Autologous APC generation from HPV+HNSCC patient PBMCs

Autologous CD40L-activated B-cell APCs were generated from specific HPV+HNSCC patients by incubating whole PBMCs with irradiated (32 Gy) K562-cell line expressing human CD40L (KCD40L) at a ratio of 4:1 (800,000 PBMCs to 200,000 irradiated KCD40Ls) in each well. The cells were maintained in B-cell media (BCM) consisting of IMDM (Gibco, USA), 10% heat inactivated human serum (Gemini Bio Products, CA, USA), Antibiotic-Antimycotic (Anti-Anti, Gibco, USA). BCM was supplemented with 10 ng/mL recombinant human IL-4 (R&D Systems, MN, USA), 2µg/mL Cyclosporin A (Sigma-Aldrich, CA, USA), 1X insulin transferrin supplement (ITES, Lonza, MD, USA). APCs were re-stimulated with fresh irradiated KCD40Ls on days 5 and 10, after washing with 1X PBS and expanding into a whole 24-well plate. After two weeks, APC purity was assessed by CD19+ CD86+ expressing cells by flow cytometry, and were generally used for T-cell stimulation after >90% purity. APCs were either restimulated upto 4 weeks or frozen and re-expanded as necessary.

B.6 HPV-CTL stimulation by autologous APCs

Antigen-specific T-cells were generated by stimulating HPV+HNSCC patient B-cell APCs by either peptide pulsing of specific HPV16-epitopes, or by transfecting whole antigen encoded in mammalian expression plasmid pCDNA3.2 (Invitrogen, CA, USA). Peptide pulsing of APCs were done under BCM 5% human serum, with recombinant IL-4. Transfection of APCs were done using the Lonza 4D Nucleofector, primary P3 buffer, program EO117 (Lonza, MD, USA) and incubated in BCM-10% human serum, IL-4 without any Anti-Anti. Twenty four hours later, on day 1, APCs were washed and incubated with thawed whole HPV+HNSCC PBMCs at a ratio of 1:2 (200,000 APCs : 400,000 PBMCs) in a 24-well plate in BCM supplemented with 20U/mL recombinant human IL-2 (R&D Systems, MN, USA), 5ng/mL IL-7 (R&D Systems, MN, USA). On day 5, partial media exchange was performed by replacing half the well with fresh B-cell media and IL-2. On day 10, fresh APCs were either peptide pulsed or transfected as described above in a new 24-well plate. On day 11, expanded T-cells were restimulated with peptide-pulsed, transfected APCs similar to day 1. T-cells were used for cytolytic assays or immunophenotyped after day 20.

B.7 HPV-CTL cytotoxicity assays

HLA-A*02:01 expressing HPV+HNSCC cell line SCC-104 was used for cytotoxicity assays. SCC-104 cells were plated at a density of 50,000 cells per well in a flat bottom 96-well sterile treated plate (Corning, USA). Twenty four hours later, cells were pre-labelled with 0.5μM CellTracker Green CMFDA (ThermoFisher Scientific, MA, USA) for one hour, washed thrice with sterile 1X PBS. HPV-specific CTLs generated by either peptide pulsing or transfected antigens were pooled by HPV-antigen, washed and resuspended in BCM supplemented with 20U/mL IL-2 along with 1μg/mL isotype IgG antibody, anti-PD1 antibody (eBosciences, USA), DMSO, and 1μM IDO-1 inhibitor Epacadostat (Selleck Chemicals, MA, USA) in various combinations as described. HPV-CTLs were added at ratio of 5:1 to SCC-104 cells and incubated for 48 hours at 37C, 5% CO_2. The cocultured cells were harvested by trypsinization, neutralized with media supernatant from

each well containing dead cells and centrifuged for 850g, 10 minutes. Cell pellets were washed twice with sterile 1X PBS, resuspended with 1mL 1X PBS, and 2uL Propidium Iodide (ThermoFisher Scientific, MA, USA) and cell death was assessed by flow cytometry. All samples were acquired with Attune flow cytometer (ThermoFisher Scientific, MA, USA) in blue-violet configuration and analyzed using Attune-software.

B.8 Flow cytometry staining for T-cell and tumor immunophenotyping

Cells were washed once in MACs buffer (containing 1X PBS, 1% BSA, 0.5mM EDTA), centrifuged at 550g, 5 minutes, and re-suspended in 200µL MACS buffer. Cells were stained in 100µL of staining buffer containing anti-CD137, conjugated with phycoerythrin (PE, clone 4B4-1; BD Biosciences, USA), anti-CD8-PC5 (clone B9.11; Beckman Coulter 1:100), anti-CD4 (clone SK3; BioLegend, 1:200), anti-CD14 (clone 63D3; BioLegend, 1:200) and anti-CD19 (clone HIB19; BioLegend,1:200), all conjugated to Fluorescein isothiocyanate (FITC) for exclusion gates, and either a combination of anti-PD1-Brilliant Violet 605 (BV605, clone EH12.2H7; BioLegend, 1:50) and anti-CD39-BV-405 (clone A1; BioLegend, 1:200) or anti-PD1-BV-605 and anti-TIM3-BV-405 (clone F38-2E2; BioLegend, 1:50) for 30 minutes on ice. PD-L1 staining on HPV+HNSCC and cervical cancer cell lines were done using 5µL PD-L1 antibody (clone MIH1, ThermoFisher Scientific, MA, USA) in 100µL MACS buffer. Samples were covered and incubated for 30 min on ice then washed twice in 1x PBS, and resuspended in 1mL 1x PBS prior to analysis.

B.9 Tetramer staining for T-cell immunophenotyping

The following HLA-A*02:01 HPV16 tetramers were obtained from NIH Tetramer Core Facility at Emory University: TLQDVSLEV E2(93-101), YICEEASVTV E2(138-147), ALQAIELQL E2(69-77), KLPQLCTEL E6(18-26), TIHDIILECV E6(29-38), FAFRDLCIV E6(52-60), YMLDLQPET E7(11-19), and YMLDLQPETT E7(11-20). Cells were washed (550g, 5 min) twice in MACS buffer with 5% human serum. After washing, cells were re-suspended in 100µL staining buffer (MACS buffer, with 5% human serum and 1mM Dasatanib (ThermoFisher Scientific, MA, USA). Each of the eight HLA-A*02:01 HPV16 tetramers (NIH Tetramer Core, Emory University, Atlanta, USA), all

conjugated with phycoerythrin (PE) was added to each respective sample at concentration of 1:100. Samples were incubated at room temperature for 30 minutes under dark. After incubation, cells were washed 2x in MACS buffer. Cells were stained in 100µL MACS buffer with anti-CD8-PC5, anti-CD4-FITC, anti-CD14-FITC and anti-CD19-FITC for exclusion gates, and either a combination of anti-PD1-BV605 and anti-CD39-BV-405 or anti-PD1-BV-605 and anti-TIM3-BV-405 for 30 minutes on ice. Samples were then washed twice in 1x PBS, and analyzed by flow cytometry. For flow cytometric analysis, all samples were acquired with Attune flow cytometer (ThermoFisher Scientific, MA, USA) and analyzed using Attune-software. Gates for expression of different markers and tetramers were determined based on flow minus one (FMO) samples for each color after doublet discrimination. Only samples with >50 CD8+Tetramer+ or CD8+CD137+ events were considered. Percentages from each of the gated population were used for the analysis.

B.10 Cell lines and immunoblotting experiments

Cervical cancer cell lines SiHa and Caski were obtained from ATCC (Manassas, VA, USA). HPV+HNSCC cell lines were obtained from the following sources: UPCI:SCC90 (SCC90) was obtained from ATCC (Manassas, USA), UM-SCC-47 (SCC47) and UM-SCC-104 (SCC104) from Merck Millipore (Billerica, MA, USA). 93-VU-147T (147T) cell line was a kind gift from Dr. Josephine Dorsman,VU Medical Center, Netherlands. All cell lines contained integrated HPV-16 DNA, and were maintained in the following media: Caski was maintained in RPMI-1640 (ATCC) with 10% heat inactivated fetal bovine serum (FBS), SiHa in Eagle's Minimum Essential Medium (EMEM, ATCC, USA) with 10% FBS, SCC90 and, SCC47 in Dulbecco's Modified Eagle's Medium (DMEM, ATCC,USA) with 10% FBS, SCC104 cells were maintained in Iscove's Modified Dulbecco's Medium (IMDM, Gibco, NY, USA) with 10% human serum. Cells were harvested by trypsinization (0.25% Trypsin, GE Healthcare, IL, USA), and resuspended in 1mL RIPA buffer (Invitrogen, CA, USA) containing a cocktail of protease inhibitors (Roche Diagnostics, IN, USA). Equal amounts of cell lysates were loaded on a 4-20 % SDS–polyacrylamide gel (Invitrogen, CA, USA) and transferred to a Polyvinylidene fluoride membrane (GE Healthcare, IL,

USA). The membrane was blocked with 5% nonfat dry milk in PBS-1% Tween (PBST) for 1 hour at room temperature. Primary antibodies and concentrations were as follows: GAPDH (Cell Signalling Technologies, MA, USA, clone 14C10, 1:2000), IDO-1 antibody (ThermoFisher Scientific, clone PA5-29819, 1:1000), anti-HPV16-E7 antibody (Fitzgerald industries, MA, USA, clone 10-7987, 1:1000). Visualization was done with Dura Western Blotting Kit (Thermo Scientific, USA) according to the manufacturer's instructions

B.11 RNASeq data alignment

RNA-Seq reads for each sample were quality checked using FastQC (version 0.10.1, Babraham bioinformatics, Babraham Institute, Cambridge, UK) and aligned to the human genome build 38 (GRCh38, GCF_000001405.33_GRCh38.p7_genomic.fna) primary assembly and HPV16 genome (GCF_000863945.1_ViralProj15505_genomic.fna) simultaneously using STAR (version 2.5.2B). After alignment, variants were discovered following GATK Best Practices workflow for RNAseq (https://gatkforums.broadinstitute.org/gatk/discussion/3892/the-gatk- best-practices- for-variant-calling- on-rnaseq- in-full- detail). Raw RNAseq reads were pre-processed by adding read groups, indexing, marking duplicates and sorting, Split'N'Trim, reassigning mapping quality and base recalibration.

B.12 HLA typing

For MSSM-cohort, HLA-typing was performed by Proimmune HLA-tissue typing services, UK. For HLA-calling from RNAseq data (TCGA and UM cohorts), PHLAT (Bai et al., 2014) was used to infer the HLA typing of the three major MHC class I (HLA-A, -B, -C) alleles (Bai, Ni, Cooper, Wei, & Fury, 2014a). The method employs a read mapping based selection of candidate allele followed by a likelihood based scoring over all pairwise combinations of selected alleles and infers the first four digits with a high accuracy (Bai, Ni, Cooper, Wei, & Fury, 2014b). For HLA-odds ratio calculations, HLA-allele typing from all 3 cohorts (MSSM, TCGA and UM) were combined resulting in 64 HPV-HNSCC (TCGA) and 77 HPV+HNSCCs.

B.13 HPV16 epitope-prediction from RNAseq data

HLA types obtained from PHLAT were used to predict the epitopes binding to patient-specific HLA alleles. Binding affinities were predicted using IEDB recommended algorithm from the Immune Epitope Database (IEDB) tool (Moutaftsi et al., 2006b; Vita et al., 2014). Reference fasta files for HPV protein sequences were used to predict peptide lengths of 8, 9, 10, 11 for each patient's allele and peptide combination. If the matching HLA allele of the patient did not exist in the current IEDB list, the closest allele was identified by keeping the first two digits the same and searching for the nearest available match for the third and fourth digit. To retain only high affinity binding epitopes with the patient-specific HLA alleles, epitopes with a binding affinity greater than 500 nM were not considered in downstream analyses.

B.14 RNASeq datasets and gene signature sources

Transcriptome data for HNSCC patient samples (n = 119) were obtained from TCGA (TCGA-cohort), and University of Michigan study (UM-cohort) (Cancer Genome Atlas Network, 2015; Zhang et al., 2016). In total, there were 34 and 18 HPV+HNSCC samples from TCGA-cohort and UM-cohort respectively. HPV-HNSCC dataset comprised of 18 tumors from UM-cohort and 49 tumors from TCGA dataset that were both HPV-negative by p16 status and HPV-FISH. For ssGSEA analysis, immune signatures, comprising of 509 genes were obtained from previous studies(Mandal et al., 2016a; Şenbabaoğlu et al., 2016a). Additional gene signatures were obtained as follows : Custom HPV gene sets were grouped into HPV (All 8 HPV genes), HPV. Onco (E6, E7) and HPV.Early (E2, E4, E5). TIGIT gene signature (50 genes) was obtained from Johnston et al (Rooney, Shukla, Wu, Getz, & Hacohen, 2015a), TIL.Treg (309 genes) and Exhaust gene sets (49 genes) were obtained from De Simone et al (De Simone et al., 2016), and CYT (GZMA, PRF) from Rooney et al (Rooney, Shukla, Wu, Getz, & Hacohen, 2015b).

B.15 ssGSEA analysis of HPV and immune gene signatures

All RNAseq analysis from HNSCC transcriptomes was performed on log transformed transcripts per million Log_2(TPM+1) from each sample, after exclusion of low expression genes (< 1 average

Log_2(TPM+1) across cohort). As opposed to GSEA, ssGSEA computes a gene set enrichment score for the relevant gene signatures on a per sample basis, without the need for a phenotypic pre-classification of the dataset. Thus, each patient's tumor can be analyzed by pre-defined immune signatures, and have been extensively validated in Şenbabaoğlu et al (Şenbabaoğlu et al., 2016b)(Şenbabaoğlu et al., 2016c)(Şenbabaoğlu et al., 2016b), Mandal et al (Mandal et al., 2016b). ssGSEA scores were computed for each tumor sample using the R package GSVA, and Z-transformed across the cohort prior to analysis. To assess impact of HPV-gene expression on immune signatures, a correlation matrix was built using the R-library Corrplot with the Z-transformed ssGSEA scores and was displayed by hierarchical clustering of correlations. Correlation values are displayed in Table S4. Individual gene expression analysis was performed by unsupervised hierarchical clustering methods and was used for heatmap analysis.

B.16 Statistical Analysis

Categorical variables, such as Elispot data, and Flow cytometric data were summarized as SFUs, and percentages. Continuous variables (RNAseq data) were presented with mean with standard error of mean (SEM). Unpaired T-test with Welch's correction was used for all categorical variable analyses, and for continuous variable analyses non-parametric Mann-Whitney's test was used. For heatmaps of T-cell frequencies and ssGSEA RNASeq analyses, Z-transformation to normalize the data across the cohorts. R statistical software V3.4.0 and Prism software (GraphPad Software) were used for data managements and statistical analyses. Significance levels were set at 0.1 (*), 0.01 (**) or 0.001 (***) for all tests as indicated.

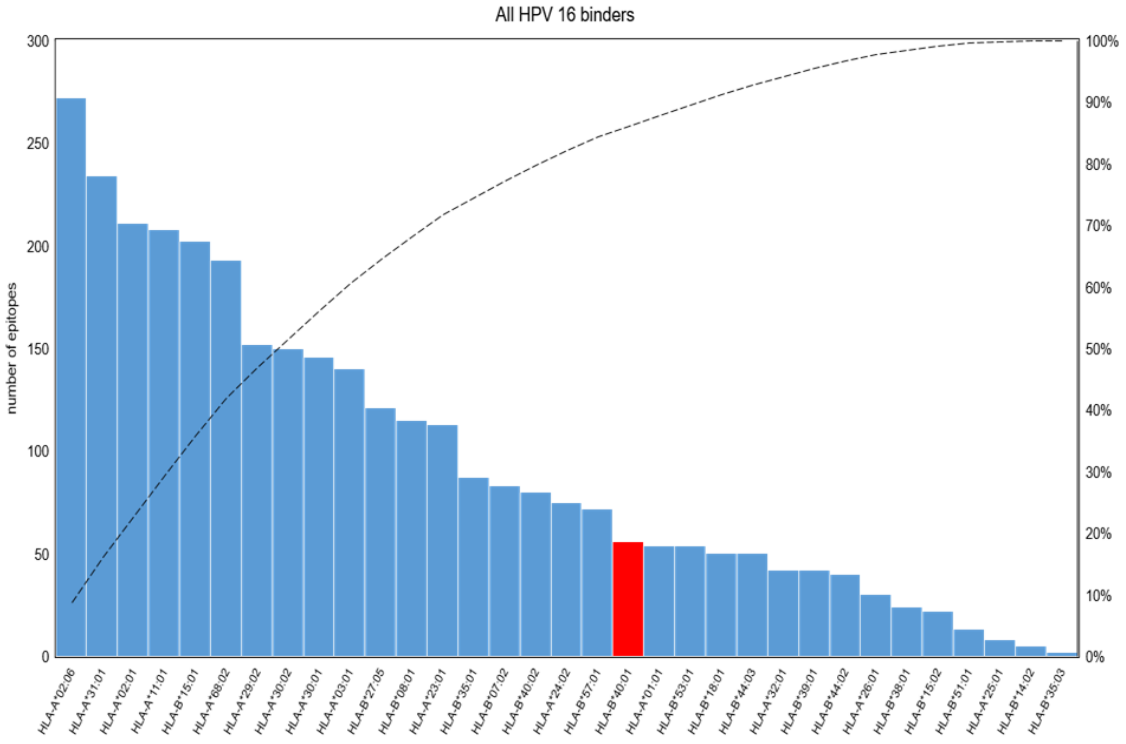

Figure. B1. **HLA-B*40:01 has low HPV16-predicted binding peptides**. Distribution of HPV16-predicted binding 9-11mer peptides from all 8 HPV16-antigens (IED-consensus IC50 < 500nM) for HLA-A, B alleles in TCGA+UM cohort (n=694 peptides), ranked by number of predicted peptides/ allele. HLA-B*40:01 is shown in red. Pareto line representing cumulative distribution of peptide frequencies is shown as dashed line.

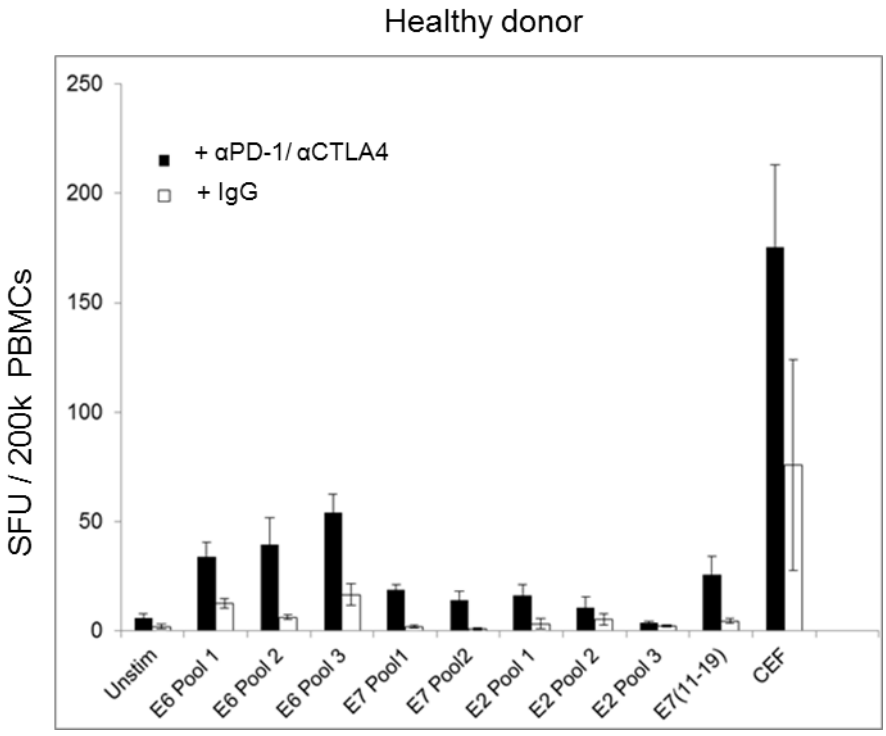

Figure. B2. PD-1/CTLA-4 blocking antibodies can enhance in vitro expansion of peripheral CTLs. Assessment of ex vivo PBMC stimulation protocol after 10 days with αPD1+αCTLA-4 blocking antibodies on day 1, compared to purified isotype IgG day 1 in healthy donor.

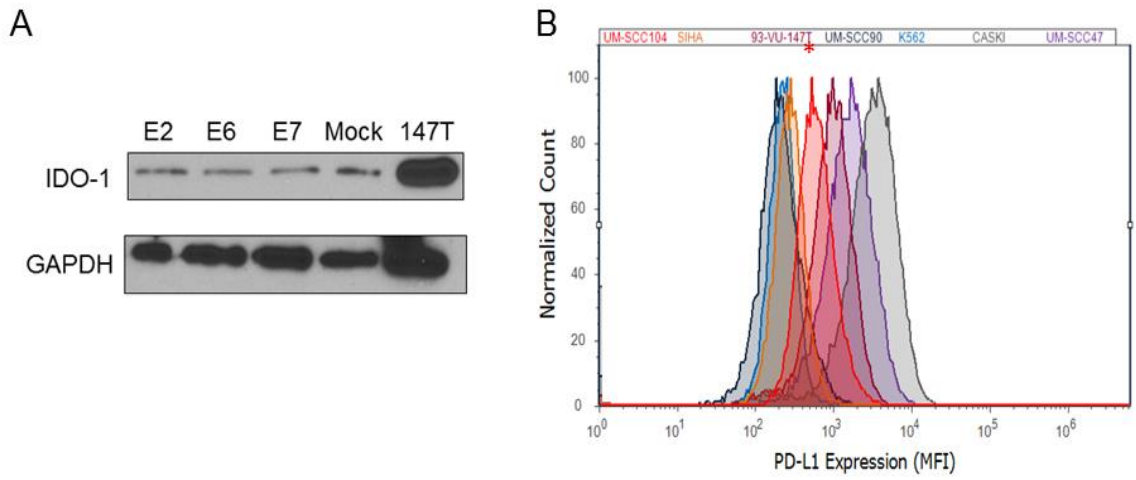

Figure. B3. (A) HPV16-E2, E6 and E7 genes do not alter IDO-1 levels in HEK-293 cells. HEK-293-LX cells were transfected using Lipofectamine 2000 with the indicated antigens and 48 hours later IDO-1 expression levels were measured by immunoblots. **(B) Cell surface PD-L1 protein expression analyzed by flow cytometry in the 6 HPV16+ cell lines**. Normalized mean fluorescence intensity of 10,000 events each indicated by different color. The erythroleukemic cell line K562 represents the negative control. Red asterisk indicates SCC-104 cell line used in cytolysis experiments.

Table B-1. Summary of all predicted candidate HPV16-E2, E6 and E7 epitopes used for experiments. Peptide pool information from each antigen, relevant information for the epitope's information from HPV T-cell antigen database (http://cvc.dfci.harvard.edu/hpv/). Highlighted epitopes represent immunogenic CTL-reactive epitopes identified in this study.

				Prediction scores, percentiles						
Position	**HLA**	**Sequence**	**Length**	**IEDB. bind**	**NetMHC. bind**	**Syfpeithi. bind**	**IEDB. Prot**	**ANN.P rot**	**Bind percentile**	**Total percentile**
E2(329-338)	HLA-B*15:01	KSAIVTLTY	9	0.8	78.78	15	0.96	0.81	82.62	82.54
E2(329-338)	HLA-B*57:01	KSAIVTLTY	9	0.55	116.59	14	1.15	0.64	81.78	81.96
E2(329-338)	**HLA-A*01:01**	**KSAIVTLTY**	**9**	**0.55**	**128.89**	**22**	**0.28**	**0.6**	**88.97**	**89.92**
E2(329-338)	**HLA-A*11:01**	**KSAIVTLTY**	**9**	**1.35**	**151.72**	**15**	**0.61**	**0.53**	**82.39**	**84.93**
E2(329-338)	**HLA-A*03:01**	**KSAIVTLTY**	**9**	**1.15**	**245.15**	**19**	**0.15**	**0.32**	**85.99**	**89.78**
E2(329-338)	**HLA-B*35:01**	**KSAIVTLTY**	**9**	**3.1**	**266.25**	**11**	**-0.09**	**0.28**	**78.12**	**85.46**
E2(93-101)	**HLA-A*02:01**	**TLQDVSLEV**	**9**	**0.5**	**9.82**	**24**	**0.2**	**0.29**	**90.88**	**92.64**
E2(138-147)	**HLA-A*02:01**	**YICEEASVTV**	**10**	**0.8**	**125.98**	**26**	**-0.73**	**-0.69**	**92.5**	**89.87**
E2(37-45)	HLA-A*11:01	RLECAIYYK	9	0.65	50.91	17	-0.34	-0.56	84.49	87.15
E2(37-45)	HLA-A*03:01	RLECAIYYK	9	0.4	152.21	24	-0.73	-1.04	90.81	87.45
E2(284-292)	**HLA-A*11:01**	**NTTPIVHLK**	**9**	**0.6**	**34.72**	**24**	**-0.12**	**-0.28**	**90.82**	**92.97**
E2(101-110)	HLA-A*24:02	VYLTAPTGCI	10	1.05	162.38	24	-1.06	-0.87	90.59	86.68
E2(207-215)	**HLA-B*07:02**	**SPEIIRQHL**	**9**	**0.7**	**73.94**	**21**	**0.13**	**0.12**	**88.06**	**91.91**
E2(303-311)	HLA-B*08:01	YRFKKHCTL	9	0.3	11.1	20	0.51	1.28	87.34	85.27
E2(303-311)	HLA-B*27:05	YRFKKHCTL	9	0.3	119.83	27	0.89	0.25	93.57	91.65
E2(303-311)	HLA-B*40:02	YRFKKHCTL	9	2.8	695.16	14	-1.2	-0.51	80.63	81.6
E2(147-155)	HLA-A*01:01	VVEGQVDYY	9	34.5	228.27	26	0.39	0.25	81.18	86.22
E2(69-77)	**HLA-A*02:01**	**ALQAIELQL**	**9**	**3.6**	**213.45**	**23**	**0.17**	**-0.27**	**88.8**	**91.59**
E2(310-318)	HLA-A*02:01	TLYTAVSST	9	3.6	656.75	21	-2.19	-2.17	86.69	74.57
E2(267-276)	**HLA-A*03:01**	**ILTAFNSSHK**	**10**	**0.35**	**25.3**	**27**	**-0.71**	**-0.55**	**93.62**	**91.19**
E2(267-276)	HLA-A*11:01	ILTAFNSSHK	10	0.95	75.17	17	-1.11	-1.02	84.37	82.15
E2(103-112)	HLA-A*11:01	LTAPTGCIKK	10	0.45	64.89	26	-0.59	-0.63	92.66	90.77
E2(103-112)	**HLA-A*03:01**	**LTAPTGCIKK**	**10**	**1.1**	**234.38**	**14**	**-1.3**	**-1.18**	**81.51**	**79.02**
E2(218-227)	HLA-B*07:02	HPAATHTKAV	10	0.35	16.59	20	0.01	0.18	87.32	91.71
E2(62-70)	HLA-B*08:01	LAVSKNKAL	9	1.3	232.92	27	-0.84	-0.38	93.16	91.08
E2(62-70)	HLA-B*35:01	LAVSKNKAL	9	4.3	666.47	13	-1.47	-0.84	79.24	78.36
E2(263-271)	HLA-B*35:01	DSAPILTAF	9	2.6	166.28	12	0.55	0.09	79.26	85.07
E2(94-102)	HLA-B*15:01	LQDVSLEVY	9	0.8	243.23	21	-0.3	0.3	87.91	90.42
E2(94-102)	HLA-A*01:01	LQDVSLEVY	9	41	259.3	27	-0.2	0.27	79.89	86.12
E2(102-110)	HLA-A*02:01	YLTAPTGCI	9	4.4	470.3	22	-1.22	-1.41	87.45	81.98
E2(191-199)	**HLA-A*02:01**	**QVILCPTSV**	**9**	**4.6**	**672.2**	**18**	**-1.86**	**-1.64**	**83.64**	**76.2**
E2(297-306)	HLA-A*02:01	TLKCLRYRFK	10	53.5	24400.31	9	-3.11	-3.07	42.84	40.94
E2(297-306)	**HLA-A*03:01**	**TLKCLRYRFK**	**10**	**0.45**	**85.36**	**21**	**-0.67**	**-0.62**	**88.14**	**87.78**
E2(297-306)	HLA-A*11:01	TLKCLRYRFK	10	1.05	86.84	16	-0.89	-0.63	83.43	84.04
E2(302-312)	HLA-A*24:02	RYRFKKHCTL	10	0.55	120.86	20	0.11	0.32	87.18	90.66
E2(302-312)	HLA-B*27:05	RYRFKKHCTL	10	10.6	122.95	15	0.73	0.32	79.32	83.46
E2(249-257)	**HLA-B*07:02**	**NPCHTTKLL**	**9**	**1.1**	**1802.85**	**21**	**-1.43**	**-1.55**	**86.74**	**80.14**
E2(163-171)	HLA-B*15:01	GIRTYFVQF	9	0.3	110.82	16	0.27	0.67	83.67	86.49
E2(163-171)	HLA-B*08:01	GIRTYFVQF	9	1.8	365.55	18	-0.35	0.15	84.79	88.95
E2(158-167)	HLA-B*15:01	YYVHEGIRTY	10	8.75	59.78	14	1.11	1.05	79.08	78.85
E2(158-167)	**HLA-B*35:01**	**YYVHEGIRTY**	**10**	**0.8**	**485.11**	**13**	**0.3**	**0.14**	**80.54**	**86.64**

Table B-1 continued

				Prediction scores, percentiles						
Position	HLA	Sequence	Length	IEDB. Bind	NetMHC. Bind	Syfpeithi. Bind	IEDB Prot	ANN. Prot	Bind Percentile	Total Percentile
E6(82-90)	HLA-A*24:02	EYRHYCYSL	9	1.3	573.3	19	-0.68	-0.86	89.14	87.52
E6(82-90)	HLA-B*08:01	EYRHYCYSL	9	1.3	588.7	18	-0.69	-0.4	88.12	88.7
E6(33-41)	HLA-A*03:01	IILECVYCK	9	0.7	265.9	23	-1.28	-1.21	93.59	86.41
E6(33-41)	**HLA-A*11:01**	**IILECVYCK**	**9**	**0.55**	**35.9**	**18**	**-0.41**	**-0.49**	**88.74**	**89.83**
E6(80-88)	HLA-A*01:01	ISEYRHYCY	9	0.25	107.9	27	0.82	1.33	97.89	90.34
E6(59-67)	HLA-B*15:01	IVYRDGNPY	9	0.7	46.1	18	0.88	0.76	88.69	86.85
E6(59-67)	HLA-B*35:01	IVYRDGNPY	9	0.9	36.9	12	0.98	1.07	82.57	81.55
E6(59-67)	HLA-B*57:01	IVYRDGNPY	9	2.75	6475.6	4	-1.27	-0.97	69.53	72.97
E6(18-26)	**HLA-A*02:01**	**KLPQLCTEL**	**9**	**1.8**	**227**	**24**	**-0.43**	**-0.1**	**94.26**	**94.61**
E6(15-24)	**HLA-B*07:02**	**RPRKLPQLCT**	**10**	**1.5**	**496.2**	**22**	**-2.03**	**-1.56**	**92.16**	**81.18**
E6(93-101)	**HLA-A*11:01**	**TTLEQQYNK**	**9**	**0.3**	**27.5**	**21**	**-0.44**	**-0.38**	**91.86**	**92.02**
E6(68-77)	**HLA-A*11:01**	**AVCDKCLKFY**	**10**	**2.5**	**533.8**	**15**	**0.03**	**0.42**	**84.73**	**89.21**
E6(87-95)	HLA-A*24:02	CYSLYGTTL	9	0.75	189.4	20	-0.21	0.03	90.59	93.56
E6(44-52)	HLA-B*15:01	LLRREVYDF	9	1.2	87.9	20	0.4	-0.11	90.51	92.44
E6(44-52)	HLA-B*08:01	LLRREVYDF	9	1.5	427	19	-0.29	-0.36	89.17	91.08
E6(65-74)	**HLA-B*07:02**	**NPYAVCDKCL**	**10**	**4.1**	**2833.6**	**21**	**-1.43**	**-1.53**	**88.7**	**81.61**
E6(29-38)	**HLA-A*02:01**	**TIHDIILECV**	**10**	**2.65**	**320.2**	**23**	**-1.2**	**-1.13**	**92.9**	**86.64**
E6(69-77)	HLA-A*01:01	VCDKCLKFY	9	0.95	4559.1	26	-0.98	-1.13	93.64	87.96
E6(67-76)	**HLA-B*35:01**	**YAVCDKCLKF**	**10**	**0.5**	**578.6**	**9**	**-0.3**	**-0.44**	**79.3**	**84.8**
E6(67-76)	HLA-B*15:01	YAVCDKCLKF	10	14.3	341.8	9	-0.07	-0.66	74.86	82.17
E6(119-126)	**HLA-B*07:02**	**CPEEKQRHL**	**9**	**2.5**	**893.3**	**20**	**-1.47**	**-1.91**	**89.54**	**80.44**
E6(37-47)	HLA-A*11:01	CVYCKQQLLR	10	1.9	348.8	24	-0.85	-1.01	94.14	89.25
E6(37-47)	HLA-A*03:01	CVYCKQQLLR	10	1.15	377.7	21	-0.88	-0.36	91.34	90.04
E6(127-135)	HLA-B*08:01	DKKQRFHNI	9	0.2	210.8	24	-1.04	-0.8	94.8	89.73
E6(52-60)	**HLA-A*02:01**	**FAFRDLCIV**	**9**	**2.3**	**115**	**20**	**-0.91**	**-1.04**	**90.13**	**86.48**
E6(106-116)	**HLA-A*03:01**	**LLIRCINCQK**	**10**	**0.9**	**79.4**	**29**	**-0.85**	**-1.19**	**99.71**	**91.88**
E6(106-116)	HLA-A*11:01	LLIRCINCQK	10	1.5	196.5	17	-1.25	-1.5	87.31	81.61
E6(82-90)	**HLA-B*35:01**	**YGTTLEQQY**	**9**	**3**	**395.1**	**11**	**-0.11**	**0.08**	**80.61**	**87.77**

Table B-1 continued

				Prediction scores, percentiles						
Position	**HLA**	**Sequence**	**Length**	**IEDB. Bind**	**NetMHC.Bind**	**Syfpeithi. Bind**	**IEDB. Prot**	**ANN. Prot**	**Bind Percentile**	**Total Percentile**
E7(15-23)	HLA-B*15:01	LQPETTDLY	9	0.8	233.2	22	0.03	0.34	91.09	93.7
E7(15-23)	HLA-A*01:01	LQPETTDLY	9	1.85	7312.4	17	-1.46	-1.5	81.2	77.08
E7(2-11)	HLA-A*01:01	HGDTPTLHEY	10	0.45	413.9	28	-0.11	-0.55	96.8	95.92
E7(2-11)	HLA-B*35:01	HGDTPTLHEY	10	2.05	3444.2	13	-1.03	-1.02	79.93	80.08
E7(2-11)	HLA-B*57:01	HGDTPTLHEY	10	31.35	11936	10	-1.57	-1.37	61.57	65.39
E7(82-90)	**HLA-A*02:01**	**LLMGTLGIV**	**9**	**0.5**	**16**	**29**	**-0.12**	**-0.31**	**98**	**97.6**
E7(82-90)	HLA-B*15:01	LLMGTLGIV	9	8	860.8	10	-1.85	-2.29	76.83	69.57
E7(51-60)	HLA-A*03:01	HYNIVTFCCK	10	2.2	4049.9	11	-2.21	-2.13	77.57	69.21
E7(51-60)	HLA-A*11:01	HYNIVTFCCK	10	3.3	1212.9	11	-1.69	-1.62	79.12	74.39
E7(88-97)	HLA-A*11:01	GIVCPICSQK	10	1.35	109	19	-1.04	-1.28	88.13	83.87
E7(88-97)	**HLA-A*03:01**	**GIVCPICSQK**	**10**	**1**	**156.6**	**24**	**-1.2**	**-1.54**	**92.98**	**85.04**
E7(5-13)	**HLA-B*07:02**	**TPTLHEYML**	**9**	**1.7**	**505.2**	**20**	**-0.97**	**-0.98**	**88.7**	**85.75**
E7(5-13)	HLA-B*08:01	TPTLHEYML	9	4.4	1881.9	17	-1.54	-1.58	84.01	78.11
E7(5-13)	HLA-B*35:01	TPTLHEYML	9	4.1	1139.5	20	-1.32	-0.58	87.47	85.25
E7(16-25)	HLA-B*44:02	QPETTDLYCY	10	4.05	647.1	12	-0.19	-0.11	80.2	87.47
E7(16-25)	HLA-B*35:01	QPETTDLYCY	10	0.6	653	22	-0.2	0.18	90.87	93.54
E7(7-15)	HLA-A*02:01	TLHEYMLDL	9	2.1	64.9	24	-0.28	-0.16	92.67	94.37
E7(7-15)	HLA-B*07:02	TLHEYMLDL	9	17	21631.5	12	-2.8	-2.6	61.73	55.34
E7(7-15)	HLA-B*08:01	TLHEYMLDL	9	6.3	3016.1	18	-1.95	-1.44	83.56	76.75
E7(7-15)	HLA-B*35:01	TLHEYMLDL	9	36	18002.2	11	-2.72	-2.79	56.88	51.96
E7(49-57)	HLA-B*07:02	RAHYNIVTF	9	3.8	1801.2	9	-0.6	-0.75	76.65	80.99
E7(49-57)	HLA-B*08:01	RAHYNIVTF	9	9	4103.7	13	-0.95	-1.12	77.17	78.33
E7(49-57)	HLA-B*15:01	RAHYNIVTF	9	0.5	64.7	11	0.85	0.75	80.82	82.48
E7(49-57)	HLA-B*57:01	RAHYNIVTF	9	0.6	233.8	14	0.29	0.25	83.53	88.48
E7(49-57)	HLA-B*35:01	RAHYNIVTF	9	0.8	283.5	11	0.21	0.35	80.58	86.61
E7(85-93)	**HLA-A*02:01**	**GTLGIVCPI**	**9**	**4.4**	**107.2**	**21**	**-1.1**	**-1.37**	**89.02**	**83.79**
E7(89-97)	**HLA-A*11:01**	**IVCPICSQK**	**9**	**0.45**	**66.9**	**21**	**-0.8**	**-0.67**	**90.36**	**88.74**
E7(89-97)	HLA-A*03:01	IVCPICSQK	9	0.55	182	31	-1.24	-1.42	99.78	89.46
E7(56-65)	HLA-A*24:02	TFCCKCDSTL	10	4.4	5243	17	-1.68	-1.92	81.75	74.76
E7(44-53)	HLA-B*35:01	QAEPDRAHY	9	4.9	462.8	12	0.13	-0.33	80.04	86.7
E7(11-19)	**HLA-A*02:01**	**YMLDLQPET**	**9**	**0.4**	**30.4**	**21**	**-0.59**	**-0.43**	**90.4**	**90.62**
E7(11-20)	**HLA-A*02:01**	**YMLDLQPETT**	**10**	**1.45**	**236.8**	**19**	**-1.62**	**-1.49**	**88.01**	**80.55**

Table B-1 continued (Cross reactive epitopes)

Patient	HLA	Epitope	Sequence	IC50 (ANN)	Predicted HLA
7007	A*68:01	E2(267-276)	ILTAFNSSHK	66.98	A*03:01/A*11:01
	A*68:01	E2(284-292)	NTTPIVHLK	4.94	A*11:01
	A*68:01	E2(297-306)	TLKCLRYRFK	182.15	A*03:01/A*11:01
	A*68:01	E2(37-45)	RLECAIYYK	711.12	A*03:01/A*11:01
	A*68:01	E6(106-115)	LLIRCINCQK	183.06	A*03:01/A*11:01
	A*68:01	E6(33-41)	IILECVYCK	803.58	A*03:01/A*11:01
	A*68:01	E6(37-46)	CVYCKQQLLR	72.47	A*03:01/A*11:01
	A*68:01	E6(93-101)	TTLEQQYNK	58.08	A*11:01
	A*68:01	E7(51-60)	HYNIVTFCCK	505.69	A*03:01/A*11:01
	A*68:01	E7(88-97)	GIVCPICSQK	469.86	A*03:01/A*11:01
	A*68:01	E7(89-97)	IVCPICSQK	72.24	A*03:01/A*11:01
7015	B*14:02	E2(147-155)	VVEGQVDYY	492.79	A*01:01
	A*32:01	E2(329-337)	KSAIVTLTY	50.93	A*01:01/A*03:01/A*11:01/B*15:01/B*35:01/B*57:01
	A*11:01	E6(68-77)	AVCDKCLKFY	219.88	A*11:01
7019	A*68:01	E2(103-112)	LTAPTGCIKK	85.73	A*03:01/A*11:01
	A*68:01	E2(284-292)	NTTPIVHLK	4.94	A*11:01
	A*68:01	E2(297-306)	TLKCLRYRFK	182.15	A*03:01/A*11:01
	A*68:01	E6(106-115)	LLIRCINCQK	183.06	A*03:01/A*11:01
	A*68:01	E6(37-46)	CVYCKQQLLR	72.47	A*03:01/A*11:01
	A*68:01	E6(93-101)	TTLEQQYNK	58.08	A*11:01
	A*32:01	E7(85-93)	GTLGIVCPI	12.13	A*02:01
7027	A*03:01	E6(93-101)	TTLEQQYNK	492.96	A*11:01
7030	A*03:01	E6(93-101)	TTLEQQYNK	492.96	A*11:01

Table B-2. Gene signatures for ssGSEA used in this study. HPV-Gene sets and immune signatures derived from for immunogenomic analyses

APM	HPV	HPV Onco	HPV Early	Exhaust	B cells	T cells	CD8 T cell	Th cells	Th1 cells	Th2 cells	Tfh Treg	TIGIT Sig	CYT	Cytotoxic	iDC	DC	Mphages	Mast cells	Eosinophil	Neutrophil	Tcm Cells	Tem Cells	Th Cells	Tgd Cells	Th17 Cells	Angiogenesis	NK Cells	pDC	NKCD56 dim	NKCD56 bright	aDC
HLA-A	E1	E6	E2	ADORA2A	ABCB4	BCL11B	ABT1	ANP32B	APBB2	ADCY1	ACAA2	CD2	GZMA	APBA2	ABCG2	CCL13	APOE	ABCC4	ABHD2	ALPL	AQP3	AKT3	B3GAT1	C1orf61	IL17A	CDH5	ADARB1	IL3RA	EDG8	DUSP4	CCL1
HLA-B	E2	E7	E4	BTLA	BACH2	BCL11A	CD2	AES	APBA2	AHI1	ACOX3	SLA2	PRF1	APOL3	BLVRB	CCL17	ATG7	ADCYAP1	ACACB	BST1	AIF1P	C7orf54	BLR1	CD160	IL17RA	ELTD1	AF107846		FLJ20699	FOXJ1	EBI3
HLA-C	E4		E5	BTNL2	BCL11A	CD28	APBA2	AIF2	ATP8A	ANK1	ACP5	SIRPG		CTSW	CARD9	CCL22	BCAT1	CALB2	C9orf156	CD83	ATM	CCR2	C18orf1	FEZ1	RORC	CLEC14A	ALDH1B1		GTF3C1	LPCAT4	INDO
B2M	E5			C10orf54	BLK	CD3D	ARHGAP8	BATF	BST2	BIRC5	ACSL4	CXCR6		DUSP2	CD1A	CD209	CCL7	C5AC4M8	CAT	CASP8	DDX17	CDKSR1	TARP			LDB2	APBB2		GZMB	MAOD	LAMP3
TAP1	E6			CD160	BLNK	CD3E	C12orf47	C13orf34	BTG3	CDC25C	ACTA2	CD247		GNLY	CD1B	HSD11B1	CD163	CMA1	CCR3	CREB5	CDC14A	EMSR1	CHGB	TRD			ECSCR	ATL2		IL21R	MPPED1	OAS3
TAP2	E7			CD200	CCR9	CD3G	C19orf6	CD28	CCL4	CDC7	ACTG2	P2RY10		GZMA	CD1C	NPR1	CD68	CPA3	CLC	CRISPLD2	CEP68	FLI1	CHI3L2	TRGV9			MYCT1	BCL2		KIR2DL3	MUC3B	
TAPBP	L1			CD200R1	CD19	CD6	C4orf15	DDX50	CD38	CENPF	ADAM10	ITK		GZMH	CD1E	PPIBP2	CD84	CTSG	CYSLTR2	CSF3R	CLUAP1	GOPD5	CXCL13			RHOJ	CDC5L		KIR2DS1	NIBP	
	L2			CD244	CD72	CD96	CAMLG	FAM111A	CD70	CXCR6	ADAT2	CXCR3		KLRB1	CLEC10A		CHI3L1	ELA2	EMR1	CYP4F3	CREBZF	CYLD	HEY1			WIF	FGF18		KIR2DS2	PLA2G6	
				CD27	COCH	GIMAP5	CD8A	FRYL	CMAH	DHFR	ADPRH	UBASH3A		KLRD1			CHIT1	GATA2	EPN2	DYSF	CYLD	MEFV	HIST1H4K			TIE1	FLT5		KIR2DS5	RRAD	
				CD274	CR2	ITM2A	CD8B	GOLGA8A	CSF2	EVI5	AHCYL2	GPR171		KLRF1	CSF1R		CLEC5A	HDC	GALC	FCAR	DOCK9	NFATC4	ICA1			KDR	FZR1		KIR3DL1	XCL1	
				CD276	DTNB	LCK	CD101P2NP	ICOS	CTLA4	GATA3	AKAP5	LOC100506776		KLRK1	CTNS		COL8A2	HPGD	GPR44	FCGR3B	FAM153B	PRKY	KCNK5			ESAM	GAGE2A		KIR3DL2		
				CD28	FCRL2	NCALD	DNAJB1	ITM2A	DGKI	GSTM4	AKIP1	SH2D1A		NKG7	F13A1		COLEC12	HPGDS	HES1	FLJ11151	FOXP1	TBC1D5	KIAA1324			PTPRB	IGFBP5		KIR3DL3		
				CD40	GLDC	PRKCQ	FLT3LG	LRBA	DOK5	HELLS	ARHGEF12	GPR174		RORA	FABP4		CXCL5	CTSK	KIT	HIST1H1C	FPR1	FYB	TBCD	MAF			GPR116	KANK2		KIR3DS1		
				CD40LG	GNG7	SH2D1A	GADD45A	NAP1L4	DPP4	IL26	ARHGEF4	SLAMF6		RUNX3			LINC01140	HRH4	FPRL1	HNRPH1	TRA				TEK	SPARCL1	LDB3		SPON2		
				CD44	HLA-DOB	IGKC	GZMM	NUP107	DUSP5	LAIR2	ARL6IP5	CD96		SIGIRR			CYBB	MAOB	IGSF2	GAS2	INPP4B	VL2	MKL2			S1PR1	PDLIM4		TMEPAI		
				CD48	HLA-DQA1	IGL	KLF9	PHF10	EGFL6	LIMA1	ARNTL2	TIGIT		WHAMMP3	GUCA1A		DNASE2B	MLPH	IL5RA	HIST1H2BC	KLF12		MYO6			A2M	PRX				
				CD70	IGHA1	TRAC	LEPROTL1	PPP2R5C	GGT1	MB	ATP13A3	IL12RB1		ZBTB16	HS3ST2		EMP1	MPO	KBTBD11	HPSE	LOC441155		MYO7A			ENG	MRC2		PSMD4		
				CD80	IGHG1	TRAT1	LIME1	RPA1	HBEGF	MICAL2	ATP2C1	LCK			LMAN2L		FDX1	MSAR2	KCNH2	IL8RA	MAP3K1		PASK			TEK	NCR1		JAM2		
				CD40LG	GNG7	SH2D1A	GADD45A	NAP1L4	DPP4	IL26	ARHGEF4	SLAMF6		RUNX3			CXCL5	LINC01140	HRH4	FPRL1	HNRPH1	TRA				S1PR1	PDLIM4				
				CD44	HLA-DOB	IGKC	GZMM	NUP107	DUSP5	LAIR2	ARL6IP5	CD96		SIGIRR			CYBB	MAOB	IGSF2	GAS2	INPP4B	VL2	MKL2			A2M	PRX				
				CD48	HLA-DQA1	TRA	KLF9	PHF10	EGFL6	LIMA1	ARNTL2	TIGIT		WHAMMP3	GUCA1A		DNASE2B	MLPH	IL5RA	HIST1H2BC	KLF12		MYO6			A2M	PRX				
				CD70	IGHA1	TRAC	LEPROTL1	PPP2R5C	GGT1	MB	ATP13A3	IL12RB1		ZBTB16	HS3ST2		EMP1	MPO	KBTBD11	HPSE	LOC441155		MYO7A			ENG	MRC2		PSMD4		
				CD80	IGHG1	TRAT1	LIME1	RPA1	HBEGF	MICAL2	ATP2C1	LCK			LMAN2L		FDX1	MSAR2	KCNH2	IL8RA	MAP3K1		PASK			TEK	NCR1		JAM2		
				ICOS	MEF2C		PRR5	UBE2L3	LRRN3	SLC39A14	C17orf66	PTPRC			PPARG		MARCO	SCG2	RRP12	MME	PCM1		SH3TC1			PECAM1	SGMS1				
				ICOSLG	MICAL3		RBM3	YME1L1	LTA	SMAD2	C5orf63	PTPN22			PREP		ME1	SIGLEC6	SIAH1	PDE4B	PCNX		SIRPG			CALCRL	SLC30A5				
				IDO1	MS4A1		SF1		SGCB	SNRPD1	CABLES1	GIMAP5			RAPGAP		MS4A4A	SLC18A2	SMPD3	S100A12	PDXDC2		SLC7A10			CLEC3B	SMEK1				
				IDO2	QSOX1	SPIB	SFRS7		SYNGR3	WDHD1	CACNB2	GIMAP1-GIMAP5			SLC26A6		MSR1	SLC24A3	SYNJ1	SIGLEC5	PHC3		SMAD1			PLVAP	SPN				
				KIR3DL1	PNOC	TGL1A	SLC16A7		ZBTB32		CADM1	SIT1			SLC7A8		PCOLCE2	TAL1	TGIF1	SLC22A4	POLR2J2		STBSA1			RGS5	TBXA2R				
				LAG3	QRSL1		TBCC				CALM3	PYHIN1			SYT17		PTGDS	TPSAB1	THBS1	SLC25A37	PSPC1		STK39			LRRC32	TCTN2				
				LAIR1	SCN3A		THUMPD1				CARD16	ARHGAP15			TACSTD2		RAI14	TPSB2	THBS4	TECPR2	REPS1		THADA			EBF1	TINAGL1				
				LGALS9	SLC15A2		TMC8				CARD17	SASH3			TMTSF4		SCARB2	VWA5A	TIPARP	INFRSF10C	RPP38		TOX			ADCY4	TRPV6				
				NRP1	SPIB		TSC22D3				CASP1	GIMAP7			VASH1		SCG5		TKTL1	VNN3	SLC7A6		TSHR			ACVRL1	XCL1				
				PD1LG2	TCL1A		VAMP2				CASQ1	GIMAP4					SGMS1				SNRPN		ZNF764			GPR124	XCL2				
				PDCD1	TNFRSF17		ZEB1				CCNB2	SLAMF1					SULT1C2				ST3GAL1					APLNR	ZNF205				
				TIGIT			ZFP36L2				CCR8	CRTAM									STX16					TM4SF18	ZNF528				
				TIMD2			ZNFZ2				CD177	PDCD1									TIMM8A					GNG11	ZNF747				
				TNFRSF14			ZNF609				CD27	GMNP1									TRAF3IP3					CNRIP1					
				TNFRSF18			ZNF91				CD274	IL2RG									TXK					ZNF423					
				TNFRSF25							CD7	CD6									TXLNGY					GIMAP8					
				TNFRSF4							CD4CA2	CCR5									USP9Y					PDGFD					
				TNFRSF8							CDH24	LOC100652927														ITGA9					
				TNFRSF9							CDK6	TESPA1														EDNRB					
				TNFSF14							CEACAM1	IL10RA																			
				TNFSF15							CENPM	GZMH																			
				TNFSF18							CEP55	PCED1B-AS1																			
				TNFSF4							CGA	CTLA4																			
				TNFSF9							CHRNA6	SCML4																			
				VTCN1							CHST2	CCL5																			
				NT5E							CHST7	CD48																			
				ENTPD1							CIT	ITGAL																			
											CLNK	PTPRCAP																			

Table B-3. Clustered correlation matrix values for each gene-set correlations

	APM	HPV	HPV Onco	HPV Early	Exhaust	B cells	T Cells	CD8. TCell	Th Cells	Th1 Cells	Th2 Cells	TIL Treg	TIGIT Sig	CYT	Cyto toxic	iDC	DC	Mphage	Mast cells	Eosino phil	Neutro phil	Tcm Cells	Tem Cells	Tfh Cells	Tgd Cells	Th17 Cells	Angio genesi	NK Cells	NKCD 56 dim	NKCD5 6 bright	aDC
APM	1.00	-0.05	-0.09	0.01	0.27	0.20	0.24	0.29	0.13	-0.07	0.09	0.18	0.23	0.32	0.28	0.13	0.28	0.33	0.11	-0.16	0.12	0.28	0.18	0.26	0.08	-0.02	0.00	0.19	0.30	0.08	-0.03
HPV	-0.05	1.00	0.97	0.97	0.32	0.37	0.50	0.32	0.39	-0.09	0.28	0.37	0.49	0.46	0.51	-0.04	0.29	-0.13	-0.20	-0.26	-0.48	0.21	0.27	0.46	-0.31	0.19	-0.26	-0.03	-0.13	-0.29	0.13
HPV.Onco	-0.09	0.97	1.00	0.89	0.31	0.35	0.49	0.30	0.38	-0.07	0.24	0.38	0.49	0.43	0.49	-0.07	0.33	-0.13	-0.17	-0.24	-0.43	0.18	0.19	0.46	-0.32	0.17	-0.28	-0.10	-0.16	-0.30	0.18
HPV.Early	0.01	0.97	0.89	1.00	0.30	0.35	0.47	0.32	0.39	-0.12	0.33	0.35	0.47	0.46	0.49	-0.01	0.23	-0.12	-0.23	-0.27	-0.51	0.22	0.33	0.42	-0.27	0.19	-0.24	0.05	-0.10	-0.29	0.08
Exhaust	0.27	0.32	0.31	0.30	1.00	0.76	0.93	0.89	0.70	0.46	0.23	0.82	0.94	0.87	0.92	0.43	0.67	0.56	0.14	-0.28	-0.18	0.63	0.55	0.82	0.04	0.17	0.36	0.33	0.60	0.02	0.56
B.cells	0.20	0.37	0.35	0.35	0.76	1.00	0.81	0.70	0.63	0.28	0.15	0.71	0.83	0.65	0.72	0.17	0.57	0.38	0.15	-0.34	-0.17	0.65	0.55	0.79	-0.06	0.08	0.28	0.13	0.32	-0.08	0.39
T.cells	0.24	0.50	0.49	0.47	0.93	0.81	1.00	0.89	0.80	0.35	0.35	0.85	0.98	0.89	0.92	0.21	0.66	0.50	0.16	-0.30	-0.30	0.65	0.53	0.86	0.02	0.17	0.32	0.29	0.48	-0.10	0.49
CD8.T.cell	0.29	0.32	0.30	0.32	0.89	0.70	0.89	1.00	0.69	0.33	0.15	0.71	0.87	0.82	0.87	0.32	0.57	0.53	0.28	-0.26	-0.22	0.59	0.51	0.83	0.10	0.17	0.34	0.38	0.62	0.18	0.45
Th.cells	0.13	0.39	0.38	0.39	0.70	0.63	0.80	0.69	1.00	0.29	0.63	0.77	0.74	0.65	0.67	-0.07	0.56	0.43	0.10	-0.14	-0.31	0.75	0.35	0.63	0.10	0.06	0.38	0.37	0.18	-0.27	0.39
Th1.Cells	-0.07	-0.09	-0.07	-0.12	0.46	0.28	0.35	0.33	0.29	1.00	0.00	0.48	0.41	0.40	0.42	0.12	0.16	0.33	0.09	0.01	0.04	0.22	-0.13	0.25	0.32	-0.01	0.18	-0.07	0.15	-0.02	0.68
Th2.Cells	0.09	0.28	0.24	0.33	0.23	0.15	0.35	0.15	0.63	0.00	1.00	0.44	0.26	0.26	0.20	-0.15	0.23	0.30	-0.14	0.09	-0.28	0.45	0.26	0.20	-0.05	-0.10	0.21	0.39	-0.09	-0.65	-0.01
Tl.Treg	0.18	0.37	0.38	0.35	0.82	0.71	0.85	0.71	0.77	0.48	0.44	1.00	0.83	0.74	0.77	0.23	0.56	0.47	0.24	-0.27	-0.08	0.62	0.34	0.78	0.20	0.04	0.36	0.28	0.35	-0.31	0.61
TIGIT.Sig	0.23	0.49	0.49	0.47	0.94	0.83	0.98	0.87	0.74	0.41	0.26	0.83	1.00	0.91	0.95	0.25	0.70	0.47	0.13	-0.36	-0.28	0.60	0.50	0.84	-0.03	0.17	0.26	0.19	0.48	-0.07	0.54
CYT	0.32	0.46	0.43	0.46	0.87	0.65	0.89	0.82	0.65	0.40	0.26	0.74	0.91	1.00	0.96	0.25	0.64	0.35	0.09	-0.32	-0.26	0.49	0.41	0.68	0.06	0.23	0.14	0.19	0.58	0.00	0.47
Cytotoxic	0.28	0.51	0.49	0.49	0.92	0.72	0.92	0.87	0.67	0.42	0.20	0.77	0.95	0.96	1.00	0.29	0.67	0.38	0.11	-0.34	-0.28	0.51	0.45	0.77	-0.03	0.20	0.15	0.19	0.52	0.06	0.55
iDC	0.13	-0.04	-0.07	-0.01	0.43	0.17	0.21	0.32	-0.07	0.12	-0.15	0.23	0.25	0.25	0.29	1.00	0.08	0.39	0.08	-0.12	0.28	-0.08	0.14	0.27	0.02	0.33	0.31	0.28	0.56	0.17	0.33
DC	0.28	0.29	0.33	0.23	0.67	0.57	0.66	0.57	0.56	0.16	0.23	0.56	0.70	0.64	0.67	0.08	1.00	0.29	0.04	-0.28	-0.15	0.40	0.30	0.53	-0.29	0.15	0.14	0.28	0.28	-0.07	0.29
Mphages	0.33	-0.13	-0.13	-0.12	0.56	0.38	0.50	0.53	0.43	0.33	0.30	0.47	0.47	0.35	0.38	0.39	0.29	1.00	0.33	0.07	0.09	0.40	0.29	0.42	0.09	0.01	0.71	0.44	0.58	0.05	0.31
Mast.cells	0.11	-0.20	-0.17	-0.23	0.14	0.15	0.16	0.28	0.10	0.09	-0.14	0.24	0.13	0.09	0.11	0.08	0.04	0.33	1.00	0.19	0.42	0.15	-0.12	0.30	0.30	-0.05	0.47	0.24	0.31	0.14	0.16
Eosinophil	-0.16	-0.26	-0.24	-0.27	-0.28	-0.34	-0.30	-0.26	-0.14	0.01	0.09	-0.27	-0.36	-0.32	-0.34	-0.12	-0.28	0.07	0.19	1.00	0.30	-0.10	-0.07	-0.25	0.19	0.20	0.22	0.16	-0.18	0.12	-0.14
Neutrophil	0.12	-0.48	-0.43	-0.51	-0.18	-0.17	-0.30	-0.22	-0.31	0.04	-0.28	-0.08	-0.28	-0.26	-0.28	0.28	-0.15	0.09	0.42	0.30	1.00	-0.14	-0.37	-0.17	0.28	0.06	0.28	-0.05	0.10	0.20	0.03
Tcm.Cells	0.28	0.21	0.18	0.22	0.63	0.65	0.65	0.59	0.75	0.22	0.45	0.62	0.60	0.49	0.51	-0.08	0.40	0.40	0.15	-0.10	-0.14	1.00	0.47	0.65	0.23	-0.15	0.31	0.43	0.14	-0.24	0.13
Tem.Cells	0.18	0.27	0.19	0.33	0.55	0.55	0.53	0.51	0.35	-0.13	0.26	0.34	0.50	0.41	0.45	0.30	0.30	0.29	-0.12	-0.25	-0.37	0.47	1.00	0.60	-0.10	0.18	0.16	0.53	0.37	-0.03	0.00
Tfh.Cells	0.26	0.46	0.46	0.42	0.82	0.79	0.86	0.83	0.63	0.25	0.20	0.78	0.84	0.68	0.77	0.27	0.53	0.42	0.30	0.19	0.28	0.65	0.60	1.00	0.05	0.14	0.26	0.38	0.37	-0.12	0.40
Tgd.Cells	0.08	-0.31	-0.32	-0.27	0.04	-0.06	0.02	0.10	0.10	0.32	-0.05	0.20	-0.03	0.06	-0.03	0.02	-0.29	0.09	0.30	0.19	0.28	0.23	-0.10	0.05	1.00	-0.08	0.21	0.23	0.10	0.12	0.09
Th17.Cells	-0.02	0.19	0.17	0.19	0.17	0.08	0.17	0.17	0.06	-0.01	-0.10	0.04	0.17	0.23	0.20	0.33	0.15	0.01	-0.05	0.20	0.06	-0.15	0.18	0.14	-0.08	1.00	0.08	0.13	0.26	0.13	0.22
Angiogenesis	0.00	-0.26	-0.28	-0.24	0.36	0.28	0.32	0.34	0.38	0.18	0.21	0.36	0.26	0.14	0.15	0.31	0.14	0.71	0.47	0.22	0.28	0.31	0.16	0.26	0.21	0.08	1.00	0.46	0.44	0.01	0.33
NK.Cells	0.19	-0.03	-0.10	0.05	0.33	0.13	0.29	0.38	0.37	-0.07	0.39	0.28	0.19	0.19	0.19	0.28	-0.09	0.44	0.24	0.16	-0.05	0.43	0.53	0.38	0.23	0.13	0.46	1.00	0.37	-0.13	-0.02
NKCD56dim	0.30	-0.13	-0.16	-0.10	0.60	0.32	0.48	0.62	0.18	0.15	-0.09	0.35	0.48	0.58	0.52	0.56	0.28	0.58	0.31	-0.18	0.10	0.14	0.37	0.37	0.10	0.26	0.44	0.37	1.00	0.35	0.31
NKCD56bright	0.08	-0.29	-0.30	-0.29	0.02	-0.08	-0.10	0.18	-0.27	-0.02	-0.65	-0.31	-0.07	0.00	0.06	0.17	-0.07	0.05	0.14	0.12	0.20	-0.24	-0.03	-0.12	0.12	0.13	0.01	-0.13	0.35	1.00	-0.07
aDC	-0.03	0.13	0.18	0.08	0.56	0.39	0.49	0.45	0.39	0.68	-0.01	0.61	0.54	0.47	0.55	0.33	0.29	0.31	0.16	-0.14	0.03	0.13	0.00	0.40	0.09	0.22	0.33	-0.02	0.31	-0.07	1.00

APPENDIX C

COPYRIGHT PERMISSIONS

Figure 1-2 Copyright permission

Sri Krishna <Srikrishna@asu.edu> Thu, Sep 28, 2017 at 11:36 AM

Dear Sri Krishna

We hereby grant you permission to reproduce the material detailed below at no charge **in your thesis, in print and on ASU Electronic Theses and Dissertations** and subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.

2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"This article was published in Publication title, Vol number, Author(s), Title of article, Page Nos, Copyright Elsevier (or appropriate Society name) (Year)."

3. Your thesis may be submitted to your institution in either print or electronic form.

4. Reproduction of this material is confined to the purpose for which permission is hereby given.

5. This permission is granted for non-exclusive world **English** rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form other than as specified above. Should you have a specific electronic project in mind please reapply for permission.

6. This includes permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission

Yours sincerely

Jennifer Jones,
Permissions Specialist